Structural base for the transfer of GPI-anchored proteins into fungal cell walls

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Chemie
der Philipps-Universität Marburg

vorgelegt von

Marian Samuel Vogt
M.Sc. Biologie

aus Marburg

Marburg/Lahn im September 2019

Vom Fachbereich Chemie

der Philipps-Universität Marburg (Hochschulkennziffer: 1180) als Dissertation am _____ angenommen.

Erstgutachter: Prof. Dr. Lars-Oliver Essen (FB Chemie)

Zweitgutachter: Prof. Dr. Hans-Ulrich Mösch (FB Biologie)

Vorsitzender d. Prüfungskommission: Prof. Dr. Norbert Hampp (FB Chemie)

Tag der mündlichen Prüfung am _____.

Meiner Familie

Eidesstattliche Erklärung
Statutory Declaraction

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel
I declare that I have authored the thesis

Structural base for the transfer of GPI-anchored proteins into fungal cell walls

selbstständig und ohne fremde Hilfe verfasst, nicht andere als die in ihr angegebenen Quellen oder Hilfsmittel benutzt, sowie alle vollständig oder sinngemäß übernommenen Zitate als solche gekennzeichnet habe.
myself without external assistance. I did not use other than the indicated references and resources and I have appropriately identified all quotations as borrowed in full or in paraphrased form.

Diese Dissertation wurde in der vorliegenden oder einer ähnlichen Form noch bei keiner anderen in- oder ausländischen Hochschule anlässlich eines Promotionsgesuchs oder zu anderen Prüfungszwecken eingereicht.
The dissertation in hand in the current or a similar form has not been handed in to any other domestic or foreign university and did not serve for a doctoral application or any other examination purposes prior to this date.

Marburg/Lahn, den 03. September 2019
Marburg/Lahn, September 3rd, 2019

_____

Marian Samuel Vogt

Verlauf des wissenschaftlichen Werdegangs

Curriculum vitae about the academic education

not part of the electronic resource

Zusammenfassung

Pilze, wie der einzellige Modellorganismus *Saccharomyces cerevisiae*, besitzen eine Zellwand aus Polysacchariden und Proteinen, die für lebensfähige und gesunde Zellen essentiell ist. Während die Synthese ihrer Komponenten entlang des sekretorischen Weges und an der Plasmamembran erfolgt, wird die korrekte Verarbeitung auf der Außenseite der Plasmamembran bewerkstelligt. Dies wird durch sogenannte Glykosidhydrolasen (GH), realisiert, welche die Spaltung und Neuverknüpfung von glykosidischen Bindungen katalysieren. In *S. cerevisiae* und dem Humanpathogen *Candida albicans* wurde gezeigt, dass Mitglieder der GH76-Familie (*Dfg5*-Subfamilie) für den Einbau von GPI-verankerten Proteinen in die Zellwand verantwortlich sind, ein Vorgang, der für Hefen überlebenswichtig ist. Obwohl bakterielle Homologe dieser Klasse bereits beschrieben wurden, sind die zugrunde liegenden Mechanismen der pilzlichen Vertreter, trotz ihres außergewöhnlichen Potenzials als Wirkstoff Zielstruktur, noch immer unbekannt.

Um diese Lücke zu schließen, wurde die GH76-Familie zunächst einer phylogenetischen Analyse unterzogen, die Einblicke in ihren multifunktionalen Charakter gewährte. Die genaue Rolle der *Dfg5*-Proteine konnte durch Struktur- und Funktionsanalysen des Orthologs *Ct*Dfg5 aus dem thermophilen Schimmelpilz *Chaetomium thermophilum* erklärt werden. Seine mit atomarer Auflösung bestimmte Kristallstruktur zeigte, dass die Faltung und das aktive Zentrum zwischen Pilz- und Bakterienhomologen konserviert ist, jedoch war es nicht möglich die bekannte α1,6-Mannanaseaktivität *in vitro* zu zeigen. Stattdessen konnte die GPI-Glykanstruktur (Manα1,2-Manα1,6-Manα1,4-GlcN) innerhalb der Bindetasche von *Ct*Dfg5 durch hochmolares Zuckerfragmentscreening assembliert werden. Dies ermöglichte nicht nur einen detaillierten Blick auf das tatsächliche Substrat der *Dfg5*-Proteine, sondern auch erste experimentell abgeleitete Erkenntnisse über die dreidimensionale Architektur des GPI-Ankerglykans. Zusammen mit der Struktur eines potentiellen Akzeptormoleküls konnte der von *Dfg5*-Proteinen katalysierte Lipid-zur-Zellwand-Transfermechanismus abgeleitet werden. Darüber hinaus ermöglichten die strukturellen Einblicke in die Substratkoordinierung eine mögliche Lenkungsfunktion verschiedener GPI-Modifikationen, mit deren Hilfe die endgültige Lokalisation von GPI-verankerten Proteinen kodiert wird. Außerdem konnte durch Docking eine Leitstruktur (FP-1) identifiziert werden, das im aktiven Zentrum von *Ct*Dfg5 bindet. FP-1 zeigt spezifische Effekte bei millimolaren Konzentrationen in Bezug auf die Lebensfähigkeit von *S. cerevisiae*. Schließlich wurde, durch die strukturelle Charakterisierung eines pilzlichen Homologs aus einer weiteren GH76-Unterfamilie, α1,6-Mannobiose als zentrales Element von GH76-Substraten identifiziert.

Summary

Fungi, such as the unicellular model organism *Saccharomyces cerevisiae*, possess a thick cell wall composed of polysaccharides and proteins, which is essential for viable and healthy cells. While the mere synthesis of its components happens along the secretory pathway and at the plasma membrane, the correct processing is established on the exterior side of the plasma membrane. This is realized by a set of enzymes, called glycoside hydrolases (GH), which act on the hydrolysis and rearrangement of glycosidic bonds. In *S. cerevisiae* and the human pathogen *Candida albicans*, it has been shown that members of the GH76 family (*Dfg5*-subfamily) carry out the incorporation of GPI-anchored proteins into the cell wall, which is essential for these organisms. Although bacterial homologs of that class were already described, our understanding of the fungal counterparts and their underlying mechanism with its exceptional potential as a drug target still lack behind.

In order to fill this gap, the GH76 family has been subjected initially to phylogenetic analysis providing insights into its multifunctional character with up to ten different subfamilies. The exact role of *Dfg5*-proteins could be explained by an in-depth structural and functional analysis of one of its members, *Ct*Dfg5, from the thermophilic mold *Chaetomium thermophilum*. Its crystal structure determined at atomic resolution showed that the overall fold and the active site motif is shared between fungal and bacterial homologs, however an annotated function as α1,6-mannanases could not be shown *in vitro*. Instead it was possible to reassemble the GPI-core glycan structure (Manα1,2-Manα1,6-Manα1,4-GlcN) within the substrate-binding pocket of *Ct*Dfg5 by screening crystals with high molar sugar-fragments. This did not only provide a detailed view on the true substrate of *Dfg5*-proteins, but also first experimentally derived insights into the three-dimensional architecture of the GPI-anchor glycan. Together with the complex structure of a putative acceptor molecule, a lipid-to-wall transfer mechanism catalyzed by *Dfg5*-proteins could be derived. Moreover, the structural insights suggested a possible way of using different GPI-modifications as a coding system to determine the final localization of GPI-anchored proteins. Furthermore, structure-based docking of a commercially available lead library helped to identify a small molecule (FP-1) that binds to the active site of *Ct*Dfg5. FP-1 shows specific effects at milli molar concentrations in terms of the viability of the model organism *S. cerevisiae,* assuming a high potential for further drug development. Finally, another fungal homolog from a so far uncharacterized subfamily showed that all GH76 family members recognize α1,6-mannobiose as a central element of cognate substrates.

Table of Content

1 Introduction

1.1 The fungal cell wall

The cell wall describes an envelope that sits just behind the plasma membrane, surrounding the whole cell like a protective armor. It is essential for the cell's morphology, its osmotic integrity, and it facilitates resistance to external harmful aggressors. Further roles of cell walls are sensing of the cell's environment, adhesion to surfaces, biofilm formation, and cell-to-cell communication, thus making it critical for pathogenicity. While absent in animals, cell walls appear in many different forms, as well as in many different classes of organisms throughout nature. For example, microalgae known as diatoms have developed a cell wall that basically consists of silicate, while bacteria have a cell wall composed of linear glycans connected by oligo-peptides, thus called peptidoglycan, whereas plants possess a cell wall made of cross-linked cellulose and other glycans (Drum and Gordon, 2003; Lampugnani et al., 2018; Vollmer et al., 2008).

Fungi possess a cell wall composed of different types of glucans, chitin and proteins (Free, 2013). The wall's outstanding importance for the organism becomes obvious when looking at the sheer mass of ~1200 genes in the genome of *Saccharomyces cerevisiae* (i.e. one fifth of all genes), which were found to be involved in its synthesis, maturation, and general maintenance (De Groot et al., 2001). A complex interplay of all the different factors enables a highly dynamic adaptation to the environment and the growing cell, while its osmotic integrity must be maintained at all time. Errors in wall assembly have fatal consequences for the fungal organism due to the high internal turgor pressure of up to 10 MPa, which would lead to spontaneous cell lysis (Money, 2001). This, together with the absence of the involved constituents in humans, is the reason, why factors involved in the biosynthesis of cell walls provide excellent targets for antifungal drugs.

In principal, fungal cell walls (FCW) from all described ascomycota are built in a similar fashion with species-specific adaptations. For the fundamental understanding, the different aspects of the FCW will be described for the model organism *S. cerevisiae* in the following. However, important differences especially to filamentous fungi will be pointed out.

1.1.1 The structural organization of the fungal cell wall

The CW of the yeast *S. cerevisiae* is responsible for up to 30% of the cell's dry weight and distributed over a 100-200 nm thick structure forming the outermost layer of the organism

(Aguilar-Uscanga and François, 2003; Dupres et al., 2010; Yamaguchi et al., 2011; Yin et al., 2007). Electron micrographs revealed a bipartite architecture with an inner layer comprised of glucans and chitin, which is covered by an outer layer composed of heavily glycosylated proteins (Figure 1) (Baba et al., 1989; Hagen et al., 2004; Kapteyn et al., 1999; Kopecka et al., 1974; Osumi, 1998; Zlotnik et al., 1984). This composition is species-specific and can change upon developmental stage changes, within different phases of the cell cycle, and in response to extracellular stimuli such as nutrients, osmotic stressors or drugs (Hopke et al., 2018).

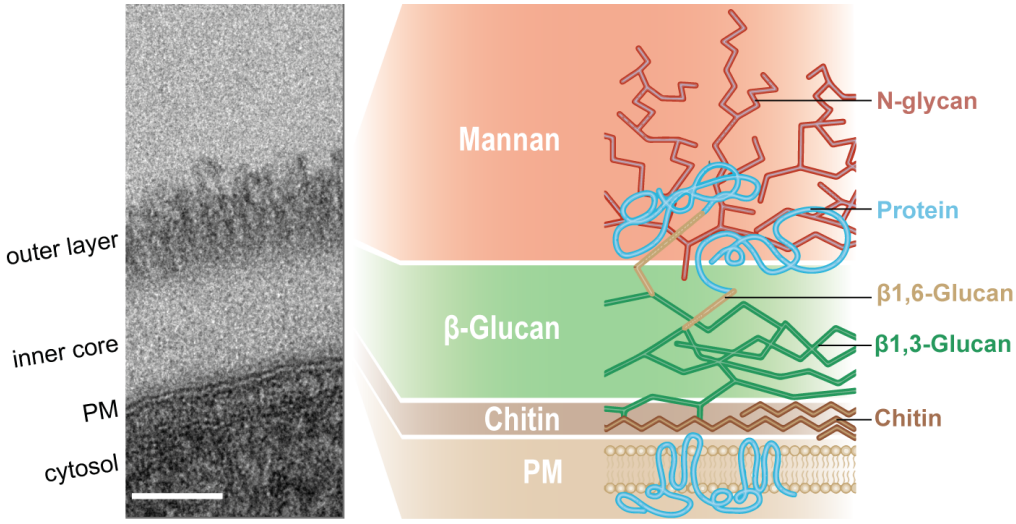


Figure 1: Architecture of the yeast cell wall. An electron microscopic picture of the fungal cell wall from *S. cerevisiae* is shown on the left. The respective structures found in the cell wall are schematically shown on the right (modified from Erwig and Gow, 2016). Scale bar=100 nm.

In yeast, the polysaccharides of the inner core consist of three different types (Kollár et al., 1995; Kopecka et al., 1974). As shown in the schematic representations linear chitin is the inner layer of the yeast cell wall (Figure 1). It is assembled of β1,4-linked N-acetylglucosamine (GlcNAc) and accounts for only 1-2% of the CW's dry weight, while filamentous fungi such as *Aspergillus fumigatus* possess up to 15% (Gastebois et al., 2009; Hu et al., 2007; Lesage and Bussey, 2006). The degree of polymerization ranges from 110 to 190 residues. In unstressed cells most chitin is found in the bud tip and the scar (Holan et al., 2004). Chitin at the bud neck comes as a chitin ring that stops cell wall growth and a chitin-based septum, which finally leads to the separation of mother and daughter cells (Lippincott and Li, 1998; Molano et al., 1980; Schmidt et al., 2002; Shaw et al., 1991). It exists in three forms within the CW: (i.) as a free form, (ii.) bound to β1,3-glucan and (iii.) linked to β1,6-glucan (Cabib, 2009; Cabib and Durán, 2005). Chitin is a stiff molecule that facilitates

rigidity to the cell wall (Kang et al., 1984). This explains elevated levels in the lateral cell wall upon cell wall related stress to compensate for loss of integrity, which can be visualized by the chitin-binding fluorophore Calcofluor white (Kapteyn et al., 1997; Ram and Klis, 2006; Ram et al., 1994).

The most abundant molecule in the CW of *S. cerevisiae* is β1,3-glucan, which accounts for half of the CW dry mass and therefore makes up most of the inner electron transparent layer (Figure 1) (Fleet, 1991). The mature molecule consists of up to 1500 linear β1,3-linked glucose units, which is generally found as a major constituent of the fungal cell wall in all ascomycota (Free, 2013; Manners et al., 1973). β1,3-glucan is an acid and alkali insoluble molecule with β1,6-side branches that interconnect the linear molecules. Due to the high abundance and the extensive cross-linking, β1,3-glucan is not only the main component responsible for the resilience of the cell wall, but also the basis for its flexibility during remodeling. Another type of glucan from the inner core of the yeast cell wall is β1,6-glucan, which accounts for 5% of the dry weight and is composed of about 140 residues. While β1,6-glucan is responsible for the attachment of cell wall proteins (CWP) in *S. cerevisiae* and *Candida albicans*, it is supposed to be absent in filamentous fungi like *A. fumigatus* and *Neurospora crassa*. Such fungi contain other types of polysaccharides described as β1,3/1,4-mixed glucan, α1,3-glucan, and galactomannans (Fontaine et al., 2000; Gastebois et al., 2009; Maddi and Free, 2010).

The outer, brush-like layer of the cell wall from *S. cerevisiae* can be digested by proteases and was found to consist of proteins highly decorated with N- and O-glycosylations (Kollár et al., 1997; Zlotnik et al., 1984). These posttranslational modifications are mannose-oligosaccharides and the eponymous element of mannoproteins forming the fibrillar structures (Figure 1). While polysaccharides are the main players for the structural framework of the cell wall, mannoproteins facilitate its functional features. There are three types of CWP: The first group contains proteins with enzymatic activities contributing to the synthesis and maintenance of the cell wall by their ability to hydrolyse and reconnect glycans. The second group has no enzymatic functions. They facilitate the capability of adhesion and flocculation in addition to further cross-linking the cell wall glucans (Dranginis et al., 2007; Goossens and Willaert, 2010; Klis et al., 2006, 2010; Kraushaar et al., 2015; Veelders and Essen, 2012; Yin et al., 2005). The third group is composed of mechanosensors that are able to detect osmotic stress, which is transmitted via a membrane helix to the cytosol to activate the cell wall integrity pathway (Rodicio and Heinisch, 2010).

Crucial for a functional cell wall is the covalent attachment of CWPs to the glycan matrix. One class is attached via an alkali labile ester bond of amino acid side chains to β1,3-glucan. These proteins contain internal repeats and are therefore referred to as PIR-CWPs (Ecker et al., 2006). Due to their attachment, PIR-proteins can be released from the CW upon β1,3-glucanase treatment (De Groot et al., 2005). Another way of CWP-attachment is based on the C-terminal posttranslational modification called the *glycosylphosphatidylinositol* (GPI)-anchor, which will be explained later in detail. A part of the GPI-glycan is linked to the β1,6-glucan in *S. cerevisiae* cell walls (Gonzalez et al., 2009). The linkage is sensitive to hydrogen fluoride (HF)/pyridine-treatment due to the hydrolysis of the phosphodiester between the ethanolamine phosphate (EtN-P) of the GPI-remnant and the peptide-bond linked C-terminal amino acid of the GPI-CWP (Yin et al., 2005). Another, indirect linkage to the CW is established by disulfide bonds, as some proteins were found to be extractable from the glycan-meshwork by reducing sulfhydryl reagents (Cappellaro et al., 1998; Orlean et al., 1986).

1.1.2 Synthesis and maturation of the fungal cell wall

The synthesis of the fungal cell wall is a multi-layered process. It includes the production of core polysaccharides at the plasma membrane, extracellular glycan remodeling, and the synthesis of mannoproteins along the secretory pathway.

The biosynthesis of β1,3-glucan and chitin is a well characterized process, although structural studies and a precise mechanistical understanding are still lacking. Both polysaccharides are synthesized by transmembrane proteins, which utilize nucleotide diphosphate sugars as a substrate for the vectorial synthesis of the polysaccharide precursor chains through the membrane into the extracellular space (Cabib et al., 1983; Duran et al., 1975; Shematek et al., 1980). In yeast, three proteins of the Fks family from the glycosyl transferase (GT) 48 class are involved in the β1,3-glucan synthesis (Bowman and Free, 2006). The functional complex requires the binding of a regulatory subunit, the lipid-bound small G-protein Rho1 that is activated upon GTP-binding (Drgonová et al., 1996; Inoue et al., 1999; Mazur and Baginsky, 1996; Mol et al., 1994). The activated complex utilizes UDP-Glc as sugar donor, which is transferred to the growing glucan chain to result in precursor chains of 60-80mers (Douglas et al., 1994; Shematek et al., 1980). Although having been described to possess redundant functions with a lethal double-mutant phenotype, Fks1 is described as the major synthase during vegetative growth, whose deletion results in a 75% reduction of synthesized β1,3-

glucan, while Fks2 is implicated in glucan-synthesis mainly under starvation, during mating and sporulation (Inoue et al., 1995; Mazur et al., 1995). The role of Fks3 is not entirely clear. It seems to be involved in ascospore wall formation, as its deletion results in abnormal spore morphology (Ishihara et al., 2007). In the filamentous pathogen *A. fumigatus* only one essential *FKS* gene is present (Mouyna et al., 2004). In contrast to the glucan synthases, the nature of chitin synthases (Chs) is more diverse. Two families subdivided into seven subclasses are described in literature, of which some fungi possess more than 20 *CHS* genes, while others contain just one (Lenardon et al., 2010; Morozov and Likhoshway, 2016). *S. cerevisiae* possesses three genes for Chs-proteins that belong to the GT2 family. Chitin synthases use UDP-GlcNAc as a substrate to produce up to 170mers *in vitro* (Kang et al., 1984; Orlean, 1987). The major synthase in yeast is Chs3, which produces more than 90% of chitin in vegetative cells with Chs2 accounting for ~5%, while no measurable contribution of chitin production could be determined for Chs1 *in vivo* (Shaw et al., 1991; Valdivieso et al., 1991).

In contrast to chitin and β1,3-glucan, the synthesis of β1,6-glucan in yeasts remains elusive. So far, no synthase could be identified to be responsible for the production of β1,6-glucan. The synthesis might be associated with the membrane as well, as crude membrane extracts from *S. cerevisiae* are described to show β1,6-glucan production *in vitro* (Vink et al., 2004). Mutants that possess a defective β1,6-glucan synthesis can be identified by their increasing K1 killer toxin resistance (Kre). The toxin is a secreted protein that uses β1,6-glucan as an acceptor and activates potassium efflux causing cell death (Sesti et al., 2001). Accordingly, a reduction of the acceptor is accompanied by an increase in resistance (Bussey, 1991; Hutchins and Bussey, 1983). Most of the discovered Kre-proteins, such as Kre5, Kre6, and Kre9, are proteins localized in the secretory pathway and possess β1,6-synthesis rates that are reduced up to 95% compared to wild type. However, a direct catalytic role is unlikely, as no intracellular β1,6-glucan could be observed so far and their role is more likely related to protein quality control (Meaden et al., 2015; Montijn et al., 1999; Shahinian and Bussey, 2000). Extracellular Kre-proteins do not seem to be directly involved in synthesis as well. Kre1 is a GPI-anchored protein (GPI-AP) with unknown function, which makes about 40% of β1,6-glucan compared to wild type upon deletion (Boone et al., 1990; Breinig et al., 2004). Mutants in the β1,3-glucan synthase Fks1 possess a K1 killer toxin phenotype accompanied by a modest reduction in β1,6-glucan. This phenomenon is explained by a reduced availability of its acceptor β1,3-glucan, although a dual function of Fks1 and Fks2 in the biosynthesis of both types of glucans is discussed (Dijkgraaf et al., 2002). Evidence for a membrane

associated synthesis is given by a recently characterized protein from the GT2 family, which was described to be a membrane embedded β1,6-glucan synthase found in the bacterium *Mycoplasma agalactiae* (Gaurivaud et al., 2016). Cell walls from filamentous fungi such as *A. fumigatus* lack β1,6-glucan, although this was recently questioned (Kang et al., 2018).

In order to generate the extensively cross-linked meshwork of polysaccharides and proteins of the fungal cell wall described above, the elongation and processing of the polysaccharide precursors, as well as the incorporation of proteins into the FCW is carried out extracellularly by carbohydrate active enzymes with hydrolyzing (and associated transglycosylation) activities. The precise temporal sequence is not entirely clear. Attempts of elucidating the sequence of events of the cell wall construction have been performed on yeast spheroplasts, which possess the ability to regenerate the removed cell wall. The regeneration process suggested β1,3-glucan as the starting component, followed by the incorporation of β1,6-glucan, mannoproteins and lastly chitin (Figure 2) (Kapteyn et al., 1997; Kreger and Kopecka, 1976; Lu et al., 1995; Roh et al., 2002; Shaw et al., 1991).

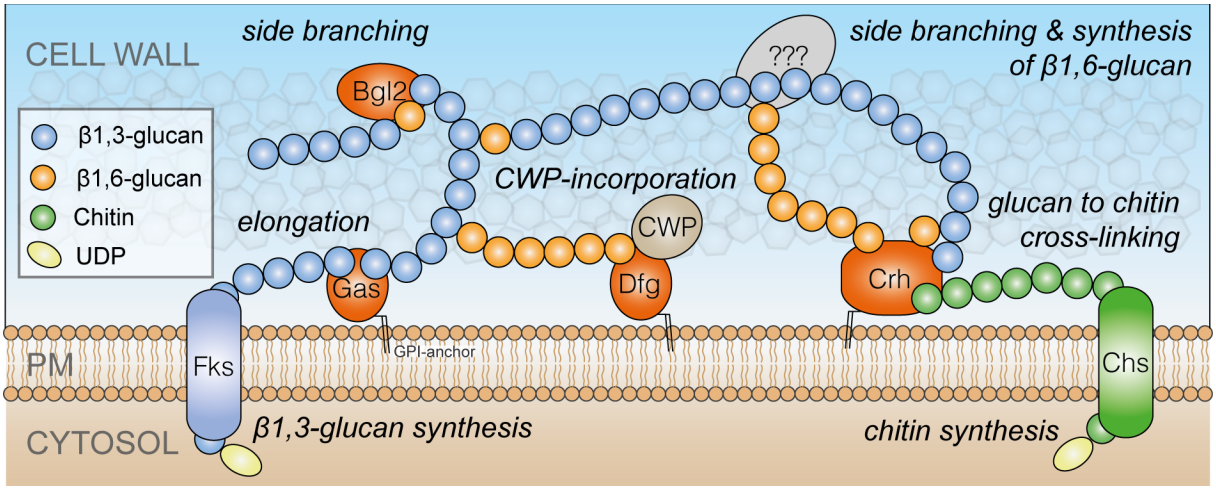


Figure 2: Schematic overview of the fungal cell wall biogenesis. Polysaccharides are produced at the plasma membrane (PM) by vectorial synthesis in the extracellular space. β1,3-glucan precursors are elongated by proteins from the Gas-family (GH72) and side branched by Bgl2-proteins (GH17). How the synthesis of β1,6-glucan happens is still unknown. In *S. cerevisiae*, CWPs are attached to the non-reducing end of β1,6-glucan by proteins from the *Dfg*5-family (GH76). Chitin cross-linking is established with variable cross-link types by Crh-proteins (GH16) (figure adapted from Cabib and Arroyo, 2013 and Teparić and Mrša, 2013).

All modifications require the cleavage of glycosidic bonds, a reaction catalyzed by glycoside hydrolases (GHs). To date, 165 different families are listed in the carbohydrate active enzyme database, of which 115 families are curator approved (www.cazy.org, June 2019) (Lombard et al., 2014). Despite this diversity, there are merely two mechanistic ways GHs follow to

hydrolyze glycosidic bonds (Figure 3). One directly displaces the leaving group by water, which leads to an inversion of the anomeric configuration, thus called *inverting* glycosidases (Figure 3A). The other type follows a *retaining* mechanism utilizing a double-displacement reaction including a covalent glycosyl-enzyme intermediate (Figure 3B) (McCarter and Withers, 1994). Both classes possess a catalytic pair of carboxylate side chains, which are further apart in inverting (~9 Å) than in retaining glycosidases (~5 Å) to allow the direct intervention of a water molecule (Wang et al., 1994). While one of the residues acts as a general acid and the other one as a general base in inverting enzymes, retaining glycosidases utilize one of the carboxylates as the nucleophile and leaving group and the other residue as a general acid/base. However, both utilize an oxocarbenium ion-like transition state (McCarter and Withers, 1994). In the transglycosylase class of retaining GH enzymes, the intercepting molecule is not water, but an acceptor molecule (usually another carbohydrate alcohol), resulting in a new glycosidic bond (Grout and Vicl, 1998).

Although sugars can adapt a broad spectrum of different conformations classified as chair (C), half-chair (H), boat (B), skew-boat (S), and envelope (E), just a couple of them favor the oxocarbenium ion transition state and are therefore utilized during the conformational distortion required for enzymatic catalysis (highlighted states in the Stoddart diagram shown in Figure 3C) (Raich et al., 2016). The Stoddart diagram is used to visualize the possible conformations of a pyranose ring, from which the distortions during catalysis (called the *catalytic itinerary*) at the -1 subsite moiety can be presented (Stoddart, 1971). The conventionally used sugar-binding subsite nomenclature labels the subsites from $-n$ to $+n$ (Davies et al., 1997). While $-n$ labels the non-reducing end with increasingly negative values towards the non-reducing end and away from the cleavage site, $+n$ accounts for moieties at the reducing end of the sugars, with increasing values away from the cleavage site towards the reducing terminus. The cleavage occurs between subsites -1 and +1.

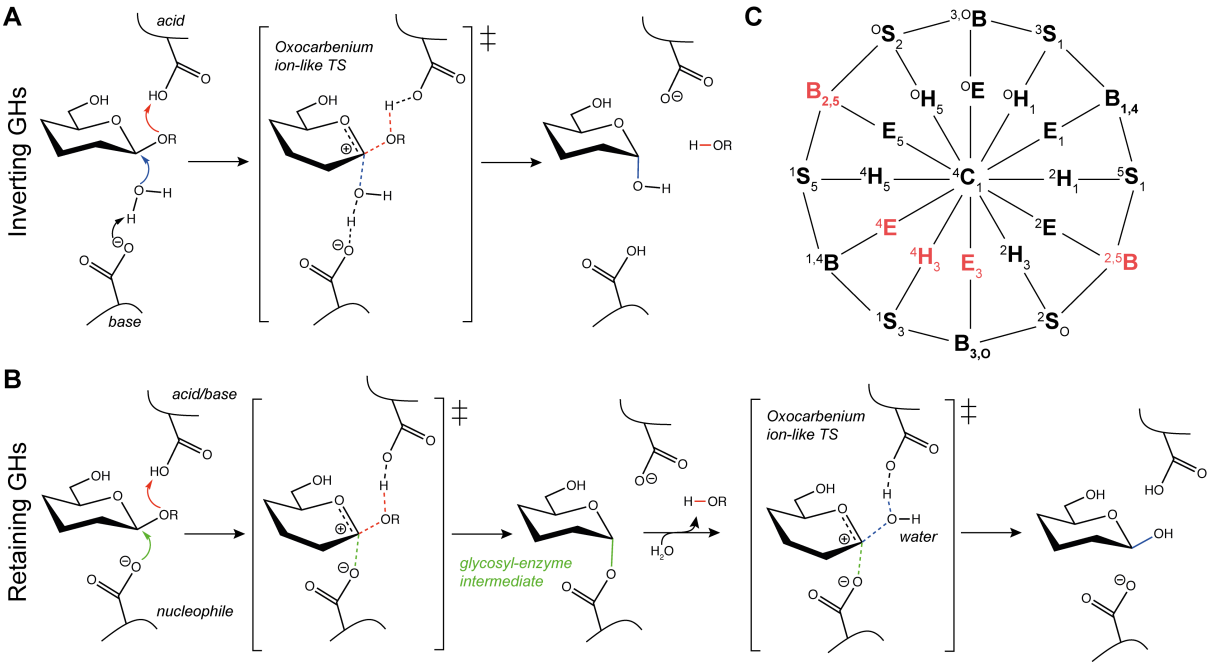


Figure 3: General mechanisms of glycoside hydrolases. The mechanisms of *inverting* (**A**) and *retaining* (**B**) GHs are shown with the characteristic oxocarbenium ion-like TS. **C** The Stoddart diagram centered at the $^{4}C_{1}$ conformation is shown. Conformations highlighted in red show the preferred distortions of the pyranoses to stabilize the oxocarbenium ion-like TS during catalysis (adapted from Raich et al., 2016).

While genetic analyses with functional predictions are numerous, biochemical *in vitro* analysis combined with structural characterizations of cell wall remodeling enzymes have been performed just for a couple of proteins (see Figure 2 and Figure 4). In yeasts, GPI-APs of the GH72 family, called Gas-proteins, have been described as β1,3-glycosyltransferases and are thought to be responsible for the β1,3-glucan elongation, as they are capable to cleave a donor β1,3-glucan chain and subsequently attach it to the non-reducing end of an β1,3-acceptor (Hurtado-Guerrero et al., 2009; Mazáň et al., 2011). Recently, a dual function of GH72 proteins has been proposed, contributing elongation and cross-linking activities to homologs from yeast and *A. fumigatus* (Aimanianda et al., 2017). There, it is suggested that GH72 and GH17 members act together on the side-branching of β1,3-glucans. GH17 proteins, such as the Bgl2 homolog from *Rhizomucor miehei*, are described as β1,3-glucan branching enzymes by exhibiting an endo-glycosidase activity coupled with the subsequent transfer of the reducing end to the 6-hydroxy group of an acceptor β1,3-glucan chain, thereby forming a β1,6-side branch (Goldman et al., 1995; Qin et al., 2015). Controversially, genetic studies in *N. crassa* suggested the GH72 family not to be involved in the β1,3-glucan processing, but to act on the incorporation of CWPs into the cell wall by cross-linking the N-linked galactomannan to lichenin (Ao and Free, 2017; Kar et al., 2019). Crh-proteins (GH16-family

members) are responsible for the chitin capping of glucan-chains by connecting the reducing end of chitin with the non-reducing end of a glucan, a notion that is strongly supported by a recently deposited crystal structure of Crh5 (PDB: 6IBW) from *A. fumigatus* in its glycosyl-enzyme intermediate state with bound chitin (Cabib, 2009).

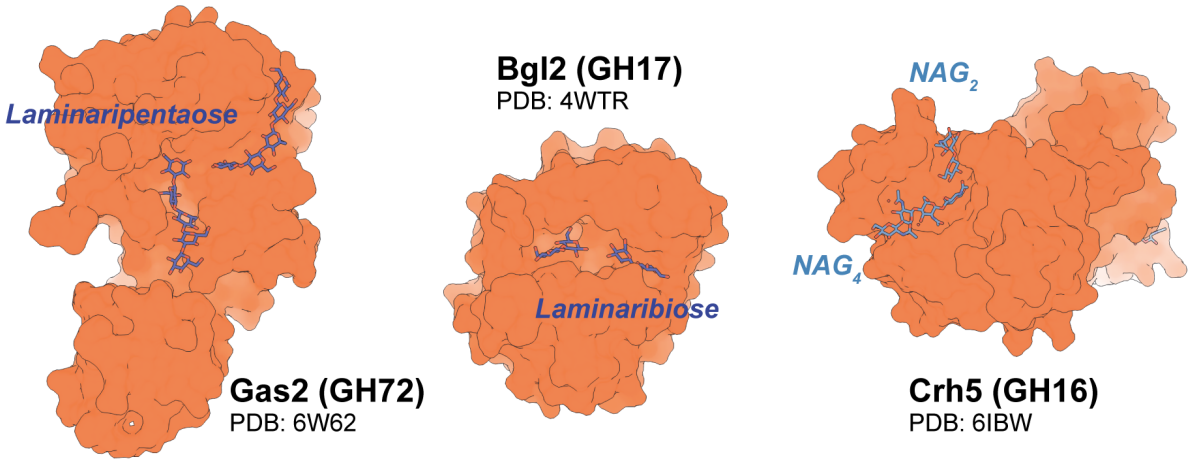


Figure 4: Overview of fungal GH-proteins involved in cell wall biosynthesis. Surface representation is shown for the structurally described GH-proteins in complex with their ligands, which are displayed as sticks. *Laminaribiose/-pentaose=β1,3-linked glucose (di-,pentamer), $NAG_{2/4}$=β1,4-linked N-acetyl glucosamine (di-/tetramer).*

1.1.4 Incorporation of proteins into the cell wall

A further essential step in the synthesis of a functional fungal cell wall is the incorporation of cell wall proteins (CWP). As described above, PIR-CWP are connected via an alkali-sensitive linkage to the cell wall. It was shown for Pir4 from *S. cerevisiae* that a specific glutamine (Q74) of the repetitive sequence $Q_{69}IGD_{72}G\mathbf{Q_{74}}VQ_{76}$ is identified as a glutamic acid upon hydrolytic release from the cell wall, while Pir4 purified from the medium still contained the glutamine (Ecker et al., 2006). It is further suggested that this conversion and the resulting ester linkage formation between the γ-carboxyl group of the glutamic acid and the hydroxyl group of a β1,3-glucan glucose happens in an autocatalytic manner assisted by the surrounding glutamines and the aspartate, as their exchange led to a loss of cell wall attachment.

The second type of CWPs attached to the cell wall via the GPI-remnant requires the action of two homologous proteins called Dfg5 (*defective* for *filamentous growth*) and Dcw1 (*defective cell wall*) from the GH76 family in *S. cerevisiae* (Kitagaki et al., 2002). Both proteins are predicted as GPI-APs and are predominantly found at the plasma membrane and to a minor extent in the cell wall (Kitagaki et al., 2002; Spreghini et al., 2003). Single deletion mutants

are viable, but possess hypersensitivity against zymolyase in case of Δ*dcw1* and are not able to invade agar via filament formation in pseudohyphal cells upon nitrogen starvation in the Δ*dfg5* background (Kitagaki et al., 2002; Mösch and Fink, 1997). In contrast to that, the double mutant is lethal and promotor shut-off cells deficient in both proteins exhibit rounded and large morphologies with an increased chitin content and released the GPI-CWP Cwp1 into the medium (Kitagaki et al., 2002). Similarly, Dfg5 and Dcw1 were identified in a screen of mutants deficient in GPI-CWP anchorage (Gonzalez et al., 2010). Conditional mutants of *dcw1* showed a cell cycle arrest in the early G2 phase and an aberrant cell wall upon elevated temperatures (Kitagaki et al., 2004). Studies in the pathogenic yeast *C. albicans* could strengthen the findings of a redundant and essential function of Dfg5 and Dcw1, as they were also found to be synthetically lethal (Spreghini et al., 2003). Here, Δ*dfg5* mutants were unable to form pH-induced hyphae, as the hyphae-specific gene (*HWT1*) was not expressed, while Δ*dcw1* did not show such a defect. In a very recent study it is shown that among seven *DFG* genes found in the genome of *A. fumigatus*, six of them are expressed under lab conditions and their complete deletion results in the absence of cell wall bound galactomannan (GM) accompanied with severe growth defects, but not lethality (Muszkieta et al., 2019). Interestingly, GM in *A. fumigatus* is a special type of polysaccharide, as it is attached to a GPI-anchor (Costachel et al., 2005).

It is suggested that Dfg5 and Dcw1 are responsible for the incorporation of GPI-CWP by cleavage of the Manα1,4-GlcN linkage within the GPI-anchor core and the subsequent transfer of the terminal mannose to β1,6-glucan, as such linkages exist in the cell wall of *S. cerevisiae* (Figure 5) (Fujii et al., 1999; Kitagaki et al., 2002; Kollár et al., 1997).

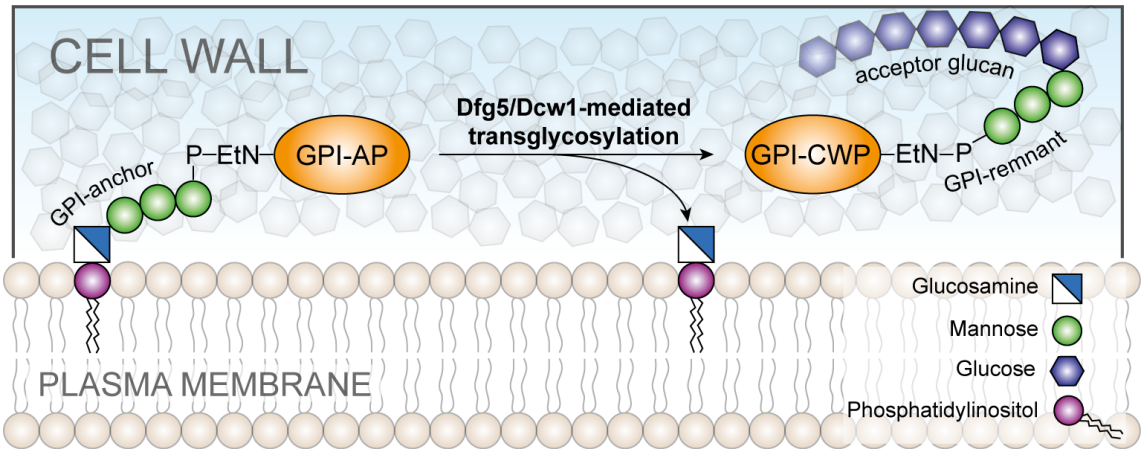


Figure 5: Principal model for CWP-incorporation by fungal GH76-proteins. GPI-APs are transported to the plasma membrane, where the recognition of the GH76-members Dfg5 and Dcw1 occurs. By cleaving the Manα1,4-GlcN linkage within the GPI-anchor, the transglycosylation is facilitated into the cell wall glycan matrix.

In contrast to this pathway, an alternative route of CWP-attachment catalyzed by Dfg5/Dcw1 is suggested for orthologs from the filamentous fungi *N. crassa*. This is based on the annotated α1,6-mannanase activity of GH76-proteins in the CAZy-database and the importance of the α1,6-mannose core structure of N-linked GM observed in Δ*och1* mutants from *N. crassa* (Maddi and Free, 2010; Maddi et al., 2012). Note that GM from *N. crassa* is the glycoprotein decorating polysaccharide, while the above-mentioned GM from *A. fumigatus* is a lipopolysaccharide. Initially it was described that mutants lacking the α1,6-mannosyltransferase Och1 show a severe defect in the incorporation of GPI-CWP (Maddi and Free, 2010). Och1 catalyzes the addition of the terminal α1,6-mannose of N-glycans on which the α1,6-mannan backbone is built on, thus leading to an immature form of the N-linked oligosaccharides (Nakanishi-Shindo et al., 1993). While *och1* mutants from a number of fungi possess a cell wall-related phenotype, it is not a general phenomenon found throughout ascomycota, as no cell wall defect was observed in *A. fumigatus* (Lambou et al., 2010; Moo et al., 2006; Schekman, 2010; Uccelletti et al., 2006; Yoko-o et al., 2001). Just two of the nine analyzed GH76-proteins from *N. crassa* were shown to possess morphological impact on the cells, of which Δ*dfg5* showed a spreading colonial phenotype, while Δ*dcw1* built thinner hyphae than the wild type (Maddi et al., 2012). The double mutant was not lethal like it is known from *S. cerevisiae*, which is explained with the remaining seven homologs that may function as cross-linking enzymes as well. Still, such cells released large amounts of cell wall proteins into the medium, a phenomenon that very well resembles the promotor shut-off cells from yeast (Kitagaki et al., 2002; Maddi et al., 2012). Further evidence for the N-glycan cross-linking hypothesis was later provided by studies of α-mannanases found in polysaccharide utilization loci (PULs) from the human gut bacterium *Bacteroides thetaiotaomicron* (Cuskin et al., 2015; Thompson et al., 2015a). They describe a sophisticated system established by this Gram-negative to degrade the yeast's outer mannan by a set of secreted mannanases including those of the GH76 family, which were shown to specifically hydrolyze the α1,6-mannose backbone (Cuskin et al., 2015). Further studies successfully cocrystallized a GH76 protein from *Bacillus circulans* in complex with α1,6-mannopentaose unveiling its substrate bound state (Thompson et al., 2015b).

1.2 GPI-anchored proteins

Aside from classical transmembrane helices, integral membrane proteins can also be produced by the attachment to lipids during their synthesis in the endoplasmic reticulum (ER). In eukaryotes, proteins can be transferred onto a glycolipid called *glycosylphosphatidylinositol anchor* (GPI-anchor) (Ikezawa, 2002). GPI-anchored proteins (GPI-AP) were discovered in the 70's upon mammalian tissue treatment with a phosphatidylinositol phospholipase C (PI-PLC), which released proteins from the plasma membrane and therefore was suggested to be covalently bound to a PI-anchor (Ikezawa et al., 1976; Low and Finean, 2015). They appear to preferentially localize at plasma membrane domains called lipid rafts, which consist of sphingolipids, sterols and certain proteins (Surma et al., 2012).

All characterized GPI-anchors from protozoa, fungi, plants, and mammalians revealed a common core structure composed of an ethanolamine phosphate (EtN-P) attached to the C-terminus of the GPI-anchored protein, followed by three mannoses (Man1, Man2, Man3), glucosamine (GlcN), and an inositol-phospholipid (PI) (Figure 6) (Fankhauser et al., 1993; Ferguson et al., 1988; Fontaine et al., 2003; Homans et al., 1988; Mukasa et al., 1995; Nakano et al., 1994; Oxley and Bacic, 1999). The glycan-core side chains as well as the lipid composition not only vary between different species but can also be different within a single organism and even the same protein, as shown for a 5'-nucleotidase from bovine liver and a surface glycoprotein from *Trypanosoma brucei* (Ferguson et al., 1988; Ikezawa, 2002; Paulick and Bertozzi, 2008; Taguchi et al., 1994).

GPI-APs can be predicted *in silico* due to their terminal amino acid composition. The number of predicted GPI-APs vary between different fungal species, ranging from just 33 in *Schizosaccharomyces pombe*, ~60 putative GPI-APs in *S. cerevisiae*, to ~100 in *N. crassa* and more than 160 putative GPI-APs in *C. albicans*. The numbers can differ significantly depending on the applied search strategies (Caro et al., 1997; Eisenhaber et al., 2004; de Groot et al., 2003). The attachment to a GPI-anchor is predicted if a protein contains (i.) a hydrophobic signal peptide at its N-terminus, which is responsible for the cotranslational translocation of the protein into the ER-lumen and (ii.) a C-terminally located GPI-attachment signal (Caro et al., 1997; Pierleoni et al., 2008). The amino acid, which is attached by a GPI-transamidase with an amide linkage between its carboxyl group and the ethanolamine of the GPI, is called the ω-residue. Preceding residues belong to the ω-minus region and residues following the ω-site are designated as ω-plus (Figure 6). All characteristic features of such a GPI-attachment signal are found in these ω-regions. While the ω-site residue itself can be an

alanine, aspartate, asparagine, cysteine, glycine, or serine, the ω-minus region from ω-10 to ω-1 consists of polar amino acids. ω+2 is mainly an alanine, glycine or serine, followed by a moderately polar spacer sequence from ω+3 to ω+9. Finally, a hydrophobic patch with a length that is able to pass the membrane completes the GPI signal sequence (Eisenhaber et al., 1998, 2004; de Groot et al., 2003).

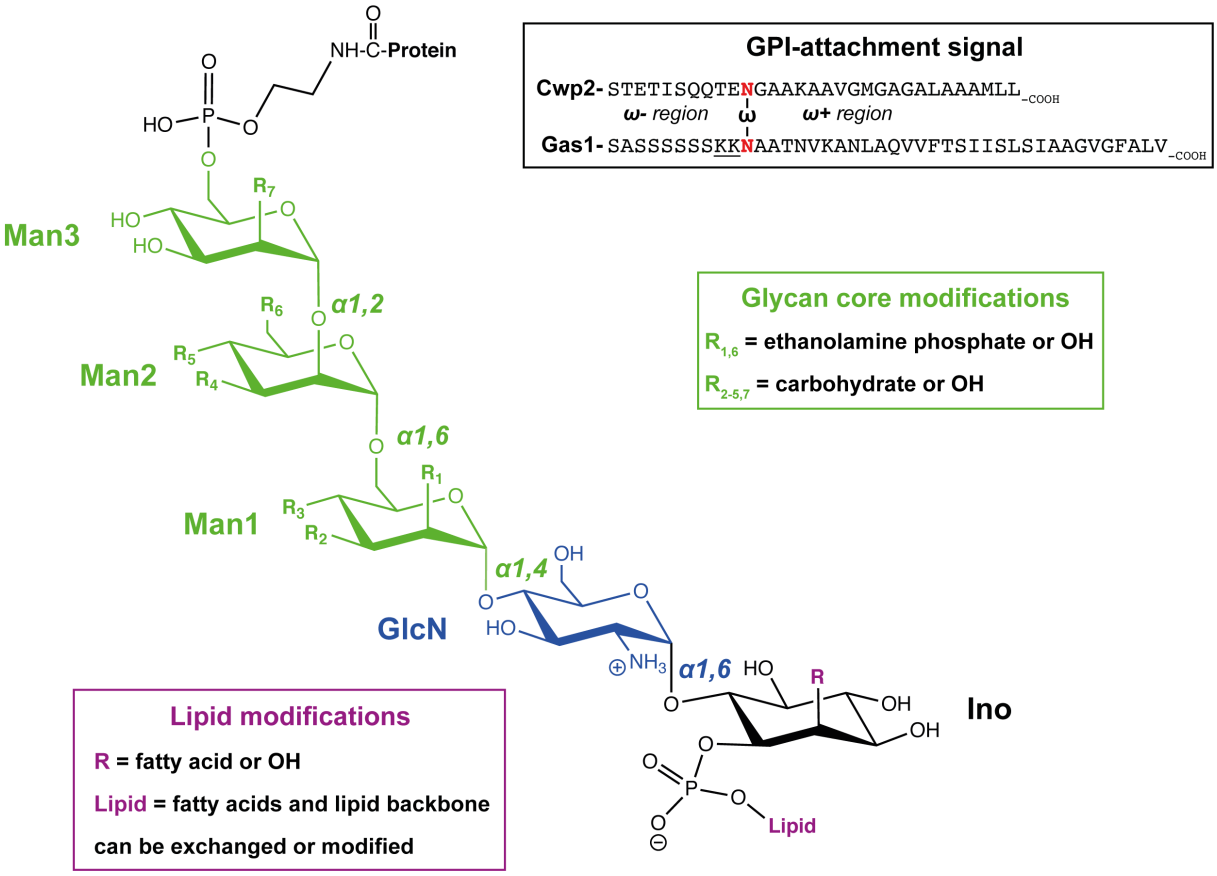


Figure 6: General GPI-anchor structure attached to a protein. The schematic structure of a GPI-anchor is shown, starting at the C-terminus of the attached protein, the bridging ethanolamine phosphate, three mannoses (Man1-3 starting from the moiety at the reducing end), a glucosamine (GlcN), and the phosphatidylinositol (Ino=Inositol). Characteristics of the different regions are highlighted in the boxes. The type of linkages between the GPI-core sugars is indicated (adapted and modified from Caro et al., 1997; Ferguson et al., 2009; Mayor and Riezman, 2004; Orlean and Menon, 2007).

The synthesis of the GPI-anchor and its transfer to the C-terminus of GPI-AP requires more than 20 genes in mammalians and yeasts, which are associated with the ER-membrane (Orlean and Menon, 2007). In principle, three major tasks must be completed. Firstly the GPI precursor assembly occurs, which is a topologically separated process including a flip of the GlcN-PI from the cytosolic to the luminal side of the ER-membrane. Secondly, the recognition, cleavage, and attachment of the C-terminal GPI-anchoring propeptide from a freshly synthesized GPI-AP onto the GPI-anchor must happen. In a third step, modifications

on the GPI-lipid and the glycan side-chains in the ER and the Golgi take place (Orlean, 2012). Due to the architecture of the GPI-anchor it is intuitive that the major steps in GPI precursor synthesis and protein attachment are highly conserved throughout eukaryotes, while the subsequent processing to the mature GPI-anchor is more diverse (Fujita and Kinoshita, 2012).

In *S. cerevisiae*, the GPI-anchor harbors four mannoses, of which three possess an EtN-P-modification attached by Mcd4, Gpi7, and Gpi13 to the 2-hydroxy group of Man1 and the 6-hydroxy-group of Man2 and Man3, respectively, and an acylated inositol at its 2-hydroxy group, which has been added in the first step of the luminal GPI assembly by the acyltransferase Gwt1 (Benachour et al., 1999; Galperin and Jedrzejas, 2001; Gaynor et al., 1999; Orlean, 2009; Tsukahara et al., 2003; Umemura et al., 2003). Upon protein attachment, the first steps in GPI-maturation consist of the removal of the inositol-linked acyl-chain by Bst1 and the remodeling of the diacylglycerol by substitution of the *sn-2* chain to a longer C26:0-fatty acid catalyzed by Per1 and Gup1. These can subsequently be replaced with a ceramide by Cwh43 (Bosson et al., 2006; Fujita et al., 2006b, 2006a; Ghugtyal et al., 2007; Martin-Yken et al., 2001; Tanaka et al., 2004; Umemura et al., 2007). After the removal of the EtN-Ps from Man1 by Cdc1 and Man2 by Ted1, the GPI-AP is transported to the Golgi where 20-30% of the GPI-glycan is extended by α1,2- or α1,3-linked mannoses to Man4, before it is finally delivered to the plasma membrane (Fankhauser et al., 1993; Haass et al., 2007; Vazquez et al., 2014). Here, a final GPI-processing step is attributed to two GPI-anchored proteins, Dfg5 and Dcw1, which were described to be responsible for GPI-CWP attachment to the CW by cleaving the GPI-glycan and its subsequent transfer onto the β1,6-glucan (Fujii et al., 1999; Kitagaki et al., 2002; Kollár et al., 1997).

Defects of GPI-anchor maturation result in delayed or false delivery of GPI-APs. Mutants deficient in *bst1*, *per1* and *gup1* that target the lipid moieties are not able to cluster into ER exit site (ERES) domains anymore, which are necessary for efficient ER to Golgi transport (Castillon et al., 2009). In *cwh43* mutants, the β1,3-glucanase Gas1, usually residing in the plasma membrane, tends to be transferred to the cell wall and the genetically related Ted1 is crucial for p24-mediated COPII vesicular from ERESs (Manzano-Lopez et al., 2015; Yoko-o et al., 2018). While not having a compromised ER-exit, conditional mutants of $cdc1^{ts}$, still containing the EtN-P at Man1, show a partial secretion block in the late secretory pathway and a fragile cell wall due to a compromised CWP-attachment to β1,6-glucans (Vazquez et al., 2014). Interestingly, the disruption of GPI-anchor synthesis in *S. cerevisiae* is lethal, while *A. fumigatus* survives with a defective cell wall, abnormal hyphae formation, rapid conidial germination, and aberrant conidiation (Leidich et al., 1995; Li et al., 2007).

1.3 Objectives of the thesis

Fungal infections are an increasing threat for human health, especially to immuno-compromised patients in clinical environments, with an estimated number of over two million infections per year and mortality rates up to 95% (Brown et al., 2012a). Due to their high similarity to humans on a cellular level, many antimycotica possess negative side effects on patients. A striking difference between mammalian and fungal cells is the fungal cell wall, which is essential for the survival of fungi and necessary for successful pathogenicity, yet absent in humans and therefore provides an excellent target for low side-effect drug development as shown for echinocandins (Perfect, 2017). An essential step in cell wall biosynthesis of yeasts such as *Saccharomyces cerevisiae* and the human pathogen *Candida albicans* is the incorporation of cell wall proteins, which is carried out by two homologous proteins called Dfg5 and Dcw1 belonging to the GH76 class (Kitagaki et al., 2002; Spreghini et al., 2003). However, a precise biochemical and structural understanding is still lacking behind. While GH76 proteins are essential in yeasts, they seem not to be essential for filamentous fungi like *Aspergillus fumigatus* and *Neurospora crassa* and may even act on other substrates (Maddi et al., 2012; Muszkieta et al., 2019).

In order to clarify the apparent contradictions and to shed light on this important step during fungal cell wall synthesis, this work will firstly address the evolutionary features of the GH76 family by utilizing *in silico* analysis to obtain a better understanding of its isofunctional subfamilies. Then, an in depth functional and structural characterization of a fungal *Dfg5*-member from the thermophilic mold *Chaetomium thermophilum* will enable us to not only allow structure-function based *in vivo* experiments, but shall also clarify the functional paradox of fungal *Dfg5*-proteins. This will be tackled by assembling its actual substrate into the binding pocket by high molecular sugar soaking. The determination of the crystal structure shall also set the basis for structure-based drug discovery. To do so, likely targets will be identified by docking and then used for soaking experiments, thereby identifying a true candidate for further drug development by combined structural and *in vivo* analysis. By expanding the structural knowledge to a further fungal GH76-homolog, the common feature of GH76-proteins shall finally be deduced.

2 Materials

2.1 Chemicals

(3-{4-imino-5,6-dimethyl-3H,4H-furo[2,3-d]pyrimidin-3-yl}propyl)dimethylamine (FP-1)	*Vitas-M Laboratory*
1H -Imidazole-2- methanamine, N ,1-dimethyl (FRAG37)	*Jena Bioscience*
2-(*N*-morpholino)ethanesulfonic acid (MES)	*Th. Geyer*
3-(N-morpholino)propanesulfonic acid (MOPS)	*Roth*
3-[3-(1H-imidazol-1-yl)propyl]-5,6-dimethyl-3H,4H-furo[2,3-d]pyrimidin-4-imine (FP-2)	*Vitas-M Laboratory*
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	*VWR*
4-Nitrophenyl α-D-mannopyranoside	*Merck*
Acetate	*VWR*
Acetonitrile	*VWR*
Agar-agar	*Roth*
Agarose	*Invitrogen*
Ammonium iodide (NH_4I)	*Fluka*
Ammonium persulfate (APS)	*Merck*
Borate	*Grüssing*
Bromphenolblue	
Calcium chloride ($CaCl_2$)	*Fluka*
cOmplete™ Protease Inhibitor Cocktail	*Merck*
Coomassie brilliant blue R-250	*Serva*
Dithiothreitol (DTT)	*Merck*
Ethanol	*VWR*
Ethylenediaminetetraacetic acid (EDTA)	*Merck*
Gadolinium(III)acetate ($Gd(OAc)_3$)	*Alfa Aesar*
Glucosamine	*TCI chemicals*
Glucose	*Roth*
Glycerol	*Roth*
Glycin	*Sigma*
Hydrochloric acid (HCl)	*VWR*
Imidazole	*Merck*
Isopropanol	*VWR*
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	*Gerbu*

Kanamycinsulfate	*VWR*
Manganese(II) chloride (Mn_2Cl)	
Mannose	*Merck*
Midori Green	*Biozym*
p-Anisaldehyde	
Peptone	*Difco*
Phenylmethylsulfonyl fluoride (PMSF)	*Fluka*
Polyethylene glycol 3350 (PEG3350)	*Fluka*
Polyethylene glycol 8000 (PEG8000)	*Fluka*
Potassium acetate	
Rotiphorese® Gel 30 (37,5:1)	*Roth*
Rubidium chloride (RbCl)	
Saccharose	*VWR*
Sodium chloride (NaCl)	*VWR*
Sodium dodecyl sulfate (SDS)	*AppliChem*
Sodiumhydroxide (NaOH)	*AppliChem*
Sulfuric acid (97%)	*VWR*
Tetramethylethylenediamine (TEMED)	*Roth*
Tris(hydroxymethyl)aminomethane (Tris)	*Roth*
Tryptone	*Th. Geyer*
Virkon	*VWR*
Yeast extract	*Th. Geyer*
α1,2-mannobiose	*Dextra*
α1,6-mannobiose	*Dextra*
β-mercaptoethanol	*Roth*
β1,3-laminaribiose	*MEGAZYMES*

Chemicals mentioned in this thesis, which are not listed here, were purchased from either *Fluka, VWR, Roth, Merck, AppliChem* or *TCI chemicals* in the highest best purity.

2.2 Equipment

Devise	**Model**	**Manufacturer**
AEC column	Bio-Scale Mini UNOsphere High Q	*Bio-Rad*
Agarose gelelectrophorese	Chamber	*Feinmechanik, Chemistry department, PUM*

Amylose resin HF	15 ml	*NEB*
Balance	PC2200	*Mettler*
Centrifuges	Centrifuge 5810 R	*Eppendorf*
	Fresco21	*Heraeus*
	Lynx 6000	*Sorvall*
	J2-HS / J2-21 M/E	*Beckman*
	L7-65	*Beckman*
Centrifuge Bottles	1l Superspeed CB with sealing	*Nalgene*
	JA-20	*Beckman*
	Polycarbonate tubes and bottles for 55.2 Ti	*Beckman*
Centrifuge Rotors	F6S 6x1000Y	*Thermo Fisher*
	JA-20 Fixed Angle Rotor	*Beckman*
	55.2 Ti	*Beckman*
Crystallization robot	Oryx8	*Douglas Instruments*
	Honeybee 963™	*Digilab*
Documentation of agarose gels	Computer E.A.S.Y.	*UVP*
	UV-transilluminator	*Herolab*
	Thermal printer UP-D 895	*Sony*
Documentation System Crystal plates	Rock Imager	*Formulatrix*
French Press	French Press	*Aminco*
	High pressure cell	*Feinmechanik, Chemistry department, PUM*
Heating block	BT3	*Grant Instuments*
IMAC column	Protino Ni-NTA Column 5mL	*Macherey-Nagel*
Incubators	BFED-53	*Binder*
	Multitron	*InforsHT*
	Innova S44i	*Eppendorf*
LC-System	NGC Chromatography System	*Bio-Rad*
Microfluidizer	Emulsiflex	
Microwave		*LG*
MilliQ water dispenser	Seralpur Pro90CN	*Seralpur*
NanoDrop	NanoDrop 200 Spectrophotometer	*Thermo Scientific*
OD-spectrometer	OD 600	*Implen*
Peristaltic pump	Pumpdrive 5201	*Heidolph*
pH meter	HI2020 edge®	*Hanna Instruments*
Pipets	Research	*Eppendorf*
Potter Homogenizer	15 ml glass cylinder and plunger	*Sartorius*
Power Boxes	EPS 301	*Amersham Biosciences*

Real-Time PCR-Cycler	Rotor-Gene Q	*Qiagen*
SDS-PAGE chamber	Mini-PROTEAN® Tetra Vertical Electrophoresis Cell	*Bio-Rad*
SEC column	HiLoad 26/600 Superdex 200pg	*GE Healthcare*
Spin concentrator	Amicon Concentrators (3-30 kDa MWCO)	*Millipore*
Thermocycler	GeneAmp PCR System 2400	*Perkin Elmer*
Thermomixer	Comfort	*Eppendorf*
Waterbath	NK22	*Haake*
X-ray sources/beamlines	Synchrotron Beamlines ID23-1/2, ID29	*ESRF, Grenoble*
	Synchrotron Beamline PXIII	*SLS, Villigen*
	In house source	*AG Klebe, Pharmacy department, PUM*

2.3 Commercial kits, enzymes, and consumables

Crystallization Screen	NeXtal Tubes JCSG Core Suite I	*Qiagen*
DNA Ladder	1 kb DNA Ladder	*NEB*
DNA-Ligase	T4-Ligase	*NEB*
	Ligase buffer 10x	*NEB*
DNA-Polymerase	Phusion Polymerase (2U/µL)	*NEB*
	Phusion HF-Buffer (5x)	*NEB*
GelEx Kit	QIAquick Gel Extraction Kit	*Qiagen*
Mini prep kit	QIAprep Spin Miniprep Kit	*Qiagen*
PCR purification	QIAquick PCR Purification Kit	*Qiagen*
Phospholipase C	PI-PLC from *Bacillus cereus*	*Merck*
Pipet tips		*Sarstedt*
Protein Ladder	Pierce Unstained Protein MW Marker	*Fermentas*
Reaction tubes	PCR-cups, 1.5 ml, 2.0 ml	*Sarstedt*
Restriction Enzymes	*Nhe*I	*NEB*
	*Xho*I	*NEB*
	*Hin*dIII	*NEB*
	*Bam*HI	*NEB*
	Cutsmart 10x	*NEB*
Sypro Orange	SYPRO™ Orange Protein Gel Stain	*Thermo Fisher*
TLC plates	TLC Silica Gel 60 F254 (5x10cm)	*Merck*
Zymolyase	100T	*Roth*

Consumables, enzymes, and kits mentioned in this thesis, which are not listed here, were purchased from either *NEB, Merck, Roth, Sarstedt* or *Qiagen.*

2.4 Oligonucleotides, vectors, and DNA

2.4.1 Oligonucleotides for gene amplification

Name	Sequence (5'-3')
*Ct*Dfg5_fwd	ATGC**GCTAGC**CAGCAGCAGTATTACAAGATCG
*Ct*Dfg5_rev	ATGC**CTCGAG**TTACCGGTCACCGGCATTAATCTCC
*Sc*Dfg5_fwd	GCATGC**GCTAGC**ATGGATTTGGATACTACTAGCAAAACG
*Sc*Dfg5_rev	ATGC**CTCGAG**TTAGCGATCACCTTTCTTAATG
*Sc*Dcw1_fwd	GCATGC**GCTAGC**GTCGAATTGGATCTAGACAACTACG
*Sc*Dcw1_rev	ATGC**CTCGAG**TTACTTTGAACCCTTAGTGATATT
*Cg*Dfg5_fwd	GCATGC**GCTAGC**ATCAATTTGCAAGACAATAAGG
*Cg*Dfg5_rev	ATGC**CTCGAG**TTAACGGTCACCACCAGTGATTTT
*Cg*Dcw1_fwd	GCATGC**GCTAGC**CTGGAGCTTGATCTGGATGATTACG
*Cg*Dcw1_rev	ATGC**CTCGAG**TTATCTAGAACCACCAGTAATATG
MBP-*Sc*Dcw1_fwd	ATCG**GGATCC**GTCGAATTGGATCTAGACAACT
MBP-*Sc*Dcw1_rev	AGTC**AAGCTT**CTTTGAACCCTTAGTGATATTT

*The **bold** characters indicate the introduced restriction sites: NheI (GCTAGC), XhoI (CTCGAG), BamHI (GGATCC), and HindIII (AAGCTT).*

The primers were used to amplify the defined regions from genomic DNA of *Saccharomyces cerevisiae* (*Sc*) and *Candida glabrata* (*Cg*), which we had available in our lab-stocks, and from cDNA of *Chaetomium thermophilum* (*Ct*), a kind gift of Dr. Patrick Pausch.

2.4.2 pET-28a(+)-Vector

The pET-system is a commonly used vector system for recombinant protein production in *E. coli.* Although the used pET-28a(+)-vector is a low copy plasmid, genes can directly be cloned in its *multiple cloning site* yielding enough plasmid for the different working steps. The plasmid provides a kanamycin resistance for selection purposes and stable maintenance within the bacterium in the presence of the antibiotic. The expression of the target gene is under the control of a T7-promotor (and terminator) as well as the lac-operator. Together with a suitable expression host (see below) an efficient protein production can be induced by lactose or its structural analog IPTG, reprogramming almost the complete metabolism of the cell towards gene expression. The vector furthermore encodes a hexa-histidine tag for efficient protein purification. The fully assembled pET-28a(+)-*Ct*Dfg5 vector used in this study is shown in Figure 7.

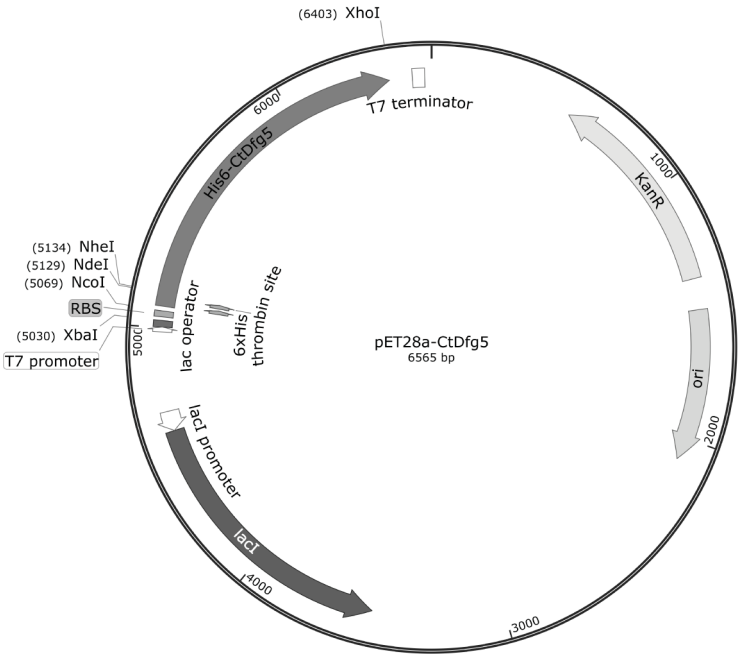


Figure 7: Vector map of pET28a-*Ct*Dfg5 used for heterologous protein production.

2.4.3 pMAL-cRI-Vector

pMAL-cRI is an old vector from *NEB* used for N-terminal MBP-fusion to the protein of interest. The plasmid provides an ampicillin resistance to the host organism for selection and maintenance. The expression is under the control of a tac-promoter and can be induced by lactose or its structural analog IPTG.

2.5 Organisms

2.5.1 *Escherichia coli* DH5α Competent Cells

The *E. coli* strain DH5α purchased from Invitrogen has been used for the efficient production and isolation of recombinant plasmid-DNA. Due to its mutations in endonucleases and recombination systems it is especially suited for that purpose, yielding high plasmid concentrations.

Genotype: F- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(rk-, mk+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ-

2.5.2 *Escherichia coli* SHuffle T7 Express Competent

E. coli SHuffle T7 cells from *NEB* were used for heterologous protein production of *Ct*Dfg5 and *Ct*GH76. The deletions of *gor* and *trxB* promote the formation of disulfide bonds, so does a constitutively expressed disulfide isomerase called DsbC. The chromosomal copy of T7 RNA-polymerase under the control of the wild type *lacZ*-promoter enables the compatibility with the T7-expression system together with the pET-vector system used in this study. Deficiencies in proteases furthermore facilitate higher yields of recombinant protein.

Genotype: *fhuA2 lacZ::T7 gene1* [lon] *ompT ahpC gal λatt::*pNEB3-r1-*cDsbC* (Spec^R, *lacI^q*) Δ*trxB sulA11 R(mcr-73::miniTn10*--Tet^S)2 [dcm] *R(zgb-210::Tn10* --Tet^S) *endA1 Δgor* Δ*(mcrC-mrr)114::IS10*

2.5.3 *Escherichia coli* BL21-Gold(DE3)

The *E. coli* BL21-Gold(DE3) strain from *NEB* is a protease deficient B strain derivative with high transformation efficiency optimized for the heterologous production of recombinant proteins using the T7 expression system.

Genotype: F^- *ompT gal dcm lon* $hsdS_B(r_B^- m_B^-)$ *λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5])* $[malB^+]_{K\text{-}12}(\lambda^S)$

2.5.4 *Saccharomyces cerevisiae* BY4741

Haploid cells of *S. cerevisiae* have been used for the preparation of soluble GPI-anchored proteins (see below). Beside common auxotrophic selection markers, they possess a deletion of the *ykl046c* locus, which encodes for the protein Dcw1. The strain was purchased from *Euroscarf* and is a S288c derivative, where many selection markers have been deleted

Genotype: BY4741; MATa; ura3Δ0; leu2Δ0; his3Δ1; met15Δ0; YKL046c::kanMX4

2.6 Software and Algorithms

EFI-EST	(Gerlt et al., 2015)
CCP4i2	(Nicholls, 2017)
Clustal Omega	(Sievers et al., 2011)
ConSurf	(Ashkenazy et al., 2016)

Coot v0.8.3	(Emsley et al., 2010)
CRANK2	(Skubák and Pannu, 2013)
Cytoscape	(Cline et al., 2007)
GraphPad Prism	(GraphPad Software)
Phaser	(McCoy et al., 2007)
PHENIX suite	(Adams et al., 2010)
SeeSAR	(BioSolveIT GmbH)
UCSF Chimera	(Pettersen et al., 2004)
WEBLOGO3	(Crooks et al., 2004)
XDS	(Kabsch, 2010)

3 Methods

3.1 Bioinformatics for biochemists

A major goal of biochemists is the comprehension of proteins and their underlying processes and mechanisms with the final understanding of the organisms' physiology. With the increasing number of genomes it is possible to combine already known principles and transfer them on so far uncharacterized proteins. While a comparison on DNA level is not practical, because of uncertainties concerning codon usage and silent mutations, it is much handier to use the amino acid sequence for global functional analysis. By comparing sequences it is possible to attribute uncharacterized proteins the functions of those already described. The whole procedure is based on the fact that proteins with similar sequences also share similar functions. Enzyme functions are classified in seven main classes characterized by the enzyme commission (EC) number classification consisting of four digits. With every digit the classification becomes more accurate. Using the famous enzyme lysozyme (EC 3.2.1.17), EC 3.-.-.- is attributed to the class of hydrolases, EC 3.2.1.- stands for glycoside hydrolases that hydrolyze O- and S-glycosyl compounds and 3.2.1.17 describes the class of lysozyme, which is responsible for the hydrolysis of β1,4-linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine in bacterial peptidoglycan. A threshold of 40% sequence identity can be used with a good confidence to transfer protein functions to a homolog concerning the first three digits of the EC number, while 60% and above are necessary for a 90% accuracy for all four digits (Tian and Skolnick, 2003). Although the actual function can vary from already characterized proteins of that class, such classifications are helpful to get a rough idea about protein properties to plan promising experiments. A good way of presenting such analyses is introduced by so called *sequence similarity networks*.

3.1.1 Protein Sequence Similarity Network (SSN)

The increasing number of available sequences requires reasonable approaches for their analysis. The *Enzyme Function Initiative* (EFI) developed a web tool called EFI-EST (*Enzyme Function Initiative-Enzyme Similarity Tool*), which allows users to analyze protein sequences according to their similarity (Gerlt et al., 2015). The aim of such sequence similarity networks (SSN) is to intuitively visualize whole protein families segregated into isofunctional subclasses by applying user-specified expectations values (E-values) on the collection of independent pairwise alignments (Atkinson et al., 2009). In such networks, *nodes* (representing one or more sequences with a certain sequence identity) are connected by *edges*,

if the applied stringency threshold is fulfilled. An example is shown in Figure 8, where it can easily be seen that moderate thresholds (or low *alignment scores*) result in large "hairball" clusters. By using more stringent cutoffs, edges between the nodes are removed and the hairball falls apart into distinct subclasses. The alignment score is increased until the family is split into isofunctional subfamilies.

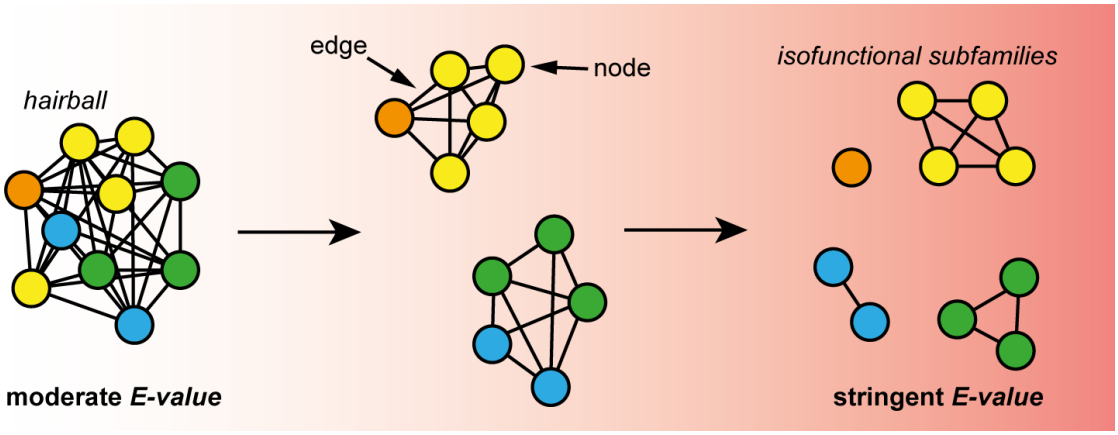


Figure 8: Effect of increasing the *E-value* stringency on the results of *sequence similarity networks*. The circles are described as *nodes*, which harbor sequences with a sequence identity of a user-defined threshold. The nodes are connected by *edges*, if the applied *E-value* cutoff is fulfilled.

For the SSN analysis, the protein family PF03663 (Glycosyl Hydrolase 76) was used to generate an EFI-EST dataset with a BLAST-derived E-value of 10^{-5} without further modifications. In a second step, different alignment scores were applied to the initial data and thereby generating the final SSN, which is produced as an `.xgmml` file composed of 7485 sequences. It can be visualized and analyzed with Cytoscape (v3.5.1) (Cline et al., 2007). For further analysis of the resulting subfamilies, multiple sequence alignments and weblogos are generated by ClustalΩ and WEBLOGO3, respectively (Crooks et al., 2004; Sievers et al., 2011).

3.2 Molecular Biology

3.2.1 Polymerase Chain Reaction (PCR)

Starting in 2015, the DNA-template based polymerase chain reaction (PCR), which has been described 30 years before, was still the method of choice to get your desired DNA in hand (Mullis et al., 1986). Within repetitive rounds of DNA denaturation, annealing, and elongation, complementary primers, dNTPs and a heat stable polymerase are utilized to

amplify a specific region from a given template. The region is defined by gene flanking primers, which also introduce restriction sites for later processing. Site-directed mutagenesis was used to generate the wanted mutant in a two-step *fusion PCR* approach with two complementary primers of ~30 bp and the gene-flanking primers. A typical PCR reaction and the respective parameters used for the experiments are described below.

	Volume or weight
Template DNA	20-30 ng
Forward Primer (10 µM)	2.0 µl
Reverse Primer (10 µM)	2.0 µl
Phusion HF-Buffer (5x)	10.0 µl
Phusion Polymerase (2U/µl)	0.5 µl
dNTP (10 mM each)	1.0 µl
MilliQ water	ad 50 µl

	Temperature	**Duration**	
Initial Denaturation	98 °C	5 min	
Denaturation	98 °C	20 s	30x
Annealing	55-58 °C	20 s	
Elongation	72 °C	15-30 s/kbp	
Final Elongation	72 °C	5 min	
Cool down	4 °C	∞	

In the meantime genes are usually ordered at companies (here BioCat), which synthesize the genes template independently. As gene synthesis became cheap in the last few years, it allows an economic access to target genes, without the need of a template DNA, that already come in the vector of choice and are codon-optimized for the respective expression host.

3.2.2 Agarose gelelectrophoresis

To analyze amplified DNA or to check the success of restriction digestions, agarose gel electrophoresis was performed. This key technique in molecular biology is used to separate DNA in an electric field according to its size following the principle: The smaller the negatively charged DNA, the higher the ability to migrate through a matrix of an agarose gel towards the anode. The samples are visualized by DNA-intercalating dyes, which can be detected under UV-light. For a typical 1% agarose gel, 0.65 g agarose were added to 70 ml TBE-buffer (0.1 M Tris, 0.1 M borate, 2 mM EDTA) and boiled in a microwave until the agarose was dissolved completely. 2.5 µl MIDORI Green were added to the solution before pouring the gel. Depending on the type of gel, 5 µl samples for an analytical run or 45 µl for a preparative gel were used.

3.2.3 PCR purification and Gel Extraction

Depending on the result of the analytical gel, a PCR product could directly be purified from the reaction by the QIAquick PCR Purification Kit from QIAGEN. Here, an isopropanol and guanidium chloride-containing buffer is added to the PCR reaction that facilitates binding of DNA to a silica matrix. While remaining nucleotides, primers, enzymes, and salts are removed with an ethanol-containing wash buffer, DNA molecules with more than 100 bp stick to the column and can finally eluted with a low salt buffer or water. To extract specific bands from a preparative gel, the QIAquick Gel Extraction Kit from QIAGEN was used. As an additional step, the DNA band must be dissolved before it is applied to the column. The exact procedures were performed according to the manufacture's guides.

3.2.4 DNA-modification: digest and ligate

In order to combine the target gene in the correct manner with the expression plasmid, compatible restriction sites are used. Usually, such sites consist of 4, 6, or 8 bp with a palindromic nature, which are recognized by endonucleases that cut the DNA in a *blunt* or *sticky end* manner. The digest is incubated at 37 °C for 1-2 h. A typical restriction digest consists of the following:

	Volume or weight
DNA	1 µg
Restriction Enzyme I	1 µl
Restriction Enzyme II	1 µl
Cutsmart (10x)	2 µl
MilliQ	ad 20 µl

When two different *sticky ends* present on the vector and the insert are used, the two DNA molecules can be combined in a specific orientation by the T4 DNA ligase. It catalyzes the formation of a phosphodiester bond between the 5'-phosphate and the 3'-OH groups of duplex DNA in an ATP-dependent manner. The ligation is carried out at RT for 30 min with a subsequent inactivation at 65 °C for 10 min using the following composition:

	Volume or weight
Vector	50 ng
Insert	*
T4-Ligase	0.5 µl
Ligase buffer (10x)	1.0 µl
MilliQ	ad 10 µl

*A molar ratio of 1:3 vector to insert is used

3.2.5 Plasmid preparation and determination of DNA concentration

When transformed into a bacterium, the respective plasmid is maintained unmodified in the organism as long as the selective pressure is present. This feature is used to multiply the number of plasmids by growing the cells in LB-medium supplemented with 35 µg/ml kanamycin until stationary phase over night at 37 °C shaking at 225 rpm. The plasmid is isolated by using the QIAprep Spin Miniprep Kit, which is based on the alkaline lysis protocol from Birnboim and Doly (Birnboim and Doly, 1979). The plasmid extraction follows a three-step protocol. The first step contains a buffer including RNase A (digestion of RNA) and EDTA (destabilizing membranes and inactivating DNases). The second buffer leads to the alkaline cell lysis and denaturation of DNA and proteins by NaOH supported by SDS, and, finally, the mixture is neutralized with potassium acetate buffer, which results in renaturation of plasmid DNA. Here lipids, proteins, and most of the genomic DNA remain insoluble and can be removed by centrifugation. The exact procedure was performed according to the manufactures guide.

LB-medium	
Tryptone	10 g/l
Yeast extract	5 g/l
NaCl	10 g/l
NaOH (10 M)	400 µl/l

3.2.6 Preparation of competent *E. coli* and plasmid transformation

The native ability of *E. coli* to take up naked DNA from its environment is limited. In 1970 it was demonstrated that the treatment with calcium chloride enhances that ability and confers the bacterium an artificial (or chemical) competence (Mandel and Higa, 1970). Later, this principle was improved by the addition of further metals to the buffer (Hanahan, 1983; Kushner, 1988). For one batch of chemical competent *E. coli* cells, 50 ml LB-medium is inoculated with 1 ml of a preculture grown over night at 37 °C. After incubation at 37 °C and 225 rpm until an OD_{600} of 0.5-0.6, the exponentially grown cells are harvested for 15 min and

3200*g* at 4 °C. The pellet is resuspended in 15 ml TBF-I buffer and incubated on ice for 2 h and subsequently pelleted again. After resuspension in 2 ml TBF-II buffer, the suspension is split up in 50 µl aliquots, flash frozen in liquid nitrogen and stored at -80 °C.

TBF-I		TBF-II	
RbCl	100 mM	RbCl	10 mM
$CaCl_2x2H_2O$	10 mM	$CaCl_2x2H_2O$	75 mM
Potassium acetate	30 mM	Glycerol	15% (v/v)
$MnCl_2$	50 mM	MOPS	10 mM
Glycerol	15% (v/v)		

For transformation, 10-100 ng plasmid-DNA is added to the cells after they thawed on ice. In case of ligations, the complete reaction was used for the transformation. After a heat shock of 45 s at 42 °C, 300 µl LB-medium are added and subsequently the cells are incubated at 37 °C shaking with 225 rpm for 50 min. The recovered cells are then pelleted at 14000 rpm, resuspended in 50 µl of the medium, and then plated on LB-agar plates supplemented with 35 µg/ml kanamycin. Colonies are grown over night at 37 °C and the plates can be stored at 4 °C for 2-3 weeks.

LB-agar	
Tryptone	10 g/l
Yeast extract	5 g/l
NaCl	10 g/l
NaOH (10 M)	400 µl/l
Agar-agar	15 g/l

3.3 Protein biochemistry

There is one major bottleneck in protein biochemistry: No protein, no experiments.

This often leads to rampant trials of test expressions to find a condition that results in soluble protein. Due to the fact that there was an established protocol for the production of other extracellular proteins from different fungi in the lab, I will directly proceed with the condition that worked for the proteins subjected in this thesis.

3.3.1 Heterologous overproduction in *E. coli* and cell lysis

The protocol for producing extracellular fungal proteins combines the ability of the *E. coli* strain SHuffle T7 Express, which is optimized for the production of proteins with multiple disulfide bonds within its cytosol, and a low temperature. For heterologous overproduction of

proteins without a solubility tag, 2 l of sterile DYT medium supplemented with 35 µg/ml kanamycin in a 5 l chicane flask are inoculated with 40 ml of an LB-grown overnight culture and incubated at 37 °C and 150 rpm until an OD_{600} of 0.2-0.3. The culture is then shifted to 12 °C and after cooling down for approx. 60 min, the production is induced with 0.1 mM IPTG. The cells are harvested at 4000 rpm and 4 °C for 20 min after an incubation of 72 h. The pellet from 2 l culture is resuspended in 25 ml lysis buffer, flash frozen in liquid nitrogen and stored at -80 °C. For cell lysis, the pellet is thawed in a water bath at RT until all ice disappeared. Steps from now on are kept on ice or at 4 °C.

For cell lysis, the suspension is poured into a pre-chilled French pressure cell press (short: French press, Aminco Inc.). A high pressure of 1000 psi is generated in the pressure cell, which is manually controlled by the outlet valve. When passing through the valve, the cells are decompressed and exposed to shear forces that finally disrupt the bacteria. The process is repeated two times. Afterwards, the lysate is centrifuged for 20 min at 4 °C and 18000 rpm, resulting in a cell-debris cleared lysate that is used for subsequent protein purification.

DYT-medium		**Lysis buffer**	
Tryptone	16 g/l	NaCl	200 mM
Yeast extract	10 g/l	HEPES	20 mM
NaCl	5 g/l		pH 8.0
NaOH (10 M)	400 µl/l		

3.3.1.1 Production of maltose-binding-ScDcw1 fusion protein

To produce MBP-*Sc*Dcw1 from *Saccharomyces cerevisiae* in *E. coli* BL21-Gold(DE3), 2 l LB-medium is inoculated with 40 ml over night culture supplemented with 50 µg/ml ampicillin and 12.5 g/l lactose for autoinduction (Studier, 2005). After 40 h at 20 °C shaking at 150 rpm in 5 l chicane flasks, cells are harvested and processed as described before.

3.3.2 Protein Purification

3.3.2.1 Immobilized Metal Affinity Chromatography (IMAC)

After cell lysis, the cleared lysate contains the overproduced protein and all soluble proteins from *E. coli*. To make life easy, the desired protein is produced with a His-tag at its N-terminus. The His-tag is utilized in the affinity chromatography technique called Ni-NTA, which is based on nitrilotriacetic acid (NTA) chelating nickel (Ni) ions. Polyhistidines are able to bind to nickel ions, while non-tagged proteins do not bind or can be removed

quantitatively during wash steps with moderate concentrations of the histidine analogon imidazole. To elute the protein competitively from the column, the imidazole concentration is simply increased. Alternatively, the pH can be shifted below 7, as the histidines become protonated and therefore loose their affinity for nickel ions.

The cleared supernatant including the His-tagged protein is loaded with a peristaltic pump using 3 ml/min flow rate onto a 5 ml Protino Ni-NTA column, which was equilibrated with 10 column volumes (CV) lysis buffer before. The column is then washed with 8 CV wash buffer and subsequently the recombinant protein is eluted by 4 CV elution buffer. Afterwards the protein is spin-concentrated with filter membranes of appropriate cutoff sizes to half of the injection loop volume used for size exclusion chromatography.

Wash buffer		**Elution buffer**	
NaCl	200 mM	NaCl	200 mM
HEPES	20 mM	HEPES	20 mM
Imidazole	25 mM	Imidazole	500 mM
	pH 8.0		pH 8.0

3.3.2.2 Size Exclusion Chromatography (SEC)

The principle of size exclusion chromatography (SEC) is based on the migration of macromolecules in a size-dependent manner through a chromatography column that consists of porous beads. Depending on the type of column, such beads can be made of dextran (Sephadex), polyacrylamide (Sephacryl) or agarose (Sepharose). While small molecules are occasionally trapped in the pores of the stationary phase, larger molecules pass directly through the column. As a consequence, bigger proteins of similar shape elute with a smaller retention volume from the column compared to smaller proteins. This is used to separate proteins of different size from each other to get the best possible monodisperse and pure protein solution in a defined buffer, which is an important aspect in protein crystallization. With the help of size standards, good analyts, i.e. proteins that do not interact with the matrix and have a similar shape as the proteins used for calibration, can be attributed an apparent size in solution, which gives information about the oligomeric state of the protein.

The concentrated elution from the IMAC is injected onto a HiLoad 26/600 Superdex 200 pg (GE Healthcare) that has been equilibrated with a sterile-filtered and degassed SEC-buffer until the conductivity reached a stable niveau prior to use. The protein containing fractions are detected at 280 nm and collected in 1.5 ml fractions. After SEC, the protein is either flash

frozen in liquid nitrogen and stored at -80 °C, directly used for crystallization or further purified by anion exchange chromatography, if necessary.

SEC-buffer	
NaCl	200 mM
HEPES	20 mM
	pH 7.5

3.3.2.3 Anion Exchange Chromatography

Proteins produced with the low-temperature protocol described above showed minor, but specific impurities, which could be removed for initial crystallization screening by a further purification method called *anion exchange chromatography* (AEC). The stationary phase of ion chromatography columns binds anions or cations depending on its matrix. The selection depends on the isoelectric point (pI, i.e. the pH value without a net charge) of the protein and the buffer system that is used. In buffers with a pH lower than the respective pI, the protein will be positively charged and *vice versa*. The working pH should be 1-2 units above or below the protein's pI. In low salt buffer systems, all proteins with the same charge will bind to the chosen column. By using increasing concentrations of salt or a change in pH, the bound proteins can be eluted from the column.

Here, the protein is loaded in AEC-1 buffer on a 1 ml Bio-Scale Mini UNOsphere High Q column from Bio-Rad. With a continuous gradient using AEC-1 and AEC-2, the protein of interest eluted at 12% of AEC-2.

AEC-1		AEC-2	
NaCl	10 mM	NaCl	1 M
HEPES	20 mM	HEPES	20 mM
	pH 8.0		pH 8.0

3.3.2.4 Amylose affinity chromatography

MBP can not only be used as a carrier protein to drive difficult, recombinant proteins in the soluble state, but it can also be used as an affinity tag for initial purification. MBP has a natural affinity for α1,4-maltodextrins, which is used in amylose resins as an affinity matrix for capturing the fusion protein. The elution occurs competitively by the addition of maltose to the buffer.

MBP-*Sc*Dcw1 is loaded after removing the cell debris on a 15 ml amylose column equilibrated with 10 CV MBP-1 buffer. The column is washed with the buffer until A_{280}

reached baseline. Subsequently, the protein is eluted with MBP-2 buffer containing 10 mM maltose.

MBP-1		MBP-2	
NaCl	300 mM	NaCl	300 M
HEPES	20 mM	HEPES	20 mM
	pH 7.5	Maltose	10 mM
			pH 7.5

3.3.3 Protein Analytics

3.3.3.1 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) describes a technique, which allows the separation of biological macromolecules according to their molecular weight under denaturing conditions (Laemmli, 1970). To ensure that, proteins are treated with SDS, which denatures proteins and simultaneously facilitates a negative charge that is proportional to the size of the protein. Furthermore, β-mercaptoethanol in the SDS-buffer reduces possible disulfide bonds. The generally described heat step can be skipped for most proteins in solutions. SDS-samples are loaded on a discontinuous SDS-PAGE. This type of SDS-PAGE allows nice results by sample focusing in a stacking gel that is based on isotachophoresis established by chloride and glycin at pH 6.8 and a size-dependent migration of the denaturated proteins in the separation gel with a pH of 8.8. By using a size standard it is possible to attribute apparent molecular weights to the protein bands and herewith control the progress of protein purification upon visualization with Coomassie brilliant blue R250, which binds to basic amino acids, especially arginine side-chains.

3-12 µl SDS-sample and 5 µl Pierce Unstained Protein MW Marker are pipetted in the pockets of an SDS-PAGE with a 4.5% (v/v) stacking gel and a 12% (v/v) separation gel. After the run, the gel is stained in hot Coomassie and subsequently destained in hot destain solution until the bands are well visible.

2x SDS-buffer		12% SDS-PAGE	Stacking Gel	Separation Gel
Tris/HCl, pH 6.8	62.5 mM	Acrylamide (30%)	6.67 ml	32 ml
Glycerol	15%	SDS (10% w/v)	500 µl	800 µl
SDS	4% (w/v)	dH_2O	29.8 ml	32 ml
Bromphenolblue	*a pinch*	Buffer*	12.5 ml	20 ml
β-mercaptoethanol	4% (v/v)	APS (10% w/v)	500 µl	800 µl
		TEMED	50 µl	80 µl

Stacking Gel Buffer*		**Separation Gel Buffer***	
Tris/HCl, pH 6.8	625 mM	Tris/HCl, pH 8.8	1.125 M
Ethanol	400 ml	Saccharose	30% (w/v)

Coomassie		**Destain**	
Coomassie brilliant Blue R250	3.2 g	Ethanol	400 ml
Ethanol	400 ml	Acetic Acid	80 ml
Acetic Acid	80 ml	dH_2O	400 ml
dH_2O	400 ml		

3.3.3.2 Determination of Protein Concentration

For the sake of reproduction and reliability in protein biochemistry, it is neccessary to know the concentration of the components including that of the protein used. To determine the concentration of the purified protein the extinction E of the respective protein at a wavelength of 280 nm was measured with a NanoDrop. It applies the law of Lambert-Beer

$$E = \varepsilon \cdot c \cdot d \qquad \textit{eq. 1}$$

E: Extinction, ε: molar extinction coefficient, c: concentration, d: thickness of irradiated object

that can also be given as

$$c_m = \frac{E \cdot MW}{\varepsilon \cdot d} \qquad \textit{eq. 2}$$

c_m: mass concentration, MW: molecular mass

The theoretical molecular mass and the extinction coefficient is calculated based on the amino acid sequence of the protein using the online tool Protparam (Gasteiger et al., 2005).

3.3.3.3 Thermal Shift Assay

The *Thermal Shift Assay* (TSA) is a biochemical method to easily determine the thermal stability of a protein due to sensitive fluorescence detection of an externally added dye, which binds to the protein upon denaturation. Interestingly, the thermal stability of proteins highly depends on the composition of the buffer cocktail. Usually, standard buffer systems are used for protein purification that can still be optimized for individual proteins in terms of pH, salts and further additives. Especially in X-ray crystallography, the optimal buffer composition can be crucial for successful protein crystallization as the protein is more stable and conformationally homogeneous (Huynh and Partch, 2015; Reinhard et al., 2013). As the

binding of small molecules can also have an impact on thermal stability, this method can also be used to screen for potential ligands.

To do so, the protein is diluted in the TSA-mixture containing the fluorescence dye Sypro Orange. The fluorescence of the dye is detected depending on the applied temperature gradient. As long as the protein is natively folded, the dye cannot bind and remains in solution. Here, the fluorescence is quenched. As the temperature increases constantly, the protein starts to denature at some point. The unfolding of the protein unmasks the unpolar core-residues to which the dye can bind and develop its full fluorescence. This increase of fluorescence is detected and the resulting curve is used to determine the melting point of the protein, which equals the inflection point of the melting peak (i.e. the maximum of the first derivation).

The buffers tested differed in their buffer substances used and the respective pH-value. All tested buffers contained 100 mM NaCl. The 5000x stock solution of Sypro Orange is 1:62.5 diluted with distilled water before usage.

The TSA-mixture and the respectively used buffers are listed below:

TSA-mixture	
Sypro Orange dilution	4 µl
*Ct*Dfg5 (100 µM)	2 µl
Buffer	ad 40 µl

Acetate buffer		**MES buffer**		**HEPES buffer**	
NaCl	100 mM	NaCl	100 mM	NaCl	100 mM
Acetate	50 mM	MES	50 mM	HEPES	20 mM
adjust to pH 4.5 or 5.5		*adjust to pH 6.0 or 6.5*		*(DTT*	*2 mM)*
				adjust to pH 7.5	

3.3.3.4 α1,6-Mannanase Assay

In order to test fungal GH76-proteins concerning their ability to hydrolyse α1,6-mannooligosaccharides, 10 µM *Ct*Dfg5 and 5 mM α1,6-linked mannotriose or mannohexaose buffered in acetate buffer (100 mM NaCl, 50 mM acetate, pH 5.5) or HEPES buffer (100 mM NaCl, 20 mM HEPES, pH 7.5) are prepared. The reactions and the negative control without protein are incubated at 30 °C for 22 h and subsequently analyzed by thin layer chromatography in comparison with standard solutions. The standards are composed of 5 mM sugar solutions from mannose, α1,6-mannobiose, α1,6-mannotriose, and α1,6-mannohexaose.

3.3.3.5 Thin layer chromatography

The biochemical properties can be determined by enzyme activity assays and the analysis of the resulting products. Unfortunately, carbohydrates lack chromophoric groups making them invisible for commonly used UV/Vis-detectors in liquid chromatography systems. However it is easy to analyze mono-, di-, and shorter oligosaccharides by thin layer chromatography (TLC). To do so, 5 µl of the sugar-containing solutions (i.e. reaction mixtures or standards) are spotted on the stationary phase (i.e. precasted silica gel 60 F_{254} plates (5 cm x 10 cm), Merck). The development of the TLC occurs by using a liquid phase consisting of 70:30 acetonitrile/water in a TLC chamber that was saturated with the mobile phase over night prior to use. The run starts from the bottom line (ca. 1 cm distance to the lower edge) sitting in <0.5 cm of the mobile phase and is terminated when the solvent front reaches a virtual line, which is about 2 cm from the top end of the plate. Then, the plate is dried with hot air and subsequently dipped in the *p*-anisaldehyde containing visualization reagent, which is subsequently developed at 110 °C for 10 min. Depending on the type of sugar, the sugar-derivates can produce blue, green and violet spots.

TLC visualization reagent

p-Anisaldehyde	1 ml
Sulfuric acid (97%)	1 ml
Ethanol	18 ml

3.3.3.6 Preparation of soluble GPI-APs from S. cerevisiae and the GPI-glycan hydrolysis assay

As a suitable GPI-anchored substrate is not commercially available and the chemical synthesis of complex sugars is extremely difficult, a promising approach could be the isolation of the substrate from the fungus directly. In order to do so, *S. cerevisiae* cells with a BY4741 background possessing a Δ*dcw1* knockout are taken from cell patches grown on YPD agar plates to inoculate 40 ml YPD medium, which is incubated over night at 30 °C shaking with 250 rpm.

YPD medium

Yeast extract	10 g/l
Peptone	20 g/l
Glucose	20 g/l
Agar-agar (for plates only)	20 g/l
dH_2O	ad 1 l

The media are autoclaved before use at 121 °C for 10 min. A preculture is taken for starting 2 l main culture in YPD medium, which grows another 24 h at 30 °C in baffled flasks at 150 rpm. Upon harvesting with 3000*g* for 20 min at room temperature, the cell pellet is resuspended in the lysis buffer (10 mM NaCl, 20 mM HEPES (pH 8.0), 2 mM DTT (added freshly prior to use), 10 mM EDTA, 2 mM PMSF (added freshly prior to use), and one tablet of cOmplete™ Protease Inhibitor Cocktail to a final volume of 50 ml. Subsequently, 10 mg Zymolyase 100T is added to the suspension and incubated for 90 min at 30 °C. Then, the cells are lysed by cycling the suspension in a microfluidizer at 200000 psi for 20 min. From now on, the steps are prepared on ice or at 4 °C. Unbroken cells and insoluble debris is removed from the cell lysate by two low speed centrifugation steps at 3000*g* for 5 min each. Subsequently, membranes are pelleted with 100000*g* for 1 h, the supernatant is discarded and the pellet is homogenized with a Potter homogenizer in the lysis buffer without DTT and the protease inhibitor cocktail prior to another centrifugation step at 100000*g* for 1 h. This wash step is repeated a second time, upon which 1 U of PI-PLC from *Bacillus cereus* is added to the homogenized and washed membrane to release GPI-anchored proteins from the membrane. The reaction is incubated over night on a roller wheel and the membrane is pelleted again at 100000*g* for 1 h. The resulting supernatant is analyzed by loading on an SDS-PAGE (see above), however the sample is just allowed to enter der separation gel. Upon Coomassie-staining, the protein bands are cut out and the proteins are identified by mass spectrometry. The presence of GPI-anchored proteins suggested the successful release and the conversion of GPI-APs to soluble GPI-APs (sGPI-AP). The supernatant, which is stored upon flash freezing in liquid nitrogen and kept at -80 °C, is then used for the final treatment with *Ct*Dfg5. 2 mg of *Ct*Dfg5 are added to 2 ml of the supernatant. The reaction is incubated at 30 °C for 6 h, the negative control is incubated without the added protein. In order to remove the proteins and other high molecular substances, the reaction is then filtered through a 3 kDa cutoff membrane and subsequently analyzed by the mass spectrometry facility and the NMR facility concerning the expected product mass and specific chemical shifts.

3.4 Protein crystallography

How proteins can be attributed certain functions according to their primary amino acid sequence was subject of the bioinformatics part above. However, such *in silico* analysis can only help to make hypotheses. They are the basis to plan experiments, which support the theories or disprove them. One major aspect in the comprehension of protein function is the

knowledge of their three-dimensional structure. The three major techniques to get atomic insights (i.e. yielding resolutions corresponding to <3 Å) of biological macromolecules are nuclear magnetic resonance spectroscopy (NMR), electron microscopy (EM), and X-ray crystallography. While NMR is limited in the size of the sample (routinely studied are proteins up to ~30 kDa), the great potential of structure determination at high resolution by cryo-EM arose in the last years and is optimal for huge complexes such as ribosomes, polymerases, and proteins that do not crystallize. However, today (April 2019) approx. 10.5% of the structures deposited in the protein data bank (PDB) were solved by NMR (~8.5%) and cryo EM (~2%), while more than 89% have been determined by X-ray crystallography. Basically, the three-dimensional structure of biological macromolecules of any size and kind (DNA, RNA, proteins, and their complexes) can be determined by X-ray crystallography. As the name indicates, the *only* requirement is their crystalline state. The steps that are necessary to determine the crystal structure of a protein will be described in the following with reference to the beautiful monographies of Jan Drenth, Bernard Rupp, and Gale Rhodes (Drenth, 2010; Rhodes, 2006; Rupp, 2010)

3.4.1 Protein crystallization

Although proteins have been crystallized on purpose since the middle of the 20th century, their individual properties (e.g. sequence, weight, and pI) still do not allow the a priori prediction of working crystallization conditions. While data collection and structure refinement is becoming more and more time efficient due to technical and computational improvements, the crystallization of proteins is still the bottleneck in X-ray crystallography. Nonetheless, a large number of commercially available screens and pipetting robots allow the screening of thousands of conditions in a straight-forward way.

The actual process of protein crystallization requires a pure and monodisperse protein solution with a sufficient concentration. This means that the solution contains the protein of interest in a distinct oligomeric state (almost) without other protein impurities. Said state is realized by consecutive steps of different purification methods up to the experimenter's satisfaction. Typical concentrations are between 5-50 mg/ml that require a reasonable yield of the protein. It can also be helpful to choose an optimal buffer system for the protein, which is determined by a thermal shift assay. The crystallization of the protein is achieved by reaching a supersaturated solution in which the protein starts to precipitate in a highly ordered manner, i.e. spontaneous formation of critical nuclei that can grow to well ordered macroscopic

crystals. This uncertain process depends not only on the protein concentration, but also on factors like pH, temperature, salt, precipitant, and other additives.

Using high-throughput screening and a pinch of luck the correct conditions can be identified. A common method to achieve supersaturation is vapor diffusion, which was also used in this work in its two variants *sitting* and *hanging drop*. Here, the undersaturated protein solution is mixed with the mother liquor in defined ratios. Within a closed system the water net diffusion is via the vapor phase from the drop into the mother liquor, as the substances from the reservoir are diluted within the drop. This results in a gradually decreasing volume within the drop and thereby an increase of the concentration (Figure 9). In an optimal experiment, this process leads from an undersaturated condition (yellow circle) to a concentration within the nucleation zone (red circle), where critical nuclei can form spontaneously by overcoming the energy barrier. Once happened, the state drops into the metastable zone, as the protein concentration is decreasing (blue circle). In that zone, crystals can grow without further nucleation. Finally, the condition reaches equilibrium of the undersaturated and saturated phase, where no further growth can happen. Within the precipitation zone the protein precipitates in an amorphous state and is therefore undesirable. However, in practice the general absence of amorphous precipitate is a good indicator that the initial protein concentration is not high enough.

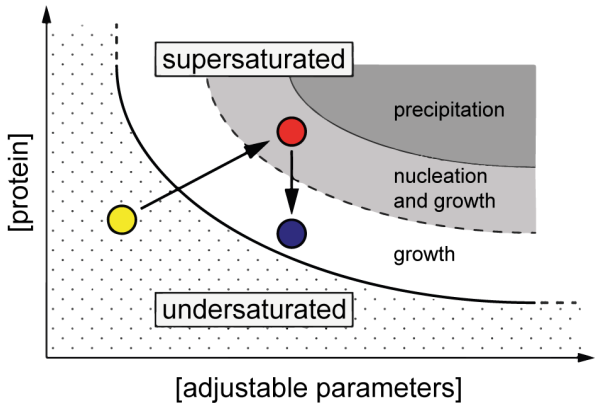


Figure 9: Phase diagram of an ideal crystallization experiment by vapor diffusion. The protein concentration is shown as a function of the precipitant concentration. Starting from the undersaturated situation (yellow circle), the concentration is shifted into supersaturation by vapor diffusion. Upon reaching the nucleation zone, critical nuclei can form and the crystal can start to grow (red circle). As the protein is removed from solution, the protein concentration decreases into the metastable growth zone, where only growth and no further nucleation can occur (blue circle) (adapted from Rhodes, 2006).

Commercially available screens were used for initial crystallization trials. Such *screening suites* contain 96 different conditions, which have been empirically shown to result in a good

number of crystal hits. The initial screens are pipetted in MRC 2 well crystallization sitting drop plates (*Swissci*) with the crystallization robot Honeybee 963^{TM} (*Digilab*) by filling 80 µl mother liquor in the reservoir and subsequently using 300 nL to fill each well. After that 300 nL of protein solution is added to each well and finally sealed with sealing foil. The incubation is carried out at 4 °C and 18 °C in a Rock Imager (*Formulatrix*) documentation system that is taking pictures of each well according to a given schedule. Based on grown protein crystals, a second type of initial screens using microseeding has been used. The advantage here is that spontaneous nucleation is not necessary as micro crystals are added to the condition directly. Herewith it is possible to find conditions that are optimal for crystal growth, but not for nucleation. To prepare a seed stock, a seed bead is chilled on ice in a 1.5 ml reaction tube with 50 µl of the respective reservoir solution. A few nice looking crystals are then crushed in its well with a suitable glass probe until no bigger crystals are visible anymore. The drop is diluted with 6 µl of the reservoir solution and then transferred into the seed bead tube. This is repeated two more times and after that the tube is vortexed for two minutes with 30 s cool down phases after every 30 s. The resulting seed stock is stored at -20 °C. For pipetting with the Oryx8 (*Douglas Instruments*), a 1:10 dilution of the seed stock with the respective reservoir condition is prepared and directly used. For each condition, 300 nL protein are mixed with 200 nL reservoir solution and 100 nL seed stock.

For reproduction of *Ct*Dfg5 crystals, the seeded JCSG core IV E7 condition has been used in a 24-well hanging drop approach. While the 3:2:1-ratio has been maintained, the final volume was increased by doubling the initial volumes. In fact, the crystallization of *Ct*Dfg5 using this approach was successful for protein concentrations between 10-30 mg/ml.

3.4.2 Soaking of protein crystals

Before going to the synchrotron for collecting diffraction data, it may be necessary to prepare the crystals in certain cases. Soaking of protein crystals describes the exposure of grown crystals to small molecules in a solution that is as close as possible to the growth condition. Soaking can be necessary to protect the crystal from ice formation, to solve the structure or to get ligand-bound states.

Ice formation can occur while flash freezing the crystals in liquid nitrogen, as the crystals consist of a remarkable amount of water (i.e. ~30-70%) and they are harvested from an aqueous solution. Ice rings result in a loss of information during data collection, as they overlay with the actual diffraction data. By adding cryogenic protection substances such as

glycerol, glucose or high molecular polyethylene glycols to the growth condition, the formation of ice can be avoided by soaking the crystal in cryo solution prior to freezing. Cryo solutions used in this work contained 30% glycerol. As they can have negative impact on crystal quality, it is advisable to check the diffraction without cryo protection and, if necessary screen for better working compounds. When such chemicals are present in the mother liquor, cryo protection may not be required at all.

In order to solve the phase problem of protein crystallography (described below), it is possible to utilize heavy metal atoms to recalculate the phases lost during data imaging. One way to introduce heavy metals in the protein structure is by substituting the native methionine with selenomethionine (i.e. the sulphur is changed to a selenium). However, the whole process of protein production, purification and crystallization needs to be repeated without a guarantee of success. Another approach for introducing heavy metals is using existing crystals and soaking them in a solution containing the respective atoms. In this work, gadolinium acetate, $Gd(OAC)_3$, was used as the acidic pI of 4.76 for *Ct*Dfg5 suggested a reasonable amount of negatively charged amino acids, which are suitable for specific Gd^{3+}-binding. The lanthanide was used as previously described in our lab by soaking with a concentration of 50 mM $Gd(OAC)_3$ with subsequent back soaking into the mother liquor containing 30% glycerol before flash freezing (Veelders and Essen, 2012).

The third purpose for the use of soaking was used in this study is to achieve substrate bound states. In principle, there are two ways. One is cocrystallization, which means that the ligand has been added to the protein prior to crystallization. This can happen by co-purification, if the ligand is present during protein production or purification and the affinity is high enough to keep it bound. Alternatively, ligands can be added prior to the crystallization screening. However, sometimes the actual ligand is not available, crystals do not grow in the presence of the ligand or the affinity is not high enough to force the protein into the bound state. Here, soaking can help. One main issue is that the addition of ligands can destroy the crystal lattice upon binding, either by interfering with the crystal contacts, or by inducing a conformational change that is also not compatible with the apo crystal form. On a macroscopic level, cracks can become visible within the crystal, the crystal can lose its sharp shape or they just dissolve. Furthermore, the crystal quality can suffer from the treatment. However, if the ligand is suitable for soaking, it enables the experimenter to utilize the apo crystals and convert them into ligand-bound states in a quite easy manner. For *Ct*Dfg5, GPI-anchor glycans were not commercially available. Therefore, fragments of the potential substrates have been used in

different concentrations and time points. Soaked crystals were directly flash frozen in liquid nitrogen from the soaking solution.

3.4.3 Principles of X-ray diffraction and data collection

The aim of the whole crystallographic procedure is to get the three-dimensional distribution of electrons from a molecule, which maps its appearance and therefore provides the structural biologist a precise picture (or model) of the protein. That the combination of crystals and X-rays is a good idea to study the atomic structure of molecules has been discovered 100 years ago, which has been awarded with two Nobel prizes in physics. The first one in 1914 honoring the work of Max von Laue for showing X-ray diffraction by crystals made of copper sulfate. One year later, in 1915, father and son William Henry Bragg and William Lawrence Bragg received the Nobel prize for extending that knowledge to analyze the actual crystal structure by means of X-ray diffraction. This is based on the periodic arrangement of the molecules within the three-dimensional crystal lattice. The repeating unit of the lattice is called *unit cell* of which the whole crystal lattice can be reconstructed by simple translation operations (Figure 10A). The unit cell itself consists of the *asymmetric unit* (AU), which is the smallest unit of a crystal. By applying rotational and translational symmetry operations on the AU, the unit cell can be generated. Basically, a diffraction pattern is the joined version of the objects that make the crystal. In the Bragg model, crystals are made up of a set of equally spaced planes that are parallel to each other and act as mirrors for the incident X-ray beam (Figure 10B).

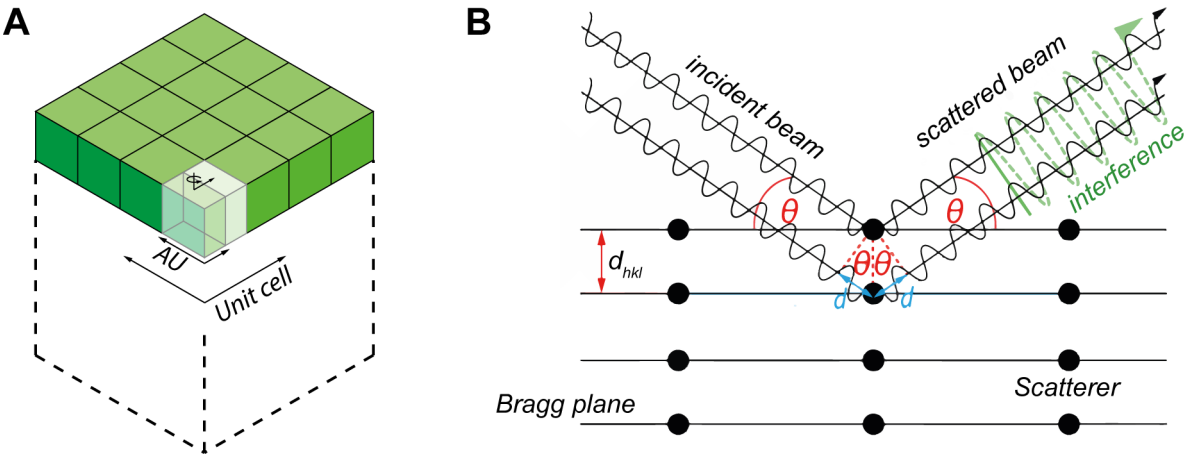


Figure 10: Schematic presentation of a three-dimensional crystal lattice and Bragg diffraction. A Crystals are made up of a periodic arrangment of its molecules. The *unit cell* is the repeating unit of the crystal, which reconstructs the crystal by translation operations in all three dimensions. The unit cell itself can consist of an *asymmetric unit* (AU). By rotational and translational operations, the unit cell can be built from the AU. **B** When applied on a crystal, the incident X-ray beam paths

on scatterers (black dots) within the crystal lattice, resulting in interaction with the scatterers. The horizontal line represents the Bragg planes, which are separated to each other by the interplanar distance called d_{hkl}. θ indicates the scattering angle, *d* is the path difference and the green dashed arrow indicates the resulting interference wave, when Bragg's law is fulfilled (see text for details).

When the beam interacts with the electrons of the atoms, it is converted into secondary spherical waves. This means that the scattered beam in principle goes into all directions, however only in one special case the resulting beam can interfere constructively and produce Bragg reflexes. This special case is mathematically described by *Bragg's law* and can be well explained by the illustration shown in Figure 10B. Here, the horizontal lines represent planes of one family of Bragg planes with incident and reflected beams that make an equal angle with the respective planes. To allow constructive interference of the two incident rays intersecting the plane, the scattered rays must be exactly in phase. This is only the case, if the lower scattered beam travels the additional distance depicted by *d* and if the traveled distance is equal to the wavelength λ, as the constructive interference requires the wave crests to be in phase. By applying simple trigonometric rules this results in *eq. 3*.

$$\lambda = d_{hkl} \sin\theta \qquad \textit{eq. 3}$$

That distance described by *eq. 3* must actually traveled twice by the beam as indicated in Figure 10B, it follows *eq. 4*.

$$\lambda = 2 \cdot d_{hkl} \sin\theta \qquad \textit{eq. 4}$$

Given the fact that for all other Bragg planes for that angle the distance to travel becomes a multiple of λ, the formula must be extended by the number of planes *n*, where *n* is an integer to allow constructive interference, resulting in *eq. 5* that is known as *Bragg's law*.

$$n\lambda = 2d_{hkl} \sin\theta \qquad \textit{eq. 5}$$

n: integer, λ: wavelength, d_{hkl}: distance between lattice planes, θ: diffraction angle

To get all the information of a crystal (i.e. expose all Bragg planes to the primary beam in the correct angle) many orientations must be analyzed during data collection.

For data collection, a protein crystal is picked with CryoLoops™ directly from the growth condition and is flash frozen in liquid nitrogen. Using the rotation method, crystals were

measured at synchrotron beamlines (ID23-1 and ID-29 at the ESRF; PXIII at the SLS) with the exception of the initial dataset needed for phase determination, which has been measured at the inhouse source of AG Klebe (Pharmacy, Philipps-Universität Marburg) using a IμS microfocus tube (*Incoatec*) and a MAR345 image plate detector (*Incoatec*) (Arndt, 1968; Nurizzo et al., 2006; De Sanctis et al., 2012). At the synchrotron, the loop is mounted automatically to a goniometer head with the crystal being protected by a cryostream running at 100 K. The goniometer enables the experimenter to rotate the crystal in any possible orientation through the primary beam, which is necessary to center the crystal in the beam in any orientation. Before collecting datasets, the crystal is characterized by two single measurements differing by 90° to assess an optimal collection strategy in terms of completeness and resolution by adjusting parameters such as exposure time, oscillation range, number of images and detector distance.

3.4.4 Processing of diffraction data

To proceed from data collection to structure determination, two major tasks must be fulfilled. The first task includes the assessment of crystal symmetry and the integration of the observed reflexes, which was done with the X-ray Detector Software (*XDS*) (Kabsch, 2010). After that data are reduced and judged by several parameters using the data reduction tool from the CCP4i2 software package (Potterton et al., 2018).

The input file `XDS.INP` is automatically created at the ESRF and SLS and is executed by the `xds_par` command in the shell. The software follows the path of eight different program steps that must be fulfilled in a fixed order, as the generated output files are used in the next steps. The software starts with the correction of geometrical distortions by using *XYCORR* (which are in fact usually already corrected and saved in the provided folder), followed by *INIT* that classifies pixels to discriminate between background and diffraction. Then diffraction spots are identified and located by *COLSPOT*, which are used in the subsequent indexing step, where *IDXREF* determines the orientation and symmetry of the crystal lattice, which explains the found spots best. As a result, every reflection of every diffraction image is assigned to their respective position in the reciprocal space using Miller indices *(h,k,l)*. After that *DEFPIX* defines a trusted region and defines pixels as untrusted, if they are obscured by hardware or excluded by a defined resolution range. The next step *XPLAN* is a tool for planning data collection, for which it estimates the completeness of new data for each starting angle and the crystals' total rotation. *INTEGRATE* then uses a two-step procedure to integrate

the reflexes and generates the INTEGRATE.HKL, which is used by *CORRECT*. It corrects intensities and standard deviations and writes out the finally integrated reflexes in XDS_ASCII.HKL including space group determination, refined unit cell constants and overall quality tables of the dataset.

The resulting XDS_ASCII.HKL file is directly imported into the X-ray data reduction and analysis module of the CCP4i2 software package, which comprises four programs. A few of these steps are redundant to some options that are possible within *XDS*, however rerunning data reduction here provides a good graphical output and careful data analysis can be done. After space group determination by *POINTLESS*, *AIMLESS* scales the data to a user defined cutoff value and *CTRUNCATE* generates amplitudes from the given intensities. Where to set the resolution cutoff to assess significant diffraction data is still a matter of debate. Current literature suggests to use the internal consistency determined by $CC_{1/2}$ as the major quality parameter (Karplus and Diederichs, 2012, 2015). This value randomly splits the data into half and gives values from 1 (for perfect correlation) to 0 (for no correlation) (see *eq. 6*). Depending on the crystallographers philosophy, data are suggested to be insignificant if the coefficient is below 0.15, while a more conservative value sets the cut-off at 0.5.

$$CC = \frac{\sum(x - \langle x \rangle)(y - \langle y \rangle)}{\sqrt{\sum(x - \langle x \rangle)^2 \sum(y - \langle y \rangle)^2}} \qquad \textit{eq. 6}$$

Historically spoken, crystallographic data were typically truncated at a resolution, where a signal-to-noise ratio reached ~2 and merging *R*-factors exceeded ~80%. R_{merge} indicates how well symmetry-related reflex intensities merge, which is defined by *eq. 7*.

$$R_{merge} = \frac{\sum_{hkl} \sum_i |I_i - \langle I \rangle|}{\sum_{hkl} \sum_i |\langle I \rangle|} \qquad \textit{eq. 7}$$

I_i=intensity for a reflex at position *hkl* in measurement i, $\langle I \rangle$=average intensity for reflex *hkl*

Typically, *R*-values are given as percentage and in principle it says that an ideal dataset would have an *R*-factor of zero. However, in reality this is of course not the case and indeed this quality parameter is vulnerable in terms of data redundancy, which results in increasing *R*-values for data with higher multiplicity (Weiss and Hilgenfeld, 1997). Unlike R_{merge}, the redundancy independent merging *R*-factor R_{rim} or R_{meas} is a quality factor, which is robust to

multiplicity, as it adds a factor for data redundancy (*eq. 8*) (Diederichs and Karplus, 1997). These values are suggested to be useful additional parameters also in terms of the correct space group assessment, however they should not play a role in determining the resolution cutoff (Karplus and Diederichs, 2015)

$$R_{meas} = \frac{\sum_{hkl}\sqrt{\frac{N}{N-1}} \times \sum_i |I_i - \langle I \rangle|}{\sum_{hkl}\sum_i |\langle I \rangle|} \qquad \textit{eq. 8}$$

N=redundancy of data

Beside the merged set of observed intensities, a second output is created by *FREERFLAG* that uses 5% of the data to generate a free-*R* set. Herewith a set of reflexes is excluded from the structure determination, which is later used for cross-validation of the refinement (see below).

3.4.5 Structure determination by solving the phase problem

The overall goal in protein structure determination is to obtain a three-dimensional map of the electron density ($\rho(x,y,z)$) in the volume of the unit cell (V), which is defined as a sum of a Fourier series of all structure factors measured in the diffraction experiment (*eq. 9*).

$$\rho(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l F_{hkl}\, e^{-2\pi i(hx+ky+lz)} \qquad \textit{eq. 9}$$

The structure factor F_{hkl} is the Fourier transform of the electron density within a unit cell and is described as an electromagnetic wave. Therefore it is necessary to know the three terms that describe a wave to get the information of the electron distribution: the wavelength, amplitude and phase (Figure 11).

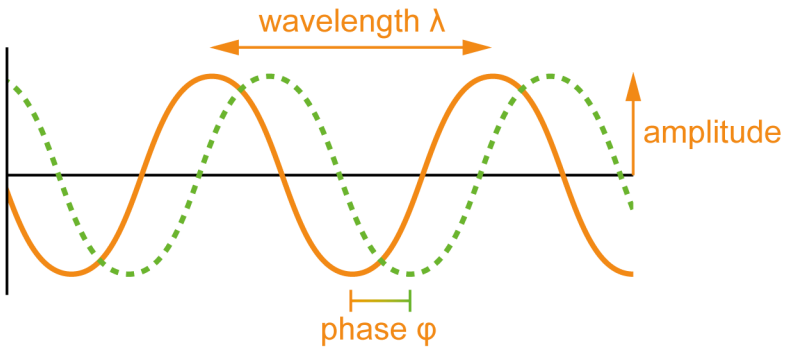


Figure 11: Characteristics of an electromagnetic wave. Two different waves are shown, which can be described by their amplitudes, wavelength λ and their phase φ.

While the wavelength is known from the experiment parameters used for data collection (no change of the incident beam upon elastic scattering) and the amplitude $|F_{hkl}|$ can be calculated as it is proportional to the measured spot intensity ${I_{hkl}}^{1/2}$, the phase can not be measured with the detection system used and is therefore lost by the experiment. This issue is called the *phase problem of protein crystallography*. A number of different methods are described to solve the phase problem (Taylor, 2010). Two of them have been used in this study: *single-wavelength anomalous diffraction* (SAD) and *molecular replacement* (MR).

The general principle of the SAD approach is based on the breakdown of *Friedel's law*, which says that the intensities of a pair of centrosymmetric reflections (the *Friedel pair*) at positions *h,k,l* and *-h,-k,-l* are equal for normal scattering. This is true for crystals, which only possess light atoms such as C, O, N, H and S, as they do not exhibit strong anomalous scattering effects (although *native SAD is maturing* (Rose et al., 2015)). However, this effect can be very strong for heavy atoms such as selenium or heavy metals, which can be incorporated into the protein crystal by amino acid derivates or crystal soaking. When the incident X-ray beam is close to the absorption edge of such atoms, the strength (*f'*) and the phase (*f''*) of the scattering changes. As the differences of changed scattering factors are not equally, Friedel's law becomes violated: i.e. that the intensities of *h,k,l* and *-h,-k,-l* are not the same anymore. This behavior is called *anomalous scattering* and depends on the type of heavy atom and its characteristic absorption maximum, which can be experimentally determined at tunable synchrotron beamlines. By using the *Patterson* method it is possible to determine the position of the heavy atom within the unit cell (Patterson, 1934). Here, the electron density function (*eq. 9*) is used, however the phases are set to zero and just the experimentally given intensity data are required, which is the square of the structure factors.

$$P(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}|F_{hkl}|^2\, e^{-2\pi i(hx+ky+lz)} \qquad \textit{eq. 10}$$

As a result, the Patterson function peaks do not correspond for the exact atom positions, but to the vectors between the atoms. This works well for substructure determination when only a few sites are present and improves the efficiency of substructure solution with *SHELXD* when Patterson seeding is used instead of random starting atoms (Schneider and Sheldrick, 2002).

For solving the structure of *Ct*Dfg5, the *CRANK2* pipeline of the ccp4i2 suite was used for gadolinium-soaked crystals from *Ct*Dfg5 including the program suite *SHELXC/D/E* (Sheldrick, 2010; Skubák and Pannu, 2013). With the obtained structural model for *Ct*Dfg5, all other data sets were solved by *molecular replacement* (MR).

In this approach, an already known structure, which has a similar primary sequence to the target (40% is high, 30% identity is usually successful, 20% can be successful) is used as a search model to determine initial phases (Dimaio et al., 2011). Basically, similar structures possess the same overall fold and by placing the search model in the correct orientation (α,β,γ) and position (x,y,z) of the unknown unit cell it is possible to obtain the correct phase. As a prerequisite, the Patterson maps from the measured intensities of the diffraction experiment (the observed structure factor amplitudes F_{obs}) and from the search model (the structure factor amplitudes F_{calc} derived from the model) must be calculated. Then, the six-dimensional problem is split in two three-dimensional problems. In the first step, the resulting maps are rotated against each other to get a best possible correlation between both maps and in the second step a translation function is used to position the search model to the correct coordinates within the asymmetric unit (Figure 12). If the positioning occurred correctly, initial phases can be calculated deriving from the information given by the input model. The software used in this study is *Phaser* implemented in the CCP4i2 suite (McCoy et al., 2007).

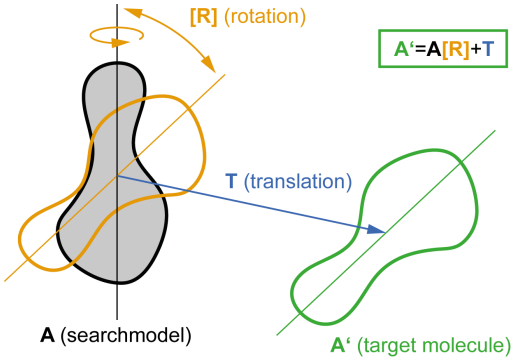


Figure 12: Visual explanation of molecular replacement. By applying rotational and translational variables on the search model A, its phases can be *borrowed* to calculated the phases of the target molecule A' without obtaining its phases experimentally.

In the optimal case, *Phaser* finds a unique solution at the end of the search. Three different parameters are considered to judge on the given solution. The *Log Likelihood Gain* (LLG) indicates how much better the model explains the diffraction pattern than a random distribution of atoms. As it is defined as the difference of the model likelihood and the Wilson distribution likelihood, the value must be positive and as high as possible. The second score is the *Translation Function Z-score* (TFZ), which is generally above 5 if the solution is likely to be correct and can clearly be distinguished from further solutions. As shown in Table 1, a definite solution can be assumed from *Z-scores* above 8.

Table 1: Probability table for a correct solution from *Phaser* after a successful molecular replacement job. The table is modified from *https://www.phaser.cimr.cam.ac.uk/index.php/Molecular_Replacement* (*July 2019*).

TFZ	Solved?
<5	no
5-6	unlikely
6-7	possibly
7-8	probably
>8	definitely

A further indication for a correct solution is the *clash score*, which should be below 10. However, looking at the clashes is in particular interesting, when a correct solution could not be obtained with the given search model. When using the default options of *Phaser,* 5% clashes are allowed for Cα-atoms in the packing step. Solutions with higher clashes are rejected from further computing, although the *Z-scores* may indicate a reasonable solution. For such cases, either the threshold could be adjusted or the input model must be modified in terms of peripheral loop regions.

3.4.6 Structure refinement

Upon structure solution, first estimated phases with an initial model are obtained, but the job is not done yet, as the model still contains errors or is simply incomplete. The actual goal is to get a final model, which explains the crystallographic data as good as possible. The process from the initial model to the final structure is called *refinement* and consists of an iterative process of (1) manual model building by visual evaluation and modification of the model and its correlation with the electron density map in *Coot* and (2) the actual computational refinement using *Phenix.refine* (Afonine et al., 2012; Emsley et al., 2010). *Phenix.refine* works automatically by introducing small changes to the model that optimize the stereochemical properties and thereby generating better model amplitudes (F_{calc}). These are compared to the observed amplitudes (F_{obs}) upon which a new cycle is started. It therefore is intuitive that a correlation factor between F_{calc} and F_{obs} is used to evaluate the obtained structures and the refinement strategies (*eq. 11*).

$$R = \frac{\sum\left||F_{obs}| - |F_{calc}|\right|}{\sum|F_{obs}|} \qquad \textit{eq. 11}$$

Accordingly, a model that perfectly explains the data would result in an *R*-factor of zero, however due to imperfections of the crystal lattice these values usually range within 0.15-0.25. Since datasets with higher resolutions contain more detailed observations, it is possible to refine such models to better *R*-values than model from low-resolution data. In fact, two *R*-values are commonly used for structure evaluation. The above described factor is calculated from the *working* model, thus called R_{work}. As mentioned earlier, 5% of the merged data are used to generate a free-*R* flag, which is now used as a cross-validation method to further evaluate the refinement quality. R_{free} is calculated with the same equation as applied for R_{work}, however using the omitted diffraction data instead of F_{obs} (Brünger, 1992). As the model amplitudes are not fitted to the excluded data, this value is always greater than the R_{work}. Significant differences of the two parameters indicate an over-fitting of the model.

3.4.7 Structural analysis and visualization

The obtained structural model coordinates are written out as `pdb`-files, which can be visualized by UCSF Chimera (Pettersen et al., 2004). All figures containing protein structures were generated with this tool. The *ConSurf*-server has been used to plot evolutionary information on the single amino acids written in a `pdb`-file, resulting in nine differently

scored conservation categories that can be visualized by different colors (Ashkenazy et al., 2016)

3.5 Structure-based drug design using *SeeSAR*

The underlying rational of finding potent antibiotics for a specific protein is to identify small molecules that are able to bind the target in such a manner that its actual biological task cannot be fulfilled anymore. In the case of essential proteins, such an interaction causes the death of the organism, which basically is the name-giving determinant of *antibiotics*. A plausible region, where such molecules should bind to enzymes is their substrate-binding region including the active site. A schematic workflow for structure-based drug development is shown in Figure 13.

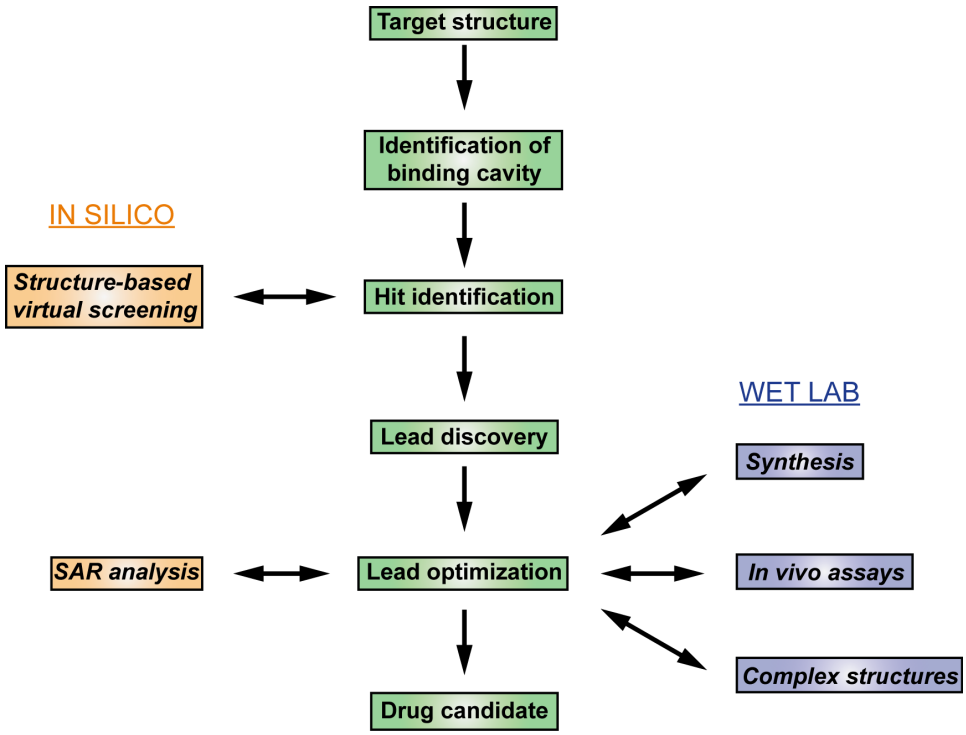


Figure 13: General workflow for structure-based drug discovery. The central green flow shows the general steps along which some tasks must be performed either *in silico* or in the wet lab (adapted from Andricopulo et al., 2009). *SAR=structure-activity relationship*

The absolute requirement of structure-based drug discovery is the three-dimensional protein structure, which is obtained in this work by X-ray crystallography. By comparing the protein to similar, already described structures, the binding cavity can be derived by structural alignments. Without prior knowledge, plausible regions can be assumed by visual inspection

of the structure concerning prominent cavities or pockets. Such a function is implemented in the software *SeeSAR*, which can automatically detect parts of the protein that are likely to bind molecules (BioSolveIT GmbH). The selected clouds are used to specify the potentially druggable binding sites, which are subsequently used for *Hit identification*. To do so, a molecule is loaded in the software, which then generates different poses with estimated binding affinities. The underlying algorithm scores the hydrogen bond and dehydration (HYDE) energies in the posed protein-ligand complex and thereby estimates the binding affinity (Reulecke et al., 2008; Schneider et al., 2013). The name-giving element of *SeeSAR* is that it allows to visually (so to say: *to see*) inspect the *structure-activity relationships* (SAR) of the leads bound to the protein. Atoms contributing to the interaction are highlighted in green, non-favoring atoms are labeled in red. Together with statistically derived torsion-angles from the CCDC small molecule crystal database, the user can modify and optimize the bound ligand with instant feedback of the applied structural changes (Guba et al., 2016; Schärfer et al., 2013).

In this work, a test license of *SeeSAR v.8.1* was used to screen for the best potential binders from the commercially available *Frag Xtal Screen* from Jena Bioscience. Ten poses for the 96 molecules were generated in the substrate-binding pocket of *Ct*Dfg5. The generated poses were then prioritized by the estimated affinities and the best twelve binders were used for lead discovery by soaking experiments as described above.

4 Results

4.1 Phylogenetic analysis of GH76 proteins

The power of bioinformatic tools to predict protein functions based on sequence similarities is commonly used in protein biochemistry to assume probable functions and to plan promising experiments for the protein in hand. To get an overview about the current knowledge of the glycoside hydrolase 76 family (GH76), the respective PFAM entry PF03663 has been subjected to an in-depth sequence similarity network (SSN) analysis with the goal to divide the family in isofunctional subfamilies (see Figure 14).

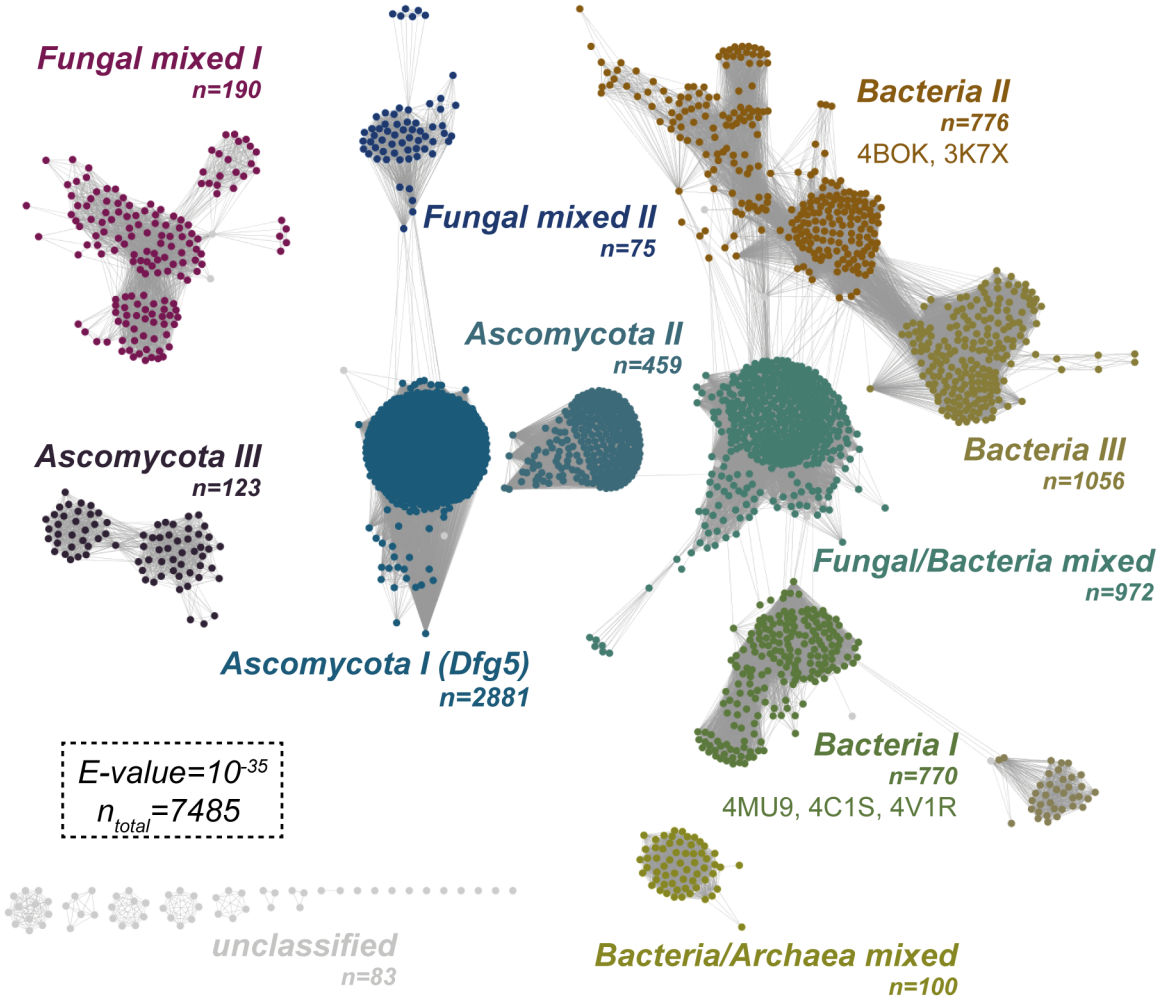


Figure 14: Sequence similarity network of GH76 proteins. The EFI EST-SSN of PF03663 shows the relationship of GH76 proteins using an E-value of 10^{-35} for all protein sequences with 300-700 amino acids. Every node harbors sequences of 40% identity. The applied parameters resulted in 7485 sequences in 2474 nodes and clustered into ten major subfamilies. The number of sequences *n* per subfamily is given. 83 sequences remained unclassified. If already deposited, PDB codes belonging to respective subfamilies are noted.

Using an E-value of 10^{-35}, the GH76 family falls apart in ten major subclasses. Five of them contain mostly bacterial proteins (*Bacteria I-III* and *mixed Bacteria/Archaea mixed*) and one subclass contains proteins of bacterial and fungal origins (*Fungal/Bacteria mixed)*. Furthermore, three different subclasses were found to harbor proteins exclusively from Ascomycota (*Ascomycota I-III*) and another two subclasses contain mixed populations of fungal species (*Fungal mixed I/II*).

Two of the bacterial subclasses (*Bacteria I* and *II*) have already been characterized on a structural and biochemical level and both subclasses were attributed an α1,6-mannanase function (Cuskin et al., 2015; Thompson et al., 2015b). In the literature, this activity has been generalized to the whole family. The largest subfamily in the SSN is the *Ascomycota I* class. Its first described member from *Saccharomyces cerevisiae* was found to have *defects in filamentous growth* and was called Dfg5 (Mösch and Fink, 1997). For clarity, the *Ascomycota I* subfamily will therefore be called *Dfg5* from here on. A few years later, Dfg5 and its homolog Dcw1 (*defective cell wall*) were described as essential factors for yeasts that are responsible for the incorporation of proteins into the fungal cell wall and it was suggested that the crosslinking happens via GPI-anchors (Kitagaki et al., 2002; Spreghini et al., 2003). Although being essential in *S. cerevisiae* and *Candida albicans*, a *Δdfg5*/*Δdcw1* double-mutant in *Neurospora crassa* has been shown to be viable (Maddi et al., 2012). The present SSN provides a suitable explanation for this phenomenon. Unlike yeasts, the filamentous ascomycete *N. crassa* possesses seven different GH76 proteins, which are members of the *Dfg5*-subfamily (see Table 2). In the context of described functional redundancy, it is plausible that a double distruptant is not lethal in organisms with more than two of such paralogs.

Interestingly, the presence of just two copies of *Dfg5*-proteins seems to be rather the exception than the rule. Other ascomycetes, such as the thermophilic fungus *Chaetomium thermophilum* or the human pathogen *Aspergillus fumigatus* each possess six *Dfg5*-orthologs. Furthermore, *N. crassa*, *C. thermophilum*, and *A. fumigatus* harbor more GH76 proteins in the *Ascomycota II* and *mixed Fungal/Bacteria mixed* subfamilies. Apart from that, *C. albicans* possesses another five GH76 proteins in an additional *Ascomycota III* class, while just two homologs of *Dfg5*-proteins are present.

Table 2: Distribution of GH76 proteins in selected ascomycetes. Uniprot-IDs are shown for the respective organism and subclass. The proteins marked with an asterisk were used for calculating the identities shown in **Table 3**.

	Saccharomyces cerevisiae	*Candida albicans*	*Neurospora crassa*	*Chaetomium thermophilum*	*Aspergillus fumigatus*
Dfg5	Q05031* P36091*	Q5ACZ2* Q5AD78*	Q1K7A8* Q7S4K4* Q1K7I4 Q7SAB2 Q7RUC1 Q7S3X3 Q1K918	G0S3F2* G0SFA3* G0SE99 G0RXS3 G0S8L9 G0S159	Q4WKP7* Q4WG09* Q4WFX5 Q4W985 Q4WFJ1 Q4WA43
Ascomycota II			Q7SGV4*	G0SHT7*	Q4WI31
Ascomycota III		Q59XS8* A0A1D8PIN7* A0A1D8PIJ7 A0A1D8PJF0 A0A1D8PFZ7			
Fungal/Bacteria mixed			Q7RYP1* Q7S863	G0S5Y9* G0S8A0	Q4WIT8

On a sequential level, selected *Dfg5*-proteins from Table 2 share sequence identities between 37-72% (see Table 3). It is suggested that proteins with a sequence identity of ~40% are similar enough that the general function can be transferred from characterized to uncharacterized proteins with a good degree of certainty (Tian and Skolnick, 2003). However, when looking at proteins from different subfamilies, the identities go down to ~20% and below. Such values have been described as the *Twilight zone of protein sequence alignments* and at this point reliable predictions of protein functions are not possible anymore (Rost, 1999). This supports the parameters applied to generate a SSN of GH76 proteins with segregated, isofunctional subclasses.

Table 3: Percent identities of selected proteins from the *Dfg5*-subfamily and the whole GH76 family

Dfg5		**GH76**	
Q4WKP7	44%	Q59XS8 *(Ascomycota III)*	14%
Q4WG09	39%	A0A1D8PIN7 *(Ascomycota III)*	13%
Q1K7A8	44%	Q9Z4P9 *(Bacteria II)*	18%
G0SFA3	48%	Q8A184 *(Bacteria I)*	17%
Q7S4K4	72%	G0S3F2 *(Dfg5)*	100%
G0S3F2	100%	Q1K7A8 *(Dfg5)*	44%
Q5ACZ2	37%	G0S5Y9 *(Fungal/Bac. mixed)*	20%
Q05031	40%	Q7RYP1 *(Fungal/Bac. mixed)*	21%
P36091	41%	Q7SGV4 *(Ascomycota II)*	20%
Q5AD78	37%	G0SHT7 *(Ascomycota II)*	23%

Taken together, the exact set of GH76 proteins in fungi is species-specific and the actual functions of the different subclasses are elusive, as for most classes *in vivo* data are not yet published. Despite the proposed contribution in cell wall protein incorporation, the mode of action is still a matter of debate and *in vitro* data still lack behind. While one side assumes that *Dfg5*-proteins hydrolyze the Manα1,4-GlcN-linkage within the GPI-anchor, the other suggests the α1,6-mannose backbone of N-linked outer chain mannan to be hydrolyzed and transferred to the CW glucan (Ao et al., 2015; Kitagaki et al., 2002; Maddi et al., 2012). The following chapter will solve the mechanistic paradoxon present in *Dfg5*-proteins.

4.2 The Dfg5-subfamily: Responsible for the incorporation of GPI-CWPs

As shown above, *S. cerevisiae* possesses only two paralogs of *Dfg5*-proteins and would be hence a good choice for *in vitro* studies, as a sophisticated model organism is available. Unfortunately, *Dfg5*-proteins from *S. cerevisiae* and from the closely related *Candida glabrata* could not be heterologously produced in the *Escherichia coli* expression system solely (SFigure 1 on p.134). A common approach to overcome this bottleneck in protein science is to use solubility tags, which can drive the fused protein of interest in a functional and soluble state. Dcw1 from *S. cerevisiae* could be produced with an N-terminally fused maltose-binding protein (MBP) using an amylose column for initial purification and subsequent SEC (SFigure 2 on p.134). MBP-*Sc*Dcw1 has been used for activity assays by Thierry Fontaine, where it was found to be inactive for all tested substrates (Muszkieta et al., 2019). As removal of the MBP-tag was not successful, it was not used for further studies.

Another approach to obtain soluble protein is to screen orthologs from thermophilic organisms, which have been described to have improved properties concerning their behavior in heterologous production and the following structural and biochemical studies due to their enhanced thermal stability (Cava et al., 2009). Unlike in bacteria and archaea, thermophilicity is less common in eukaryotes, but some fungi adopted an enhanced heat tolerance. Originally isolated from dung, the filamentous ascomycete *Chaetomium thermophilum* tolerates temperatures up to 60 °C and grows in standard media with a temperature optimum of 50-55 °C (Bock et al., 2014). The genome of *C. thermophilum* was sequenced in 2011 in order to provide a platform to study the nuclear pore complex (Amlacher et al., 2011). Therefore proteins from this fungus were chosen to be the subjects of an in-depth analysis.

4.2.1 Cloning, production and purification of recombinant *Ct*Dfg5

Suggesting an isofunctional subpopulation within the GH76-family, G0S3F2 from *Chaetomium thermophilum* (now termed as *Ct*Dfg5) found in the *Dfg5*-cluster is subjected to a detailed analysis in order to characterize the functional aspects of this subfamily. According to the annotations in the Uniprot entry, *Ct*Dfg5 is synthesized as type-I membrane protein containing an N-terminal signal peptide, which guides the nascent chain into the ER-lumen, and a C-terminal transmembrane helix (TMH). The presence of a TMH differs from the architecture of its yeast homologs, as the latter were described to harbor an ω-signal at the C-terminus that is attached to a GPI-anchor during protein maturation. In the final expression plasmid used for recombinant protein production in *Escherichia coli*, the terminal pro-peptide and TMH have been removed to obtain the protein in a soluble and matured state. *Ct*Dfg5 was amplified from cDNA of *Chaetomium thermophilum* and cloned into pET28a cut with *Nhe*I/*Xho*I. Production of soluble, recombinant protein could be achieved with the disulfide bond-promoting strain *E. coli* Shuffle T7 Express for three days at 12 °C induced with 0.1 mM IPTG (see Table 4 for protein chemical parameters).

Table 4: Theoretical properties of *Ct*Dfg5 calculated with Protparam (Gasteiger et al., 2005)**.**

Name	Uniprot-ID	Native amino acid range	Properties	Length	pI	MW	Ext. coefficient
*Ct*Dfg5	G0S3F2	30-449	His_6, Thrombin	443 aa	5.9	49.5 kDa	110.2 mM^{-1} cm^{-1}

The protein was purified from pellets of 2 l liquid cultures that were lysed either with a microfluidizer or with a French press. The initial purification started with a Ni-NTA affinity chromatography of *Ct*Dfg5. To achieve a best pure and homogenous protein solution for further experiments, the elution of the IMAC was concentrated and further purified with size exclusion chromatography (SEC) (see Figure 15). About 8 mg per liter cell culture could finally be obtained after SEC. Although a slight delay could be observed compared to standards, the elution volume of *Ct*Dfg5 suggests a monomeric state of the protein in solution.

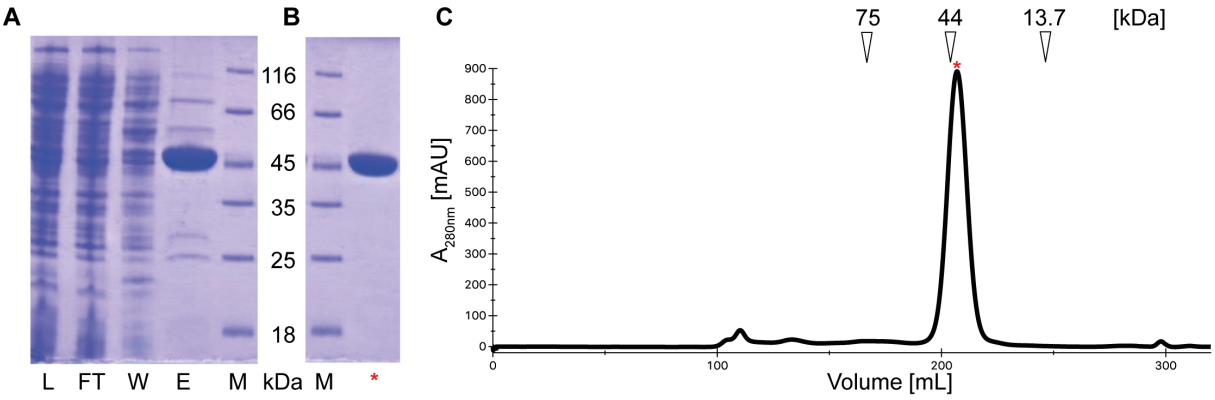


Figure 15: Purification of *Ct*Dfg5. A shows the IMAC purification with L=Load, FT=Flowthrough, W=Wash, E=Elution and M=Marker. **B** shows the peak of the size exclusion indicated by a red asterisk. **C** The chromatogram of the SEC is shown with the peak fraction from the SDS-PAGE indicated by the red asterisk and the arrows marking the elution volume of standard proteins with the respective molecular weight.

For the first crystallization screening that yielded the crystals shown in Figure 17, a further purification step was performed after the SEC step, as minor contaminations were still present. For that purpose, an anion exchange chromatography (AEC) has been performed, upon which the remaining contaminations were removed (SFigure 3 on p.135). The protein fractions indicated were subsequently concentrated to 3 ml and dialyzed over night with a 3 kDa cutoff dialysis membrane against 1 l AEC-1 buffer and then directly used for crystallization (see below).

As literature describes enzymatic assays for extracellular fungal proteins in acidic conditions to around pH 5, a buffer optimization screening was performed (Hurtado-Guerrero et al., 2009; Qin et al., 2015). Interestingly, the highest melting temperature was observed in the presence of 50 mM acetate buffered at pH 5.5, which is close to the pI of *Ct*Dfg5 (see Figure 16). Furthermore the addition of 2 mM DTT leads to a decrease in stability of 5 °C. This supports the suggestion that disulfide bonds are present in the protein, as seven cysteines are present in the amino acid sequence.

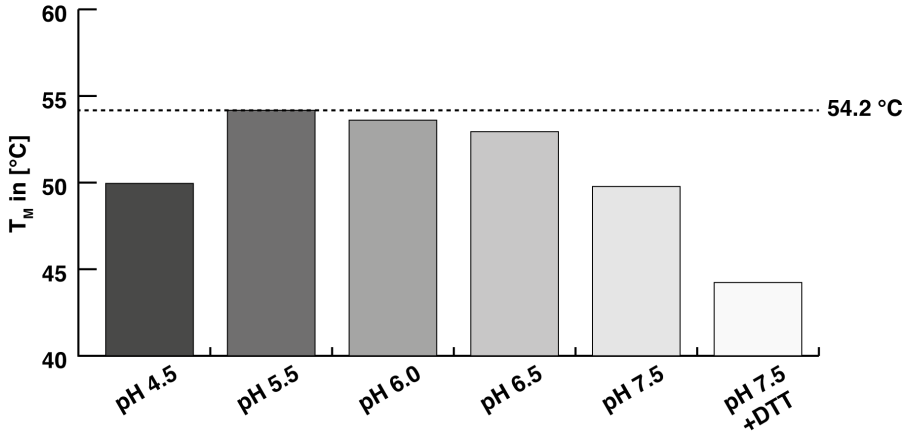


Figure 16: Thermal shift assay of *Ct*Dfg5. Different buffers with indicated pH-values have been used to determine the optimal conditions for *Ct*Dfg5. The highest melting temperature is indicated with the dashed line at 54.2 °C in the presence of pH 5.5.

4.2.2 Crystallization and structure determination of *Ct*Dfg5

In order to determine the first crystal structure of a fungal GH76 ortholog, *Ct*Dfg5 was concentrated to 30 mg/ml in 10 mM NaCl, 20 mM HEPES at pH 7.5 (AEC-1 buffer) and subsequently used for initial crystallization screening. Macroscopic crystals appeared after two days at 18 °C in the presence of 20 mM NH_4I and 20% (w/v) PEG3350 (JCSG Core IE4). The crystals appeared with a monoclinic shape and were sufficient after two weeks of growth for diffraction studies. The course of growth is shown in Figure 17.

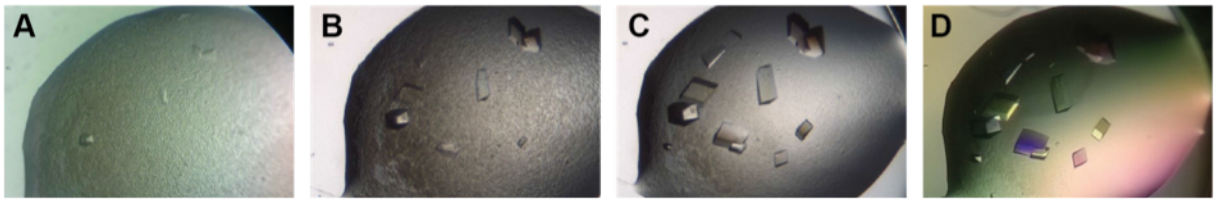


Figure 17: Crystallization of *Ct*Dfg5 in JSCG IE4. First crystals appeared after two days within amorphous precipitate (**A**). With advancing crystallization (**B** after four days, **C** after six days), the precipitate disappeared. **D** shows the crystals directly before freezing after two weeks of growth.

First diffraction studies resulted in a dataset with a ranking resolution of 1.2 Å, however with a rather low completeness (~84%). The space group as suggested from these data was the monoclinic C2 (complete: C 1 2 1). As the available homologs from the *bacterial* subfamilies in the PDB have low identities (3K7X 22%; 4BOK 20%; 4MU9 18%; 4C1S 17%; 4V1R ~16%), molecular replacement (MR) was not successful at this point. For phase

determination, the apo crystals were soaked in mother liquor with 50 mM $Gd(OAc)_3$ for 60 minutes, before they were shortly back soaked in mother liquor supplemented with 30% (v/v) glycerol as cryo protectant and flash frozen in liquid nitrogen. The diffraction studies were performed close to the L-II absorption edge of gadolinium (i.e. 1.56 Å) at a rotating CuKα-anode with a wavelength of 1.5418 Å, resulting in a dataset with significant anomalous signal to 1.96 Å and an overall resolution of 1.7 Å (see Table 5). A first model could be solved by SAD-phasing using the CRANK2 pipeline of the CCP4i2 suite (Potterton et al., 2018). A total number of 17 gadolinium atoms with an occupancy of at least 25% were found during substructure determination, which led to 404 amino acids that could automatically be build.

Table 5: Crystallographic data collection statistics of gadolinium soaked *Ct*Dfg5 crystals. Values in parentheses are for the highest resolution shell.

	***Ct*Dfg5$_{Gd}$**
Data collection	
X-ray source	Iμ^{TM} Mikrofocus source (Bruker)
Space group	C 1 2 1
Unit-cell parameters (Å, °)	*a*=83.65, *b*=54.82, *c*=80.51, *α*=90.00, *β*=91.14, *γ*=90.00
Wavelength (Å)	1.54
Resolution range (Å)	40.03-1.8 (1.84-1.8)
Completeness (%)	99.9 (99.9)
Observed reflections	207832 (12113)
Unique reflections	33956 (2030)
Multiplicity	6.1
Wilson *B* factor (Å^2)	19.15
R_{merge} (%)	7.9 (38.4)
Mean *I/σ(I)*	12.8 (3.3)
CC(½)	99.5 (89.9)
Anomalous completeness	99.8 (99.9)
Anomalous multiplicity	3.1 (3.0)
Anomalous CC(½)	53.7 (4.7)

As the pH has a huge influence on the stability of *Ct*Dfg5, further crystallization trials have been performed in the presence of 50 mM acetate, 100 mM NaCl at pH 4.5 to bias the final pH-values towards the protein's optimum. The initial crystals were used as a seed stock to increase the hit rate. *Ct*Dfg5 crystallized at RT in JCSG Core IE7 (0.2 M Ca-acetate, 0.1 M MES, pH 6.0, and 20% (w/v) PEG 8000) and IF8 (0.1 M Na-citrate, pH 5.5, 20% (w/v) PEG 3000) by mixing 300 nl of 30 mg/ml protein solution with 200 nl mother liquor and 100 nl of the seed stock (1:10 dilution) (Figure 18). The crystals could be reproduced by using the same seed stock for the entire duration of this work in hanging and sitting drop conditions with variable final volumes (0.6-2 µl). Like it was observed for the initial condition JCSG Core IE4, the protein crystallized in three-dimensional, monoclinically shaped crystals with very

good diffraction properties. Both conditions crystallized in space group C 1 2 1 with one monomer per asymmetric unit.

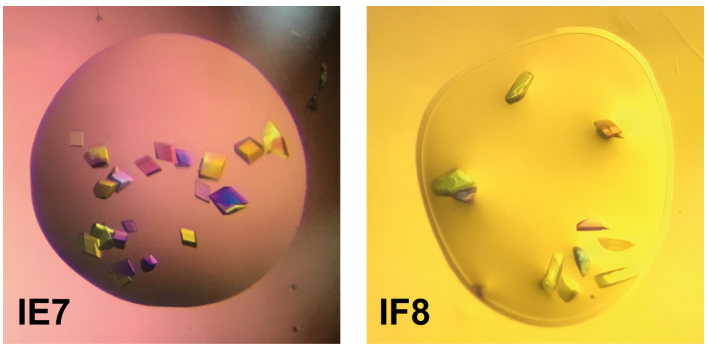


Figure 18: Crystallization of *Ct*Dfg5 in the presence of acetate buffer in JCSG Core IE7 and IF8.

The final dataset that resulted to the apo structure discussed in the following was obtained from soaking experiments of IE7-crystals with Frag37 from the *Frag Xtal Screen* (Jena Bioscience) at the European Synchrotron Radiation Facility (ESRF, Grenoble) (see Table 6). The structure has been solved by using the initial model from above for direct molecular substitution in *Phenix.refine*. Manual model building and refinement were performed with *Coot* and *Phenix.refine*, respectively (Afonine et al., 2012; Emsley et al., 2010).

Table 6: Crystallographic data collection and refinement statistics of the apo structure from *Ct*Dfg5. Values in parentheses are for the highest resolution shell.

	***Ct*Dfg5$_{apo}$**		
PDB code	6RY0		
Data collection		**Refinement**	
X-ray source	ID23-1, ESRF	Resolution range (Å)	30.28-1.05
Space group	C 1 2 1	R_{work}/R_{free} (%)	10.6/12.3
Unit-cell parameters (Å, °)	a=83.32, b=54.99, c=80.41, α=90.00, β=90.27, γ=90.00	Average B factor (Å^2)	13.4
Wavelength (Å)	0.972	No. of atoms	
Resolution range (Å)	30.28-1.05 (1.09-1.05)	Total	3935
Completeness (%)	97.20 (97.36)	Protein	3348
Observed reflections	297179 (28157)	Ligand	15
Unique reflections	164339 (16424)	Solvent	572
Multiplicity	1.8 (1.7)	R.m.s.d., bond lengths (Å)	0.013
Wilson B factor (Å^2)	9.66	R.m.s.d., bond angles (°)	1.62
R_{merge} (%)	1.9 (22.3)	Ramachandran plot	
R_{meas} (%)	2.7	Favored (%)	99
Mean $I/\sigma(I)$	17.06 (3.29)	Allowed (%)	1
$CC_{1/2}$ (%)	99.9 (88.4)	Disallowed (%)	0

4.2.3 The first crystal structure of a fungal GH76 protein

While bacterial GH76 proteins were shown to share their overall fold, general substrate specificity, and the catalytic DD-motif, the fungal *Dfg5*-subfamily remained uncharacterized (Cuskin et al., 2015; Thompson et al., 2015b). To gain detailed insights into that subclass, the crystal structure from the thermophilic fungus *Chaetomium thermophilum* has been determined at 1.05 Å resolution in space group C 1 2 1 with a continuous density map from residue Q32 till F441. The solvent content of the crystal is estimated to a low value of ~33%. Therefore, the tight packing of protein molecules within the crystal as shown in Figure 19B is not surprising. Ca^{2+} from the condition is traceable in electron density and responsible for crystal contacts between symmetry mates (Figure 19B). The ion is well coordinated by D79 from one molecule and E285 from its neighbor.

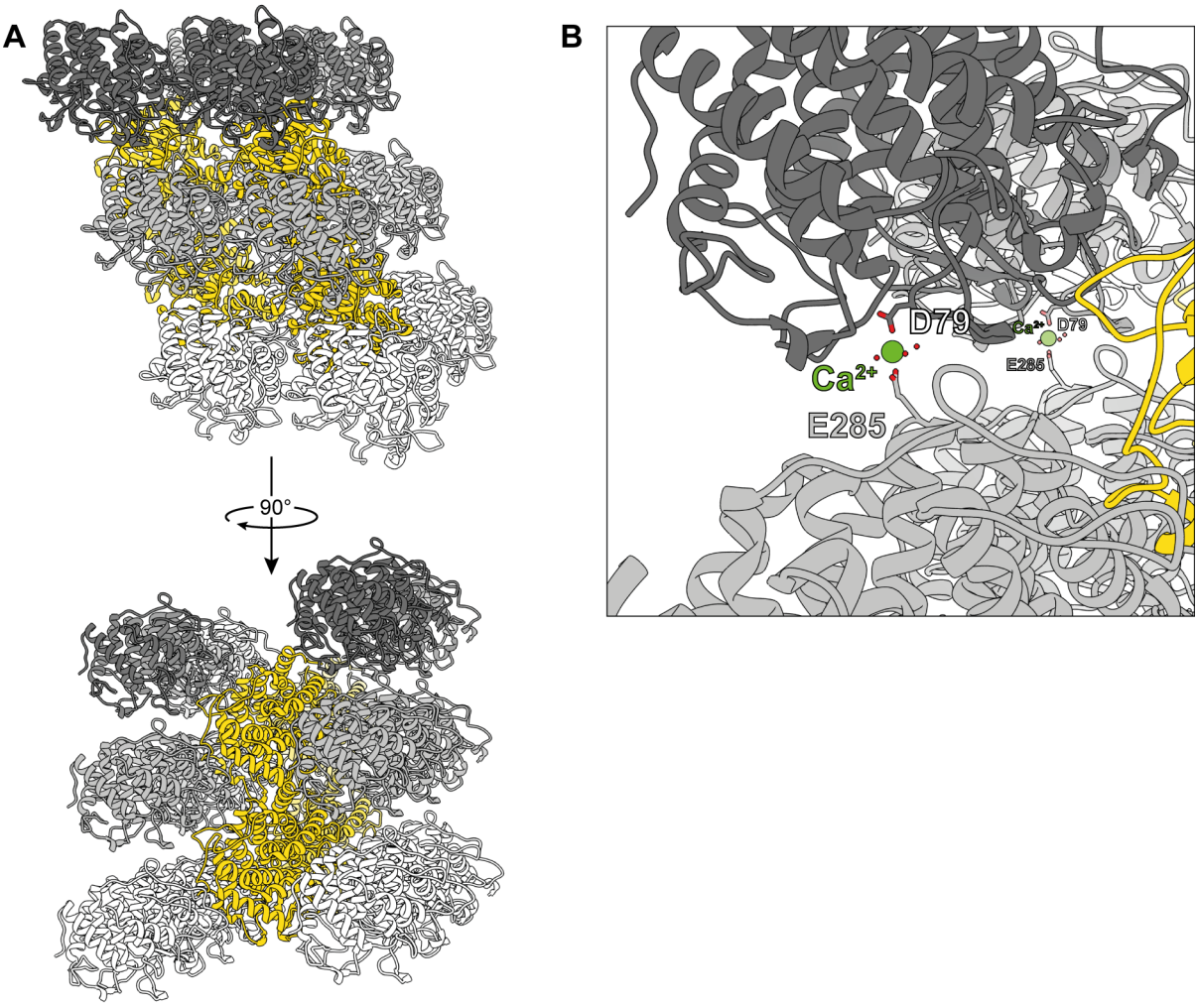


Figure 19: Crystal lattice of *Ct*Dfg5. A The arrangement of 16 molecules from *Ct*Dfg5 in the crystal lattice is shown from two sides as indicated by the arrow. The respective molecules are shown as ribbons. **B** Ca^{2+} (green sphere) establishes crystal

contacts between two molecules of *Ct*Dfg5 by being coordinated from D79 from one molecule and E285 from the neighboring lattice mate. Waters are shown as red dots.

The overall structure of *Ct*Dfg5 revealed a monomeric protein comprised of twelve GH76-characteristic α-helices arranged in an $(\alpha/\alpha)_6$-helical barrel fold (Figure 20).

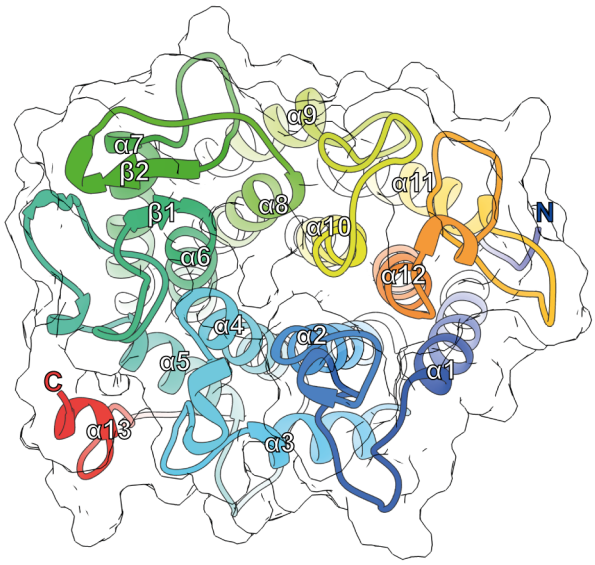


Figure 20: Overall structure of *Ct*Dfg5. The crystal structure of *Ct*Dfg5 is represented as a ribbon, which is colored from N- to C-terminus in rainbow colors from blue to red surrounded by white, transparent surface. The secondary structure elements are indicated.

A short alpha turn (α13) is formed at the C-terminus from A437 to the terminal F441. Beside the helical part, two β-strands are found between α5/6 and α7/8. As suggested from the thermal shift experiment (Figure 16), three disulfide bonds formed by cysteine pairs C182:C258, C315:C322, and C366:C375 stabilize the tertiary structure of *Ct*Dfg5.

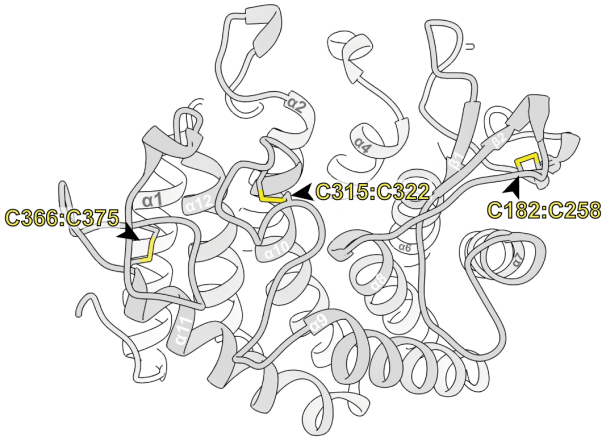


Figure 21: Disulfide bridges in *Ct*Dfg5. The three stabilizing disulfide bonds indicated by black arrows are shown as sticks in the context of the proteins secondary structure. Sulfur atoms are colored in yellow.

As mentioned above, the identities of already described bacterial homologs and *Ct*Dfg5 from the *Dfg5*-subfamily are at low values between ~15-20%. However, the superposition of *Ct*Dfg5 and *Bc*GH76 (PDB: 4BOK) as the representative for bacterial GH76 proteins shows a well conserved overall fold with an r.m.s.d. of 3.2 Å for 400 aligned Cα-atoms (see Figure 22).

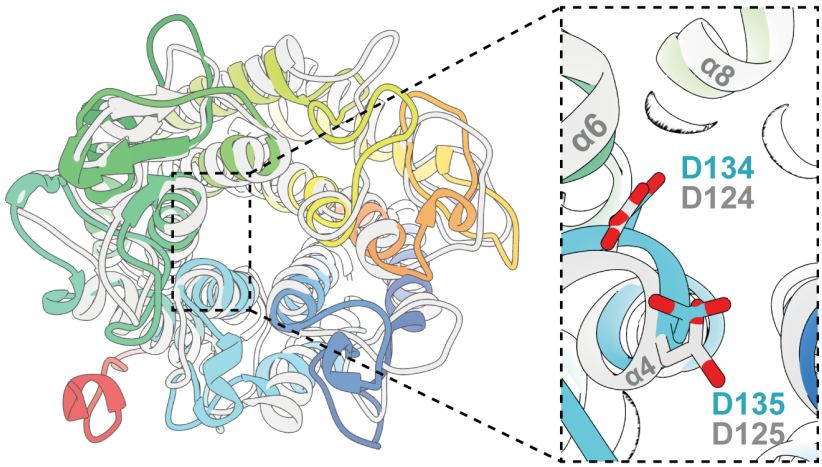


Figure 22: Comparison of bacterial and fungal GH76 proteins. A The conserved overall fold of GH76 proteins is represented by the superposition of *Ct*Dfg5 (rainbow) and *Bc*GH76 (grey, PDB: 4BOK) shown as ribbons. The dashed box shows the conserved position of the putative active site DD-motif represented as sticks and colored according to the secondary structure. The oxygen's are red.

The major difference of the overall structure lays in the extended C-terminus of *Ct*Dfg5. While the structure of the bacterial protein terminates at α12, *Ct*Dfg5 spans A413 to G436 over its back before it continues with α13. The active site of *Bc*GH76 was identified as a DD-motif consisting of the catalytic nucleophile D124 and the general acid/base residue D125 (Thompson et al., 2015b). In the structural alignment shown above, the active site of the bacterial protein and D134/D135 of *Ct*Dfg5 superimpose very well (box in Figure 22). Finding such significant overlapping features of *bacterial* and *Dfg5*-proteins, it raises the question, how the identified putative active site mediates the function of *Dfg5*-proteins, and how the conserved active site is consistent with different substrate specificity.

4.2.4 Bacterial and Fungal GH76 proteins act on different substrates

One hypothesis of how fungal GH76 members are involved in the incorporation of cell wall-resident proteins is based on the annotated α1,6-mannanase activity of the described bacterial homologs (Cuskin et al., 2015; Maddi et al., 2012; Thompson et al., 2015b). This model combines similar defects of the cell wall in *Δoch1* mutants, which is responsible for the transfer of the initial mannose of the N-glycan α1,6-mannose backbone, and *Δdcw1/Δdfg5* deletion mutants in *Neurospora crassa* and *Candida albicans* (Ao et al., 2015; Maddi and Free, 2010; Maddi et al., 2012). The authors suggested that *Dfg5*-proteins manage the cell wall incorporation of GPI-APs by cleaving the α1,6-mannose backbone of N-linked outer chain mannan, as it is known from the bacterial homologs, and subsequently transfer them onto the cell wall glycan by a so far unknown mechanism. Accordingly, *Ct*Dfg5 should be able to catalyze the first reaction step, i.e. the cleavage of an α1,6-mannooligomer with a minimum length of a triose, as found for the bacterial member from *Bacillus circulans, Bc*GH76 (Thompson et al., 2015b).

In order to test that hypothesis *in vitro*, *Ct*Dfg5 reaction mixtures were incubated with α1,6-linked manno-oligosaccharides (*kind gift of Dr. Daniel Varón Silva, MPI Potsdam*) and analyzed by thin layer chromatography (TLC) (Figure 23A). However, aside from a mannose contamination within the mannotriose sample that is present in both, the negative control and the respective reaction, no migration shift or a formed product could be observed.

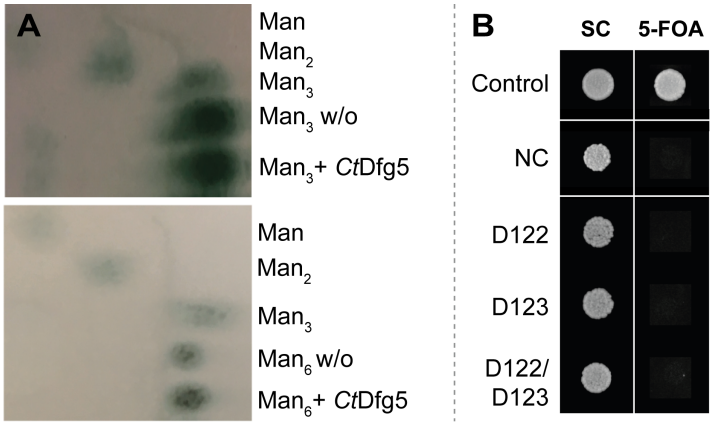


Figure 23: TLC of manno-oligosaccharides and *in vivo* test of the putative active site. A Upper panel: Mannose (Man), α1,6-linked mannobiose (Man_2), and α1,6-linked mannotriose (Man_3) are shown as standards. The incubated samples without (w/o) and with (+*Ct*Dfg5) protein are shown below. Lower panel: Size standards are shown as above. Here, an α1,6-linked mannohexaose (Man_6) has been analyzed. **B** A plasmid-loss assay shows the essential role of the putative active site residues D122 and D123 from *S. cerevisiae*. Gesa Schmitz (AG Mösch) performed the *in vivo* work. SC: minimal medium for *S. cerevisiae*. 5-FOA: SC-medium supplemented with 5-Fluoroortic acid. NC: negative control.

The second route using the GPI-anchor as the unit to transfer proteins into the cell wall is based on the findings of GPI-remnants found in HF-threated cell walls and the simultaneous release of proteins, which are predicted to possess an ω-signal at their C-terminus that is necessary for posttranslational modification of a GPI-anchor (Fujii et al., 1999; Kollár et al., 1997). Together with the findings of GPI-anchored proteins released into the medium in double knockouts of *Δdcw1*/*Δdfg5* in *Saccharomyces cerevisiae*, it is proposed that the transfer is catalyzed by *Dfg5*-proteins cleaving the Manα1,4-GlcN linkage within the GPI-anchor (see Figure 24) (Kitagaki et al., 2002). Unfortunately, such a substrate is not commercially available and the synthesis is difficult to realize, even for experts. Consequently, it was not possible to test hydrolysis or defective mutants *in vitro* due to the lack of a suitable substrate. However, it has been suggested by the SSN analysis (Figure 14 and Table 2) that the proteins involved in the lipid-to-wall transfer are all resident in the *Dfg5*-subfamily, including Dcw1 and Dfg5 from *S. cerevisiae*. Working in the living cell would not only enable us to analyze structure-function relationships in the context of the living cell and thereby bypass the lack of substrate, but also make a more general statement for the whole *Dfg5*-subfamily, as the predictions made on the basis of a protein from *Chaetomium thermophilum* are examined in another fungal species. Therefore, a collaboration with the group of Prof. Hans-Ulrich Mösch has been started. Gesa Schmitz established the double knockout strains of *Δdcw1*/*Δdfg5* in *S. cerevisiae* and performed the *in vivo* work shown in this thesis.

As shown in Figure 22, the DD-motif identified as the active site in bacterial GH76 proteins is also highly conserved in *Ct*Dfg5. Thus, the residues in Dfg5 from *S. cerevisiae* (*Sc*Dfg5) have been specified as residues D122/D123 (see Figure 28A on p.74) and the respective active-site mutants D122N, D123N and D122N/D123N were analyzed concerning their ability to complement the wildtype protein in a plasmid-loss assay (Figure 23B). As expected, this is not possible as their function should be absolutely essential for the organism in a double-knockout background.

Resulting from the first functional results from *Dfg5*-proteins, the active site seems to be highly conserved throughout the whole GH76 family, however at least the so far characterized bacterial homologs and those from the *Dfg5*-subfamily do not share the same substrate specificity. Following this, the next part of this work will deal with the identification of the real substrate to shed light on the last step of cell wall protein maturation.

4.2.5 Mapping the GPI-core glycan by sugar-fragment screening

To learn more from a structural perspective, substrate bound states are sought to be determined. The true substrate of *Dfg5*-proteins might be the GPI-anchor, as its mannose core structure is found covalently attached to the glucan matrix in the fungal cell wall and the delivery of GPI-AP into the CW is compromised in *Δdcw1/Δdfg5* promoter shutoff cells (Kitagaki et al., 2002; Kollár et al., 1997). As GPI-anchors or derived carbohydrates are not commercially available, the primary goal was to utilize sugar fragments for cocrystal structures with *Ct*Dfg5 (Figure 24).

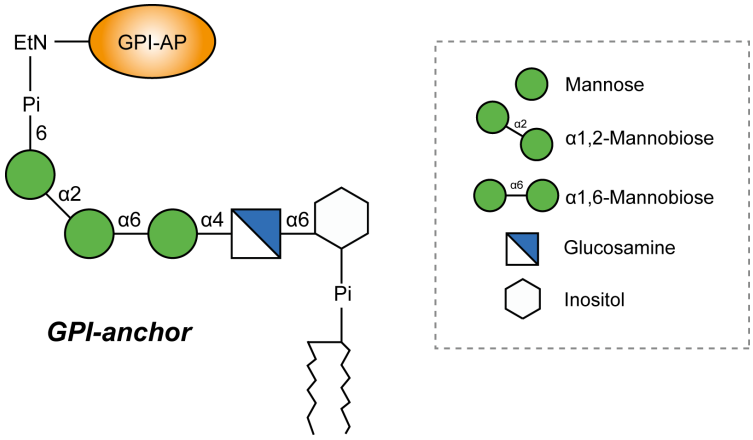


Figure 24: Schematic representation of the GPI-anchor and its tested fragments. The core structure of the GPI-anchor is shown with the conventionally used symbols for glycan representation with the respective glycosidic bonds being indicated. The edged line shows the lipid moieties. The box on the right shows the respectively used fragments in the crystallographic experiment. *GPI-AP: glycosylphosphatidylinositol-anchored protein*; *EtN: Ethanolamine phosphate*; *Pi: inorganic phosphate*.

To elucidate the structure bound state of *Ct*Dfg5, the protein was crystallized like described for the apo structure and used after one week of growth for further soaking experiments with different mono-, di-, and higher oligosaccharides. The sugars were solved in a 1:1 mixture composed of JCSG Core IE7 and 100 mM NaCl buffered with 50 mM acetic acid at pH 4.5 to final concentrations of 0.5-1 M. The apo crystals were transferred in 1 µl of the respective solution and incubated for several hours. The exact concentrations and durations for the individual structures are shown in Table 7.

Table 7: Sugar concentrations and soaking times used for successful cocrystal structures of *Ct*Dfg5.

	Man	**α1,2-Man$_2$**	**α1,6-Man$_2$**	**GlcN**
Concentration	0.5 M	0.5 M	1 M	0.5 M
Duration	4 h	4 h	8 h	4 h

A critical point for successful soaking was to omit backsoaking after the incubation. Due to the similar crystal lattice parameters, the structures were also solved by molecular substitution using the apo structure as template. The collection and refinement statistics are shown in Table 8. While the soaking was successful for the GPI-glycan fragments listed above, soaking with inositol under the similar conditions did not result in additional electron density. The data collection statistics are given in Table S 1 on p.144.

Table 8: Crystallographic data collection and refinement statistics of the fragment bound states from *Ct*Dfg5. Values in parentheses are for the highest resolution shell.

	***Ct*Dfg5$_{Man}$**	***Ct*Dfg5$_{α1,2M}$**	***Ct*Dfg5$_{α1,6M}$**	***Ct*Dfg5$_{GlcN}$**
PDB code	6RY1	6RY2	6RY5	6RY6
Data collection				
X-ray source	ID23-1, ESRF	ID23-1, ESRF	ID29, ESRF	ID23-2, ESRF
Space group	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
Unit-cell parameters (Å, °)	*a*=83.86, *b*=55.02, *c*=80.49, *α*=90.00, *β*=90.45, *γ*=90.00	*a*=83.72, *b*=54.86, *c*=80.13, *α*=90.00, *β*=90.45, *γ*=90.00	*a*=83.65, *b*=54.82, *c*=80.51, *α*=90.00, *β*=91.14, *γ*=90.00	*a*=83.42, *b*=55.02, *c*=80.12, *α*=90.00, *β*=90.37, *γ*=90.00
Wavelength (Å)	0.972	0.972	1.54	0.873
Resolution range (Å)	41.93-1.3 (1.35-1.3)	41.86-1.3 (1.35-1.3)	45.93-1.3 (1.35-1.3)	40.25-1.3 (1.35-1.3)
Completeness (%)	99.76 (99.89)	94.32 (91.56)	96.40 (95.07)	96.27 (94.52)
Observed reflections	175788 (17387)	161441 (15866)	159977 (15490)	166168 (15984)
Unique reflections	89845 (8959)	84221 (8154)	85994 (8464)	86433 (8481)
Multiplicity	2.0 (1.9)	1.9 (1.9)	1.9 (1.8)	1.9 (1.9)
Wilson *B* factor (Å^2)	10.47	10.59	13.03	10.66
R_{merge} (%)	2.7 (16.2)	6.0 (11.6)	2.0 (9.4)	2.5 (14.0)

R_{meas} (%)	3.8	8.4	2.8	3.5
Mean *I/σ(I)*	15.10 (4.25)	8.63 (6.01)	20.02 (6.93)	15.54 (4.51)
CC(½)	99.9 (92.8)	99.0 (94.7)	99.9 (97.5)	99.9 (94.6)
Refinement				
Resolution range (Å)	41.93-1.3	41.86-1.3	45.93-1.3	40.25-1.3
R_{work}/R_{free} (%)	9.6/12.5	10.5/12.9	10.2/12.3	11.3/12.6
Average *B* factor (Å^2)	16.0	15.8	17.5	15.1
No. of atoms				
Total	4070	4078	3811	3947
Protein	3335	3335	3410	3331
Ligand	28	39	35	23
Solvent	707	704	366	593
R.m.s.d., bond lengths (Å)	0.010	0.011	0.013	0.011
R.m.s.d., bond angles (°)	1.51	1.52	1.62	1.44
Ramachandran plot				
Favored (%)	99	99	99	99
Allowed (%)	1	1	1	1
Disallowed (%)	0	0	0	0

As shown in Figure 24, the core component of the GPI-anchor is composed of three mannoses, glucosamine and inositol. While the latter one could not be observed in the respective soaking experiment, mannose alone could be unambiguously identified at GPI-glycan binding site -2 in the substrate binding canyon of *Ct*Dfg5 (see Figure 25A). It is well bound by electrostatic and hydrophobic interactions mainly from residues of helices α6, α8, and α10. Soaking with α1,2-mannobiose led to a complex structure that also occupies subsite -2, which is now extended to subsite -3 (Figure 25B). The -3 mannose is bound by fewer interactions than the -2 sugar and it is more surface exposed, however kept in place by an H-bond between its OH3-group and N262. By soaking the crystals with α1,6-mannobiose, the central mannose could further be expanded to subsite -1 (Figure 25C). The -1 sugar sits deeply in the binding pocket of *Ct*Dfg5 above the active site residue D134 and is stacked by W83. Interestingly, its OH1 and OH2 groups replace waters that used to coordinate a Ca^{2+} ion, which is coordinated above D135 and surrounded by three waters roughly on the plane of the hydroxyl groups of the sugar and a water opposing to D135. To fit the stereochemical properties of the -1 moiety, a β-mannose had to be modeled into the electron density, while all other mannoses were found in the α-state. The rearrangement probably happens due to the coordinated calcium. In fact, this kind of coordination seems to provide no further insights into the true distortion of the -1 position during catalysis, as this would also require the presence of the +1-sugar. This position however could be attributed to glucosamine, which has been found at GPI-glycan subsite +1 above the general acid/base residue D135 (Figure 25D). The sugar is bound by H-bonds with Q189 from the front and stacking interactions of Y81 from the back. Further electrostatic interactions are established by D325 and S130.

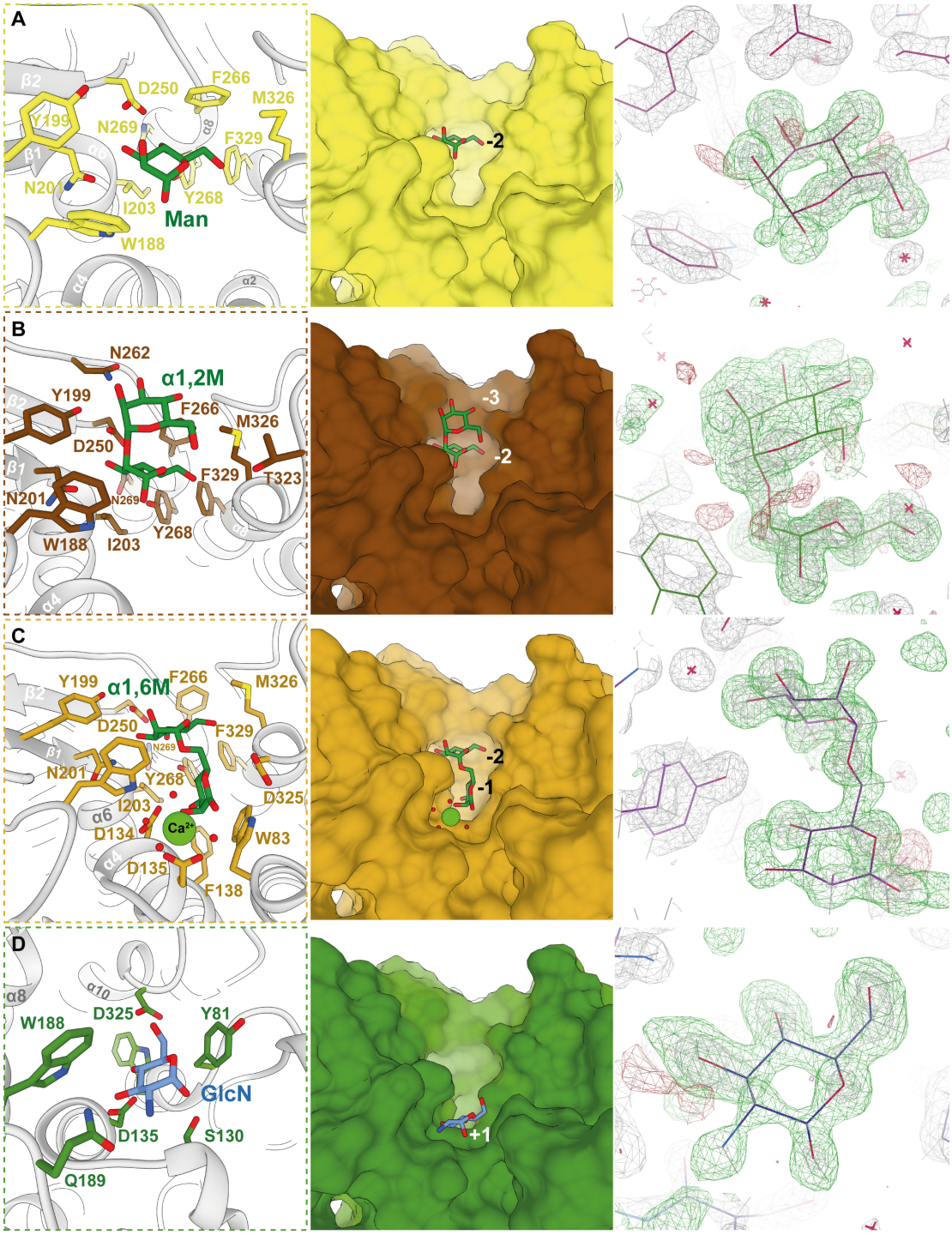


Figure 25: Sugar fragments resemble the GPI-anchor core in the binding pocket of *Ct*Dfg5. **A** The monosaccharide mannose (Man) is bound to the sugar-binding site -2 in the binding pocket of *Ct*Dfg5. **B** That central mannose could be extended by α1,2-mannobiose (α1,2M) to subsite -3 and with **C** α1,6-mannobiose (α1,6M) to the -1 site on top of the active site residues D134/D135 (C). **D** Glucosamine (GlcN) is bound to the +1 sugar-binding site. On the left, the sugar and its coordinating residues are presented as sticks colored with oxygens in red, nitrogens in blue, sulphur in yellow, and carbon atoms of Man in green, of GlcN in light blue and of the protein in the color given. The arrangement is shown in the context of the secondary structure elements, which are colored in grey. Ca^{2+} is shown as a light green sphere with its coordinating waters as red dots. In the middle panel, the sugar is shown as in the left panel in the context of the surface from *Ct*Dfg5 in the respective color scheme. The sugar-binding sites are indicated. In the right panel the OMIT-maps of the respective carbohydrate are shown with the difference map colored in green and contoured at σ-level 3.0.

By superimposing the four different states shown in Figure 25 it is possible to reconstruct the GPI-core structure from glucosamine to mannose 3 within the binding pocket of *Ct*Dfg5 (Figure 26). The GPI-core occupies the GPI-glycan binding sites from -3 to +1 and is found in a *U-* or *C-shaped* conformation that very well fills out the complete binding pocket.

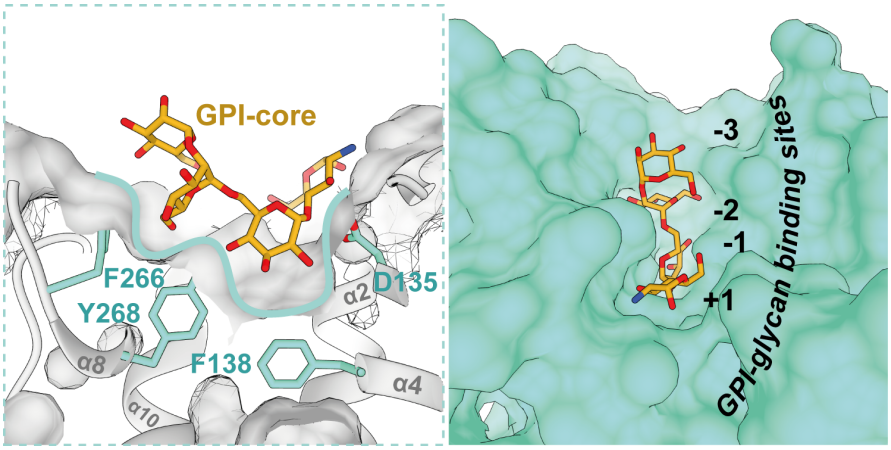


Figure 26: Assembled GPI-core bound to the binding pocket of *Ct*Dfg5. From the top view on the right with all four GPI-glycan binding subsites labeled, the GPI-core is shown on the left from the left site with the surface being capped in front of the binding pocket, which is highlighted by the turquoise line. Residues (turquoise sticks) and secondary structure elements (grey ribbon) are shown for orientation, the GPI-core is presented as orange sticks with red oxygens and blue nitrogen.

With this finding, the exact role of *Dfg5*-proteins seem to be clear, i.e. utilization of the GPI-anchor as a donor substrate to transfer CW-resident proteins from the plasma membrane into the wall. The differences and similarities between GH76 proteins from the *Bacteria II* and *Dfg5*-subfamilies concerning their substrate binding mode are shown in the following and allow an explanation, why *Ct*Dfg5 is incapable of α1,6-mannan hydrolysis.

4.2.6 Structural differences between bacterial and fungal GH76 proteins

In 2015, the first crystal structure of a GH76 protein in complex with its substrate α1,6-mannopentaose spanning over the whole binding canyon from subsites -4 to +1 has been published (PDB: 5AGD) (Thompson et al., 2015b). This structure can now be used to compare the substrate bound states of two members of different GH76 subfamilies. In Figure 27A, *Ct*Dfg5 is shown in a similar manner as in Figure 26, but superimposed with the complex structure of *Bc*GH76. It clearly shows a quite similar binding mode of the α1,6-linked mannoses at subsites -2 and -1 with a slightly tilted mannose 2, however it reveals huge differences in the binding modes of the peripheral glycan-binding sites. This becomes

already obvious looking at the predicted clash of the mannopentaose with the surface of *Ct*Dfg5 at subsites -4 and -3 (Figure 27A).

A closer look on that region is given in Figure 27B. The upper panel shows the binding mode of the GPI-core in *Ct*Dfg5, the lower panel that of *Bc*GH76. While the interactions at subsite -2 are conserved by the tyrosine interaction from below (Y268 and Y243) and from the front by an aspartate (D250 and D228), they differ in the interactions involved in the binding site close to C6 of mannose 2. One striking example is the already mentioned clash of the mannopentaose with the surface of *Ct*Dfg5. Responsible for this clash is F266, which stacks the C6 atom of mannose 2. This leads to the OH6-group swiveling away from F266. Consequently, an α1,2-linkage to the mannose 3 of the GPI-core must be formed to circumvent the hydrophobic block that is different in *Bc*GH76, where a threonine (T241) is present instead of the phenylalanine. In the latter enough space for the α1,6-linkage is provided to continue the mannopentaose in a linear manner. Furthermore, residues from α-helix 10 also play an important role. The arrangement in *Ct*Dfg5 shows a highly conserved F329, which is guiding M326 around the C6-OH6 configuration as described above. In *Bc*GH76, this phenylalanine is substituted by a glycine and the methionine position shows a leucine (L295). The free space given by the glycine is used by L295 to stack the C6 atom of the -2 mannose, while it clashes with OH6 when superimposing the GPI-core glycan.

The other site of the binding pocket above the active site is compared in Figure 27C. Again, *Ct*Dfg5 is shown in the upper and *Bc*GH76 in the lower panel. In both structures, an H-bond is formed by an aspartate (D325 and D294) sitting above α10 with the sugar moiety at subsite +1. In *Ct*Dfg5, the α1,4-linked glucosamine is bound from the front by electrostatic interactions of Q189, whereas from the back the C6 region is stacked by Y81. This binding mode is reversed in *Bc*GH76, as the +1 mannose is α1,6-linked with the -1 mannose. Accordingly, the stacking interaction is established by F122 from the front, while an electrostatic interaction is realized from the back by D71.

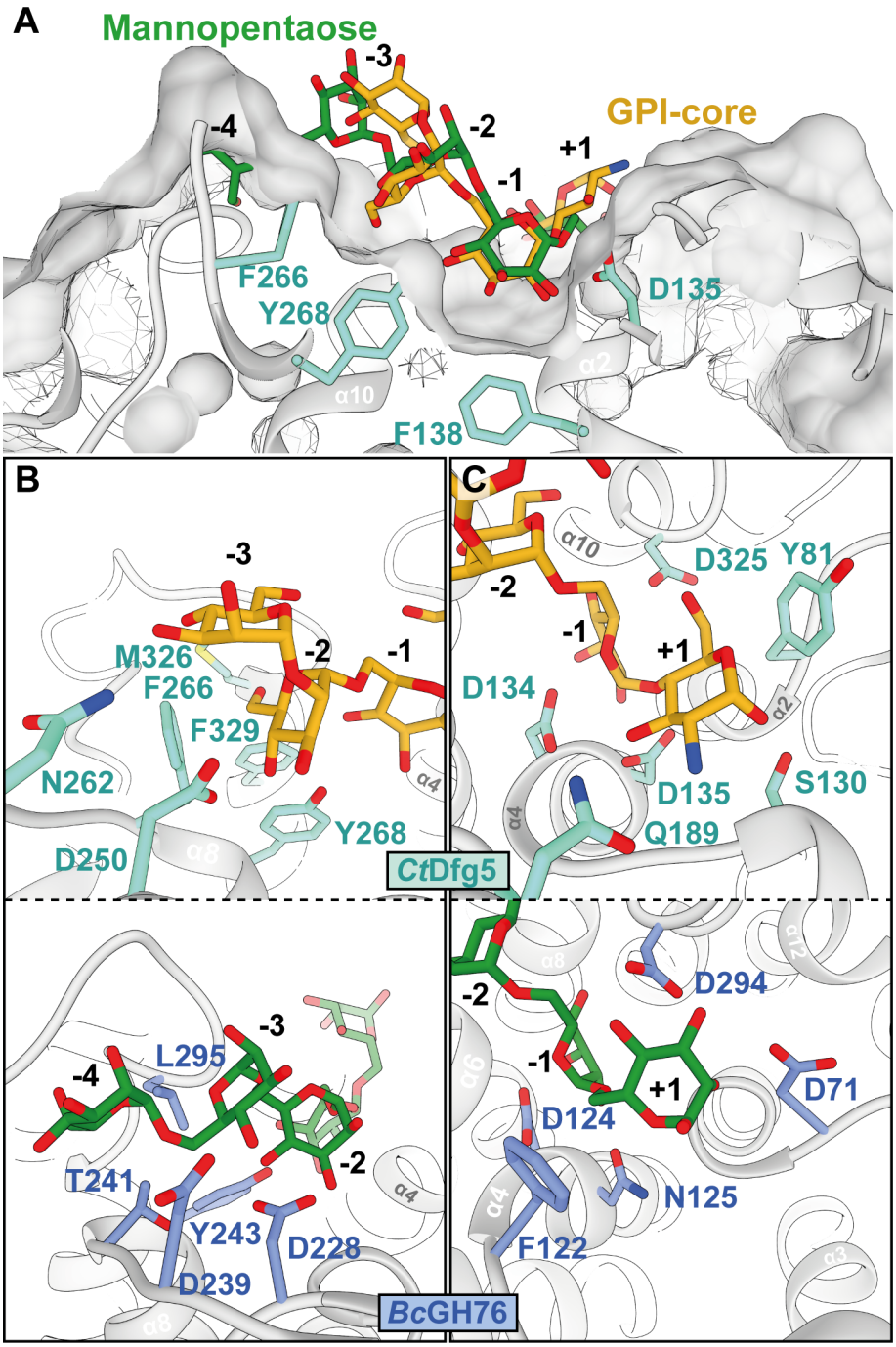


Figure 27: Comparison of *Ct*Dfg5 and *Bc*GH76 in their substrate bound states. A The structure of *Bc*GH76 in complex with α1,6-mannopentaose (PDB: 5AGD) is superimposed with *Ct*Dfg5 and just the substrate (green sticks) is shown in the context of *Ct*Dfg5. The capped surface and secondary structure (both in grey), turquoise sticks for the residues and the orange GPI-core shown as sticks are presented of *Ct*Dfg5. **B** The upper panel shows the -1 to -3 sugar binding site of *Ct*Dfg5, the lower panel the -2 to -4 region in *Bc*GH76. **C** The same arrangement as in **B** shows the substrate coordination at subsite +1. The shown secondary structure elements, the residues, and the sugar binding sites are indicated. The coloring in **B** and **C** is like in **A** with bluish sticks for the residues of *Bc*GH76.

4.2.7 The conserved substrate binding pocket of *Ct*Dfg5

The analysis of evolutionary conserved and variable regions of proteins is a handy way to predict important residues and regions of a protein. On the left side of Figure 28A, a *ConSurf*-analysis of the *Dfg5*-family is plotted on the surface of *Ct*Dfg5 (Ashkenazy et al., 2016). It clearly shows the importance of the substrate-binding pocket, as almost all residues are attributed the highest conservational score possible (SFigure 4 on p.136). Further mutational studies based on this should strengthen the role of the substrate-coordinating residues, however as already explained above *in vitro* studies with defined substrates could not be performed due to a lack of suitable substrate. Like it was performed for the active site motif, structure-based mutants were planned according to the homology model of *Sc*Dfg5 (right side of Figure 28A) and analyzed *in vivo* by Gesa Schmitz (Figure 28B).

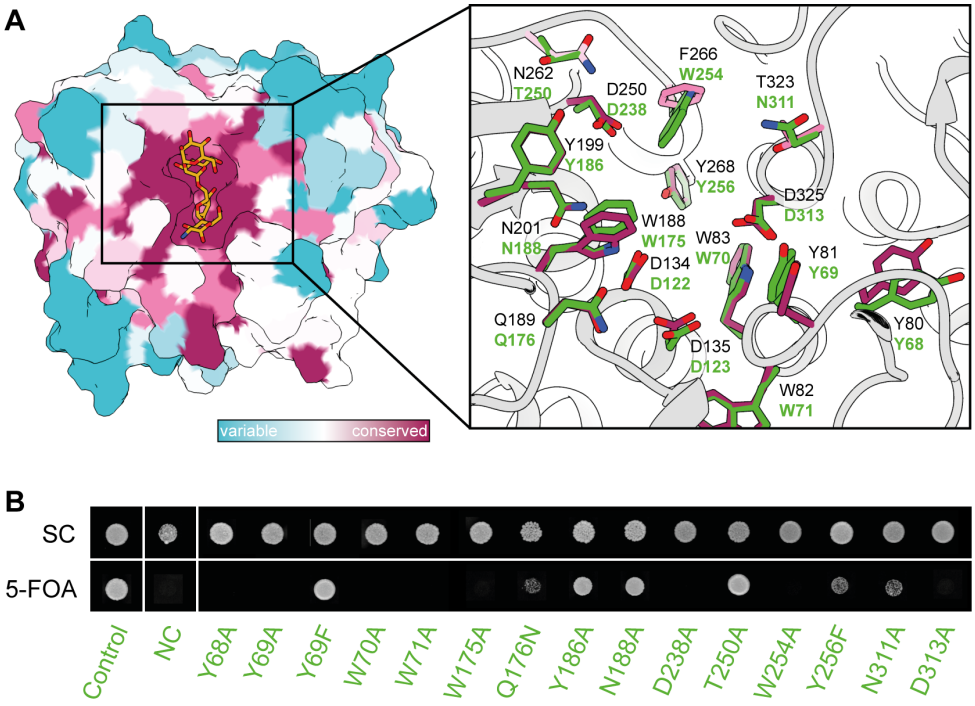


Figure 28: Structure-based mutants show the important role of residues in the binding pocket of *Dfg5*-proteins. A The surface of *Ct*Dfg5 is shown and colored according to the legend from variable (cyan) to conserved (dark magenta) residues. The assembled GPI-core is shown as orange sticks. The box on the right shows the binding pocket residues of *Ct*Dfg5 (*ConSurf* colors, indicated in black letters) superimposed with a structural model of *Sc*Dfg5 (green, indicated in green letters) in the context of the secondary structure of *Ct*Dfg5 (grey). **B** The structure-based mutants are analyzed in a plasmid-loss assay in *S. cerevisiae* on SC-medium with and without 5-FOA. The type of mutant is indicated. *In vivo* work has been performed by Gesa Schmitz (AG Mösch).

The plasmid-loss assay identified eight of 15 tested mutants as essential (Y68A, Y69A, W70A, W71A, W175A, D238A, W254A, D313A) for the viability of *Saccharomyces*

cerevisiae. Interestingly, the essential phenotype of Y69A is not given for the Y69F mutant, strengthening the importance of the glucosamine interaction described above by *Ct*Dfg5's Y81. Apparently, the additional OH-group does not seem to be necessary, although tyrosine is fully conserved in the *Dfg5*-subfamily. On the other side of the glucosamine interaction, the rather conservative mutation of Q176N (*Ct*Dfg5: Q189) leads to a reduced growth rate, pointing out the importance of correct placements of the functional group of glutamine. In contrast to Y69, the OH group of Y256 (*Ct*Dfg5: Y268) seems to play a more important role, as the Y256F mutant exhibits a reduced growth. Interestingly, the hydrophobic block at mannose 2, here established by W254 instead of F266 in *Ct*Dfg5, is essential for correct protein function, while the N188A mutant, which is also involved in the binding of mannose 2, lacks any effect. Two residues coordinating at subsite -3 have been mutated as well. While T250A shows unaltered growth behavior in the spot assay, the growth rate of N311A is reduced.

4.2.8 *Ct*Dfg5 hydrolyze GPI-core of soluble GPI-APs from *S. cerevisiae*

Taken the structural findings of *Ct*Dfg5 and its structure-based analyses together, it becomes obvious why the hydrolysis of α1,6-linked oligomannoses failed. Unfortunately, in terms of GPI-core glycan cleavage, any type of potential donor substrate is commercially not available. To show GPI-hydrolysis anyway, an *in vitro* assay with the following rational has been elaborated (see Figure 29):

As the GPI-core state provides no hint for the requirement of species-specific GPI-anchor modifications due to a lack of further space in the binding pocket of *Ct*Dfg5, it can be assumed that any GPI-anchor with that core structure should be suitable as substrate. Such GPI-anchors are synthesized by *S. cerevisiae*, where 80% contain an additional α1,2-linked mannose at Man3 and some anchors possess a fifth mannose (Fankhauser et al., 1993). Nevertheless, these additional mannoses should not affect the binding within the pocket. Therefore, crude membrane extracts from *S. cerevisiae* were prepared, followed by a *Bacillus cereus*-PI-PLC treatment, which is known to release GPI-APs from membranes in a Ca^{2+}-independent manner and thereby producing a cyclic phosphate (see Figure 29, step 1) (Heinz et al., 1996). Upon centrifugation the supernatant contained soluble GPI-APs (sGPI-AP) such as Gas1 (Uniprot: P22146), Gas3 (Uniprot: Q03655) Gas5 (Uniprot: Q08193), Ecm33 (Uniprot: P38248) and YJL171C (Uniprot: P46992), as identified by tryptic digests using ESI-MS (Table S 2 on p.144). In step 2, *Ct*Dfg5 is incubated with the supernatant,

which is subsequently (step 3) filtered through a 3 kDa MWCO cutoff filter membrane. In this step, the hydrolysis of the Manα1,4-GlcN linkage within the GPI-core should occur. The flow through with molecules smaller than 3 kDa was again analyzed by ESI-MS that should find the resulting GlcN-Ino-Pi with a theoretical mass of 402.08 Da. Indeed, such a molecular weight could be identified (step 4, SFigure 5 on p.137). Furthermore, the flow through has been analyzed by Dr. Xiulan Xie from the NMR-facility concerning characteristic chemical shifts of GlcN-Ino-Pi atoms (step 5). One is a single anomeric carbon, which is known to show up in a certain region in *2-D* ^{1}H ^{13}C spectra, while cyclic phosphates are identified by a chemical shift around 16 ppm with ^{31}P-NMR in the literature (Bubb, 2003; Chakraborty and d'Alarcao, 2005; Hecht et al., 2010; Schofield et al., 2002). Both, the specific region within the *2-D* spectrum and the cyclic phosphate could be observed in this analysis (SFigure 6 and SFigure 7 on pp.138-139). This strongly suggests that *Ct*Dfg5 is able to hydrolyze the Manα1,4-GlcN linkage in the GPI-core for the subsequent transfer onto an acceptor sugar in the fungal cell wall.

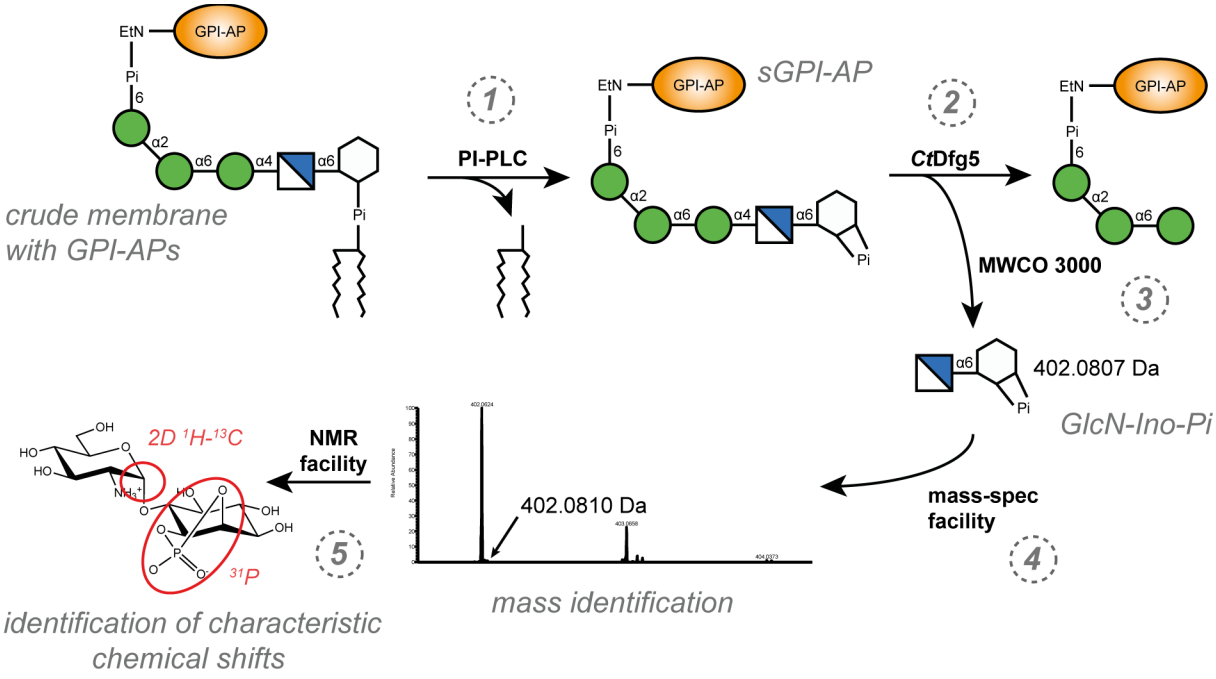


Figure 29: The rational principle of the *in vitro* assay. The workflow to obtain the product is shown. After membrane purification, PI-PLC (1) is used to release GPI-AP from the lipid moiety and thereby producing a cyclic phosphate. Upon ultracentrifugation, the supernatant undergoes a *Ct*Dfg5 treatment (2) that releases a 402.08 Da large disaccharide-phosphate. Subsequently the reaction is filtered with a 3 kDa cut off filter membrane (3) and the flow through is analyzed by mass spectrometry (4) and NMR-facilities (5).

4.2.9 Structural base of the transfer mechanism catalyzed by *Dfg5*-proteins

From *Saccharomyces cerevisiae* it is known that GPI-CWPs are attached to the wall via β1,6-glucan (Gonzalez et al., 2009). However, the situation in other fungi is less clear, as the CW architecture differs from species to species. While *Candida* spp. seem to have a similar ensemble, *Aspergillus fumigatus* apparently lacks the presence of β1,6-glucans, but contains differently linked glycans. Therefore it is hard to predict, what the actual attachment-glycan of CWPs in the respective species is.

In order to answer the question of the acceptor identity for *Ct*Dfg5, possible acceptor-molecules were used again for soaking experiments that were performed as described before. While soaking with 0.5 M gentiobiose (Glcβ1,3-Glc) did not result in additional density, soaking of 1 M laminaribiose (Glcβ1,3-Glc) for 8 h resulted in another complex structure of *Ct*Dfg5 (see Table 9 and Figure 30A).

Table 9: Crystallographic data collection and refinement statistics of *Ct*Dfg5 soaked with laminaribiose. Values in parentheses are for the highest resolution shell.

	***Ct*Dfg5$_{β1,3G}$**		
PDB code	6RY7		
Data collection		**Refinement**	
X-ray source	ID23-1, ESRF	Resolution range (Å)	40.05-1.3
Space group	C 1 2 1	R_{work}/R_{free} (%)	11.9/9.4
Unit-cell parameters (Å, °)	*a*=83.36, *b*=54.88, *c*=80.10, *α*=90.00, *β*=90.45, *γ*=90.00	Average *B* factor (Å^2)	16.1
Wavelength (Å)	0.972	No. of atoms	
Resolution range (Å)	40.05-1.3 (1.35-1.3)	Total	3866
Completeness (%)	97.41 (97.52)	Protein	3332
Observed reflections	169436 (16736)	Ligand	34
Unique reflections	86591 (8648)	Solvent	500
Multiplicity	2.0 (1.9)	R.m.s.d., bond lengths (Å)	0.009
Wilson *B* factor (Å^2)	11.99	R.m.s.d., bond angles (°)	1.41
R_{merge} (%)	2.6 (9.9)	Ramachandran plot	
R_{meas} (%)	3.7	Favored (%)	99
Mean *I/σ(I)*	16.58 (6.88)	Allowed (%)	1
$CC_{1/2}$ (%)	99.8 (97.4)	Disallowed (%)	0

Laminaribiose could unambiguously be identified in the electron density, although the +1 subsite is much clearer than the +2 acceptor-subsite (Figure 30A+B). However, the β-linkage as well as C1, C2 and its hydroxy-group can be well seen and lowering the σ-level leads to the appearance of the full pyranose ring. This can be explained by partial occupation of the +2 subsite, which is only loosely bound by two loop regions termed as *acceptor binding loops* (ABL) 1 and 2 (Figure 30B+C). These regions are not conserved in *Dfg5*-proteins, but shape the conserved *canyon* that is suggested to harbor the PI-binding site. Strikingly, the non-

reducing sugar moiety of laminaribiose adapted a conformation quite similar to the observed +1-subsite conformation of glucosamine (Figure 30C).

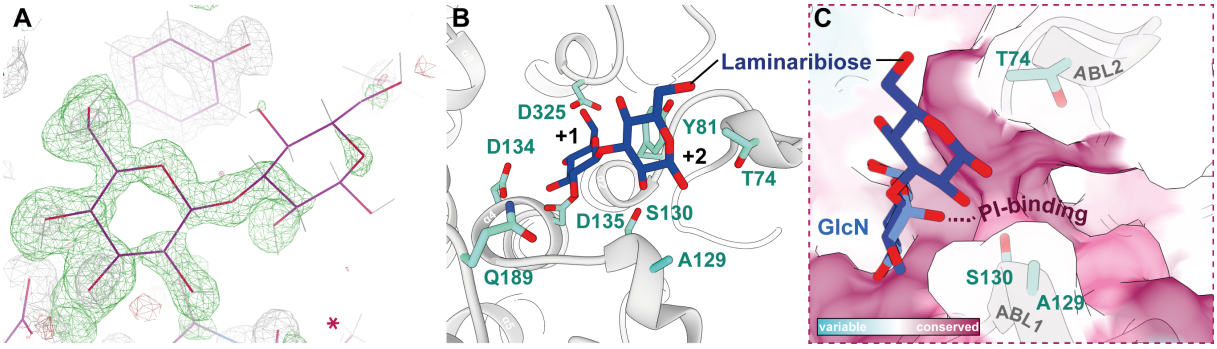


Figure 30: Laminaribiose bound state of *Ct*Dfg5. A OMIT-map of laminaribiose with the Fo-Fc map shown in green and contoured at σ-level 3.0. Laminaribiose is shown as blue sticks bound to +1 GPI-glycan binding site with its non-reducing end in the context of secondary structure elements of *Ct*Dfg5. The second moiety defines an acceptor-specific subsite +2. The coordinating residues are shown as turquoise sticks. **B** A superposition with the GlcN-bound state shows the highly similar coordination at subsite +1. The coloring of the surface is according to the *ConSurf* analysis, highlighting the conserved site where PI-binding is suggested and the ABL-regions, which are neither conserved nor variable. Oxygens are colored in red, nitrogens in blue.

4.3 The Dfg5-subfamily: A potential drug target against fungal infections

Most diseases, which are in the focus of the public are associated with multi-resistant bacterial or epidemic viral infections, however the influence of fungal-based infections in humans cannot be underestimated. The *Global Action Fund for Fungal Infections* (https://www.gaffi.org) suggests that fungal infections are the 4th most common illness on earth. As a functional human immune system is usually able to cope with fungal challenges, primary infections are rare. Due to the increasing number of modern medication especially in clinical environments, invasive systemic mycoses are an increasing problem that require effective and safe drugs (Brown et al., 2012b). While antibiotics against bacteria can easily target structures that are absent in humans and therefore work quite specifically, eukaryotic targets usually come with serious side effects. It is therefore desirable to develop antifungal drugs, which target components of the fungal cell that are absent in humans. Accordingly, the fungal cell wall is an excellent target, since the structural components as well as the synthesizing factors are unique to fungi. It is not surprisingly that current antimycotica that are well tolerated by the patient belong to the class of echinocandins, which target the β1,3-glucan synthase (Perfect, 2017). However, due to the frequent use in hospitals the number of resistant fungi, in particular *Candida* spp. is an increasing phenomenon resulting in

therapeutic failure (Perlin, 2015). As *Dfg5*-proteins are essential for these organisms and they do not share homology with human proteins, it might be a good starting point for treating fungal infections. Therefore, a structure-based search for potential inhibitors of *Dfg5*-proteins has been started.

4.3.1 Structure-based prioritization of a commercial fragment screen library

If a known protein structure is available, virtual screening is the method of choice to identify potential inhibitors. Usually this requires screening of large libraries with millions of possible lead structures. In order to find a suitable molecule for *Ct*Dfg5, a commercially available fragment library was chosen for initial studies. The molecules from the Frag Xtal Screen (*Jena Bioscience*) were further prioritized by docking analysis. For that, its 96 fragments were docked with *SeeSAR* to the GPI-glycan binding pocket of *Ct*Dfg5 (BioSolveIT GmbH). Ten poses of each fragment were calculated using the default parameter of the software. The twelve best fragments were chosen regarding their estimated binding affinities to *Ct*Dfg5. Three of the fragments (34, 35, and 37) were estimated to bind in the low to mid µM-range, for six fragments (2, 24, 32, 48, 51, and 85) a mid µM to low mM-range have been calculated and another three molecules were attributed binding affinities in the mM region (Figure 31).

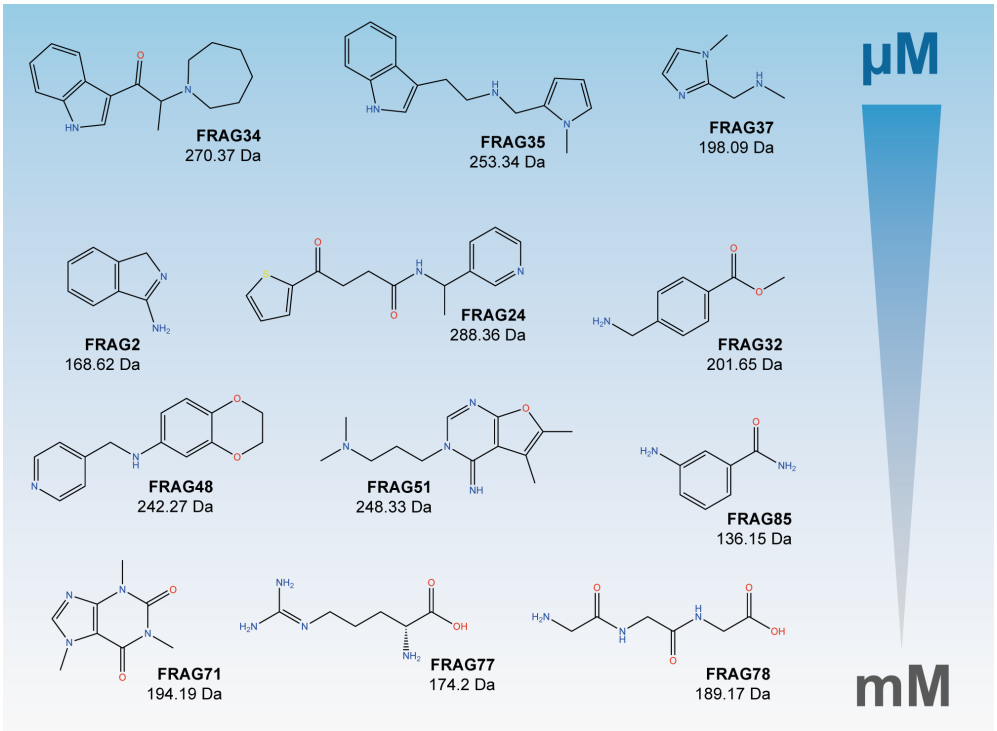


Figure 31: Twelve most promising hits from virtual ligand screening using *SeeSAR*. The chosen fragments are sorted according to their estimated binding affinities (see bar on the right). The fragments are presented as structural formulas and named according to their position in the Frag Xtal Screen. The respective molecular weight is given.

4.3.2 Crystal structure of *Ct*Dfg5 in complex with a potential inhibitor lead structure

The fragments identified above were diluted from 1 M DMSO stocks (*kind gift from AG Klebe*) to 100 mM concentrations in the already outlined buffer used for previous soaks. Interestingly, the condition with FRAG37, which could not be found in the crystal structure after 3 h of soaking, enhanced the diffraction quality of the crystal up to 1.05 Å resolution as mentioned above. Luckily, a positive hit for fragment binding could be found for FRAG51 (from now on referred to as FP-1 due to the furo[2,3-*d*]pyrimidin substructure that constitutes the majority of the molecule) after 3 h soaking time. The additional electron density from a high resolution data set (Table 10) could be unambiguously assigned to the fragment (see Figure 32A).

The binding within the GPI-glycan pocket of *Ct*Dfg5 is established by the furo[2,3-d]pyrimidin ring of FP-1, which is stacked by Y81. Furthermore, it forms a salt bridge between D325 and its propane-amine tail, which extends to subsite -1 and makes an electrostatic interaction with D134 (see Figure 32B). A superimposition with the GPI-core clearly reveals the excellent position at the active site of *Ct*Dfg5 giving a good initial hit for a competitive inhibitor lead structure (Figure 32C). Noteworthy, the observed orientation is contrary to the orientation that was predicted by *SeeSAR* (SFigure 8 on p.140).

Table 10 Crystallographic data collection and refinement statistics of *Ct*Dfg5 soaked with FP-1. Values in parentheses are for the highest resolution shell.

	***Ct*Dfg5$_{FP-1}$**		
PDB code	6RY8		
Data collection		**Refinement**	
X-ray source	ID23-1, ESRF	Resolution range (Å)	30.27-1.3
Space group	C 1 2 1	R_{work}/R_{free} (%)	14.1/13.1
Unit-cell parameters (Å, °)	*a*=83.31, *b*=55.04, *c*=80.35, *α*=90.00, *β*=90.23, *γ*=90.00	Average *B* factor (Å^2)	15.5
Wavelength (Å)	0.972	No. of atoms	
Resolution range (Å)	30.27-1.3 (1.35-1.3)	Total	4057
Completeness (%)	95.47 (96.61)	Protein	3346
Observed reflections	151167 (15034)	Ligand	33
Unique reflections	85345 (8647)	Solvent	678
Multiplicity	1.8 (1.7)	R.m.s.d., bond lengths (Å)	0.007
Wilson *B* factor (Å^2)	10.36	R.m.s.d., bond angles (°)	1.31
R_{merge} (%)	3.6 (25.3)	Ramachandran plot	
R_{meas} (%)	5.1	Favored (%)	99
Mean *I/σ(I)*	11.16 (2.77)	Allowed (%)	1
CC(½)	99.8 (88.1)	Disallowed (%)	0

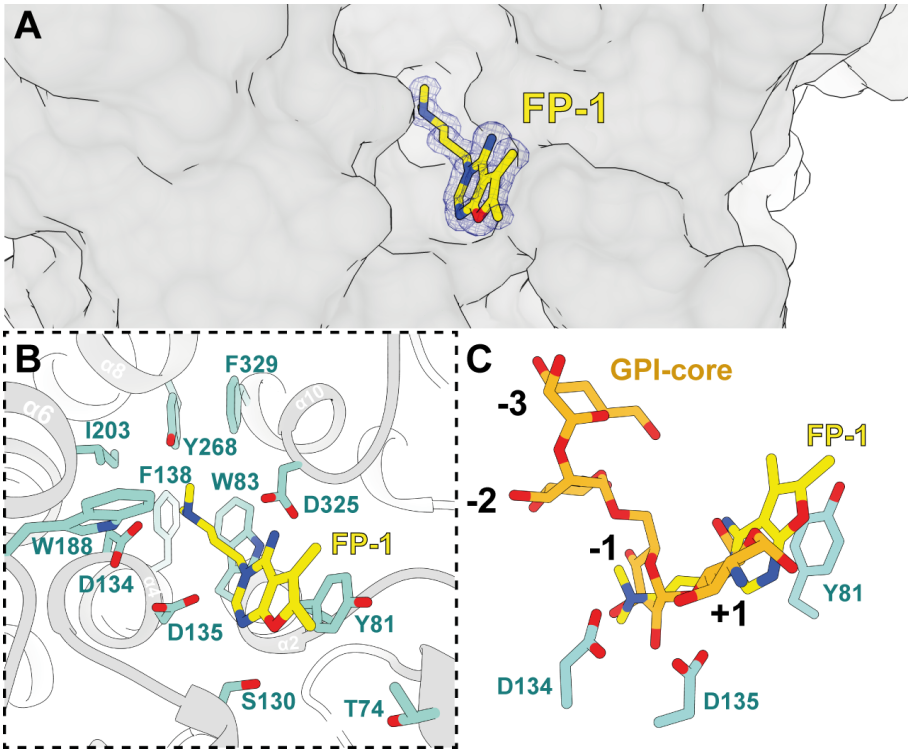


Figure 32: FP-1 bound above the active site of *Ct*Dfg5. A FP-1 is shown with its experimental electron density contoured at σ-level 1.0 within the binding pocket of *Ct*Dfg5. **B** FP-1 shown in the context of its interacting residues and the secondary structure of *Ct*Dfg5. **C** Superimposition of FP-1 and the GPI-core with Y81, D134, and D135 shown for orientation. The GPI-glycan binding sites are indicated. Residues and ligands are represented as sticks, all amino acids are colored in turquoise, FP-1 in yellow and GPI-core in orange, with the oxygens in red and nitrogens in blue.

As the binding mode of FP-1 suggests an inhibitory effect on the protein and non-functional *Dfg5*-proteins are lethal for *S. cerevisiae*, the effect of FP-1 supplementation in the yeast-growth medium will be analyzed in the following.

4.3.3 FP-1 impacts the viability of *S. cerevisiae*

In order to get an idea about a useful working concentration for *in vivo* tests, biolayer inferometry measurements were performed with the help of Dr. Sven-Andreas Freibert. The K_D of the FP-1:*Ct*Dfg5 interaction could be determined at 2.71±0.96 mM (SFigure 9 on p.141). Based on that, Gesa Schmitz (AG Mösch) tested the effect of FP-1 on living cells. She tested three different strains: the wild type, and two single mutant strains of which one produced just Dfg5, while the other one produced only Dcw1. The growth behavior in YEPD medium alone of all three strains was almost identical (Figure 33A). Initially, 10 mM FP-1 was added to the YEPD medium, however no growth could be observed. Therefore, the concentration was reduced to 5 mM FP-1 (Figure 33B). As a result, wild type cells could not reach exponential growth phase within the observed time and finally reached a cell density

approx. half the cell density without FP-1. While this already shows an impact on the fitness of treated cells, the single mutants possessing only one *Dfg5*-protein were barley able to grow. Small differences can be observed between the Dfg5 and Dcw1 mutant, as a slight increase in optical density can be seen for Dfg5. In order to test, whether this effect indeed relies on the interaction between the supplemented FP-1 and the protein, a structural derivative called FP-2 has been analyzed as well. Instead of the two methyl groups, FP-2 possesses an imidazole ring at the furo[2,3-*d*]pyrimidin-distal end of the propyl-tail (Figure 33C). FP-2 failed to produce a complex structure upon soaking experiments and the K_D could not be determined by biolayer inferometry (SFigure 9 on p.141). Wild type cells could grow in the presence of 5 mM FP-2 until saturation, which is quite convincing in this context (Figure 33C). The Dfg5 mutant also reached the saturation 2 h after the wild type. The Dcw1 mutant clearly suffered from the FP-2 treatment, however it could also reach the exponential phase at the end of the observed period of time.

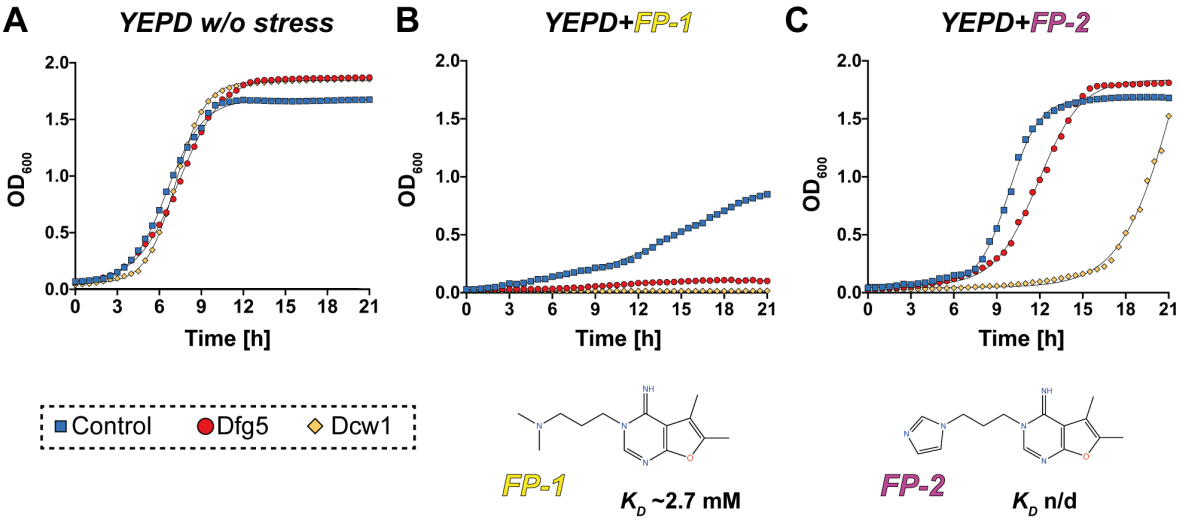


Figure 33: Effect of FP-1 and a derivate on the growth of *S. cerevisiae*. Growth curves of three different strains (indicated in panel **A**) are shown in YEPD (**A**), YEPD supplemented with 5 mM FP-1 (**B**), and YEPD supplemented with 5 mM FP-2 (**C**). The structural formulas of the respective supplement are given below the respective growth curve. The *in vivo* experiments were performed by Gesa Schmitz (AG Mösch). The binding studies were performed with the help of Dr. Sven-Andreas Freibert, who analyzed the data.

These findings indicate a promising starting point for an efficient drug against ascomycetes.

4.3.4 Combined fragment state as the structural basis for lead structure optimization

In order to optimize a lead structure in terms of its binding properties, one promising approach can be the combination of different binders. As we have seen in the sugar-fragment screen before, mannose is able to bind to GPI-glycan subsite -2 when *Ct*Dfg5 crystals are soaked in 500 mM mannose containing mother liquor. This should allow a combination with FP-1, which was identified at subsites -1/+1.

To show simultaneous binding of mannose and FP-1, both substances were used for soaking experiments. The final dataset was collected from a *Ct*Dfg5 crystal soaked in the mother liquor as described above supplemented with 1 M mannose and 100 mM FP-1 at the PXIII beamline at the Swiss Light Source (SLS) in Villigen, Switzerland (see Table 11).

Table 11: Crystallographic data collection and refinement statistics of *Ct*Dfg5 soaked with FP-1 and mannose. Values in parentheses are for the highest resolution shell.

	***Ct*Dfg5$_{FP-1+Man}$**		
Data collection		**Refinement**	
X-ray source	PXIII, SLS	Resolution range (Å)	30.28-1.2
Space group	C 1 2 1	R_{work}/R_{free} (%)	12.4/13.8
Unit-cell parameters (Å, °)	*a*=83.66, *b*=55.06, *c*=80.68, *α*=90.00, *β*=90.60, *γ*=90.00	Average *B* factor (Å^2)	18.73
Wavelength (Å)	0.999	No. of atoms	
Resolution range (Å)	30.28-1.05 (1.09-1.2)	Total	3929
Completeness (%)	96.56 (88.12)	Protein	3335
Observed reflections	207437 (16167)	Ligand	53
Unique reflections	110541 (10023)	Solvent	541
Multiplicity	1.9	R.m.s.d., bond lengths (Å)	0.015
Wilson *B* factor (Å^2)	14.34	R.m.s.d., bond angles (°)	1.68
R_{merge} (%)	2.2 (30.4)	Ramachandran plot	
R_{meas} (%)	3.1 (43.0)	Favored (%)	98.53
Mean *I*/*σ*(*I*)	14.56 (2.47)	Allowed (%)	1.47
CC(½)	99.9 (75.2)	Disallowed (%)	0

As it was expected from the individual structures, the simultaneous binding of both fragments together is not interfered (see Figure 34). Both molecules are found in the same conformation like observed before for the individual complexes. The shortest distance between them is 3.3 Å from the 1-hydroxy group of the mannose at subsite -1 (Man1) and the respective methyl group of FP-1.

Another interesting point of this complex structure is additional density of a mannose moiety (Man2) above Y199, which is located in the β1-strand. The presence of that sugar can be explained due to the higher mannose concentration used for the soaking experiment compared to *Ct*Dfg5$_{Man}$ (1 M instead of 500 mM). The closest contact is 3.8 Å established between

3-hydroxy group of Man1 and the 4-hydroxy group of Man2, which may provide another possible linker between to fragments to increase the binding properties of a *Dfg5*-subfamily specific inhibitor. Furthermore, it raises the possibility of another GPI-glycan binding site (i.e. subsite -4), as the common core structure of the GPI-anchor in fungi harbors another α1,2-linked mannose moiety that branches off mannose 3.

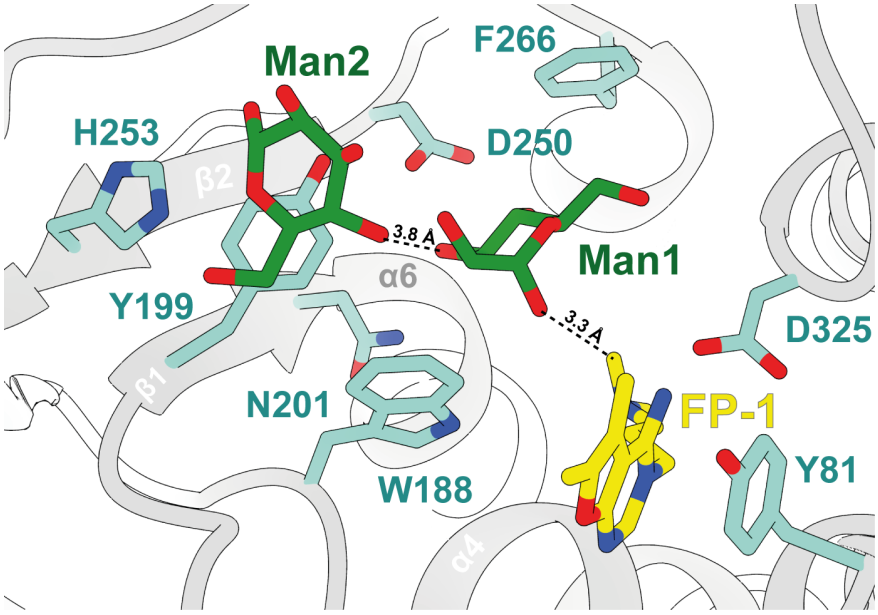


Figure 34: Combined soaking of mannose and FP-1. The bound ligands and their coordinating residues are shown as sticks in the context of the secondary structure elements from *Ct*Dfg5 colored in grey. Amino acid carbons are colored in turquoise, mannose carbons in green and FP-1 carbons in yellow. Oxygens are red and nitrogens blue. The closest distances between FP-1 and Man1, and Man1 and Man2 are indicated.

4.4 Expanding the knowledge to the *Fungal/Bacteria mixed* subfamily

To this point, bacterial and fungal GH76 proteins have been described for three different subfamilies and it was shown that their function differs drastically. After elucidating the functional background of *Dfg5*-proteins and their potential role as a drug target, the last part of this thesis will deal with another subfamily of GH76 proteins trying to emphasize the common features of the whole family.

So far, the knowledge about the *Fungal/Bacteria mixed* subfamily is rather limited. One of its members, which is in the following referred to as *Ct*GH76 (Uniprot-ID G0S5Y9), is composed of 439 amino acids and in contrast to *Ct*Dfg5, no further annotations regarding its topology or posttranslational modifications are given. No signal peptide or ω-signal sequence are predicted by common *in silico* tools such as SignalP-5.0 or the *GPI Fungal Prediction*

Server. However, the transmembrane prediction tool TMHMM2.0 provides a hint about its topology (see Figure 35) (Eisenhaber et al., 2004; Sonnhammer et al., 1998).

Although the general output does not predict any transmembrane helix (TMH), the resulting expectation number is 17.6. This is close to the cutoff criterion 18, which is thought to predict a TMH to be very likely and, furthermore, a warning of a possible N-terminal signal sequence is produced. As a TMH at the N-terminal region cannot be excluded from this initial analysis, the first 50 amino acids were dismissed for soluble heterologous protein production in *E. coli* SHuffle, which was established by Markus Friedrich during his master thesis using the same expression condition as described in this thesis.

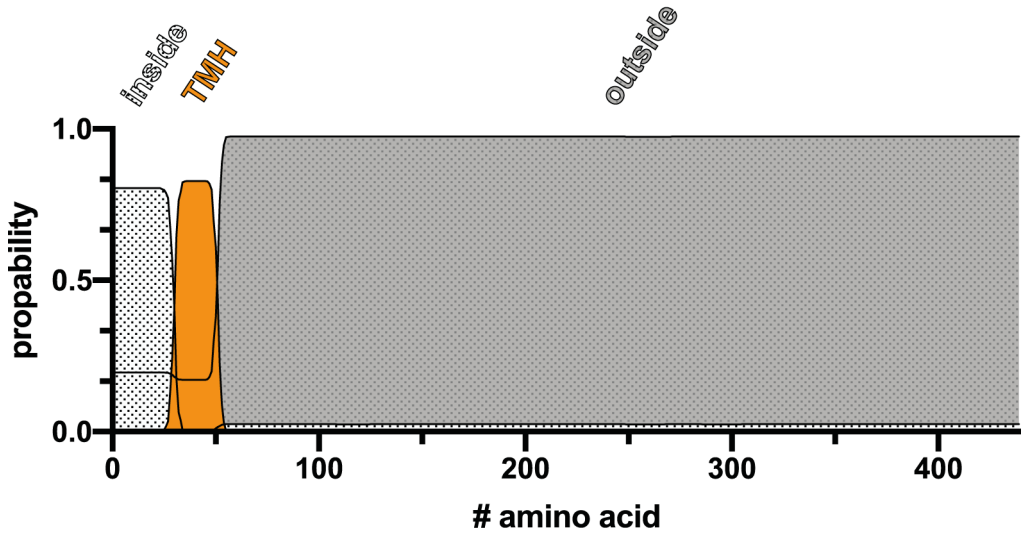


Figure 35: TMHMM probability plot of *Ct*GH76. The topology of *Ct*Dfg5 is shown with a propability plotted against the amino acids. From the TMHMM predicition tool analysis a TMH can be predicted thus giving rise to a putative type-II membrane protein.

Although the phylogenetic analysis of the GH76 class suggested a quite close relationship of the *Fungal/Bacteria mixed* subfamily to the bacterial homologs, Markus Friedrich could not show hydrolysis of α1,6-manno oligosaccharides by catalysis of *Ct*GH76. This goes along with the previous findings for *Ct*Dfg5, though it was a little bit surprising due to the close relationship to the *bacterial* subfamilies (see Figure 14 on p.53).

Markus Friedrich was able to obtain the crystal structure of *Ct*GH76. It crystallized in space group P 31 2 1 and diffracted to 2.0 Å resolution. As shown before for *bacterial* GH76 and *Dfg5*-proteins, this *mixed* subfamily has the GH76-typical $(\alpha/\alpha)_6$-helical barrel fold as well (Figure 36). The most significant differences compared to the already described homologs are the regions indicated as β1/2-wing and β3/4-wing. In the superposition with *Ct*Dfg5 it is easy

to see that these regions are prolonged in *Ct*GH76. This difference is unique for *Ct*GH76 and not found in other GH76 structures, which are excluded from superposition for clarity. These extended regions lead to a break of the antiparallel β-sheet between α-helices 5/6 and 7/8 described for *Ct*Dfg5, however forming two antiparallel sheets in the loop between α-helices 7/8 resulting in the β1/2-wing. Noteworthy, the corresponding β-strands (i.e. β1 of *Ct*GH76 and β2 from *Ct*Dfg5) are tilted to each other. Furthermore, the loop extension between α-helices 11 and 12 results in the formation of two additional β-strands forming the antiparallel β3/4-wing region in *Ct*GH76, a structural element that is completely absent in the other GH76 proteins.

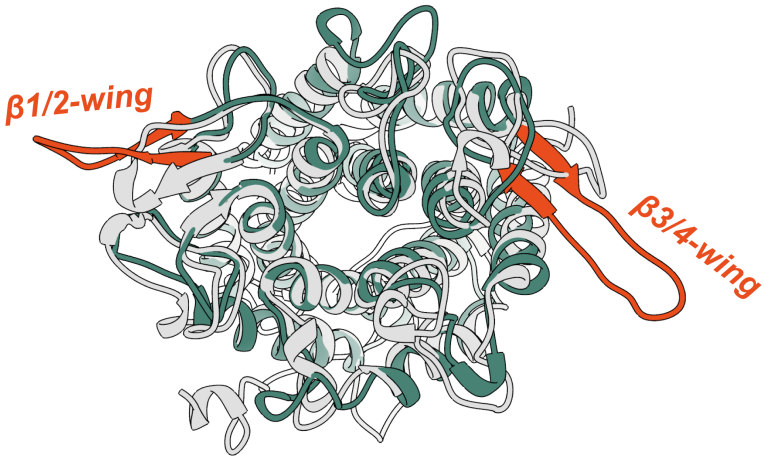


Figure 36: Superposition of *Ct*GH76 and *Ct*Dfg5. Both structures are shown with their secondary structure elements. *Ct*Dfg5 is shown in grey, *Ct*GH76 is shown in blue-green and its indicated wing regions are colored in red.

The biological role of the *Fungal/Bacteria mixed* subfamily in fungi is still elusive. With the initial *in silico* findings it is even hard to predict the actual destination of such proteins making any functional prediction highly speculative. In the following, *Ct*GH76 is subjected to sugar-fragment soaking experiments as described for *Ct*Dfg5 to get more insights into its functional properties.

Unlike it was found for *Ct*Dfg5, glucosamine could neither be identified at subsite +1, nor at any other possible glycan-binding site of *Ct*GH76. An explanation might be that the glucosamine-coordinating residues Y81 and Q189 from *Ct*Dfg5 are substituted by a proline and an aspartate in *Ct*GH76. However, this binding site does not only seem to exclude the GPI-glycan binding ability described for *Dfg5*-proteins, but also does not support an α1,6-linked mannose residue like it is known for the *bacterial* homologs.

As shown above the *bacterial* and *Dfg5*-subfamilies apparently share their binding mode at subsites -2 to -1, where an α1,6-linked mannobiose structure can be observed. However, the α1,6-mannobiose-bound states per se differ, as the biose is coordinated at subsites -3/-2 in the bacterial ortholog, while *Ct*Dfg5 coordinates the disaccharide at -2/-1 (see Figure 48 on p.108). In order to achieve that state for *Ct*GH76, α1,6-mannobiose is dissolved to a final concentration of 1 M in a solution consisting of equal volumes of the protein buffer (200 mM NaCl, 20 mM HEPES, pH 8.0) and the crystallization condition from the JCSG Core IV suite, E5 (3.6 M Sodium formate, 10% (v/v) glycerol). The soaking was performed at room temperature for three hours. The crystals were directly flash frozen in liquid nitrogen and diffraction studies resulted in a dataset that was solved at 2.1 Å resolution (Table 12).

Table 12: Crystallographic data collection and refinement statistics of *Ct*GH76 soaked with α1,6-mannobiose. Values in parentheses are for the highest resolution shell.

	***Ct*GH76$_{\alpha 1,6M}$**		
Data collection		**Refinement**	
X-ray source	PXIII, SLS	Resolution range (Å)	49.24-2.25
Space group	P 31 2 1	R_{work}/R_{free} (%)	20.4/21.8
Unit-cell parameters (Å, °)	*a*=106.84, *b*=106.84, *c*=127.0, *α*=90.00, *β*=90.00, *γ*=120.00	Average *B* factor (Å^2)	68.1
Wavelength (Å)	0.999	No. of atoms	
Resolution range (Å)	49.24-2.25 (2.33-2.25)	Total	2905
Completeness (%)	99.74 (99.57)	Protein	2806
Observed reflections	80464 (7980)	Ligand	50
Unique reflections	40269 (3990)	Solvent	49
Multiplicity	2.0	R.m.s.d., bond lengths (Å)	0.010
Wilson *B* factor (Å^2)	47.78	R.m.s.d., bond angles (°)	1.32
R_{merge} (%)	3.1 (24.5)	Ramachandran plot	
R_{meas} (%)	4.4 (34.6)	Favored (%)	98.33
Mean *I*/*σ(I)*	9.3 (2.0)	Allowed (%)	1.11
CC(½)	99.9 (94.0)	Disallowed (%)	0.56

Unlike in the bacterial homolog, the additional density could be observed at glycan-binding sites -2 and -1 corresponding to the two α1,6-linked mannoses (Figure 37). The overall binding appears as it was observed for the respective state of *Ct*Dfg5, however a striking difference becomes obvious in that context. In all described GH76 structures so far, a tryptophan (W188 in *Ct*Dfg5, W172 in *Bc*GH76) residue is directing the α1,6-linkage and the -1 moiety in the correct position due to its steric presence. Conversely, *Ct*GH76 possess a valine (V210) at that position, which does not result in a narrow channel but a more open binding pocket. However, a functional relevance of that substitution is questionable, as the majority of *Fungal/Bacteria mixed* proteins possess a tryptophan at that position.

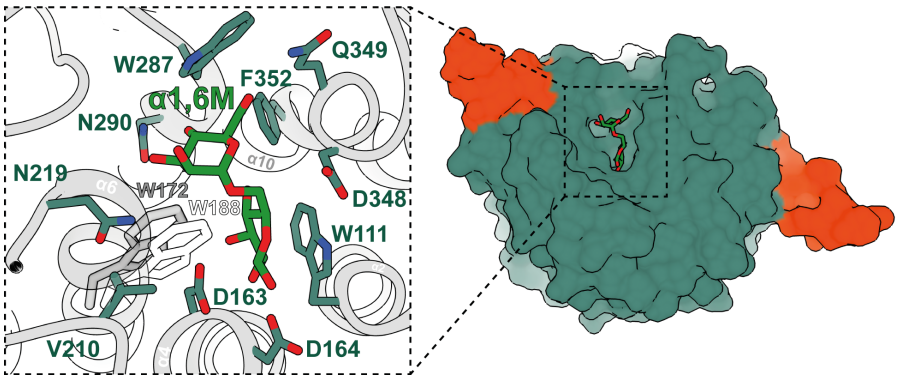


Figure 37: *Ct*GH76 bound to α1,6-mannobiose. On the left side, the coordination of α1,6-mannobiose (α1,6-M) at subsites -2 to -1 is shown in context of the secondary structure elements of *Ct*GH76. The coordinating residues are shown as sticks as well as the residues W188 (white) and W172 (transparent grey) from *Ct*Dfg5 and *Bc*GH76, respectively. On the right, the ligand is shown in the context of a surface representation of *Ct*GH76, with the majority of the surface colored in blue-green and the above-explained wing-regions colored in red. Oxygens are shown in red, nitrogens in blue.

Due to the extended wing regions it seems to be possible that this region may be involved in carbohydrate coordination. Therefore a trisaccharide was tested in the soaking screen as described before, which is composed of three mannoses that resemble a core component of the high mannose N-glycan from fungi (Figure 38A). Here, two mannoses are α1,3- and α1,6-linked to a central mannose. The soaked crystals were treated as before with a 1 M soaking solution and a complex structure could be determined at 2.6 Å resolution (Table 13).

Table 13: Crystallographic data collection and refinement statistics of *Ct*GH76 soaked with mannotriose. Values in parentheses are for the highest resolution shell.

	***Ct*GH76$_{Man3}$**		
Data collection		**Refinement**	
X-ray source	PXIII, SLS	Resolution range (Å)	49.1-2.6
Space group	P 31 2 1	R_{work}/R_{free} (%)	16.7/19.6
Unit-cell parameters (Å, °)	*a*=106.59, *b*=106.59, *c*=126.24, *α*=90.00, *β*=90.00, *γ*=120.00	Average *B* factor (Å^2)	61.73
Wavelength (Å)	0.999	No. of atoms	
Resolution range (Å)	49.1-2.6 (2.69-2.6)	Total	2893
Completeness (%)	99.83 (99.92)	Protein	2806
Observed reflections	51928 (5077)	Ligand	46
Unique reflections	25983 (2539)	Solvent	41
Multiplicity	1.8	R.m.s.d., bond lengths (Å)	0.010
Wilson *B* factor (Å^2)	55.40	R.m.s.d., bond angles (°)	1.31
R_{merge} (%)	4.2 (30.6)	Ramachandran plot	
R_{meas} (%)	6.0 (43.2)	Favored (%)	98.05
Mean $I/\sigma(I)$	8.2 (2.0)	Allowed (%)	1.11
CC(½)	99.7 (83.0)	Disallowed (%)	0.84

Indeed, the additional density appeared close to the described β1/2-wing region interacting with its proximal residues A271 and Y280 (Figure 38B). The C6-carbon involved in the α1,6-linkage is stacked by Y217 and L284 guides the orientation of C2-C1-O1 from the α1,3-linked mannose. Electrostatic interactions are facilitated by T216 with the 2-hydroxy group of the α1,6-linked mannose and N282 to the reducing end of the central mannose. Additional density also appeared in the -2 subsite, which may be attributed to a single mannose that might be a contamination due to partial hydrolysis of the triose.

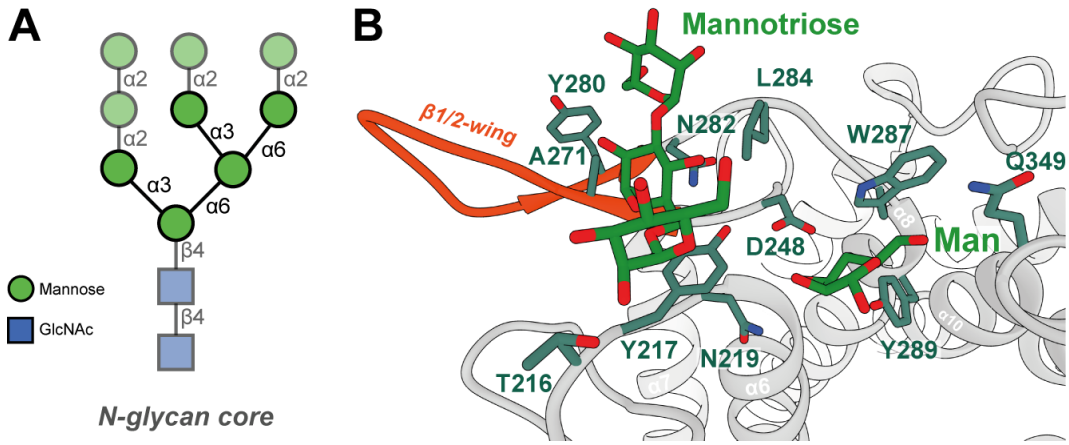


Figure 38: Mannotriose bound state of *Ct*GH76. A The schematic representation of the high mannose core of fungal N-glycans is shown. The mannotriose is presented twice in that structure highlighted by its opaqueness. The symbols are explained at the bottom left and the respective linkages are indicated. **B** Mannotriose, mannose and their coordinating residues are shown as sticks in the context of the secondary structure elements from *Ct*GH76. The β1/2-wing is emphasized in red, while the remaining cartoon is in grey with labeled α-helices. Amino acid-carbons are colored in blue-green, carbohydrate-carbons in green, oxygens are red, and nitrogens are colored in blue. *GlcNAc=N-acetylglucosamine; Man=mannose.*

5 Discussion

The cell wall is a special part of the fungal cell. Although not being part of the interior, the outermost layer must be built and maintained in such a fine-tuned manner to simultaneously ensure the protective and rigid role, as well as its adaptable and highly dynamic character. While the mere synthesis of the polysaccharides and proteins of the fungal cell wall (FCW) happens along the secretory pathway either soluble or embedded in the plasma membrane, the further glycan-processing and the wall's maturation is established by a set of proteins that act extracellularly. An essential element of that system is the incorporation of GPI-anchored proteins (GPI-AP) into the FCW (Kitagaki et al., 2002). Glycoside hydrolases of the GH76 class were identified to be responsible for the lipid-to-wall transfer, however the exact mode of action was not clear, as already described bacterial homologs were attributed an enzymatic activity that contradicts the previously described way of GPI-CWP attachment (Ao et al., 2015; Cuskin et al., 2015; Fujii et al., 1999; Kitagaki et al., 2002; Kollár et al., 1997; Thompson et al., 2015b).

Initial sequence similarity analyses from this study gave a hint that GH76-proteins are not an isofunctional group of proteins, but possess a still elusive number of different functions in bacteria, fungi and archaea. One of the occurring subfamilies exclusively contained proteins from ascomycota, to which a cell wall defective phenotype had already been ascribed. A member of that *Dfg5*-subfamily from the thermophilic filamentous fungi *Chaetomium thermophilum*, *Ct*Dfg5, was then assigned to an in-depth analysis. This allowed insights into the substrate binding mode and the way of how CWP are transferred to the glucan meshwork. This work also showed the potential of *Dfg5*-proteins as a drug target and provides an idea about the general feature of GH76-proteins by expanding the structural repertoire of this family with complex structures of an additional member from the *Fungal*/*Bacteria mixed* subfamily, called *Ct*GH76.

The different aspects of this thesis mentioned above will be discussed in the following.

5.1 A subfamily of GH76-proteins act on fungal cell wall biogenesis

Due to the technological advances in recent years, sequencing of entire genomes has become a routine, which provides us with millions of sequences. As plain sequences are useless for researchers, functional assessment is required. The first stage of protein domain identification is fulfilled by protein family databases such as InterPro (http://www.ebi.ac.uk/interpro/) or

Pfam (https://pfam.xfam.org). The developers are constantly challenged with the increasing number of sequences, trying to improve the classification and functional annotation based on existing experimentally characterized proteins (El-Gebali et al., 2019; Mitchell et al., 2019). A key problem of sequence-based protein function assignment to a complete protein family is the uncertainty in the low-identity region, which may still give an idea about superior features of the respective protein class, however the exact mechanistic function is often disguised in the *Twilight zone* and the general assignment of a few homologs to a complete family can result in misleading hypotheses (Rost, 1999).

In order to clarify the controversy in the current literature regarding the function of fungal GH76-proteins, the first aspect that was addressed was, whether this protein family is isofunctional per se, or not. This was tackled by generating a sequence similarity network (SSN) with the intention to produce isofunctional clusters. In the presented SSN (see Figure 14 on p.53), the family already falls apart at moderate cutoff criteria into ten major subfamilies. This behavior indicates that GH76 have diverse functional properties. It happens that all fungal proteins, which were previously described to be involved in cell wall biogenesis, are exclusively found in the *Ascomycota I* subfamily (Table 14). Therefore, this subfamily has been renamed to *Dfg5* with respect to the first ortholog that was described (Mösch and Fink, 1997). Consequently, it must be concluded that proteins from this cluster follow the same functional principles and thereby disagree with the current annotation as α1,6-mannanase. Notably, the bacterial subfamilies well separate from the *Dfg5*-subfamily, strengthening the assumption that the latters catalytic function is indeed different.

Table 14: Fungal *Dfg5*-proteins described in the literature regarding their role in the cell wall biogenesis. (#) In the cited publication, single disruptants were analyzed concerning CW-related phenotype for all nine GH76 proteins found in *N. crassa*, of which just the two shown in the table were further characterized. (*) Q4WIT8 (called Dfg2 in the cited publication) does not belong to the *Dfg5*-subfamily, but to the *Fungal/Bacteria mixed* cluster. A further GH76-protein belonging to the *Ascomycota II* subfamily (Q4WI31) was not analyzed in this study.

Organism	Uniprot	Publication
S. cerevisiae	Q05031, P36091	(Kitagaki et al., 2002, 2004)
C. albicans	Q5ACZ2, Q5AD78	(Ao and Free, 2017; Spreghini et al., 2003)
N. crassa	Q1K7A8, Q7S4K4	(Maddi et al., 2012)#
A. fumigatus	Q4WKP7, Q4WIT8, Q4WG09, Q4WFX5, Q4W985, Q4WFJ1, Q4WA43	(Muszkieta et al., 2019)*

5.2 How GPI-AP's are attached to the cell wall glycan

In order to determine the catalytic mechanism within the *Dfg5*-subfamily, the crystal structure of a homolog from *Chaetomium thermophilum* was determined. The structure alone revealed the conserved $(\alpha/\alpha)_6$-helical barrel fold, which is known from the already described bacterial GH76-proteins, including the conserved position of the active site DD-motif (Cuskin et al., 2015; Thompson et al., 2015b). While the active site motif was found to be essential in structure-based mutants of *S. cerevisiae,* the α1,6-mannanase activity on a mannotriose and a mannohexaose could not be observed for *Ct*Dfg5. Interestingly, this was suggested as a key feature regarding the hypothesis of CWP-attachment via N-linked outer chain mannans (Maddi et al., 2012).

For a better understanding of the functional properties of *Ct*Dfg5, complex structures were sought to be determined. As suitable substrates of complex carbohydrates are not commercially available, *Ct*Dfg5 crystals were soaked with sugar fragments of the N-linked outer chain mannan and the GPI-anchor. Surprisingly and very convincingly at the same time, complex structures of *Ct*Dfg5 with mannose, α1,2-mannobiose, α1,6-mannobiose, and glucosamine could be obtained, which together reassembled the core structure of the GPI-anchor within the binding pocket of the protein from GPI-glycan binding sites -3 to +1. This not only sets the basis for the true role of *Dfg5*-proteins in the lipid-to-wall transfer of GPI-CWPs, it also provides first insights into the conformational properties of GPI-anchors as discussed below. In addition to the assembled substrate-bound state of *Ct*Dfg5, the protein was also successfully soaked with laminaribiose, the disaccharide unit of the major cell wall polysaccharide β1,3-glucan. This structure revealed the non-reducing end bound in the same manner coordinated to subsite +1, as it was observed for the glucosamine (Figure 39).

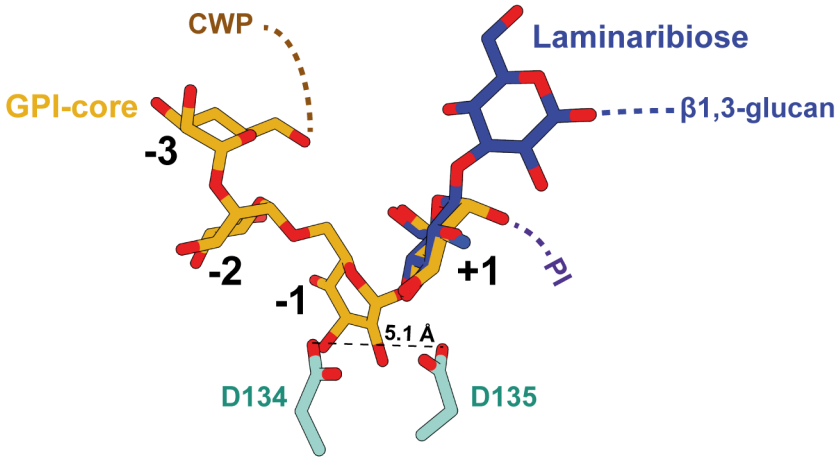


Figure 39: Superposition of the reassembled GPI-core and laminaribiose bound to *Ct*Dfg5. The active site DD-motif (turquoise), the GPI-core (orange) and laminarib

iose (blue) are shown as sticks. The *native* extensions are added to the respective termini by the colored dashed lines. The dashed line between the carboxy-groups of D134 and D135 shows their distance as indicated. The GPI-glycan binding subsites are indicated from -3 to +1. *CWP=cell wall protein; PI=phospatidylinositol.*

With this state of an acceptor sugar ready to attack a covalent glycosyl-enzyme intermediate and the typical distance of the active site residues of ~5 Å for retaining glycosidases, it is possible to propose a structure-based mechanism for the two-step transglycosylation reaction of *Dfg5*-proteins exemplary for *Ct*Dfg5 (see Figure 40). Upon binding of the GPI-core, the nucleophilic D134 attacks the anomeric center of Man1, while a proton donated by D135 supports the departure of the GlcN-PI leaving group. This results in a covalent glycosyl-enzyme intermediate, typical for retaining reaction mechanisms (Figure 40). After the release of the leaving group, the acceptor polysaccharide enters the scenery with its non-reducing end occupying subsite +1. In this particular case, the binding of a β1,3-linked laminaribiose was observed, however from the coordination it seems to be important that the +1-penetrating sugar-moiety is the non-reducing end of a glucose-polymer. In the following step, the 4-hydroxy group is activated by the deprotonated D135, allowing a nucleophilic attack on the glycosyl-enzyme intermediate, which results in the formation of a new α1,4-glycosidic bond established between mannose and glucose. Thompson and colleagues proposed a reaction coordination with a $B_{2,5}$ transition state following a ${}^{O}S_2 \leftrightarrow B_{2,5}^{\ddagger} \leftrightarrow {}^{1}S_5$ itinerary for GH76 retaining α-mannanases (Thompson et al., 2015b). To comment on the conformational distortion of the -1 subsite moiety of *Ct*Dfg5 in order to support the glycosidic bond breakage, a -2 to +1 spanning molecule would probably be the minimum requirement to allow the necessary distortion. Unfortunately, such a state could not be achieved, moreover the -2/-1 subsite state mapped by a bound α1,6-mannobiose presumably does not reflect the catalytically competent conformation of Man1, as the 1- and 2-hydroxy groups coordinate a Ca^{2+} together with D135. Anyway, the presence of calcium is an interesting fact regarding the preference of *Dfg5*-proteins to catalyze transglycoslyation rather than hydrolysis observed for the bacterial homologs. The negatively charged plasma membrane is suggested to increase the proximal calcium concentrations up to 30 times compared to the bulk concentration and that the binding of calcium to the bilayer is an entropy-driven process due to the displacement of interfacial water (Binder and Zschörnig, 2002; Seelig, 1990; Sinn et al., 2006; Vernier et al., 2009). This seems to be a perfect environment for *Dfg5*-catalyzed transglycosylation due to proximity to the plasma membrane with an increased amount of calcium that inhibits

hydrolysis, a low water availability in general and the supply of acceptor molecules by their membrane embedded synthesis.

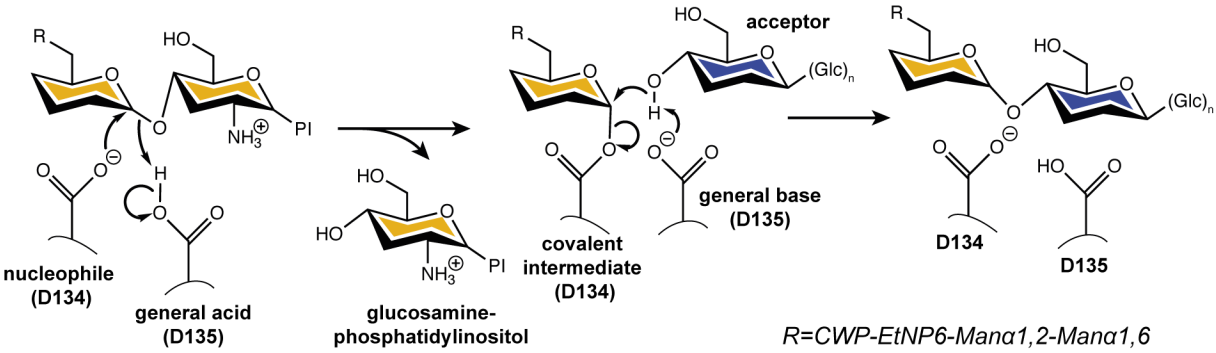


Figure 40: Mechanism of the transglycosylation mediated by *Dfg5*-proteins. A classical retaining mechanism can be suggested with D134 (*Ct*Dfg5-nomenclature) acting as the nucleophile and D135 as the general acid/base residue. Upon the first cleavage step of the glycosidic bond, the glycosyl-enzyme intermediate is formed. In the second step, the acceptor molecule (instead of the water from the general hydrolysis mechanism) intercepts the anomeric carbon and thereby forms a new glycosidic bond. *PI=phosphatidylinositol; $(Glc)_n$=glucose polymer, here β1,3-glucan*.

In the context of the incorporation of cell wall proteins, a model for the lipid-to-wall transfer can be suggested (Figure 41). The first step is the binding of the GPI-AP and the subsequent cleavage and release of the GPI-lipid, while the protein is covalently attached to Dfg5. This is a crucial step in the transglycosylation event, as it prevents the GPI-AP to be lost to bulk solvent. Dfg5 then presents the covalently bound GPI-AP to the cell wall glycan and allows the binding of a suitable acceptor molecule, which fits to the +1 GPI-glycan binding subsite. Subsequently, the new glycosidic bond is established and the GPI-AP becomes a GPI-CWP bound with its GPI-core to the non-reducing end of the cell wall glycan. The transfer essentially requires the GPI-core structure. Strikingly, the recent findings in *A. fumigatus* described that *Dfg5* proteins are responsible for the transfer of GPI-anchored galactomannans, which goes along with the findings within this work (Muszkieta et al., 2019).

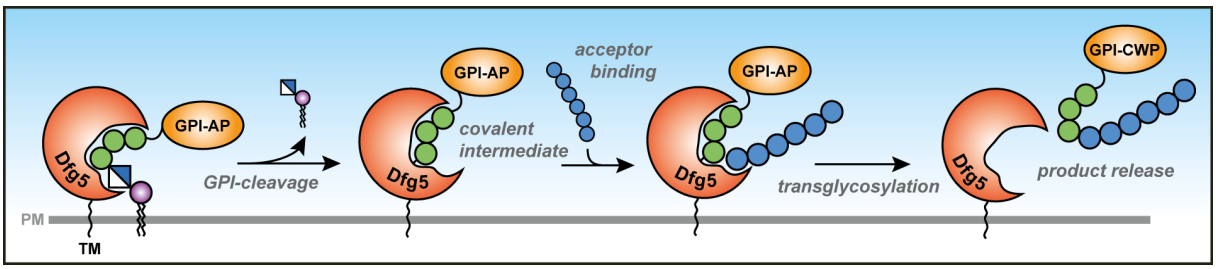


Figure 41: Model of *how GPI-APs are attached to the cell wall glycan*. Upon GPI-AP recognition, *Dfg5*-proteins cleave the GPI-anchor core by forming a covalent intermediate. After binding of the acceptor molecule, the GPI-remnant is transferred on the non-reducing end of the cell wall glycan. Finally, a GPI-CWP is released from the enzyme. *TM=transmembrane helix; PM=plasma membrane.*

5.3 *To be or not to be* sorted into the fungal cell wall

A major question in the context of GPI-CWPs is how the sorting into the cell wall occurs. In principle all GPI-APs end up at the plasma membrane. Some of these proteins can be classified as GPI-plasma membrane proteins (GPI-PMPs), others are transferred into the cell wall ending up as GPI-CWPs. However, this division seems to be not absolutely strict, as several GPI-APs end up in both locations (Pittet and Conzelmann, 2007). The transglycosidases Dfg5 and Dcw1 themselves are usually produced as GPI-APs in ascomycota, so it must be ensured that they remain at the plasma membrane to incorporate GPI-CWPs as described above.

In yeast, temperature-sensitive mutants of a gene involved in the cell-division cycle (*cdc1*^{ts}) were shown to stop growth in early G2 phase having small buds and replicated DNA, a phenomenon that was found in *dcw1*^{ts} mutants as well (Hartwell et al., 1970; Kitagaki et al., 2004). More recently it has been shown that Cdc1 removes the EtN-P added to Man1 by Mcd4 during GPI-anchor biosynthesis and that it can be deleted in a Δ*mcd4* background (Vazquez et al., 2014). It is not clear, what the consequences of EtN-P removal by Cdc1 exactly are regarding intracellular trafficking. Interestingly, Ted1, the phosphodiesterase that removes the EtN-P from Man2, is shown to act together with Emp24 and Erv25 at the ERES for facilitating efficient cargo exit from the ER (Haass et al., 2007). With the reassembled GPI-core structure of *Ct*Dfg5, however, it is possible to explain the consequences caused by the presence of an EtN-P at Man1. By adding an EtN-P group to the 2-hydroxy group of Man1 it becomes obvious that there is no space left at subsite -1 for suitable binding as it penetrates the proteins surface (Figure 42A). Such a modification most likely results in the non-binding of the GPI-anchor core, in any case there is no catalytically relevant conformation (and distortion) of Man1 anymore possible, thus preventing cleavage of the Manα1,4-GlcN linkage and subsequent transfer onto the cell wall glycan. This explains the connected phenotypes of *cdc1* and *dfg5*/*dcw1* mutants, which basically show the same behavior for complete deletion (lethal) and the temperature-sensitive mutants (G2 arrest) as both result in an abolished or compromised lipid-to-wall transfer of GPI-CWPs, respectively.

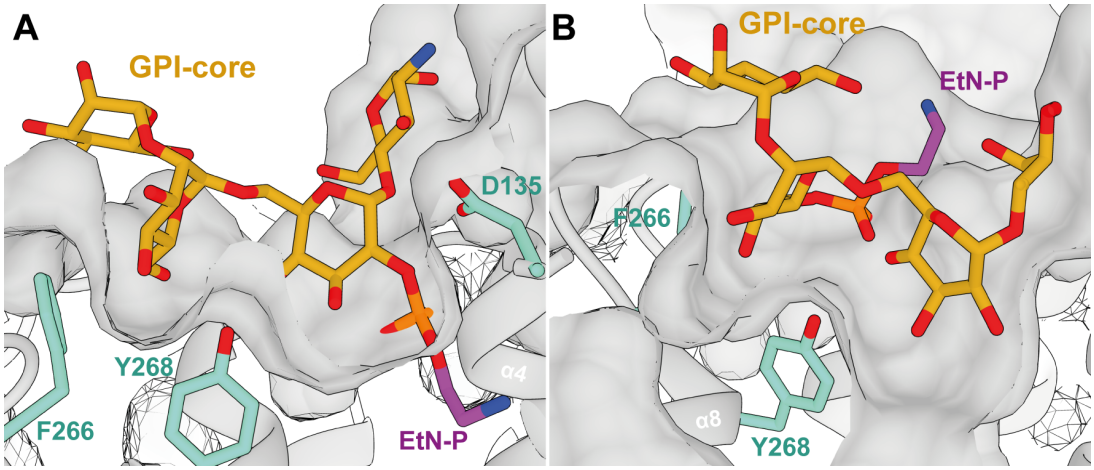


Figure 42: Hypothetical positions of EtN-P at Man1 and Man2. The additional, hypothetical EtN-P (purple) added to Mannose 1 (**A**) and Mannose 2 (**B**) of the GPI-core (orange) are shown in context of the clipped surface of *Ct*Dfg5 (grey) with *marker* residues (turquoise) shown for orientation.

In contrast to Cdc1, the removal of EtN-P from Man2 by Ted1 is not essential in *S. cerevisiae*, though it was shown that its deletion results in a delayed ER-export of GPI-APs, as it works in a concerted manner with the p24-complex that is responsible for efficient COPII traffic from ER to Golgi by interaction with the mature GPI-glycan (i.e. without the Man2-EtN-P) (Manzano-Lopez et al., 2015). The findings of Δ*ted1* mutants resemble the deletion mutants of the p24-complex components *emp24* and *erv25*, which show the same defect of ER-exit (Belden and Barlowe, 1996; Schimmöller et al., 1995). The consequences for the final cell wall delivery by *Dfg5*-proteins in Δ*ted1*-cells (i.e. the presence of EtN-P at the 6-hydroxy group of Man2) may not be as fatal as for the Δ*cdc1* mutant, as the 6-hydroxy group of Man2 is not buried in the binding pocket of the transglycosidase (Figure 42B). A slightly distorted coordination of Man2 is very likely, however it cannot be excluded that catalysis can still happen if the Man1-coordination is not substantially affected. This goes along with Δ*ted1* mutants that show a growth behavior in rich medium as observed for the wild type (Yoko-o et al., 2018). The latter study investigated the GPI-lipid remodeling mediated by Cwh43 from PI (phosphatidylinositol) to IPC (inositolphosphorylceramide), which is important for the final destination of GPI-APs. Interestingly, the EtN-Ps of Man1 and Man2 are required for ceramide-remodeling by Cwh43 (Benachour et al., 1999; Zhu et al., 2006). While the CWP Cwp2 possesses a PI in wild type cells, GPI-PMPs such as Gas1 harbor an IPC lipid. In Δ*cwh43* however, Gas1 with the PI-form lipid favored to be localized to the cell wall. Like observed for Δ*ted1*, the Δ*cwh43* mutant possessed a doubling time like the wild type, while the double mutant Δ*ted1*/Δ*cwh43* showed a reduced growth rate (Yoko-o et al., 2018). It is

not entirely clear, when and why Cwh43 acts on lipid remodeling. In yeast, two different models are proposed. A sequential model is suggested as the main pathway, where Cwh43 utilizes a PI containing diacylglycerol (pG2) after Gup1, while the divergent model suggests an alternative route utilizing lysoPI for the ceramide remodeling and thereby separating the reactions of Gip1 and Cwh43 (Ghugtyal et al., 2007; Umemura et al., 2007). A recent study suggests the absence of an alternative route in *A. fumigatus* (Li et al., 2018).

As a consequence, the fate of GPI-CWPs and GPI-PMPs is encoded during the maturation of the GPI-anchor in the secretory pathway, while *Dfg5*-proteins decode the made decision at the plasma membrane. Based on the structural model of the complex between the GPI-core glycan and *Ct*Dfg5, a sequence of events can be suggested in the ER of yeasts (and probably other fungi possessing two paralogs of the mammalian PGAP5, i.e. Cdc1 and Ted1), upon which the destination of GPI-APs is determined (Figure 43). As a start the remodeling of the lipid by Bst1, Per1, and Gup1 occurs in the first three steps of the sequential pathway directly after the transamidase added the GPI-AP to the GPI-precursor. The GPI-lipid remodeling (and not the GPI-glycan remodeling) is the crucial step for the concentration at ERESs, suggesting that this step is required for the presorting and concentration at the exit sites (Manzano-Lopez et al., 2015). The next step determines the destination. Future GPI-CWP must be recognized by Cdc1 at this point (upper route), which not only removes the EtN-P at Man1, but also prevents further lipid-modification by Cwh43, as this modification is required for ceramide-transfer (Vazquez et al., 2014; Zhu et al., 2006). Subsequently Ted1 removes the EtN-P from Man2, which results in the mature ER-form of the GPI-glycan and facilitates efficient COPII-vesicular traffic to the Golgi initiated by the p24-complex (Haass et al., 2007; Manzano-Lopez et al., 2015). Upon export to the plasma membrane, the glycan of such GPI-APs containing a PI-lipid are recognized by *Dfg5*-proteins and further transferred to the cell wall glycan. Note that further modifications such as the fourth and fifth glycan added during GPI-biosynthesis and remodeling in the Golgi, respectively, are omitted in this model for clarity. GPI-PMPs however are directly recognized by Cwh43 (lower route), which remodels the lipid from PI to IPC. As the preferred substrate of Cdc1 are GPI-anchors with a diacylglycerol lipid this step is omitted and the GPI-ceramide is directly handed over to Ted1 (Vazquez et al., 2014). Ted1 proceeds with the removal of the EtN-P from Man2 and thereby facilitates the efficient binding of the p24-complex for cargo export to the Golgi. This results in a secreted GPI-AP possessing an IPC and an EtN-P at Man1. This is the modification that is suggested to be incompatible with the efficient binding and the catalytic function of *Dfg5*-proteins, which finally leaves the GPI-PMP in the plasma membrane. As mammalians do not have to

discriminate between GPI-PMPs and GPI-CWPs, this might also explain why they possess just one PGAP5-homolog for GPI-glycan processing, which ensures efficient ER to Golgi transport by removing EtN-P from Man2, while the EtN-P at Man1 is commonly found in those GPI-anchors (Fujita et al., 2009; Gonzalez et al., 2009).

It must be mentioned at this point that initial studies of the GPI-composition in *S. cerevisiae* proposed a structure totally lacking EtN-Ps moieties at the Man1 and Man2 positions using the GPI-anchors purified from crude membrane preparations standing in stark contrast to the proposed model, as the EtN-P at Man1 is an obligatory requirement for the glycan based discrimination by *Dfg5*-proteins (Fankhauser et al., 1993). Analyses in *A. fumigatus* used an approach including an HF-treatment that did not allow any statements concerning the EtN-P modifications as they were removed before analysis, while further studies in *S. cerevisiae* questioned the Fankhauser model by showing the presence of EtN-P at the α1,4-linked mannose (i.e. Man1) (Canivenc-Gansel et al., 1998; Flury et al., 2000). To date it is clear that the addition of EtN-P to Man1-3 is essential for the synthesis and functionality of the GPI-anchor, suggesting a methodical problem in early fungal GPI-anchor studies due to harsh isolation conditions, however the EtN-P at Man3 linking the protein to the anchor remained untouched. In current literature reviewing the GPI-structure, no unambiguous model is proposed and recent attempts to explain the role of the GPI-anchor in intracellular trafficking indeed included the EtN-P, while claiming a lack of knowledge concerning the side-branches upon their cross-linking (Kinoshita and Fujita, 2016; Orlean and Menon, 2007; Pittet and Conzelmann, 2007).

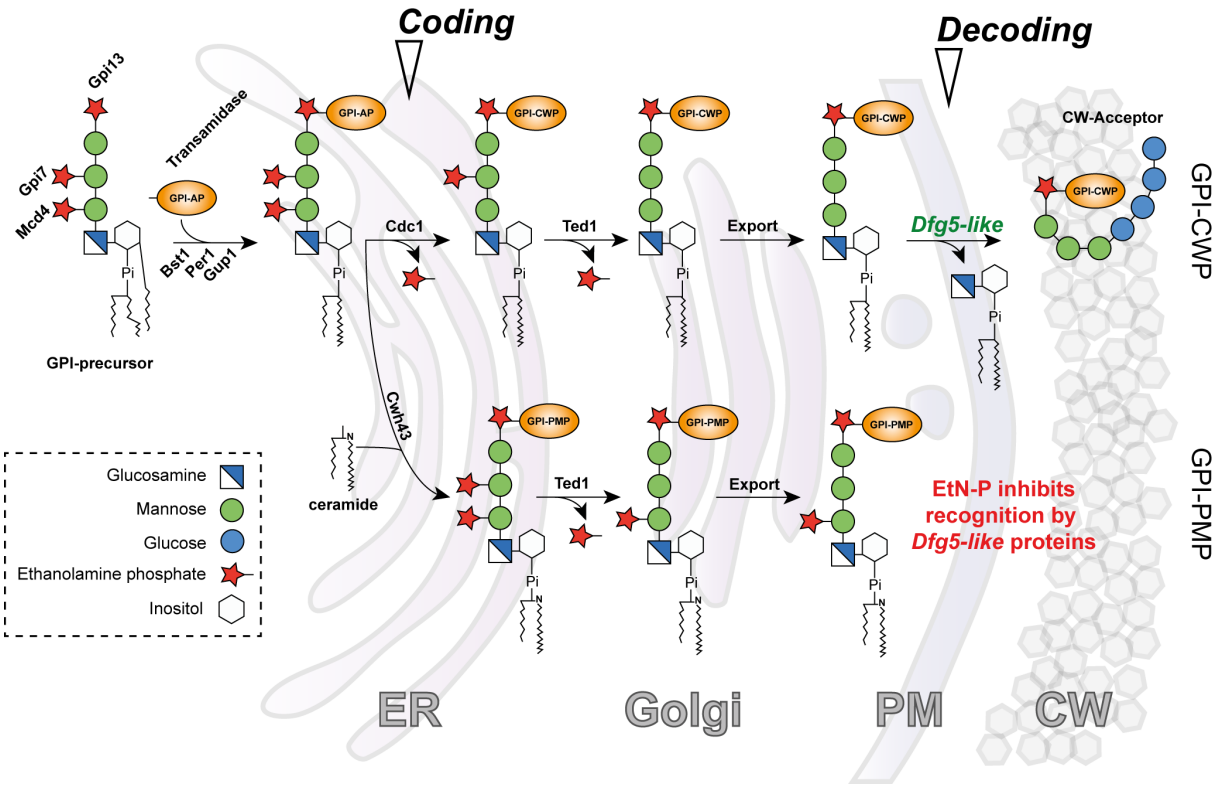


Figure 43: Presence of EtN-P at Man1 prevents GPI-based CW-incorporation by *Dfg5*-proteins. After GPI-precursor biosynthesis, the transamidase complex transfers the protein onto the EtN-P of Man3, upon which Bst1, Per1, and Gup1 remodel the GPI-lipids. Depending on the destination, either Cdc1 removes the EtN-P from Man1, or Cwh43 catalyzes the exchange from PI to IPC. Thereby, the EtN-P *code* for the final localization of the GPI-AP is written. Subsequently, Ted1 removes the EtN-P from Man2 to facilitate efficient ER-exit mediated by the p24 complex. Upon further processing in the Golgi, *Dfg5*-proteins are able to *decode* the presence or absence of an EtN-P within the GPI-core and thereby facilitating the transport of correctly processed GPI-CWPs into the CW, while GPI-PMPs remain bound to the membrane (adapted from Kinoshita and Fujita, 2016).

What remains is the question, how the discrimination between the Cdc1-route and the Cwh43-route works. Structural studies on the GPI-anchor and its interacting partners would be a way to shed light on this issue, but to my knowledge this thesis presents the first insights into experimentally supported GPI-glycan conformations, if only based on the assembly of its glycan-fragments. This is not really surprising, as the high degree of flexibility of glycosidic bonds suggest a highly dynamic behavior and simultaneously makes it hard to analyze it in solution (Vliegenthart, 2011). However, a number of computational simulations exist, which try to suggest likely conformations and dynamic behaviors by searching for energetically favorable ensembles optima and highly occupied states. A tetrasaccharide consisting of Manα1,2-Manα1,6-Manα1,4-GlcNAc-αOMe (simulating the GPI-core) was observed to adopt frequently a *C-like shaped* conformation with distances between C1 of GlcNAc (mimicking

GlcN) and C4 of Man3 (*r1'4-distance*) of about 12.7 Å (Figure 44A) (Wehle et al., 2012). Distances in other snapshots were further found to be as short as 5 Å up to 16 Å in a stretched conformation. Interestingly, the GPI-core as predicted to be bound to *Ct*Dfg5 adopts a conformation that is surprisingly similar to the described *C-like shape* with an *r1'4*-distance of ~11 Å (Figure 44B).

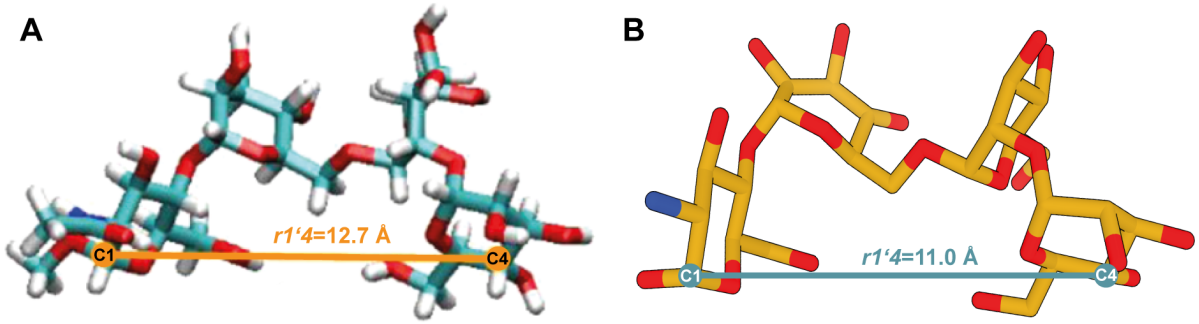


Figure 44: Comparison of simulated and fragment-derived GPI-shape. A *C-shaped* GPI-core state observed frequently in MD-simulations with an *r1'4*-distance of 12.7 Å (modified from Wehle et al., 2012). **B** The reassembled GPI-core from *Ct*Dfg5 is shown with an indicated *r1'4*-distance of 11.0 Å.

GPI-anchors are thought to exist in between two opposing states, a *lollipop* and a *flop-down* conformation, which have impact on the distance of the attached protein and its interaction with the bilayer as the *lollipop* stands straight up from the membrane while the *flop-down* remains in close contact with the lipid-head groups concerning all core-glycans (Lehto and Sharom, 2002). A recent study used the complete GPI-anchor inserted in a lipid bilayer with and without an attached GFP for their simulations (Banerjee et al., 2018). They describe the *lollipop* conformation as a rather unlikely configuration for the GPI-anchor alone and with the GFP attached to it, while their simulations clearly favor the *flop-down* state. However, due to its flexibility, additional modifications at the GPI-core itself and residues close to the C-terminal ω-site are good candidates to drive the GPI-anchor configuration in a way that favors certain shapes and thus the ability of proteins to bind more or less likely. As mentioned above, Cwh43 requires the presence of both charged EtN-Ps at Man1 and Man2, while the ER-exit by the p24-complex is highly enhanced upon removal of the Man2-EtN-P. While an explanation can be the direct steric hindrance of the modifications with the binding pockets of the respective proteins, another idea might be their influence on the GPI-shape. This may happen due to the negatively charged surface of the lipid bilayer and the positive amino group of the EtN-Ps. That positive charges indeed have an effect on the final localization has been shown for amino acids at the ω-minus region in *S. cerevisiae* and *A. fumigatus* (Frieman and

Cormack, 2003; Hamada et al., 1998, 1999; Ouyang et al., 2013). It was observed that basic amino acids at ω-1 or ω-2 lead to a residence in the plasma membrane, while tyrosine, valine, isoleucine, leucine or asparagine favor the attachment to the cell wall. The final destination could be modified by respective point mutations or chimeric proteins with exchanged ω-regions, thereby targeting plasma membrane proteins into the cell wall or *vice versa*. Hydrophobic residues may allow less stringent contact with the plasma membrane leading to a conformation that preferentially allows binding of the GPI-core glycan modifying enzyme Cdc1, while lysine at ω-1 or ω-2 helps to flatten the glycan on the lipid head groups, hiding it from Cdc1 and allowing the action of Cwh43, which anyway has to act on the lipid moiety, and not the glycan itself.

Clearly the GPI-AP/lipid interactions would not absolutely fix a certain conformation, but rather lead to a preferred state that still has intrinsic dynamics. Actually this kind of system would explain, why the analyses of GPI-PMPs and GPI-CWPs suggest a more relative than an absolute distribution, as several GPI-PMPs can be also found covalently attached to the cell wall.

5.4 Fungal cell wall biogenesis as a drug target

Secondary infections caused by fungal pathogens are an increasingly recognized threat for immunocompromised patients suffering for example from cancer, HIV or neutropenia. Ascomycota are commonly identified as the cause of such systemic fungal infections, among which *Candida* spp., *Aspergillus* spp., and *Cryptococcus* spp. belong to the most frequently found members (Enoch et al., 2006). Echinocandin-class drugs target the cell wall biosynthesis by inhibiting the β1,3-glucan synthases. However, due to their good tolerance in patients, the high number of their usage led to an increased number of echinocandin-resistant strains, especially among *Candida* spp. (Perlin, 2015). Anyway, the cell wall of fungi seems to be a promising target for antimycotica, as its structures are absent in humans, thus implying a low degree of negative side effects. Recently, the development of potent inhibitors against the β1,3-glycosyltransferases from the GH72 family has been published (Delso et al., 2018). The authors described an initially low-affinity laminaritriose-based binder, which was modified at its flexible linker at the reducing end to a final binding affinity in the low µM-region (3.1 µM for *Sc*Gas2, 1.5 µM for *Af*Gel4). However, the substances have not been tested *in vivo*. As the incorporation of proteins into the cell wall is an essential process for fungal life, the development of a potent *Dfg5*-inhibitor seems to be a promising approach. For

bacterial *endo*-α1,6-mannanases, O- and S-linked isofagomine based inhibitors are known (ManIFG and ManSIFG), which bind to the -1/-2 subsites with affinities of about 1 µM (Belz et al., 2017; Thompson et al., 2015b). The addition of a second mannoside to the non-hydrolyzable S-linked variant ($(ManS)_2IFG$) increased the affinity about 37-fold (Belz et al., 2017).

Compared to these molecules, the binding of FP-1 to *Ct*Dfg5 (K_D=2.71 mM±0.96) is rather low, however a clear impact on the viability of *S. cerevisiae* could be observed, suggesting a good starting point for further drug development. An obvious step for optimizing the binding properties would be the rationalized binding optimization by the substitution of functional groups, which would contribute to an energetically favored ligand-protein interaction and thus lowered binding enthalpies. Another common approach during drug optimization is called *SAR by NMR* (i.e. structure-activity relationships obtained by nuclear magnetic resonance), where NMR-based screens identify two reasonable binders, which are then covalently linked and result in high affinity ligands (Shuker et al., 1996). For example, it was shown that the combination of a 17 mM and a 0.02 mM binder of stromelysin result in a potent inhibitor with a K_D=15 nM when linked together (Hajduk et al., 1997). This is achieved by a decrease of freedom in their free states and accordingly a lower entropy penalty upon binding, which results in better ΔG values. In fact, the simultaneous binding of FP-1 and mannose has been shown. The closest distance between FP-1 and the -2 subsite mannose is 3.3 Å, however when superimposed with the α1,2-mannobiose state the actual distance is only 2.8 Å, which is quite close to the 2.77 Å needed for a thio ether linkage as used in the *endo*-α1,6-mannanase inhibitors described above (Figure 45). The combination of the identified fragment together with a partial structure of the real substrate resulting in a ligand spanning the GPI-glycan binding site from subsite -3 to +1 is a promising candidate to inhibit the lipid-to-wall transfer already at low inhibitor concentrations.

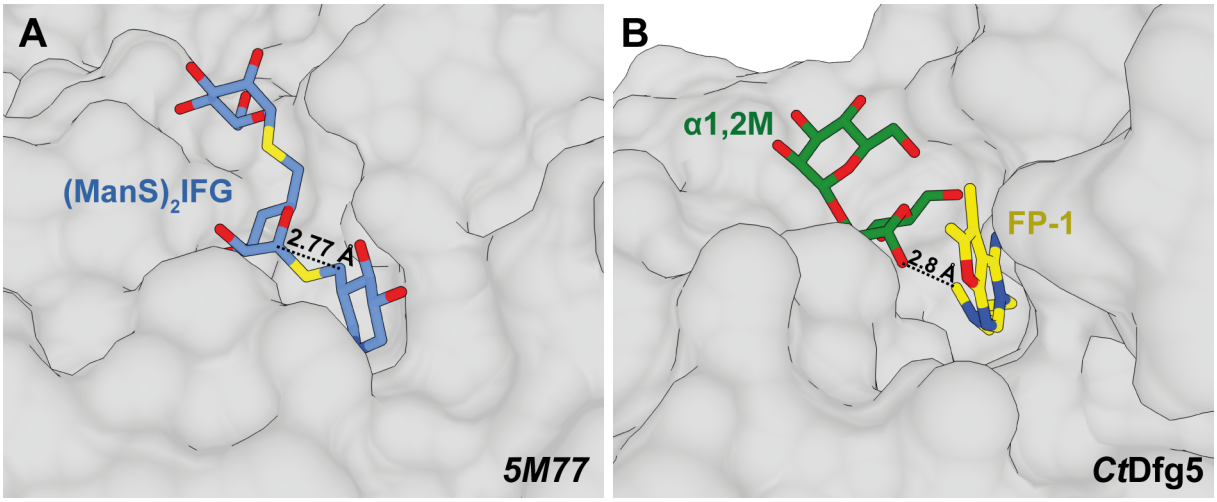


Figure 45: Comparison of the $(ManS)_2IFG$ state of bacterial α1,6-mannanases with a superposition of FP-1 and α1,2-mannobiose states from *Ct*Dfg5. A The inhibitor is shown as blue sticks in the context of the protein surface. The distance between the adjacent carbons of the thio-ether linkage is indicated. The PDB code is given at the bottom right. **B** The *Ct*Dfg5-complex states of FP-1 (yellow sticks) and α1,2-mannobiose (green sticks) are superimposed in the context of the protein surface and the resulting distance between the closest atoms is indicated. Surfaces are colored in grey.

Aside from the high potential as a specific antifungal drug, such a molecule can be an excellent tool to study the behavior of *Dfg5*-proteins *in vivo*. As discussed in the following, *Dfg5*-paralogs exhibit a number of individual preferences on their substrate. FP-1 may enable to shut off or reduce the action of *Dfg5*-single mutants to directly observe differences in the attachment of proteins onto the cell wall without changing the gene expression.

5.5 The functional diversity of *Dfg5*-paralogs

The SSN analysis not only revealed distinct clusters of bacterial and fungal homologs, but also provided insights into the diverse number of *Dfg5*-proteins within different organisms ranging from just two copies in *S. cerevisiae* up to seven paralogs in *Neurospora crassa* (see Table 2 on p.55). As only the deletion of all members is expected to cause a lethal phenotype in yeasts, the paralogs can complement the loss of the other protein at least to some extent, which implies a redundancy in the molecular function of *Dfg5*-proteins. This goes along with the proposed mechanism, as all *Dfg5*-proteins use the same GPI-structure as a substrate. Emphasizing this notion, the six *Dfg5*-members of *A. fumigatus* expressed under *in vitro* conditions were recently shown to incorporate the GPI-anchored galactomannan to β1,3-glucan of the cell wall, expanding the range of substrates to lipid-anchored carbohydrates (Muszkieta et al., 2019). There it was necessary to delete all expressed *Dfg5*-paralogs to

abolish the presence of GM in the cell wall. Astonishingly, *AfDFG3* possessing the largest impact on cell morphology in *Aspergillus fumigatus*, could complement a temperature sensitive double mutant in *S. cerevisiae*, showing that although the GPI-GM is suggested as the substrate in the mold, it is able to transfer GPI-AP to the cell wall in the yeast, as the lipid-anchored galactomannan is a unique feature for *Aspergillus* spp. and the only plausible target in yeasts are GPI-APs (Muszkieta et al., 2019). However, unlike in yeasts, the function of *Dfg5*-proteins in *A. fumigatus* does not seem to be essential, which might be explained due to different cell wall architectures.

In yeasts more stringent culture conditions such as pH stress or nutritional starvation showed morphological differences between the two *Dfg5*-paralogs (Mösch and Fink, 1997; Spreghini et al., 2003). Both, *Candida albicans* and *S. cerevisiae*, were shown to have defects in their (pseudo)hyphae formation in the Δ*dfg5* background, but not in Δ*dcw1*. A different effect was also observed regarding the adhesive ability of Flo11-expressing *S. cerevisiae* cells, where the Δ*dfg5* strain could be removed by a wash test indicating non-adhesive cells due to a compromised incorporation of Flo11 into the cell wall, while Δ*dcw1* cells are resistant to the wash test and stick to the solid medium (*personal communication with Gesa Schmitz, AG Mösch*). Furthermore, *N. crassa* exhibits a colonial phenotype with highly irregularly grown hyphae in the Δ*dfg5* background, while Δ*dcw1* has just a subtle phenotype (Maddi et al., 2012). Taken together the different paralogs possess specificity for certain targets, however the discrimination on the molecular level is still elusive, due to the apparent similarity of the used substrate. This is even true on an interspecies level as *Dfg5*-orthologs from filamentous fungi can complement in yeast (Muszkieta et al., 2019).

In this context, a sequence feature could be observed in the acceptor-binding loop (ABL1) described for the laminaribiose state. By looking at the overall conservation as it is displayed there (see Figure 30 on p.78), this loop region does not seem to have a certain function, as the level of preservations is *strikingly inconspicuous*. However, ABL1 was the only region of the protein showing an increased degree of flexibility. While this loop was missing in the electron density maps of the initial data used for structure determination, the loop became structured in the seeded screening upon lowering the overall pH. Surprisingly, this loop region was able to move up to 4 Å towards the active site when comparing different states, while the overall conformation of the protein remained unchanged (Figure 46).

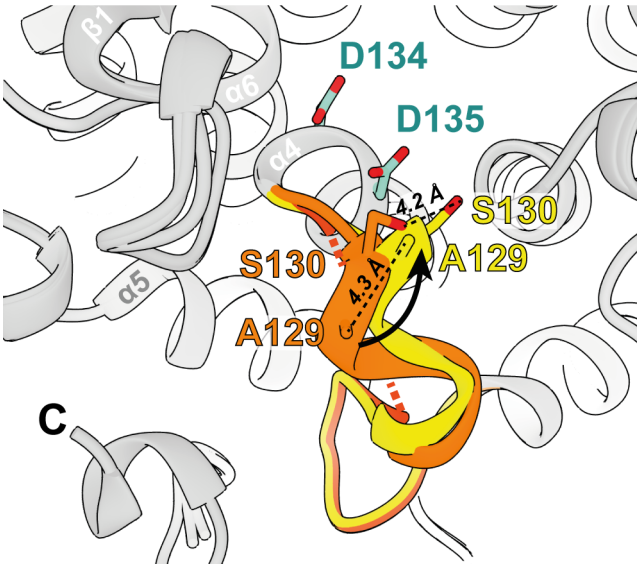


Figure 46: Dynamics within the ABL1-region. Three different ABL1-conformations were observed in the structures of *Ct*Dfg5, which are superimposed here. The first one lacks density for P125-S130 (colored in red, indicated by dotted line). The second conformation is shown in orange with density for the complete loop, and the third state colored in yellow is moved by approx. 4 Å indicated by the distances between the respective residues A129 and S130 and the arrow showing the *direction* towards the active site DD-motif. Secondary structure is indicated for orientation, residues are shown as sticks, oxygens are colored in red. The C-terminus is labeled.

When looking at the individual sequences of the ABL1 in the multiple sequence alignment, a surprising pattern can be observed, where several subpopulations emerge from the *Dfg5*-subfamily as shown in Figure 47 (SFigure 10 and SFigure 11 on pp.143-143 for the full ABL1-alignment). Intriguingly, some of these subpopulations exist only in filamentous fungi, others are exclusively found in Saccharomycetes, while being absent in the other subphylum. Surprisingly, it is possible to distinguish between the paralogs Dfg5 and Dcw1 in Saccharomycetes just by comparing their ABL1 motifs. All characterized and annotated Dfg5 homologs possess an invariable methionine together with valine or threonine (MV/T), while Dcw1 homologs usually possessed a TT-motif with few variations in the first threonine. Undeniably, the situation in Pezizomycotina is more complex, as the *Dfg5*-subfamily is more diverse. However, a pattern of conserved ABL1-motifs suggests a certain degree of substrate specialization within the different taxonomic classes, where some can be found in several classes (R/KT-motif), while others seem to be specific for a certain class (QW and N/SY for Eurotiomycetes, AS for Sordariomycetes). It is worth mentioning that by excluding apparent outliers of the shown ABL1-subpopulations, all protein sequences from Taphrinomycotina (e.g. the fission yeast *Schizosaccharomyces pombe)* disappeared from the final alignment. However, a similar pattern can be observed for these proteins as well, but the number of

sequences is rather limited (just 31 hits in the Uniprot database when searching *PF03663+Taphrinomycotina*) and therefore they are removed from this analysis.

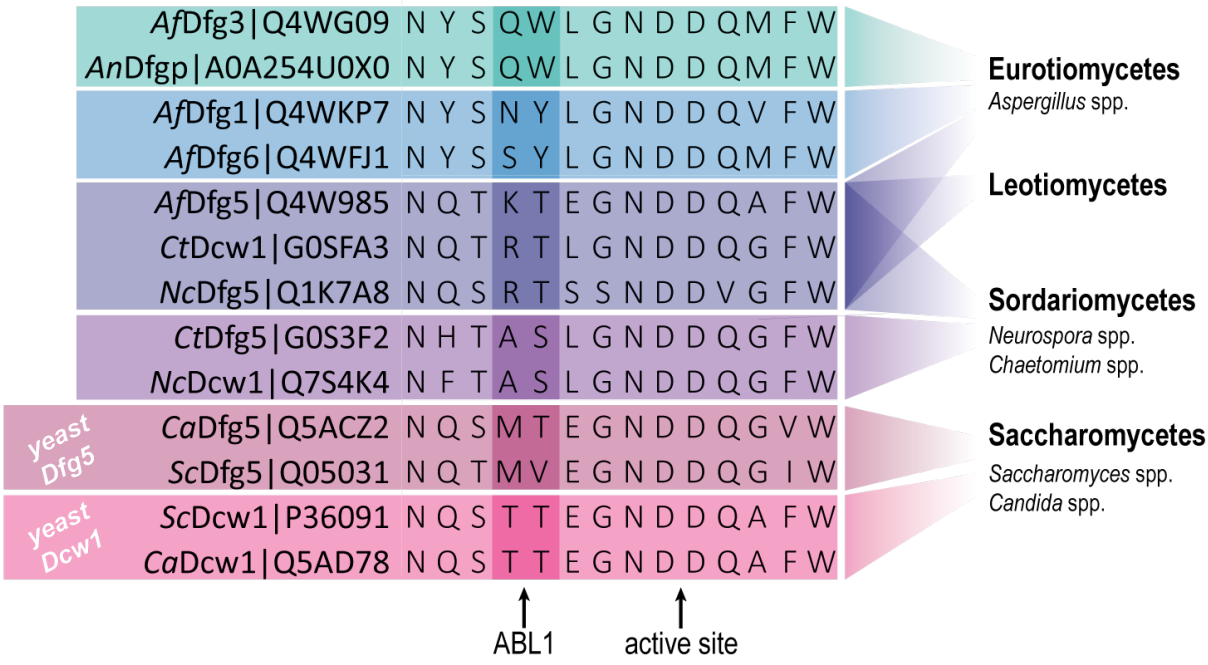


Figure 47: Multiple sequence alignment of *Dfg5*-proteins highlighting subpopulations in the ABL1-region. The ABL1 is highlighted by the boxes and the arrow, active site DD-residues are indicated by an arrow as well. The respective Uniprot-IDs and names according to the literature are given. The emerging subpopulations of the two homologs in yeast are indicated. *Af Aspergillus fumigatus*, *An Aspergillus niger*, *Ct Chaetomium thermophilum*, *Nc Neurospora crassa*, *Ca Candida albicans*, *Sc Saccharomyces cerevisiae*.

Since ABL1 is associated with acceptor binding (i.e. the cell wall polysaccharides), a correlation between the number of *Dfg5*-paralogs and the complexity of the cell wall could be assumed in order to optimize the lipid-to-wall transfer for certain proteins or in a special cell morphotype. However, whether other regions of *Dfg5*-proteins discriminate between different substrates cannot be concluded at this point. Anyway, as *Af*Dfg3 is able to complement the temperature-sensitive phenotype in yeast, the general function seems to be highly conserved throughout the *Dfg5*-subfamily, also supporting the notion that the principle type of the acceptor β-glucan is not depending on the protein, but on the supply at the plasma membrane (Muszkieta et al., 2019). It can also be speculated, whether the ABL1-loop may act like a flap for the efficient binding of the GPI-anchor, as this region clearly is in close contact to the membrane, if not even dipped into it. A dynamic, shovel-like behavior with different modifications at the shovel may contribute to a preferred or less favored binding of GPI-APs,

depending on their individual (most likely C-terminally influenced) shape of the GPI-core glycan.

5.6 The common feature of GH76-proteins

After describing the different aspects of the *Dfg5*-subfamily of GH76-proteins it becomes clear that this family separates from the already described bacterial homologs not only in the SSN analysis, but also on a functional level. Generally, different members of individual glycoside hydrolase families can have a broad spectrum of substrates (Naumoff, 2011). Apart *Dfg5*-proteins, another subfamily of GH76-proteins has been studied on a structural level in the scope of this thesis. This cluster found in the SSN is comprised of proteins from bacteria and fungi, therefore termed as the *Fungal/Bacteria mixed* subfamily. Only a fungal members have been assessed, however a further characterization concerning their cellular function is missing, as no phenotype could yet be observed (Maddi et al., 2012; Muszkieta et al., 2019). From the SSN analysis, this class is apparently closer related to the bacterial homologs, suggesting also a similar function. However, an α1,6-mannanase activity could not be observed *in vitro* for *Ct*GH76 (Markus Friedrich, Master thesis). He also obtained the crystal structure of the apo-state, which showed differences in the two *wing* domains described above, while the overall $(\alpha/\alpha)_6$-helical barrel fold is conserved (see Figure 36 on p.86).

To extend the functional knowledge on GH76-proteins, a similar glycan-fragment screening approach has been applied to the apo-crystals of *Ct*GH76 as done before for *Ct*Dfg5. The crystal structure in complex with an α1,3-α1,6-mannotriose showed that the extended wing region is involved in carbohydrate binding, however unlike for *Dfg5*-proteins no glucosamine could be observed at subsite +1 upon molar soaking. A striking difference between *Bc*GH76 and *Ct*Dfg5 is the coordination of α1,6-linked mannobiose. The bacterial protein utilizes α1,6-oligomers as substrate, which enables it to bind the biose to its -3/-2 subsites (Figure 48B). A coordination at subsite -1 could only be achieved by a combination of a substrate-binding canyon spanning substrate and an active-site variant, as the protein provides a better binding site for the disaccharide at -3/-2, while at subsite -1 the sugar moiety is supposed to be distorted for catalysis (Thompson et al., 2015b). This is a further indication that the substrates differ between *bacterial* and *Dfg5*-subfamilies, as seen by the α1,6-mannobiose state in *Ct*Dfg5 (Figure 48A). Interestingly, the approach applied to *Ct*GH76 revealed the same occupied subsites as in *Ct*Dfg5, suggesting that linear α1,6-mannooligomers are not the prime candidate for a substrate of the *Fungal/Bacteria mixed* subfamily (Figure 48C). This explains

the findings in Markus Friedrich's master thesis that *Ct*GH76 is incapable to hydrolyze α1,6-mannooligomers.

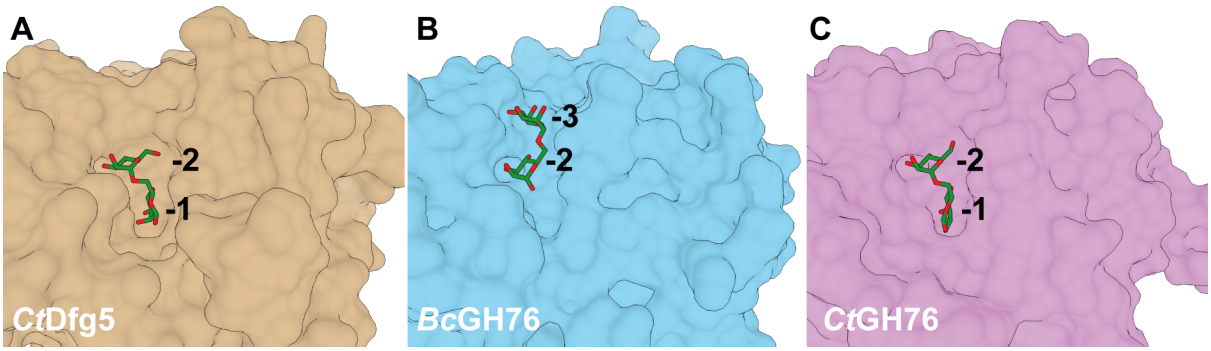


Figure 48: Known α1,6-mannobiose states of GH76-proteins. GH76 family members from three different subfamilies are shown in complex with α1,6-mannobiose with the respective protein surfaces and the ligand shown as green sticks with red oxygens bound to either subsites -2 to -1 in the fungal proteins (**A**+**C**) or -3 to -2 in the bacterial structure (**B**).

Due to the lack of *in vivo* work on this type of GH76-proteins it is currently hard to predict their function. The complex structures combined do not assemble a plausible glycan candidate on which *Ct*GH76 may act, although the α1,3-α1,6-Mannotriose exists within the high mannose N-glycan core structure known from yeast mannoproteins.

The α1,6-mannobiose state helps to understand the common feature of that protein family. By comparing the complex states of *Ct*Dfg5, *Ct*GH76 and *Bc*GH76, which occupy the -2/-1 subsites, the similar binding mode of that part of the binding site clearly stands out (Figure 49). It can therefore be hypothesized that GH76-proteins act on substrates that contain an α1,6-mannobiose core unit, which fits to the -2/-1 subsites, while coordinating the neighboring glycan residues in a subfamily-specific manner.

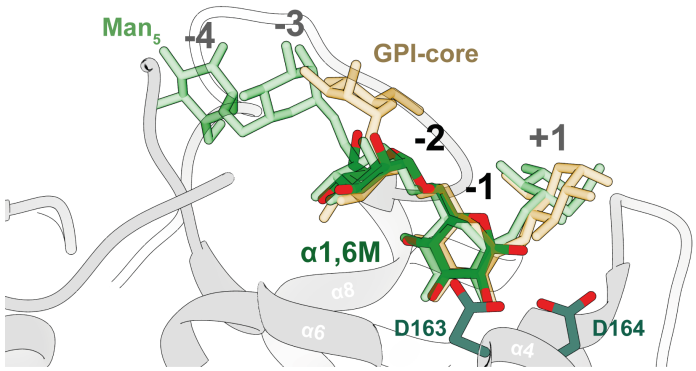


Figure 49: Comparison of substrate states from GH76-proteins in the context of *Ct*GH76. A superposition of the mannopentaose (Man_5) bound state of *Bc*GH76 (PDB 5AGD), the assembled GPI-state of *Ct*Dfg5 and the α1,6-mannobiose (α1,6M) state of *Ct*GH76 is shown in the context of *Ct*GH76. The active site residues (D163/D164) of *Ct*GH76 are shown as sticks for orientation to assign the substrate binding sites. The ligands are shown as sticks, Man_5 is colored in transparent

green, GPI-core in transparent orange and α1,6M is shown as solid sticks to highlight the conserved substrate element of GH76-proteins.

5.7 Outlook

Within the scope of this work the role of *Dfg5*-proteins was analyzed and clarified by in depth structural analysis of a heterologously produced thermophilic homolog from *Chaetomium thermophilum* in terms of its underlying mechanism to transfer GPI-anchored substrates from the plasma membrane into the fungal cell wall. First experimentally derived insights into the spatial arrangement of the GPI-anchor core are provided. Furthermore, the suggested great potential as a drug target was pointed out by identifying a small molecule bound to the active site of the protein, which was shown to inhibit the growth of *Saccharomyces cerevisiae*. Finally, the common feature of GH76-proteins was suggested by combining the known insights of substrate coordination with a further member of another GH76-subfamily showing a conserved α1,6-mannobiose coordination at the -2/-1 subsites.

These findings only opened the gate for further studies. Still a robust method for an *in vitro* assay showing the hydrolysis of the GPI-core and the transglycosylation onto the acceptor remains to be established. A major problem for a successful transglycosylation might be the absence of the membrane, where the transfer is supposed to happen. From a structural point of view, there is no reason why *Dfg5*-proteins should not act like their bacterial homologs. A functional assay may require a membrane-based system including embedded glucan synthases, correctly processed GPI-anchored substrates, membrane associated *Dfg5*-proteins in an aqueous solution containing native concentrations of calcium. An *in vivo* system with suitable reporters (such as described for *Aspergillus fumigatus* in Donat et al., 2012) and a sophisticated shutoff system in combination with a potent inhibitor such as an improved FP-1 might work as well. Such a system may also enable the analysis of the specificities of the two yeast-paralogs on different targets. Combined with MS-based cell wall compositional analyses it may show which GPI-APs are preferentially transferred by the respective paralogs. A further aspect in this context is the discrepancy of paralog numbers in individual organisms, which range from two in yeasts to nine in *Neurospora crassa*. This number may depend on different developmental stages with variable acceptors in the cell wall, and an adaptation to specific donor substrates.

The sorting of GPI-APs along the secretory pathway and the decision at the plasma membrane for the final destination by decoding the EtN-P modifications remains to be validated as well.

While it is clear that the cleavage of the GPI-core glycan by *Dfg5*-proteins with the Man1-EtN-P cannot occur, molecular dynamic (MD-) simulations may answer the question if binding in principle can happen with such modifications, or not. A full molecular understanding of the processing factors within the secretory pathway that are responsible for the coding (i.e. Cdc1, Cwh43, Ted1) still lacks behind. In order to understand the precise sequence of the GPI-sorting, this aspect must be addressed as well.

The initial hit of FP-1 by structure-based fragment screening shows principal biological activity against yeast and is a good starting point for further drug development. However, it requires optimization due to the low binding properties. The simultaneous binding of FP-1 and mannose has been shown experimentally and the combination of two fragments is discussed above. Also a combination with α1,2-mannobiose seems attractive. This might be a promising approach to combine these two fragments in order to obtain an optimal inhibitor for *Dfg5*-proteins. Such inhibitors may be able to compete with native substrates already at low concentrations.

Finally, several subfamilies of the GH76-class remain to be characterized for a full understanding of these glycoside hydrolases. A start for the *Fungal/Bacteria mixed* class is already provided and a very likely common feature for GH76 proteins is discussed, however the actual function needs to be explained, also in terms of the function in archaea, where a few organisms contain such proteins.

6 Literature

Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 213–221.

Afonine, P. V, Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. Sect. D Biol. Crystallogr. *68*, 352–367.

Aguilar-Uscanga, B., and François, J.M. (2003). A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation. Lett. Appl. Microbiol. *37*, 268–274.

Aimanianda, V., Simenel, C., Garnaud, C., Clavaud, C., Tada, R., Barbin, L., Mouyna, I., Heddergott, C., Popolo, L., Ohya, Y., et al. (2017). The dual activity responsible for the elongation and branching of β-(1,3)-glucan in the fungal cell wall. MBio *8*, e00619-17.

Amlacher, S., Sarges, P., Flemming, D., van Noort, V., Kunze, R., Devos, D.P., Arumugam, M., Bork, P., and Hurt, E. (2011). Insight into Structure and Assembly of the Nuclear Pore Complex by Utilizing the Genome of a Eukaryotic Thermophile. Cell *146*, 277–289.

Andricopulo, A., Salum, L., and Abraham, D. (2009). Structure-Based Drug Design Strategies in Medicinal Chemistry. Curr. Top. Med. Chem. *9*, 771–790.

Ao, J., and Free, S.J. (2017). Genetic and biochemical characterization of the GH72 family of cell wall transglycosylases in *Neurospora crassa*. Fungal Genet. Biol. *101*, 46–54.

Ao, J., Chinnici, J.L., Maddi, A., and Free, S.J. (2015). The N-linked outer chain mannans and the Dfg5p and Dcw1p Endo-α-1,6-mannanases are needed for incorporation of *Candida albicans* glycoproteins into the cell wall. Eukaryot. Cell *14*, 792–803.

Arndt, U.W. (1968). The optimum strategy in measuring structure factors. Acta Crystallogr. B. *24*, 1355–1357.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. *44*, W344–W350.

Atkinson, H.J., Morris, J.H., Ferrin, T.E., and Babbitt, P.C. (2009). Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. PLoS One *4*, e4345.

Baba, M., Baba, N., Ohsumi, Y., Kanaya, K., and Osumi, M. (1989). Three-dimensional analysis of morphogenesis induced by mating pheromone alpha factor in *Saccharomyces cerevisiae*. J. Cell Sci. *94*, 207–216.

Banerjee, P., Wehle, M., Lipowsky, R., and Santer, M. (2018). A molecular dynamics model for glycosylphosphatidyl-inositol anchors: “flop down” or “lollipop”? Phys. Chem. Chem. Phys. *20*, 29314–29324.

Belden, W.J., and Barlowe, C. (1996). Erv25p, a component of COPII-coated vesicles, forms a complex with Emp24p that is required for efficient endoplasmic reticulum to Golgi transport. J. Biol. Chem. *271*, 26939–26946.

Belz, T., Jin, Y., Coines, J., Rovira, C., Davies, G.J., and Williams, S.J. (2017). An atypical interaction explains the high-affinity of a non-hydrolyzable S-linked 1,6-α-mannanase inhibitor. Chem. Commun. *53*, 9238–9241.

Benachour, A., Sipos, G., Flury, I., Reggiori, F., Canivenc-Gansel, E., Vionnet, C., Conzelmann, A., and Benghezal, M. (1999). Deletion of GPI7, a yeast gene required for addition of a side chain to the glycosylphosphatidylinositol (GPI) core structure, affects GPI protein transport, remodeling, and cell wall integrity. J. Biol. Chem. *274*, 15251–15261.

Binder, H., and Zschörnig, O. (2002). The effect of metal cations on the phase behavior and hydration characteristics of phospholipid membranes. Chem. Phys. Lipids *115*, 39–61.

BioSolveIT GmbH SeeSAR v8.1.

Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. *7*, 1513–1523.

Bock, T., Chen, W.-H., Ori, A., Malik, N., Silva-Martin, N., Huerta-Cepas, J., Powell, S.T., Kastritis, P.L., Smyshlyaev, G., Vonkova, I., et al. (2014). An integrated approach for genome annotation of the eukaryotic thermophile *Chaetomium thermophilum*. Nucleic Acids Res. *42*, 13525–13533.

Boone, C., Sommer, S.S., Hensel, A., and Bussey, H. (1990). Yeast KRE genes provide evidence for a pathway of cell wall β-glucan assembly. J. Cell Biol. *110*, 1833–1843.

Bosson, R., Jaquenoud, M., and Conzelmann, A. (2006). GUP1 of *Saccharomyces*

cerevisiae Encodes an O-Acyltransferase Involved in Remodeling of the GPI Anchor. Mol. Biol. Cell *17*, 2636–2645.

Bowman, S.M., and Free, S.J. (2006). The structure and synthesis of the fungal cell wall. BioEssays *28*, 799–808.

Breinig, F., Schleinkofer, K., and Schmitt, M.J. (2004). Yeast Kre1p is GPI-anchored and involved in both cell wall assembly and architecture. Microbiology *150*, 3209–3218.

Brown, G.D., Gow, N.A.R., Denning, D.W., Netea, M.G., Levitz, S.M., and White, T.C. (2012a). Hidden Killers: Human Fungal Infections. Sci. Transl. Med. *4*, 1–10.

Brown, G.D., Denning, D.W., and Levitz, S.M. (2012b). Tackling human fungal infections. Science *336*, 647.

Brünger, A. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472–475.

Bubb, W.A. (2003). NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. Concepts Magn. Reson. Part A Bridg. Educ. Res. *19*, 1–19.

Bussey, H. (1991). K1 killer toxin, a pore-forming protein from yeast. Mol. Microbiol. *5*, 2339–2343.

Cabib, E. (2009). Two Novel Techniques for Determination of Polysaccharide Cross-Links Show that Crh1p and Crh2p Attach Chitin to both β(1-6)- and β(1-3)Glucan in the *Saccharomyces cerevisiae* Cell Wall. Eukaryot. Cell *8*, 1626–1636.

Cabib, E., and Arroyo, J. (2013). How carbohydrates sculpt cells: chemical control of morphogenesis in the yeast cell wall. Nat. Rev. Microbiol. *11*, 648–655.

Cabib, E., and Durán, A. (2005). Synthase III-dependent chitin is bound to different acceptors depending on location on the cell wall of budding yeast. J. Biol. Chem. *280*, 9170–9179.

Cabib, E., Bowers, B., and Roberts, R.L. (1983). Vectorial synthesis of a polysaccharide by isolated plasma membranes. Proc. Natl. Acad. Sci. *80*, 3318–3321.

Canivenc-Gansel, E., Imhof, I., Reggiori, F., Burda, P., Conzelmann, A., and Benachour, A. (1998). GPI anchor biosynthesis in yeast: Phosphoethanolamine is attached to the α1,4-linked mannose of the complete precursor glycophospholipid. Glycobiology *8*, 761–770.

Cappellaro, C., Mrsa, V., and Tanner, W. (1998). New Potential Cell Wall Glucanases of *Saccharomyces cerevisiae* and Their Involvement in Mating. J. Bacteriol. *180*, 5030–5037.

Caro, L.H.P., Tettelin, H., Vossen, J.H., Ram, A.F.J., Van Den Ende, H., and Klis, F.M. (1997). In silicio identification of glycosyl-phosphatidylinositol-anchored plasma-membrane and cell wall proteins of *Saccharomyces cerevisiae*. Yeast *13*, 1477–1489.

Castillon, G.A., Watanabe, R., Taylor, M., Schwabe, T.M.E., and Riezman, H. (2009). Concentration of GPI-anchored proteins upon ER exit in yeast. Traffic *10*, 186–200.

Cava, F., Hidalgo, A., and Berenguer, J. (2009). *Thermus thermophilus* as biological model. Extremophiles *13*, 213–231.

Chakraborty, N., and d'Alarcao, M. (2005). An anionic inositol phosphate glycan pseudotetrasaccharide exhibits high insulin-mimetic activity in rat adipocytes. Bioorg. Med. Chem. *13*, 6732–6741.

Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Avila-Campilo, I., Creech, M., Gross, B., et al. (2007). Integration of biological networks and gene expression data using cytoscape. Nat. Protoc. *2*, 2366–2382.

Costachel, C., Coddeville, B., Latgé, J.P., and Fontaine, T. (2005). Glycosylphosphatidylinositol-anchored fungal polysaccharide in *Aspergillus fumigatus*. J. Biol. Chem. *280*, 39835–39842.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: A sequence logo generator. Genome Res. *14*, 1188–1190.

Cuskin, F., Lowe, E.C., Temple, M.J., Zhu, Y., Cameron, E.A., Pudlo, N.A., Porter, N.T., Urs, K., Thompson, A.J., Cartmell, A., et al. (2015). Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. Nature *517*, 165–169.

Davies, G.J., Wilson, K.S., and Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. Biochem. J. *321*, 557–559.

Delso, I., Valero-Gonzalez, J., Gomollón-Bel, F., Castro-López, J., Fang, W., Navratilova, I., van Aalten, D.M.F., Tejero, T., Merino, P., and Hurtado-Guerrero, R. (2018). Inhibitors against Fungal Cell Wall Remodeling Enzymes. Chem. Med. Chem. *13*, 128–132.

Diederichs, K., and Karplus, P.A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat. Struct. Biol. *4*, 269–275.

Dijkgraaf, G.J.P., Abe, M., Ohya, Y., and Bussey, H. (2002). Mutations in FskIp affect the cell wall content of β-1,3- and β-1,6-glucan in *Saccharomyces cerevisiae*. Yeast *19*, 671–690.

Dimaio, F., Terwilliger, T.C., Read, R.J., Wlodawer, A., Oberdorfer, G., Wagner, U., Valkov, E., Alon, A., Fass, D., Axelrod, H.L., et al. (2011). Improved molecular replacement by density- and energy-guided protein structure optimization. Nature *473*, 540–543.

Donat, S., Hasenberg, M., Schäfer, T., Ohlsen, K., Gunzer, M., Einsele, H., Löffler, J., Beilhack, A., and Krappmann, S. (2012). Surface display of *Gaussia princeps* luciferase allows sensitive fungal pathogen detection during cutaneous aspergillosis. Virulence *3*, 50–61.

Douglas, C.M., Foor, F., Marrinan, J.A., Morin, N., Nielsen, J.B., Dahl, A.M., Mazur, P., Baginsky, W., Li, W., and El-Sherbeini, M. (1994). The *Saccharomyces cerevisiae* FKS1 (ETG1) gene encodes an integral membrane protein which is a subunit of 1,3-beta-D-glucan synthase. Proc. Natl. Acad. Sci. U. S. A. *91*, 12907–12911.

Dranginis, A.M., Rauceo, J.M., Coronado, J.E., and Lipke, P.N. (2007). A Biochemical Guide to Yeast Adhesins: Glycoproteins for Social and Antisocial Occasions. Microbiol. Mol. Biol. Rev. *71*, 282–294.

Drenth, J. (2010). Principles of Protein X-Ray Crystallography.

Drgonová, J., Drgon, T., Tanaka, K., Kollár, R., Chen, G.C., Ford, R.A., Chan, C.S.M., Takai, Y., and Cabib, E. (1996). Rho1p, a yeast protein at the interface between cell polarization and morphogenesis. Science *272*, 277–279.

Drum, R.W., and Gordon, R. (2003). Star Trek replicators and diatom nanotechnology. Trends Biotechnol. *21*, 325–328.

Dupres, V., Dufreêne, Y.F., and Heinisch, J.J. (2010). Measuring cell wall thickness in living yeast cells using single molecular rulers. ACS Nano *4*, 5498–5504.

Duran, A., Bowers, B., and Cabib, E. (1975). Chitin synthetase zymogen is attached to the yeast plasma membrane. Proc. Natl. Acad. Sci. *72*, 3952–3955.

Ecker, M., Deutzmann, R., Lehle, L., Mrsa, V., and Tanner, W. (2006). Pir proteins of *Saccharomyces cerevisiae* are attached to β-1,3-glucan by a new protein-carbohydrate linkage. J. Biol. Chem. *281*, 11523–11529.

Eisenhaber, B., Bork, P., and Eisenhaber, F. (1998). Sequence properties of GPI-anchored proteins near the omega-site: constraints for the polypeptide binding site of the putative transamidase. Protein Eng. *11*, 1155–1161.

Eisenhaber, B., Schneider, G., Wildpaner, M., and Eisenhaber, F. (2004). A sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans, Candida albicans, Neurospora crassa, Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. J. Mol. Biol. *337*, 243–253.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., et al. (2019). The Pfam protein families database in 2019. Nucleic Acids Res. *47*, D427–D432.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Cryst *66*, 486–501.

Enoch, D.A., Ludlam, H.A., and Brown, N.M. (2006). Invasive fungal infections: A review of epidemiology and management options. J. Med. Microbiol. *55*, 809–818.

Erwig, L.P., and Gow, N. a. R. (2016). Interactions of fungal pathogens with phagocytes. Nat. Rev. Microbiol. *14*, 163–176.

Fankhauser, C., Homans, S.W., Thomas-Oates, J.E., McConville, M.J., Desponds, C., Conzelmann, A., and Ferguson, M.A.J. (1993). Structures of glycosylphosphatidylinositol membrane anchors from *Saccharomyces cerevisiae*. J. Biol. Chem. *268*, 26365–26374.

Ferguson, M.A.J., Homans, S.W., Dwek, R.A., and Rademacher, T.W. (1988). Glycosyl-Phosphatidylinositol Moiety That Anchors *Trypanosoma brucei* Variant Surface Glycoprotein to the Membrane. Science *239*, 753–759.

Ferguson, M.A.J., Kinoshita, T., and Hart, G.W. (2009). Glycosylphosphatidylinositol Anchors. In Essentials of Glycobiology, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, and M.E. Etzler, eds. (New York: Cold Spring Harbor Labaratory Press), pp. 143–162.

Fleet, G.H. (1991). Cell Walls. In The Yeasts, pp. 199–277.

Flury, I., Benachour, A., and Conzelmann, A. (2000). YLL031c belongs to a novel family of membrane proteins involved in the transfer of ethanolaminephosphate onto the core structure of glycosylphosphatidylinositol anchors in yeast. J. Biol. Chem. *275*, 24458–24465.

Fontaine, T., Simenel, C., Dubreucq, G., Adam, O., Delepierre, M., Lemoine, J., Vorgias, C.E., Diaquin, M., and Latgé, J.P. (2000). Molecular organization of the alkali-insoluble fraction of *Aspergillus fumigatus* cell wall. J. Biol. Chem. *275*, 27594–27607.

Fontaine, T., Magnin, T., Melhert, A., Lamont, D., Latgé, J.P., and Ferguson, M.A.J. (2003). Structures of the glycosylphosphatidylinositol membrane anchors from *Aspergillus fumigatus* membrane proteins. Glycobiology *13*, 169–177.

Free, S.J. (2013). Fungal Cell Wall Organization and Biosynthesis. In Advances in Genetics, (Elsevier Inc.), pp. 33–82.

Frieman, M.B., and Cormack, B.P. (2003). The omega-site sequence of glycosylphosphatidylinositol-anchored proteins in *Saccharomyces cerevisiae* can determine distribution between the membrane and the cell wall. Mol. Microbiol. *50*, 883–896.

Fujii, T., Shimoi, H., and Iimura, Y. (1999). Structure of the glucan-binding sugar chain of Tip1p, a cell wall protein of *Saccharomyces cerevisiae*. Biochim. Biophys. Acta - Gen. Subj. *1427*, 133–144.

Fujita, M., and Kinoshita, T. (2012). GPI-anchor remodeling: Potential functions of GPI-anchors in intracellular trafficking and membrane dynamics. Biochim. Biophys. Acta - Mol. Cell Biol. Lipids *1821*, 1050–1058.

Fujita, M., Umemura, M., Yoko-o, T., and Jigami, Y. (2006a). PER1 Is Required for GPI-Phospholipase A 2 Activity and Involved in Lipid Remodeling of GPI-anchored Proteins. Mol. Biol. Cell *17*, 5253–5264.

Fujita, M., Yoko-o, T., and Jigami, Y. (2006b). Inositol Deacylation by Bst1p Is Required for the Quality Control of Glycosylphosphatidylinositol-anchored Proteins. Mol. Biol. Cell *17*, 834–850.

Fujita, M., Maeda, Y., Ra, M., Yamaguchi, Y., Taguchi, R., and Kinoshita, T. (2009). GPI Glycan Remodeling by PGAP5 Regulates Transport of GPI-Anchored Proteins from the ER to the Golgi. Cell *139*, 352–365.

Galperin, M.Y., and Jedrzejas, M.J. (2001). Conserved core structure and active site residues in alkaline phosphatase superfamily enzymes. Proteins *45*, 318–324.

Gastebois, A., Clavaud, C., Aimanianda, V., and Latgé, J.P. (2009). *Aspergillus fumigatus*: Cell wall polysaccharides, their biosynthesis and organization. Future Microbiol. *4*, 583–595.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook, pp. 571–607.

Gaurivaud, P., Baranowski, E., Pau-Roblot, C., Sagné, E., Citi, C., and Tardy, F. (2016). *Mycoplasma agalactiae* Secretion of b-1,6-Glucan, a Rare Polysaccharide in Prokaryotes, Is Governed by High-Frequency Phase Variation. Appl. Environ. Microbiol. *82*, 3370–3383.

Gaynor, E.C., Mondesert, G., Grimme, S.J., Reed, S.I., Orlean, P., and Emr, S.D. (1999). MCD4 Encodes a Conserved Endoplasmic Reticulum Membrane Protein Essential for Glycosylphosphatidylinositol Anchor Synthesis in Yeast. Mol. Biol. Cell *10*, 627–648.

Gerlt, J.A., Bouvier, J.T., Davidson, D.B., Imker, H.J., Sadkhin, B., Slater, D.R., and Whalen, K.L. (2015). Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim. Biophys. Acta - Proteins Proteomics *1854*, 1019–1037.

Ghugtyal, V., Vionnet, C., Roubaty, C., and Conzelmann, A. (2007). CWH43 is required for the introduction of ceramides into GPI anchors in *Saccharomyces cerevisiae*. Mol. Microbiol. *65*, 1493–1502.

Goldman, R.C., Sullivan, P.A., Zakula, D., and Capobianco, J.O. (1995). Kinetics of β-1,3 Glucan Interaction at the Donor and Acceptor Sites of the Fungal Glucosyltransferase Encoded by the BGL2 Gene. Eur. J. Biochem. *227*, 372–378.

Gonzalez, M., Lipke, P.N., and Ovalle, R. (2009). GPI proteins in biogenesis and structure of yeast cell walls. In The Enzymes, A.K. Menon, T. Kinoshita, P. Orlean, and F. Tamanoi, eds. (New York: Academic Press), pp. 321–356.

Gonzalez, M., Goddard, N., Hicks, C., Ovalle, R., Rauceo, J.M., Jue, C.K., and Lipke, P.N. (2010). A screen for deficiencies in GPI-anchorage of wall glycoproteins in yeast. Yeast *27*, 583–596.

Goossens, K., and Willaert, R. (2010). Flocculation protein structure and cell-cell adhesion mechanism in *Saccharomyces cerevisiae*. Biotechnol. Lett. *32*, 1571–1585.

de Groot, P.W.J., Hellingwerf, K.J., and Klis, F.M. (2003). Genome-wide identification of fungal GPI proteins. Yeast *20*, 781–796.

De Groot, P.W.J., Ruiz, C., De Vázquez Aldana, C.R., Duenas, E., Cid, V.J., Rey, F. Del, Rodríquez-Peña, J.M., Pérez, P., Andel, A., Caubín, J., et al. (2001). A genomic approach for the identification and classification of genes involved in cell wall formation and its regulation in *Saccharomyces cerevisiae*. Comp. Funct. Genomics *2*, 124–142.

De Groot, P.W.J., Ram, A.F., and Klis, F.M. (2005). Features and functions of covalently

linked proteins in fungal cell walls. Fungal Genet. Biol. *42*, 657–675.

Grout, D.H.G., and Vicl, G. (1998). Glycosidases and glycosyl transferases oligosaccharide synthesis. Curr. Opin. Chem. Biol. *2*, 98–111.

Guba, W., Meyder, A., Rarey, M., and Hert, J. (2016). Torsion Library Reloaded: A New Version of Expert-Derived SMARTS Rules for Assessing Conformations of Small Molecules. J. Chem. Inf. Model. *56*, 1–5.

Haass, F.A., Jonikas, M., Walter, P., Weissman, J.S., Jan, Y.-N., Jan, L.Y., and Schuldiner, M. (2007). Identification of yeast proteins necessary for cell-surface function of a potassium channel. Proc. Natl. Acad. Sci. *104*, 18079–18084.

Hagen, I., Ecker, M., Lagorce, A., Francois, J.M., Sestak, S., Rachel, R., Grossmann, G., Hauser, N.C., Hohelsel, J.D., Tanner, W., et al. (2004). Sed1p and Srl1p are required to compensate for cell wall instability in *Saccharomyces cerevisiae* mutants defective in multiple GPI-anchored mannoproteins. Mol. Microbiol. *52*, 1413–1425.

Hajduk, P.J., Sheppard, G., Nettesheim, D.G., Olejniczak, E.T., Shuker, S.B., Meadows, R.P., Steinman, D.H., Carrera, G.M., Marcotte, P.A., Severin, J., et al. (1997). Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. J. Am. Chem. Soc. *119*, 5818–5827.

Hamada, K., Terashima, H., Arisawa, M., and Kitada, K. (1998). Amino acid sequence requirement for efficient incorporation of glycosylphosphatidylinositol-associated proteins into the cell wall of *Saccharomyces cerevisiae*. J. Biol. Chem. *273*, 26946–26953.

Hamada, K., Terashima, H., Arisawa, M., Yabuki, N., and Kitada, K. (1999). Amino acid residues in the ω-minus region participate in cellular localization of yeast glycosylphosphatidylinositol-attached proteins. J. Bacteriol. *181*, 3886–3889.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. *166*, 557–580.

Hartwell, L.H., Culotti, J., and Reid, B. (1970). Genetic control of the cell-division cycle in yeast. I. Detection of mutants. Proc. Natl. Acad. Sci. *66*, 352–359.

Hecht, M.L., Tsai, Y.H., Liu, X., Wolfrum, C., and Seeberger, P.H. (2010). Synthetic inositol phosphoglycans related to GPI lack insulin-mimetic activity. ACS Chem. Biol. *5*, 1075–1086.

Heinz, D.W., Ryan, M., Smith, M.P., Weaver, L.H., Keana, J.F.W., and Hayes Griffith,

O. (1996). Crystal structure of phosphatidylinositol-specific phospholipase C from *Bacillus cereus* in complex with glucosaminyl(α1→6)-D-myo-inositol, an essential fragment of GPI anchors. Biochemistry *35*, 9496–9504.

Holan, Z., Pokorny, V., Beran, K., Gemperle, A., Tuzar, Z., and Baldrian, J. (2004). The glucan-chitin complex in *Saccharomyces cerevisiae*. Arch. Microbiol. *130*, 312–318.

Homans, S.W., Ferguson, M.A.J., Dwek, R.A., Rademacher, T.W., Anand, R., and Williams, A.F. (1988). Complete structure of the glycosyl phosphatidylinositol membrane anchor of rat brain Thy-1 glycoprotein. Nature *333*, 269–272.

Hopke, A., Brown, A.J.P., Hall, R.A., and Wheeler, R.T. (2018). Dynamic Fungal Cell Wall Architecture in Stress Adaptation and Immune Evasion. Trends Microbiol. *26*, 284–295.

Hu, X., Du, Y., Tang, Y., Wang, Q., Feng, T., Yang, J., and Kennedy, J.F. (2007). Solubility and property of chitin in NaOH/urea aqueous solution. Carbohydr. Polym. *70*, 451–458.

Hurtado-Guerrero, R., Schüttelkopf, A.W., Mouyna, I., Ibrahim, A.F.M., Shepherd, S., Fontaine, T., Latgé, J.P., and van Aalten, D.M.F. (2009). Molecular mechanisms of yeast cell wall glucan remodeling. J. Biol. Chem. *284*, 8461–8469.

Hutchins, K., and Bussey, H. (1983). Cell wall receptor for yeast killer toxin: involvement of (1-6)-beta-D-glucan. J. Bacteriol. *154*, 161–169.

Huynh, K., and Partch, C.L. (2015). Analysis of protein stability and ligand interactions by thermal shift assay. Curr. Protoc. Protein Sci. *79*, 28.9.1-28.9.14.

Ikezawa, H. (2002). Glycosylphosphatidylinositol (GPI)-Anchored Proteins. Biol. Pharm. Bull. *25*, 409–417.

Ikezawa, H., Yamanegi, M., Taguchi, R., Miyashita, T., and Ohyabu, T. (1976). Studies on phosphatidylinositol phosphodiesterase (phospholipase C type) of *Bacillus cereus*. I. Purification, properties and phosphatase-releasing activity. Biochim. Biophys. Acta *450*, 154–164.

Inoue, S.B., Takewaki, N., Takasuka, T., Mio, T., Adachi, M., Fujii, Y., Miyamoto, C., Arisawa, M., Furuichi, Y., and Watanabe, T. (1995). Characterization and gene cloning of 1,3-beta-D-glucan synthase from *Saccharomyces cerevisiae*. Eur. J. Biochem. *231*, 845–854.

Inoue, S.B., Qadota, H., Arisawa, M., Watanabe, T., and Ohya, Y. (1999). Prenylation of Rho1p is required for activation of yeast 1,3-β-glucan synthase. J. Biol. Chem. *274*, 38119–

38124.

Ishihara, S., Hirata, A., Nogami, S., Beauvais, A., Latge, J.-P., and Ohya, Y. (2007). Homologous Subunits of 1,3-Beta-Glucan Synthase Are Important for Spore Wall Assembly in *Saccharomyces cerevisiae*. Eukaryot. Cell *6*, 143–156.

Kabsch, W. (2010). XDS. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 125–132.

Kang, M., Elango, N., Mattia, E., Au, J., Robbins, P., and Cabib, E. (1984). Isolation of Chitin Synthetase from *Saccharomyces cerevisiae*. J. Biol. Chem. *259*, 14966–14972.

Kang, X., Kirui, A., Muszyński, A., Widanage, M.C.D., Chen, A., Azadi, P., Wang, P., Mentink-Vigier, F., and Wang, T. (2018). Molecular architecture of fungal cell walls revealed by solid-state NMR. Nat. Commun. *9*, 2747.

Kapteyn, J.C., Ram, A.F.J., Groos, E.M., Kollar, R., Montijn, R.C., Van Den Ende, H., Llobell, A., Cabib, E., and Klis, F.M. (1997). Altered extent of cross-linking of β1,6-glucosylated mannoproteins to chitin in *Saccharomyces cerevisiae* mutants with reduced cell wall β1,3-glucan content. J. Bacteriol. *179*, 6279–6284.

Kapteyn, J.C., Van Den Ende, H., and Klis, F.M. (1999). The contribution of cell wall proteins to the organization of the yeast cell wall. Biochim. Biophys. Acta - Gen. Subj. *1426*, 373–383.

Kar, B., Patel, P., Ao, J., and Free, S.J. (2019). *Neurospora crassa* family GH72 glucanosyltransferases function to crosslink cell wall glycoprotein N-linked galactomannan to cell wall lichenin. Fungal Genet. Biol. *123*, 60–69.

Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. Science *336*, 1030–1033.

Karplus, P.A., and Diederichs, K. (2015). Assessing and maximizing data quality in macromolecular crystallography. Curr. Opin. Struct. Biol. *34*, 60–68.

Kinoshita, T., and Fujita, M. (2016). Biosynthesis of GPI-anchored proteins: special emphasis on GPI lipid remodeling. J. Lipid Res. *57*, 6–24.

Kitagaki, H., Wu, H., Shimoi, H., and Ito, K. (2002). Two homologous genes, DCW1 (YKL046c) and DFG5, are essential for cell growth and encode glycosylphosphatidylinositol (GPI)-anchored membrane proteins required for cell wall biogenesis in *Saccharomyces cerevisiae*. Mol. Microbiol. *46*, 1011–1022.

Kitagaki, H., Ito, K., and Shimoi, H. (2004). A temperature-sensitive dcw1 mutant of

Saccharomyces cerevisiae is cell cycle arrested with small buds which have aberrant cell walls. Eukaryot. Cell *3*, 1297–1306.

Klis, F.M., Boorsma, A., and De Groot, P.W.J. (2006). Cell wall construction in *Saccharomyces cerevisiae*. Yeast *23*, 185–202.

Klis, F.M., Brul, S., and De Groot, P.W.J. (2010). Covalently linked wall proteins in ascomycetous fungi. Yeast *27*, 489–493.

Kollár, R., Petrakova, E., Ashwell, G., Robbins, P.W., and Cabib, E. (1995). Architecture of the yeast cell wall. The linkage between chitin and β(1→3)-glucan. J. Biol. Chem. *270*, 1170–1178.

Kollár, R., Reinhold, B.B., Petráková, E., Yeh, H.J.C., Ashwell, G., Drgonová, J., Kapteyn, J.C., Klis, F.M., and Cabib, E. (1997). Architecture of the yeast cell wall: β(1→6)glucan interconnects mannoprotein, β(1→3)-glucan, and chitin. J. Biol. Chem. *272*, 17762–17775.

Kopecka, M., Phaff, H.J., and Fleet, G.H. (1974). Demonstration of a fibrillar component in the cell wall of the yeast *Saccharomyces cerevisiae* and its chemical nature. J. Cell Biol. *62*, 66–76.

Kraushaar, T., Brückner, S., Veelders, M., Rhinow, D., Schreiner, F., Birke, R., Pagenstecher, A., Mösch, H.-U., and Essen, L.-O. (2015). Interactions by the Fungal Flo11 Adhesin Depend on a Fibronectin Type III-like Adhesin Domain Girdled by Aromatic Bands. Structure *23*, 1005–1017.

Kreger, D.R., and Kopecka, M. (1976). On the Nature and Formation of the Fibrillar Nets Produced by Protoplasts of *Saccharomyces cerevisiae* in Liquid Media: An Electronmicroscopic, X-ray Diffraction and Chemical Study. J. Gen. Microbiol. *92*, 207–220.

Kushner, S. (1988). An improved method for transformation of *Escherichia coli* with colEI derived plasmid. In Genetic engineering. Genet. Eng. (N. Y). 17–24.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Lambou, K., Perkhofer, S., Fontaine, T., and Latge, J.P. (2010). Comparative functional analysis of the OCH1 mannosyltransferase families in *Aspergillus fumigatus* and *Saccharomyces cerevisiae*. Yeast *27*, 625–636.

Lampugnani, E.R., Khan, G.A., Somssich, M., and Persson, S. (2018). Building a plant

cell wall at a glance. J. Cell Sci. *131*, 1–6.

Lehto, M.T., and Sharom, F.J. (2002). Proximity of the protein moiety of a GPI-anchored protein to the membrane surface: A FRET study. Biochemistry *41*, 8368–8376.

Leidich, S.D., Kostova, Z., Latek, R.R., Costello, L.C., Drapp, D.A., Gray, W., Fassler, J.S., and Orlean, P. (1995). Temperature-sensitive Yeast GPI Anchoring Mutants gpi2 and gpi3 Are Defective in the Synthesis of N-Acetylglucosaminyl Phosphatidylinositol. J. Biol. Chem. *270*, 13029–13035.

Lenardon, M.D., Munro, C.A., and Gow, N.A. (2010). Chitin synthesis and fungal pathogenesis. Curr. Opin. Microbiol. *13*, 416–423.

Lesage, G., and Bussey, H. (2006). Cell Wall Assembly in *Saccharomyces cerevisiae*. Microbiol. Mol. Biol. Rev. *70*, 317–343.

Li, H., Zhou, H., Luo, Y., Ouyang, H., Hu, H., and Jin, C. (2007). Glycosylphosphatidylinositol (GPI) anchor is required in *Aspergillus fumigatus* for morphogenesis and virulence. Mol. Microbiol. *64*, 1014–1027.

Li, J., Mouyna, I., Henry, C., Moyrand, F., Malosse, C., Chamot-Rooke, J., Janbon, G., Latgé, J.-P., and Fontaine, T. (2018). Glycosylphosphatidylinositol Anchors from Galactomannan and GPI-Anchored Protein Are Synthesized by Distinct Pathways in *Aspergillus fumigatus*. J. Fungi *4*, 1–14.

Lippincott, J., and Li, R. (1998). Sequential assembly of myosin II, an IQGAP-like protein, and filamentous actin to a ring structure involved in budding yeast cytokinesis. J. Cell Biol. *140*, 355–366.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. *42*, D490–D495.

Low, M.G., and Finean, J.B. (2015). Release of alkaline phosphatase from membranes by a phosphatidylinositol-specific phospholipase C. Biochem. J. *167*, 281–284.

Lu, C.F., Montijn, R.C., Brown, J.L., Klis, F., Kurjan, J., Bussey, H., and Lipke, P.N. (1995). Glycosyl phosphatidylinositol-dependent cross-linking of α-agglutinin and β1,6-glucan in the *Saccharomyces cerevisiae* cell wall. J. Cell Biol. *128*, 333–340.

Maddi, A., and Free, S.J. (2010). α-1,6-mannosylation of N-linked oligosaccharide present on cell wall proteins is required for their incorporation into the cell wall in the filamentous

fungus *Neurospora crassa*. Eukaryot. Cell *9*, 1766–1775.

Maddi, A., Fu, C., and Free, S.J. (2012). The *Neurospora crassa* dfg5 and dcw1 genes encode α-1,6-mannanases that function in the incorporation of glycoproteins into the cell wall. PLoS One *7*, e38872.

Mandel, M., and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. J. Mol. Biol. *53*, 159–162.

Manners, D.J., Masson, A.J., and Patterson, J.C. (1973). The structure of a β-(1→3)-d-glucan from yeast cell walls. Biochem. J. *135*, 19–30.

Manzano-Lopez, J., Perez-Linero, A.M., Aguilera-Romero, A., Martin, M.E., Okano, T., Silva, D.V., Seeberger, P.H., Riezman, H., Funato, K., Goder, V., et al. (2015). COPII coat composition is actively regulated by luminal cargo maturation. Curr. Biol. *25*, 152–162.

Martin-Yken, H., Dagkessamanskaia, A., De Groot, P., Ram, A., Klis, F., and François, J. (2001). *Saccharomyces cerevisiae* YCRO17c/CWH43 encodes a putative sensor/transporter protein upstream of the BCK2 branch of the PKC1-dependent cell wall integrity pathway. Yeast *18*, 827–840.

Mayor, S., and Riezman, H. (2004). Sorting GPI-anchored proteins. Nat. Rev. Mol. Cell Biol. *5*, 110–120.

Mazáň, M., Ragni, E., Popolo, L., and Farkaš, V. (2011). Catalytic properties of the Gas family β-(1,3)-glucanosyltransferases active in fungal cell-wall biogenesis as determined by a novel fluorescent assay. Biochem. J. *438*, 275–282.

Mazur, P., and Baginsky, W. (1996). In vitro activity of 1,3-β-D-glucan synthase requires the GTP-binding protein Rho1. J. Biol. Chem. *271*, 14604–14609.

Mazur, P., Morin, N., Baginsky, W., El-Sherbeini, M., Clemas, J.A., Nielsen, J.B., and Foor, F. (1995). Differential Expression and Function of Two Homologous Subunits of Yeast 1,3-β-D-Glucan Synthase. Mol. Cell. Biol. *15*, 5671–5681.

McCarter, J.D., and Withers, S.G. (1994). Mechanisms of enzymatic glycoside hydrolysis. Curr. Opin. Struct. Biol. *4*, 885–892.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Meaden, P., Hill, K., Wagner, J., Slipetz, D., Sommer, S.S., and Bussey, H. (2015). The yeast KRE5 gene encodes a probable endoplasmic reticulum protein required for (1-6)-β-D-

glucan synthesis and normal cell growth. Mol. Cell. Biol. *10*, 3013–3019.

Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.Y., El-Gebali, S., Fraser, M.I., et al. (2019). InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res. *47*, D351–D360.

Mol, P.C., Park, H.M., Mullins, J.T., and Cabib, E. (1994). A GTP-binding protein regulates the activity of (1→3)-β-glucan synthase, an enzyme directly involved in yeast cell wall morphogenesis. J. Biol. Chem. *269*, 31267–31274.

Molano, J., Bowers, B., and Cabib, E. (1980). Distribution of chitin in the yeast cell wall: An ultrastructural and chemical study. J. Cell Biol. *85*, 199–212.

Money, N.P. (2001). Biomechanics of Invasive Hyphal Growth. In Biology of the Fungal Cell, pp. 237–249.

Montijn, R.C., Vink, E., Müller, W.H., Verkleij, A.J., Van Den Ende, H., Henrissat, B., and Klis, F.M. (1999). Localization of synthesis of β1,6-glucan in *Saccharomyces cerevisiae*. J. Bacteriol. *181*, 7414–7420.

Moo, W.K., Eun, J.K., Kim, J.Y., Park, J.S., Oh, D.B., Shimma, Y.I., Chiba, Y., Jigami, Y., Sang, K.R., and Hyun, A.K. (2006). Functional characterization of the *Hansenula polymorpha* HOC1, OCH1, and OCR1 genes as members of the yeast OCH1 mannosyltransferase family involved in protein glycosylation. J. Biol. Chem. *281*, 6261–6272.

Morozov, A.A., and Likhoshway, Y. V (2016). Evolutionary history of the chitin synthases of eukaryotes. Glycobiology *26*, 635–639.

Mösch, H.-U., and Fink, G.R. (1997). Dissection of filamentous growth by transposon mutagenesis in *Sacchromyces cerevisiae*. Genetics *145*, 671–684.

Mouyna, I., Henry, C., Doering, T.L., and Latgé, J.P. (2004). Gene silencing with RNA interference in the human pathogenic fungus *Aspergillus fumigatus*. FEMS Microbiol. Lett. *237*, 317–324.

Mukasa, R., Umeda, M., Endo, T., Kobata, A., and Inoue, K. (1995). Characterization of glycosylphosphatidylinositol (GPI)-anchored NCAM on mouse skeletal muscle cell line C2C12: The structure of the GPI glycan and release during myogenesis. Arch. Biochem. Biophys. *318*, 182–190.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. *51*, 263–273.

Muszkieta, L., Fontaine, T., Beau, R., Mouyna, I., Vogt, M.S., Trow, J., Cormack, B.P., Essen, L.-O., Jouvion, G., and Latgé, J.-P. (2019). The Glycosylphosphatidylinositol-Anchored DFG Family Is Essential for the Insertion of Galactomannan into the β-(1,3)-Glucan–Chitin Core of the Cell Wall of *Aspergillus fumigatus*. MSphere *4*, e00397-19.

Nakanishi-Shindo, Y., Nakayama, K.I., Tanaka, A., Toda, Y., and Jigami, Y. (1993). Structure of the N-linked oligosaccharides that show the complete loss of α-1,6-polymannose outer chain from och1, och1 mnn1, and och1 mnn1 alg3 mutants of *Saccharomyces cerevisiae*. J. Biol. Chem. *268*, 26338–26345.

Nakano, Y., Noda, K., Endo, T., Kobata, A., and Tomita, M. (1994). Structural Study on the Glycosyl-Phosphatidylinositol Anchor and the Asparagine-Linked Sugar Chain of a Soluble Form of CD59 in Human Urine. Arch. Biochem. Biophys. *311*, 117–126.

Naumoff, D.G. (2011). Hierarchical classification of glycoside hydrolases. Biochem. *76*, 622–635.

Nicholls, R.A. (2017). Ligand fitting with CCP4. Acta Crystallogr. Sect. D, Struct. Biol. *73*, 158–170.

Nurizzo, D., Mairs, T., Guijarro, M., Rey, V., Meyer, J., Fajardo, P., Chavanne, J., Biasci, J.C., McSweeney, S., and Mitchell, E. (2006). The ID23-1 structural biology beamline at the ESRF. J. Synchrotron Radiat. *13*, 227–238.

Orlean, P. (1987). Two Chitin Synthases in *Saccharomyces cerevisiae*. J. Biol. Chem. *262*, 5732–5739.

Orlean, P. (2009). Phosphoethanolamine addition to glycosylphosphatidylinositols. Enzym. *26*, 117–132.

Orlean, P. (2012). Architecture and biosynthesis of the *Saccharomyces cerevisiae* cell wall. Genetics *192*, 775–818.

Orlean, P., and Menon, A.K. (2007). Thematic review series: Lipid Posttranslational Modifications. GPI anchoring of protein in yeast and mammalian cells, or: how we learned to stop worrying and love glycophospholipids. J. Lipid Res. *48*, 993–1011.

Orlean, P., Ammer, H., Watzele, M., and Tanner, W. (1986). Synthesis of an O-

glycosylated cell surface protein induced in yeast by factor. Proc. Natl. Acad. Sci. *83*, 6263–6266.

Osumi, M. (1998). The ultrastructure of yeast: cell wall structure and formation. Micron *29*, 207–233.

Ouyang, H., Chen, X., Lü, Y., Wilson, I.B.H., Tang, G., Wang, A., and Jin, C. (2013). One single basic amino acid at the ω-1 or ω-2 site is a signal that Retains glycosylphosphatidylinositol-anchored protein in the plasma membrane of *Aspergillus fumigatus*. Eukaryot. Cell *12*, 889–899.

Oxley, D., and Bacic, A. (1999). Structure of the glycosylphosphatidylinositol anchor of an arabinogalactan protein from *Pyrus communis* suspension-cultured cells. Proc. Natl. Acad. Sci. *96*, 14246–14251.

Patterson, A.L. (1934). A fourier series method for the determination of the components of interatomic distances in crystals. Phys. Rev. *46*, 372–376.

Paulick, M.G., and Bertozzi, C.R. (2008). The glycosylphosphatidylinositol anchor: A complex membrane-anchoring structure for proteins. Biochemistry *47*, 6991–7000.

Perfect, J.R. (2017). The antifungal pipeline: A reality check. Nat. Rev. Drug Discov. *16*, 603–616.

Perlin, D.S. (2015). Echinocandin Resistance in Candida. Clin. Infect. Dis. *61*, S612–S617.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Pierleoni, A., Martelli, P., and Casadio, R. (2008). PredGPI: A GPI-anchor predictor. BMC Bioinformatics *9*, 1–11.

Pittet, M., and Conzelmann, A. (2007). Biosynthesis and function of GPI proteins in the yeast *Saccharomyces cerevisiae*. Biochim. Biophys. Acta - Mol. Cell Biol. Lipids *1771*, 405–420.

Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R., Jenkins, H.T., Keegan, R., Krissinel, E., Stevenson, K., et al. (2018). CCP4i2: The new graphical user interface to the CCP4 program suite. Acta Crystallogr. Sect. D Struct. Biol. *74*, 68–84.

Qin, Z., Yan, Q., Lei, J., Yang, S., Jiang, Z., and Wu, S. (2015). The first crystal structure of a glycoside hydrolase family 17 β-1,3-glucanosyltransferase displays a unique catalytic

cleft. Acta Crystallogr. Sect. D Biol. Crystallogr. *71*, 1714–1724.

Raich, L., Nin-Hill, A., Ardèvol, A., and Rovira, C. (2016). Enzymatic Cleavage of Glycosidic Bonds: Strategies on How to Set Up and Control a QM/MM Metadynamics Simulation. In Methods in Enzymology, pp. 159–183.

Ram, A.F.J., and Klis, F.M. (2006). Identification of fungal cell wall mutants using susceptibility assays based on Calcofluor white and Congo red. Nat. Protoc. *1*, 2253–2256.

Ram, A.F.J., Wolters, A., Hoopen, R. Ten, and Klis, F.M. (1994). A new approach for isolating cell wall mutants in *Saccharomyces cerevisiae* by screening for hypersensitivity to calcofluor white. Yeast *10*, 1019–1030.

Reinhard, L., Mayerhofer, H., Geerlof, A., Mueller-Dieckmann, J., and Weiss, M.S. (2013). Optimization of protein buffer cocktails using Thermofluor. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *69*, 209–214.

Reulecke, I., Lange, G., Albrecht, J., Klein, R., and Rarey, M. (2008). Towards an integrated description of hydrogen bonding and dehydration: Decreasing false positives in virtual screening with the HYDE scoring function. ChemMedChem *3*, 885–897.

Rhodes, G. (2006). Crystallography Made Crystal Clear.

Rodicio, R., and Heinisch, J.J. (2010). Together we are strong - Cell wall integrity sensors in yeasts. Yeast *27*, 531–540.

Roh, D.-H., Bowers, B., Riezman, H., and Cabib, E. (2002). Rho1p mutations specific for regulation of beta(1-3)glucan synthesis and the order of assembly of the yeast cell wall. Mol. Microbiol. *44*, 1167–1183.

Rose, J.P., Wang, B.-C., and Weiss, M.S. (2015). Native SAD is maturing. IUCrJ *2*, 431–440.

Rost, B. (1999). Twilight zone of protein sequence alignments. Protein Eng. *12*, 85–94.

Rupp, B. (2010). Biomolecular Crystallography - Principles, Practice, and Application to Structural Biology.

De Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B., et al. (2012). ID29: A high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. J. Synchrotron Radiat. *19*, 455–461.

Schärfer, C., Schulz-Gasch, T., Ehrlich, H.C., Guba, W., Rarey, M., and Stahl, M.

(2013). Torsion angle preferences in druglike chemical space: A comprehensive guide. J. Med. Chem. *56*, 2016–2028.

Schekman, R. (2010). Editorial Expression of Concern: Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast Pichiapastoris. Proc. Natl. Acad. Sci. *107*, 7113–7113.

Schimmöller, F., Singer-Krüger, B., Schröder, S., Krüger, U., Barlowe, C., and Riezman, H. (1995). The absence of Emp24p, a component of ER-derived COPII-coated vesicles, causes a defect in transport of selected proteins to the Golgi. EMBO J. *14*, 1329–1339.

Schmidt, M., Bowers, B., Varma, A., Roh, D.-H., and Cabib, E. (2002). In budding yeast, contraction of the actomyosin ring and formation of the primary septum at cytokinesis depend on each other. J. Cell Sci. *115*, 293–302.

Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. Acta Crystallogr. Sect. D Biol. Crystallogr. *58*, 1772–1779.

Schneider, N., Lange, G., Hindle, S., Klein, R., and Rarey, M. (2013). A consistent description of HYdrogen bond and DEhydration energies in protein-ligand complexes: Methods behind the HYDE scoring function. J. Comput. Aided. Mol. Des. *27*, 15–29.

Schofield, L., Hewitt, M.C., Evans, K., Siomos, M.-A., and Seeberger, P.H. (2002). Synthetic GPI as a candidate anti-toxic vaccine in a model of malaria. Nature *418*, 785–789.

Seelig, J. (1990). Interaction of phospholipids with Ca2+ions. On the role of the phospholipid head groups. Cell Biol. Int. Rep. *14*, 353–360.

Sesti, F., Shih, T.M., Nikolaeva, N., and Goldstein, S.A.N. (2001). Immunity to K1 killer toxin: Internal TOK1 blockade. Cell *105*, 637–644.

Shahinian, S., and Bussey, H. (2000). β-1,6-glucan synthesis in *Saccharomyces cerevisiae*. Mol. Microbiol. *35*, 477–489.

Shaw, J.A., Mol, P.C., Bowers, B., Silverman, S.J., Valdivieso, M.H., Duran, A., and Cabib, E. (1991). The function of chitin synthases 2 and 3 in the *Saccharomyces cerevisiae* cell cycle. J. Cell Biol. *114*, 111–123.

Sheldrick, G.M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 479–485.

Shematek, E.M., Braatz, J.A., and Cabib, E. (1980). Biosynthesis of the yeast cell wall. I.

Preparation and properties of beta-(1 leads to 3)glucan synthetase. J. Biol. Chem. *255*, 888–894.

Shuker, S.B., Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. Science *274*, 1531–1534.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Sinn, C.G., Antonietti, M., and Dimova, R. (2006). Binding of calcium to phosphatidylcholine-phosphatidylserine membranes. Colloids Surfaces A Physicochem. Eng. Asp. *282–283*, 410–419.

Skubák, P., and Pannu, N.S. (2013). Automatic protein structure solution from weak X-ray data. Nat. Commun. *4*, 1–6.

Sonnhammer, E.L., von Heijne, G., and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. Proceedings. Int. Conf. Intell. Syst. Mol. Biol. 1–8.

Spreghini, E., Davis, D.A., Subaran, R., Kim, M., and Mitchell, A.P. (2003). Roles of *Candida albicans* Dfg5p and Dcw1p cell surface proteins in growth and hypha formation. Eukaryot. Cell *2*, 746–755.

Stoddart, J.F. (1971). Stereochemistry of Carbohydrates (Wiley-Interscience).

Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. Protein Expr. Purif. *41*, 207–234.

Surma, M.A., Klose, C., and Simons, K. (2012). Lipid-dependent protein sorting at the trans-Golgi network. Biochim. Biophys. Acta - Mol. Cell Biol. Lipids *1821*, 1059–1067.

Taguchi, R., Hamakawa, N., Harada-Nishida, M., Fukui, T., Nojima, K., and Ikezawa, H. (1994). Microheterogeneity in Glycosylphosphatidylinositol Anchor Structures of Bovine Liver 5′-Nucleotidase. Biochemistry *33*, 1017–1022.

Tanaka, S., Maeda, Y., Tashima, Y., and Kinoshita, T. (2004). Inositol Deacylation of Glycosylphosphatidylinositol-anchored Proteins Is Mediated by Mammalian PGAP1 and Yeast Bst1p. J. Biol. Chem. *279*, 14256–14263.

Taylor, G.L. (2010). Introduction to phasing. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 325–338.

Teparić, R., and Mrša, V. (2013). Proteins involved in building, maintaining and remodeling of yeast cell walls. Curr. Genet. *59*, 171–185.

Thompson, A.J., Cuskin, F., Spears, R.J., Dabin, J., Turkenburg, J.P., Gilbert, H.J., and Davies, G.J. (2015a). Structure of the GH76 α-mannanase homolog, BT2949, from the gut symbiont *Bacteroides thetaiotaomicron*. Acta Crystallogr. Sect. D Biol. Crystallogr. *71*, 408–415.

Thompson, A.J., Speciale, G., Iglesias-Fernández, J., Hakki, Z., Belz, T., Cartmell, A., Spears, R.J., Chandler, E., Temple, M.J., Stepper, J., et al. (2015b). Evidence for a boat conformation at the transition state of GH76 α-1,6-mannanases - Key enzymes in bacterial and fungal mannoprotein metabolism. Angew. Chemie - Int. Ed. *54*, 5378–5382.

Tian, W., and Skolnick, J. (2003). How well is enzyme function conserved as a function of pairwise sequence identity? J. Mol. Biol. *333*, 863–882.

Tsukahara, K., Hata, K., Nakamoto, K., Sagane, K., Watanabe, N.-A.A., Kuromitsu, J., Kai, J., Tsuchiya, M., Ohba, F., Jigami, Y., et al. (2003). Medicinal genetics approach towards identifying the molecular target of a novel inhibitor of fungal cell wall assembly. Mol. Microbiol. *48*, 1029–1042.

Uccelletti, D., Farina, F., Rufini, S., Magnelli, P., Abeijon, C., and Palleschi, C. (2006). The *Kluyveromyces lactis* α1,6-mannosyltransferase KlOch1p is required for cell-wall organization and proper functioning of the secretory pathway. FEMS Yeast Res. *6*, 449–457.

Umemura, M., Okamoto, M., Nakayama, K. ichi, Sagane, K., Tsukahara, K., Hata, K., and Jigami, Y. (2003). GWT1 gene is required for inositol acylation of glycosylphosphatidylinositol anchors in yeast. J. Biol. Chem. *278*, 23639–23647.

Umemura, M., Fujita, M., Yoko-o, T., Fukamizu, A., and Jigami, Y. (2007). *Saccharomyces cerevisiae* CWH43 Is Involved in the Remodeling of the Lipid Moiety of GPI Anchors to Ceramides. Mol. Biol. Cell *18*, 4304–4316.

Valdivieso, M.H., Mol, P.C., Shaw, J.A., Cabib, E., and Duran, A. (1991). CAL1, a gene required for activity of chitin synthase 3 in *Saccharomyces cerevisiae*. J. Cell Biol. *114*, 101–109.

Vazquez, H.M., Vionnet, C., Roubaty, C., and Conzelmann, A. (2014). Cdc1 removes the ethanolamine phosphate of the first mannose of GPI anchors and thereby facilitates the integration of GPI proteins into the yeast cell wall. Mol. Biol. Cell *25*, 3375–3388.

Veelders, M., and Essen, L.O. (2012). Complex Gadolinium-Oxo Clusters Formed along Concave Protein Surfaces. ChemBioChem *13*, 2187–2190.

Vernier, P.T., Ziegler, M.J., and Dimova, R. (2009). Calcium binding and head group dipole angle in phosphatidylserine- phosphatidylcholine bilayers. Langmuir *25*, 1020–1027.

Vink, E., Rodriguez-Suarez, R.J., Gérard-Vincent, M., Ribas, J.C., de Nobel, H., van den Ende, H., Durán, A., Klis, F.M., and Bussey, H. (2004). An in vitro assay for (1→6)-β-D-glucan synthesis in *Saccharomyces cerevisiae*. Yeast *21*, 1121–1131.

Vliegenthart, J.F.G. (2011). Introduction to NMR Spectroscopy of Carbohydrates. In NMR Spectroscopy and Computer Modeling of Carbohydrates, pp. 1–19.

Vollmer, W., Blanot, D., and De Pedro, M.A. (2008). Peptidoglycan structure and architecture. FEMS Microbiol. Rev. *32*, 149–167.

Wang, Q., Withers, S.G., Graham, R.W., Trimbur, D., and Warren, R.A.J. (1994). Changing Enzymic Reaction Mechanisms by Mutagenesis: Conversion of a Retaining Glucosidase to an Inverting Enzyme. J. Am. Chem. Soc. *116*, 11594–11595.

Wehle, M., Vilotijevic, I., Lipowsky, R., Seeberger, P.H., Varon Silva, D., and Santer, M. (2012). Mechanical compressibility of the glycosylphosphatidylinositol (GPI) anchor backbone governed by independent glycosidic linkages. J. Am. Chem. Soc. *134*, 18964–18972.

Weiss, M.S., and Hilgenfeld, R. (1997). On the use of the merging R factor as a quality indicator for X-ray data. J. Appl. Crystallogr. *30*, 203–205.

Yamaguchi, M., Namiki, Y., Okada, H., Mori, Y., Furukawa, H., Wang, J., Ohkusu, M., and Kawamoto, S. (2011). Structome of *Saccharomyces cerevisiae* determined by freeze-substitution and serial ultrathin-sectioning electron microscopy. J. Electron Microsc. (Tokyo). *60*, 321–335.

Yin, Q.Y., de Groot, P.W.J., Dekker, H.L., de Jong, L., Klis, F.M., and de Koster, C.G. (2005). Comprehensive proteomic analysis of *Saccharomyces cerevisiae* cell walls. J. Biol. Chem. *280*, 20894–20901.

Yin, Q.Y., De Groot, P.W.J., De Jong, L., Klis, F.M., and De Koster, C.G. (2007). Mass spectrometric quantitation of covalently bound cell wall proteins in *Saccharomyces cerevisiae*. FEMS Yeast Res. *7*, 887–896.

Yoko-o, T., Tsukahara, K., Watanabe, T., Hata-Sugi, N., Yoshimatsu, K., Nagasu, T.,

and Jigami, Y. (2001). *Schizosaccharomyces pombe* och1(+) encodes α-1,6-mannosyltransferase that is involved in outer chain elongation of N-linked oligosaccharides. FEBS Lett. *489*, 75–80.

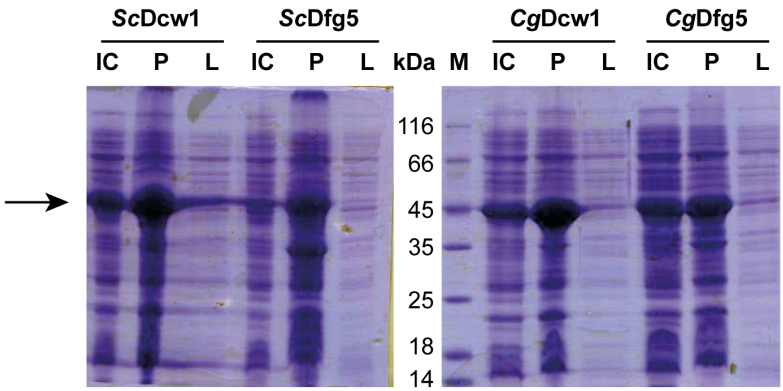
Yoko-o, T., Umemura, M., Komatsuzaki, A., Ikeda, K., Ichikawa, D., Takase, K., Kanzawa, N., Saito, K., Kinoshita, T., Taguchi, R., et al. (2018). Lipid moiety of glycosylphosphatidylinositol-anchored proteins contributes to the determination of their final destination in yeast. Genes to Cells *23*, 880–892.

Zhu, Y., Vionnet, C., and Conzelmann, A. (2006). Ethanolaminephosphate side chain added to glycosylphosphatidylinositol (GPI) anchor by Mcd4p is required for ceramide remodeling and forward transport of GPI proteins from endoplasmic reticulum to Golgi. J. Biol. Chem. *281*, 19830–19839.

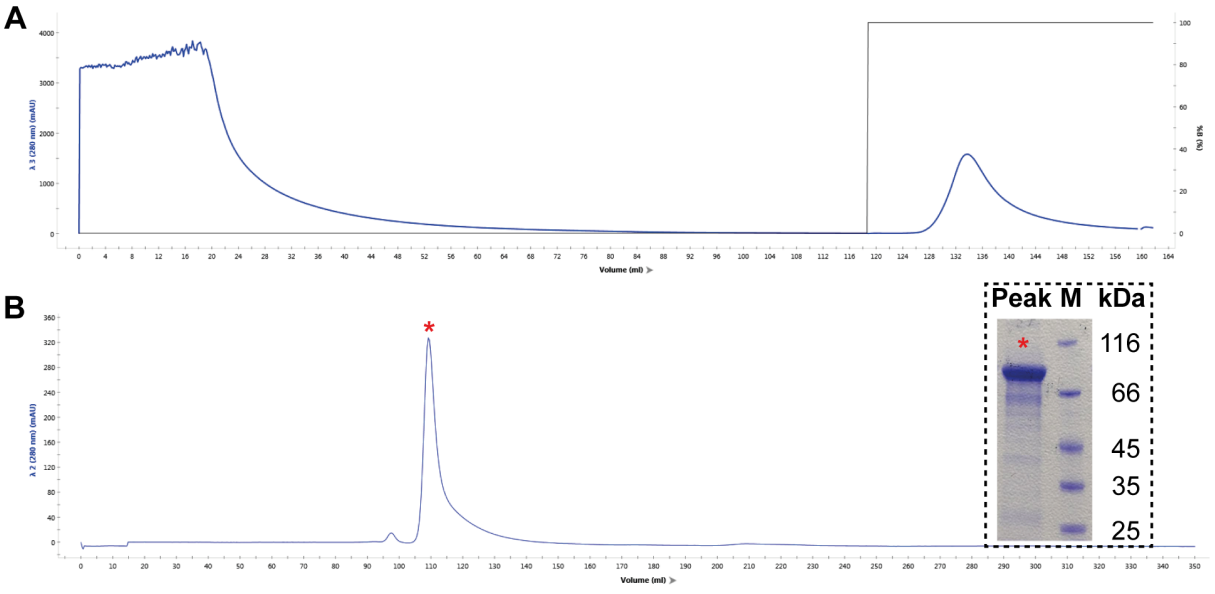
Zlotnik, H., Pilar Fernandez, M., Bowers, B., and Cabib, E. (1984). *Saccharomyces cerevisiae* mannoproteins form an external cell wall layer that determines wall porosity. J. Bacteriol. *159*, 1018–1026.

7 Supplement

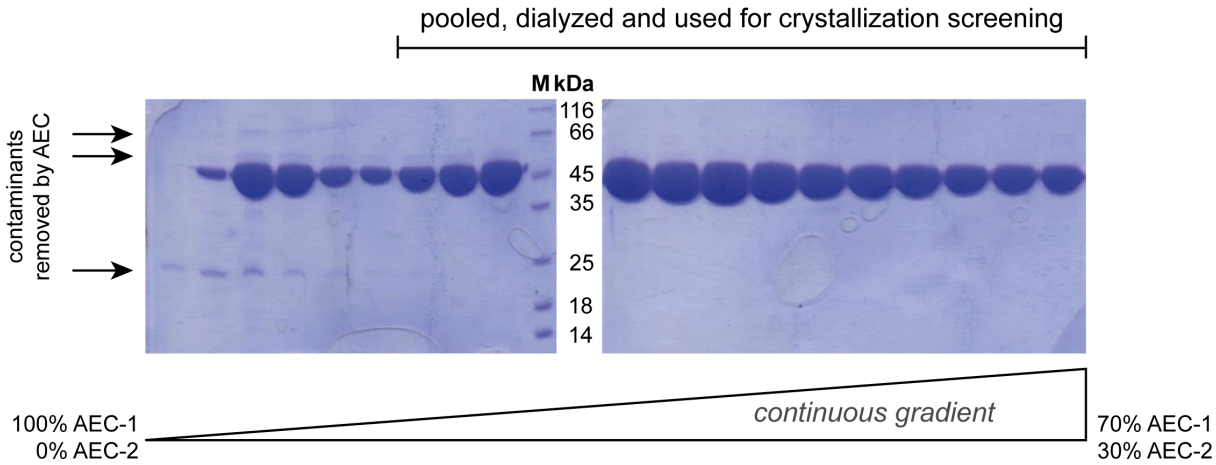
7.1 Figures



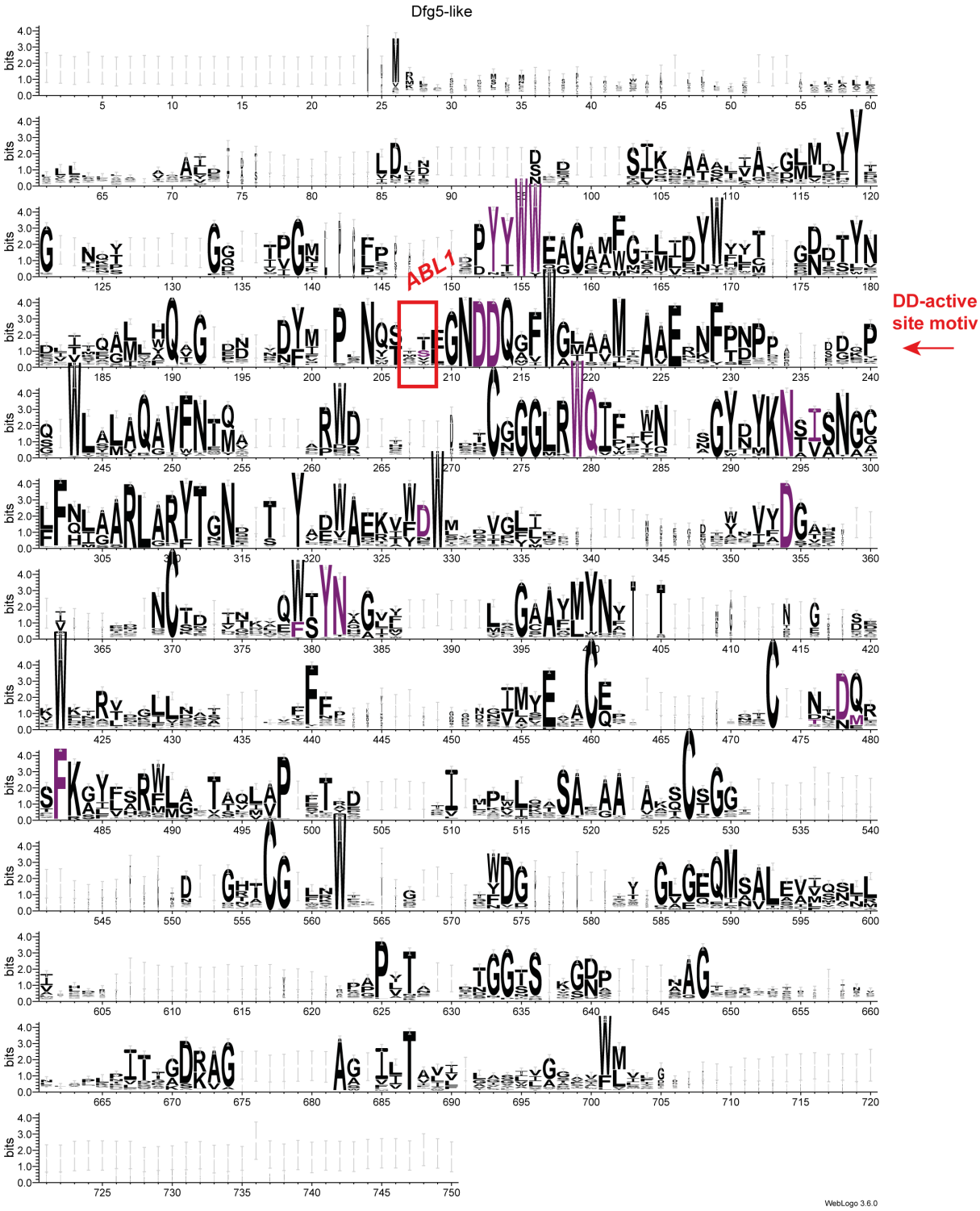
SFigure 1: Test expressions of Dfg5-homologs from mesophilic fungi. Heterologous production of Dfg5-orthologs from *Saccharomyces cerevisiae* (*Sc*) and *Candida glabrata* (*Cg*) were overexpressed, but only found insoluble. The arrow indicates the apparent size of over produced proteins, *IC=induced cells*, *P=pellet*, *L=Load* (for affinity chromatography, i.e. the supernatant of the lysis upon centrifugation), *M=Molecular weight marker*. The marker is labeled with the respective molecular weights in kDa.



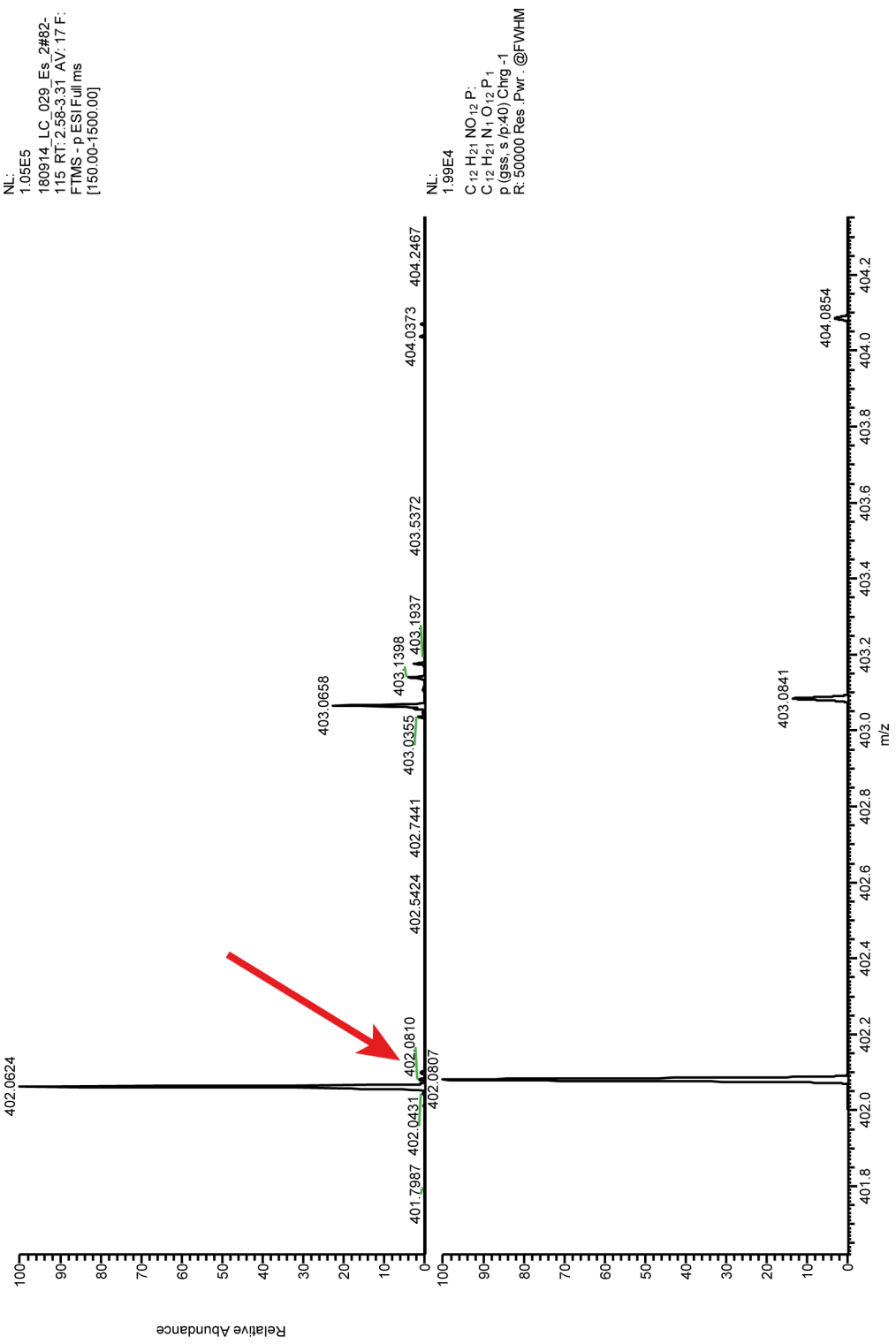
SFigure 2: Purification of MBP-*Sc*Dcw1. A The amylose affinity purification with a single step gradient is shown. $A_{280\ nm}$ is shown as a blue line and the percentage of MBP-2 buffer („B“) is shown as a black line. **B** The SEC run of MBP-*Sc*Dcw1 is shown. The peak fraction of the 87.3 kDa fusion protein is loaded on SDS-PAGE as indicated by the red asterisk. *M=molecular weight marker*



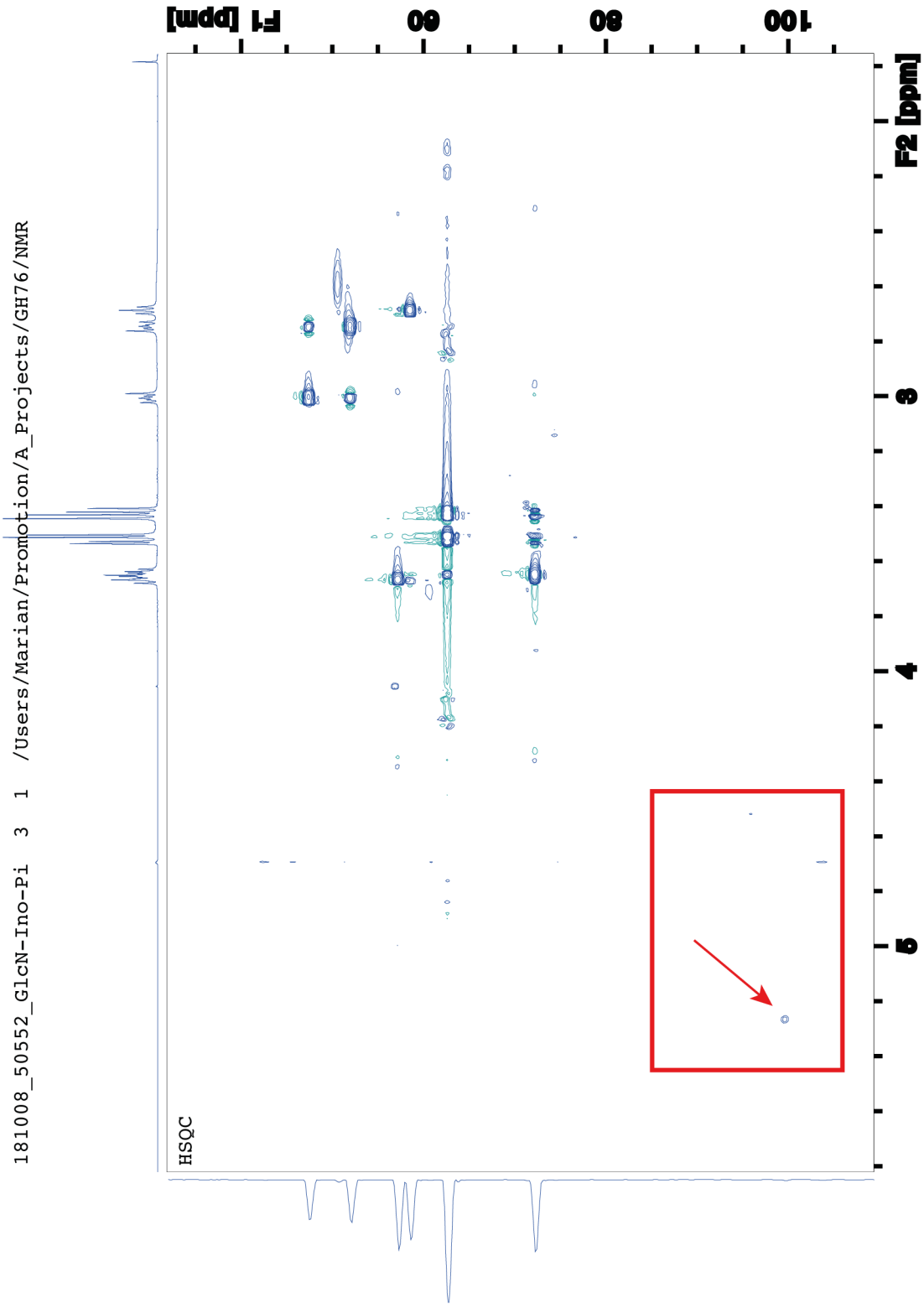
SFigure 3: Anion exchange chromatography of *Ct*Dfg5 for first crystal screen. *M=Molecular weight marker*.



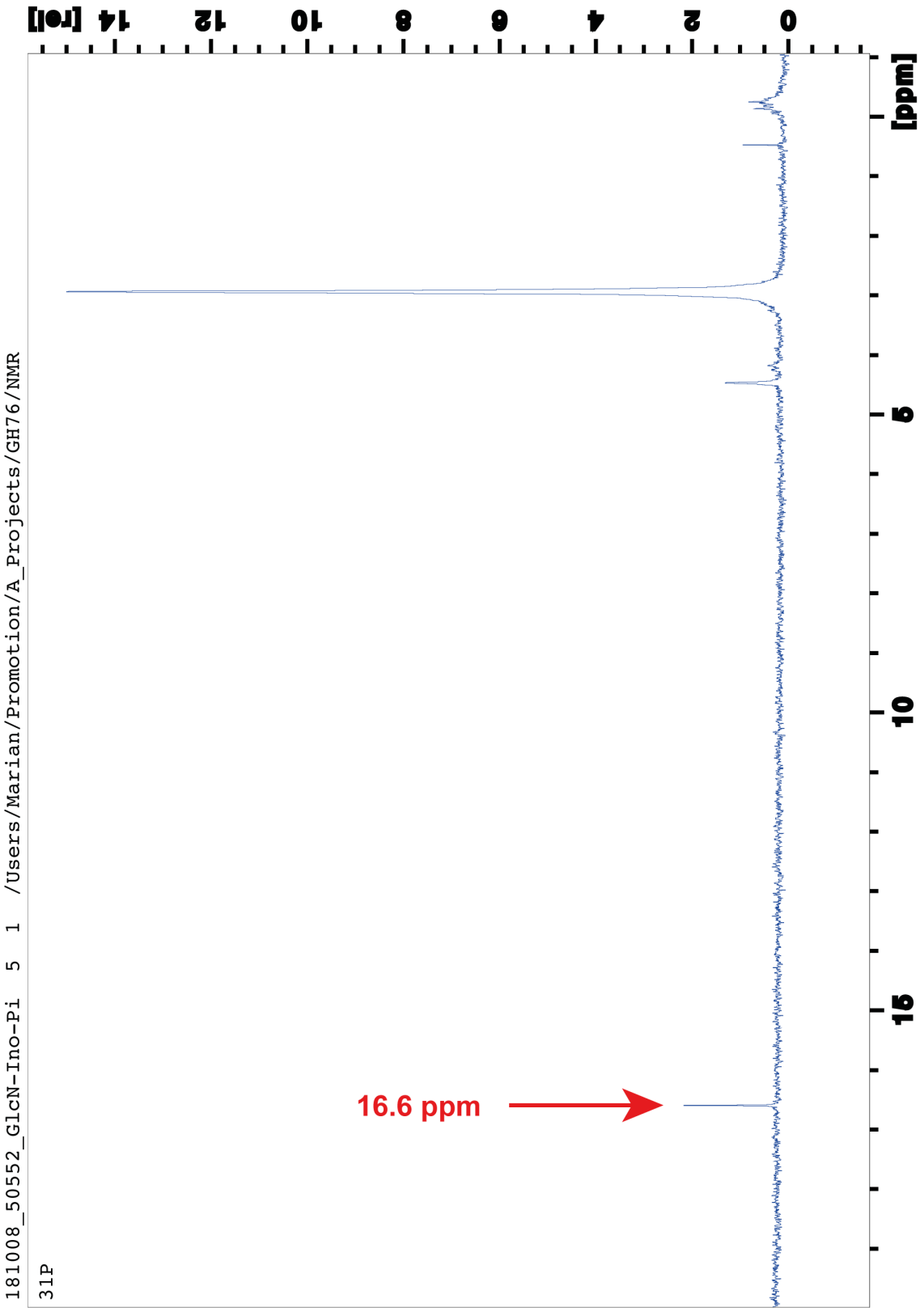
SFigure 4: WEBLOGO of the *Dfg5*-subfamily calculated from the MSA that was used for the *ConSurf* analysis. The conserved residues are highlighted, as well as the ABL1 region. The full MSA comprising 250 *Dfg5*-sequences is only part of the electronic resource due to its length.



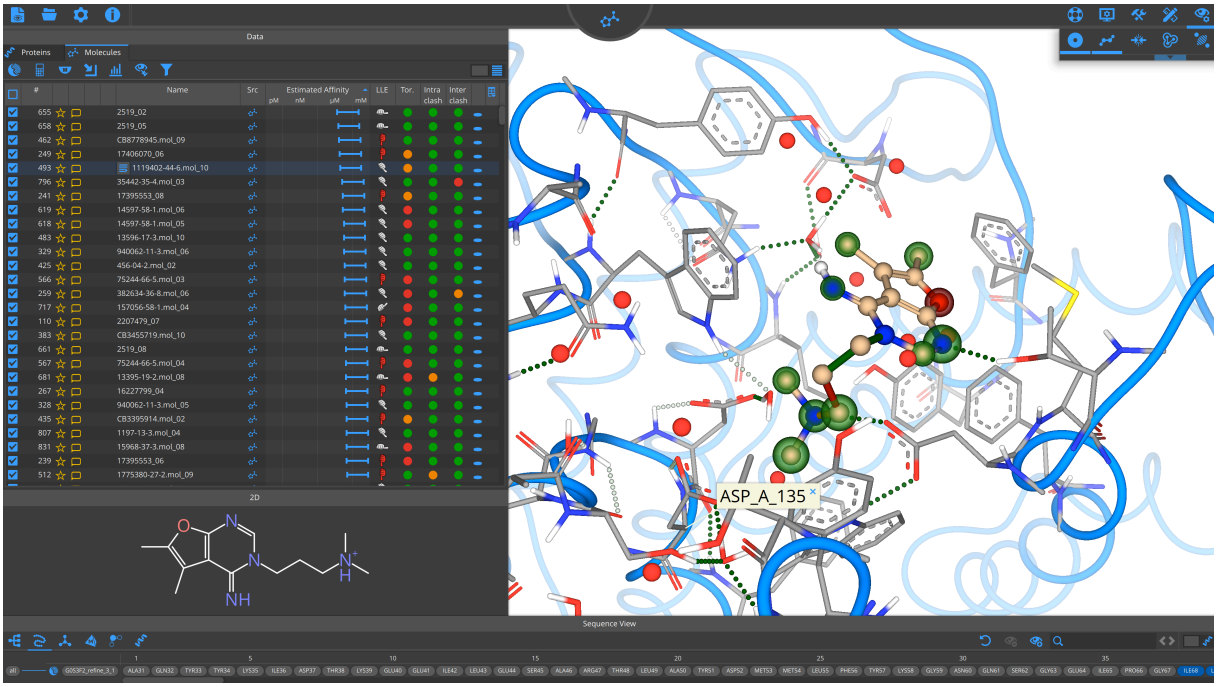
SFigure 5: Identification of target mass calculated for GlcN-Ino-Pi (402.08 Da). MS analysis was performed by the MS-facility led by Dr. Uwe Linne.



SFigure 6: Identification of anomeric carbon by 2D-1H-13C NMR. The region marked in red indicates the typical chemical shift region for anomeric carbons in carbohydrates. A singlet peak corresponding to anomeric carbons was observed for the analyzed sample. NMR analysis was performed by Dr. Xiulan Xie.

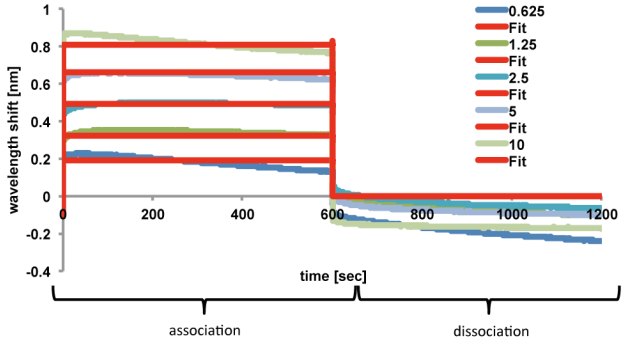


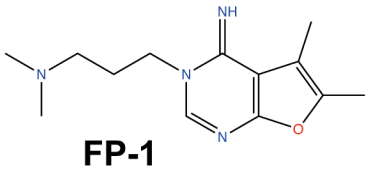
SFigure 7: Identification of specific chemical shift of cyclic phosphate by 31P NMR around 16 ppm in form of a singlet. NMR analysis was performed by Dr. Xiulan Xie.



SFigure 8: Docked pose of FP-1 generated by *SeeSAR*. The conformation from molecular docking is inversed to the real conformation found in the crystal structure.

FP-1 + *Ct*Dfg5

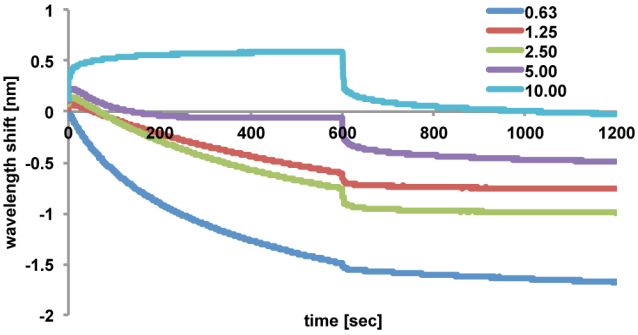


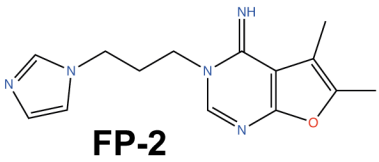


FP-1

Loading Sample	Loading Conc. (µg/ml)	c(F51) (mM)	KD (M)	KD Error	ka (1/Ms)	ka Error	kdis (1/s)	kdis Error	Full R^2
Dfg5	49.5	0.625	2.71E-03	9.56E-04	1.08E+04	2.70E+03	2.94E+01	7.32E+00	0.9331
Dfg5	49.5	1.25	2.71E-03	9.56E-04	1.08E+04	2.70E+03	2.94E+01	7.32E+00	0.9331
Dfg5	49.5	2.5	2.71E-03	9.56E-04	1.08E+04	2.70E+03	2.94E+01	7.32E+00	0.9331
Dfg5	49.5	5	2.71E-03	9.56E-04	1.08E+04	2.70E+03	2.94E+01	7.32E+00	0.9331
Dfg5	49.5	10	2.71E-03	9.56E-04	1.08E+04	2.70E+03	2.94E+01	7.32E+00	0.9331

FP-2 + *Ct*Dfg5





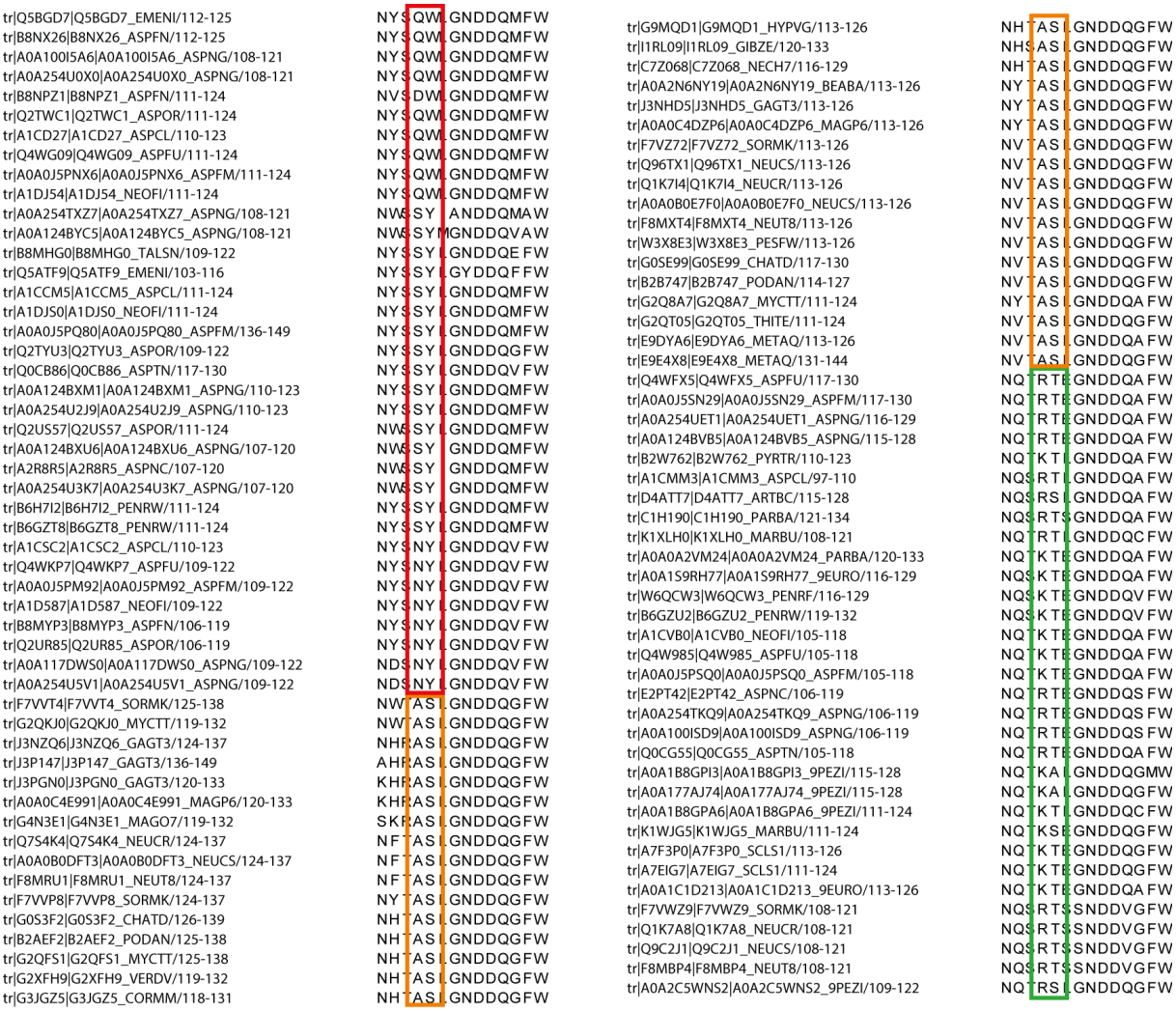
FP-2

Loading Sample	Loading Conc. (µg/ml)	c(F51a) (mM)	KD (M)	KD Error	ka (1/Ms)	ka Error	kdis (1/s)	kdis Error	Full R^2
Dfg5	49.5	0.625	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Dfg5	49.5	1.25	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Dfg5	49.5	2.5	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Dfg5	49.5	5	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Dfg5	49.5	10	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

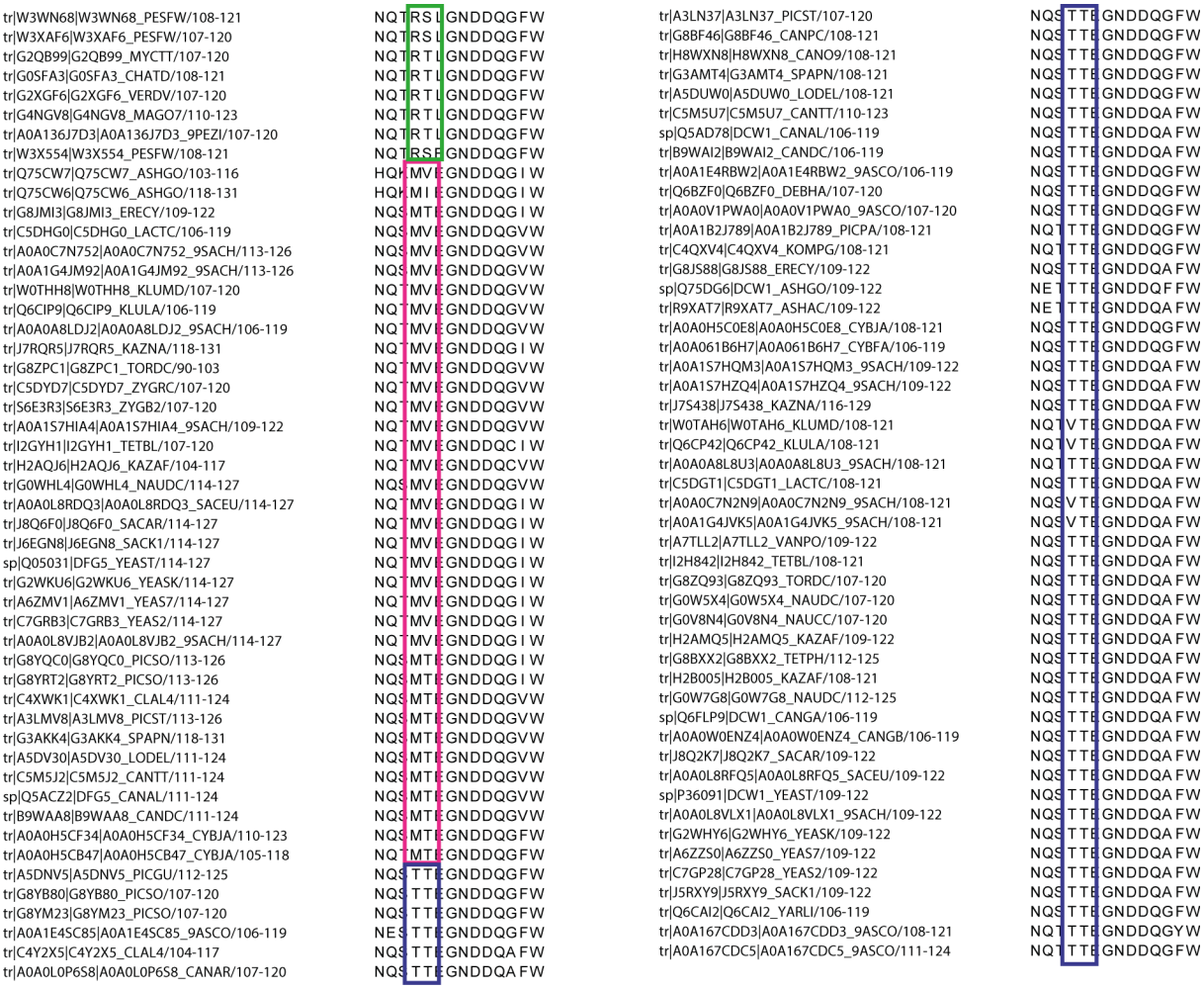
Summary Biolayer Inferometry

		cFrag (mM)	KD (M)	KD Error	ka (1/Ms)	ka Error	kdis (1/s)	kdis Error	Full R^2
Dfg5	FP-1	0.625	2.71E-03	9.56E-04	1.08E+04	2.70E+03	2.94E+01	7.32E+00	0.9331
	FP-2	10	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

SFigure 9: Binding analyses of FP-1 and FP-2 with *Ct*Dfg5 using biolayer inferometry. The analysis was performed together with Dr. Sven-Andreas Freibert, who evaluated the data.



SFigure 10: Multiple sequence alignment of ABL2-region



SFigure 11: Multiple sequence alignment of ABL2-region generated by Clustal Omega using 200 sequences.

7.2 Tables

Table S 1: Crystallographic data collection statistics of *Ct*Dfg5 crystals soaked with 500 mM inositol for 4 h. Values in parentheses are for the highest resolution shell.

	***Ct*Dfg5$_{Ino}$**
Data collection	
X-ray source	ID 23-2, ESRF
Space group	C 1 2 1
Unit-cell parameters (Å, °)	*a*=83.41, *b*=55.07, *c*=80.34, *α*=90.00, *β*=90.31, *γ*=90.00
Wavelength (Å)	0.87
Resolution range (Å)	45.96-1.2 (1.22-1.2)
Completeness (%)	98.2 (84.1)
Observed reflections	479751 (14185)
Unique reflections	111409 (4656)
Multiplicity	4.3
Wilson *B* factor (Å^2)	10.01
R_{merge} (%)	7.6 (110.5)
Mean *I/σ(I)*	8.1 (1.0)
CC(½)	98.2 (84.1)

Table S 2: Identified GPI-APs upon PI-PLC treatment in the supernatant. The analytics were performed by the MS-facility led by Dr. Uwe Linne.

Uniprot-ID	Description	Coverage [%]	# Peptides	# PSMs	# Unique Peptides
P22146	1,3-beta-glucanosyltransferase GAS1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=GAS1 PE=1 SV=2	14	16	105	16
Q03655	Probable 1,3-beta-glucanosyltransferase GAS3 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=GAS3 PE=1 SV=1	29	13	53	13
Q08193	1,3-beta-glucanosyltransferase GAS5 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=GAS5 PE=1 SV=1	14	5	26	5
P38248	Cell wall protein ECM33 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=ECM33 PE=1 SV=3	39	25	145	25
P46992	Cell wall protein YJL171C OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YJL171C PE=1 SV=2	28	7	11	7

Table S 3: List of abbreviations used in this thesis.

ABL	Acceptor binding loop
Af	*Aspergillus fumigatus*
An	*Aspergillus niger*
Bc	*Bacillus circulans*
C-terminus	Carboxyl-terminus
Ca	*Candida albicans*
Ct	*Chaetomium thermophilum*
CV	Column volume
CWP	Cell wall protein
Dcw1	Defective cell wall
DNA	Deoxyribonucleic acid
Dfg5	Defective in filamentous growth
e.g.	For example
et al.	*Latin: Et alii/aliae* – „and others“

EtN-P	Ethanolamine phosphate
ERES	Endoplasmatic reticulum exit site
FCW	Fungal cell wall
FP-1	(3-{4-imino-5,6-dimethyl-3H,4H-furo[2,3-d]pyrimidin-3-yl}propyl)dimethylamine
FP-2	3-[3-(1H-imidazol-1-yl)propyl]-5,6-dimethyl-3H,4H-furo[2,3-d]pyrimidin-4-imine
GH	Glycosyl hydrolase
Glc	Glucose
GlcN	Glucosamine
GlcNAc	N-acetylglucosamine
GPI-AP	Glycosylphosphatidylinositol-anchored protein
GPI-CWP	Glycosylphosphatidylinositol-cell wall protein
GPI-PMP	Glycosylphosphatidylinositol-plasma membrane protein
GT	Glycosyl transferase
GTP	Guanosine triphosphate
i.e.	*Latin: id est* – „that is to say“
IMAC	Immobilized metal ion affinity chromatography
Ino	Inositol
IPC	Inositolphosphorylceramide
JCSG	Joint Center for Structural Genomics
Man	Mannose
MBP	Maltose-binding protein
MR	Molecular replacement
MSA	Multiple sequence alignment
n	Integer
N-terminus	Amino-terminus
Nc	*Neurospora crassa*
Ni-NTA	Nickel-nitrilotriacetic acid
p.	Page
PDB	Protein data bank
PI	Phosphatidylinositol
PIR-CWP	Protein with internal repeats-cell wall protein
PM	Plasma membrane
SAD	Single-wavelength anomalous diffraction
SAR	Structure activity relationship
Sc	*Saccharomyces cerevisiae*
SEC	Size exclusion chromatography
SSN	Sequence similarity network
TM(H)	Transmembrane helix
UDP	Uridine diphosphate

Table S 4: List of amino acids and their one-letter codes.

Amino acid	One-letter symbol
Alanine	A
Arginine	R
Asparagine	N
Aspartic acid	D
Cysteine	C
Glutamic acid	E
Glutamine	Q
Glycine	G

Histidine	H
Isoleucine	I
Leucine	L
Lysine	K
Methionine	M
Phenylalanine	F
Proline	P
Serine	S
Threonine	T
Tryptophan	W
Tyrosine	Y
Valine	V

8 Acknowledgement – Danksagung

An erster Stelle gilt mein ganz besonderer Dank Herrn Prof. Dr. Lars-Oliver Essen, der mir die Möglichkeit gab, an diesem interessanten Projekt zu arbeiten und diese Arbeit anzufertigen. Neben seiner herausragenden Unterstützung auf wissenschaftlicher Ebene möchte ich mich vor allem auch für die Möglichkeit bedanken, mich auch in anderen Themengebieten austoben zu können.

Herrn Prof. Dr. Hans-Ulrich Mösch danke ich für die Übernahme des Zweitgutachtens meiner Arbeit und die erfolgreiche Dfg5-Kooperation, zusammen mit Gesa Schmitz, die unsere zahlreichen *in vivo* Ideen in die Realität umgesetzt hat.

Für die Übernahme des Vorsitzes der Prüfungskommission danke ich Herrn Prof. Dr. Norbert Hampp.

Mein Dank gilt natürlich auch den vielen Helfern, ohne die eine solche Arbeit niemals funktionieren kann: Herrn Prof. Dr. Andreas Heine danke ich für die Hilfe beim Messen des Gadolinium Datensatzes. Dr. Patrick Pausch danke ich für das letzte Aliquot *Chaetomium* cDNA, die Investition hat sich gelohnt! Dr. Ankan Banerjee, I thank you for the different projects we performed together, it was always a pleasure! Dr. Daniel Varón Silva and Dr. Thierry Fontaine, thank you for the material support and the exchange of ideas. Ralf Pöschke, Dr. Uwe Linne, Dr. Thomas Heimerl und Dr. Xiulan Xie danke ich stellvertretend für die Mitarbeiter in unseren Servicelabors. An dieser Stelle gilt mein Dank auch Dr. Sven-Andreas Freibert für die unkomplizierten Messungen am Octet.

Laura Werel und Christian Mayer danke ich für die Durchsicht und Korrektur meiner Arbeit.

Meinen Bachelor- und Masterstudenten Thomas Klaus, Lea-Sophie Hubert, Simon Huber, Simon Völpel, Kai Krämer, Markus Friedrich und Roi Rostand Ngouoko danke ich für ihr blindes Vertrauen und ihre Unterstützung im Labor.

Meiner Arbeitsgruppe danke ich für die tolle Zeit in den vergangenen Jahren. Insbesondere Petra Gnau, Thomas Marcellino (for being my favorite baguette and english buddy), Viktoria Reithofer, Laura Werel, Maximilian Biermeier, Elisabeth Ignatz, und Filipp Bezolt.

Pietro, thanks for the countless coffees we enjoyed together and Wieland, dir danke ich für die zahlreichen, natürlich rein wissenschaftlichen Gespräche, die wir in all den Jahren hatten.

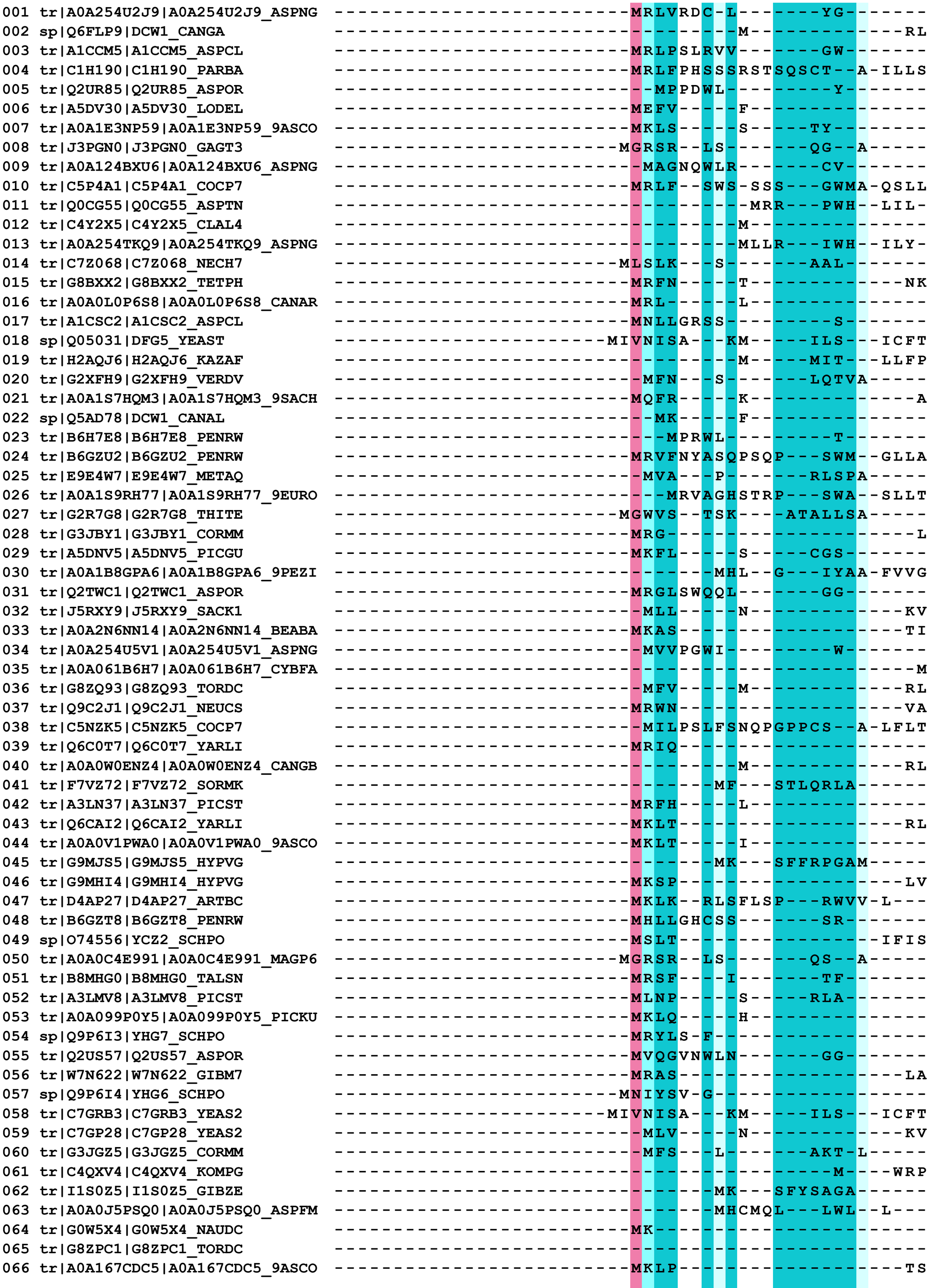
Meiner Familie und meinen Freunden danke ich dafür, dass sie immer da sind und ich mich auf ihre Unterstützung verlassen kann.

Karlotta, dir hier zu danken wird der Sache zwar sowieso nicht gerecht. Dennoch gilt dir hier mein größter Dank. Mein Dank für deine Liebe. Mein Dank mit mir dein Leben zu teilen. Mein Dank für deine Unterstützung. Mein Dank für deine Stärke als Mutter. Mein Dank für dein richtiges Gefühl. Mein Dank für die zwei Menschlein, die einem zeigen, worauf es tatsächlich ankommt. Mein Dank für alles, was schon war. Mein Dank für alles, was noch kommt. Ich kann es kaum erwarten!

Supplement of the electronic resource

Complete multiple sequence alignment used for *Consurf* analysis of *Ct*Dfg5

ConSurf Color-Coded MSA



```
067 tr|R9XAT7|R9XAT7_ASHAC ------------------------------------------------------------M----IAV
068 tr|B2B747|B2B747_PODAN ------------------------------------------------------MI---SPIRHTA------
069 tr|G0SE99|G0SE99_CHATD ----------------------------------------------------ML--P--LSL---SSPRWLA------
070 tr|J8Q6F0|J8Q6F0_SACAR ----------------------------------------------------MIVDNSM--KM-----IPL---ICFT
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH ------------------------------------------------------------MK
072 tr|J3K6E2|J3K6E2_COCIM ------------------------------------------------------------MA-QSLL
073 tr|G4N3E1|G4N3E1_MAGO7 ----------------------------------------------------MALHK--LA-------RC--L------
074 tr|C5M5J2|C5M5J2_CANTT ------------------------------------------------------MRFF----N---------------
075 tr|A0A254U0X0|A0A254U0X0_ASPNG ----------------------------------------------------MRSHHI-------TA-------
076 sp|Q75DG6|DCW1_ASHGO ------------------------------------------------------------M----LAV
077 tr|Q6BZF0|Q6BZF0_DEBHA ------------------------------------------------------MKLT-----I---------------
078 tr|E9E3Q1|E9E3Q1_METAQ ------------------------------------------------------MFS---S-------LVSPA------
079 sp|P36091|DCW1_YEAST ------------------------------------------------------MLV-----N-------------KV
080 tr|W0TAH6|W0TAH6_KLUMD ------------------------------------------------------------M-------------KC
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA ------------------------------------------------------MRFE-----V---------------
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI ------------------------------------------------------MRFS-------------VA
083 tr|Q2TYU3|Q2TYU3_ASPOR ------------------------------------------------------MRMPRGG-L-------LG-------
084 tr|G2XGF6|G2XGF6_VERDV ------------------------------------------------------MRLL-------------ST
085 tr|G3AMT4|G3AMT4_SPAPN ------------------------------------------------------MRLY-----Q---------------
086 tr|A0A1C1D213|A0A1C1D213_9EURO ----------------------------------------------------MRPFRR---RPLL-HLTL
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS ------------------------------------------------------------------
088 tr|A0A100I5A6|A0A100I5A6_ASPNG ----------------------------------------------------MRSHHL-------TA-------
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH ----------------------------------------------------MRV-------SF------AEF---WALT
090 tr|G2Q8A7|G2Q8A7_MYCTT ------------------------------------------------------MV---WT--RFA------
091 tr|B8NPZ1|B8NPZ1_ASPFN ------------------------------------------------------MRCLSFCQV-------GL-------
092 tr|Q7SAB2|Q7SAB2_NEUCR ------------------------------------------------------------------------
093 tr|A2R8R5|A2R8R5_ASPNC ------------------------------------------------------MAGNQWLR-------CV-------
094 tr|C5DHG0|C5DHG0_LACTC ------------------------------------------------------------M----LAQS
095 tr|G0SFA3|G0SFA3_CHATD ------------------------------------------------------MRLF-------------EG
096 tr|K1WJG5|K1WJG5_MARBU ------------------------------------------------------MKGT---RILA-AGVA
097 tr|H2B005|H2B005_KAZAF ------------------------------------------------------MVI-----N-------------NL
098 tr|C5DYD7|C5DYD7_ZYGRC ------------------------------------------------------------M----LLHN
099 tr|G8YM23|G8YM23_PICSO ------------------------------------------------------MKIF-----P---------------
100 tr|W3X8E3|W3X8E3_PESFW ------------------------------------------------------MA---SFVSRAA------
101 tr|A0A254U3K7|A0A254U3K7_ASPNG ------------------------------------------------------MAGNQWLR-------CV-------
102 tr|Q6CER8|Q6CER8_YARLI ----------------------------------------------------MRPST-----Y-------------IW
103 tr|G2QT05|G2QT05_THITE ------------------------------------------------------------MV---SA--ALS------
104 sp|Q5ACZ2|DFG5_CANAL ------------------------------------------------------MVSL-----Q---------------
105 tr|G2WKU6|G2WKU6_YEASK ----------------------------------------------------MIVNISA--KM-----ILS---ICFT
106 tr|A0A0C7N752|A0A0C7N752_9SACH ----------------------------------------------------MGI-------EI------ALG---WPSQ
107 tr|W3WN68|W3WN68_PESFW ------------------------------------------------------MLGV-------------------LR
108 tr|G2QFS1|G2QFS1_MYCTT ----------------------------------------------------MALSP--PSN---SPARWLLS------
109 tr|G3J9G4|G3J9G4_CORMM ------------------------------------------------------MKAS-------------------AI
110 tr|Q2UJ03|Q2UJ03_ASPOR ------------------------------------------------------------MRT---PWC--AVL-
111 tr|G8YRT2|G8YRT2_PICSO ------------------------------------------------------MRFI-----T------RYD-----C--
112 tr|W3X554|W3X554_PESFW ------------------------------------------------------MRSI-------------------LR
113 tr|Q1K7A8|Q1K7A8_NEUCR ------------------------------------------------------MRWN-------------------VA
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM ------------------------------------------------------MKLSS-RT-------------W-------
115 tr|A7F3P0|A7F3P0_SCLS1 ------------------------------------------------------------MRVSGG---KALA-YTVA
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA ------------------------------------------------------MRA-------------------L
117 tr|Q6CIP9|Q6CIP9_KLULA ------------------------------------------------------------------------MK
118 tr|H2AMQ5|H2AMQ5_KAZAF ------------------------------------------------------MI-------P-------------R-
119 tr|G8JMI3|G8JMI3_ERECY ------------------------------------------------------------------------MRII
120 tr|C5DGT1|C5DGT1_LACTC ------------------------------------------------------------M-------------RL
121 tr|K1XJR4|K1XJR4_MARBU ----------------------------------------------------MSL--P--TLV---SG---AA------
122 tr|A0A124BYC5|A0A124BYC5_ASPNG ----------------------------------------------------MVFAHYI-------AA-------
123 tr|A7TLL2|A7TLL2_VANPO ------------------------------------------------------MLFH-----K-------------V
124 tr|Q0CB86|Q0CB86_ASPTN ------------------------------------------------------MRPSTPR-L-------SS-------
125 tr|K1XLH0|K1XLH0_MARBU ------------------------------------------------------------M---KWSS-VIYT
126 tr|J7S438|J7S438_KAZNA ------------------------------------------------------MQFS-----S-----CKM-----GTS
127 tr|Q4WG09|Q4WG09_ASPFU ------------------------------------------------------MRMLSLRQL-------CG-------
128 tr|H8WZ09|H8WZ09_CANO9 ------------------------------------------------------MFGL-----N---------------
129 tr|I1RL09|I1RL09_GIBZE ----------------------------------------------------ML-FS---S-------RGFIA------
130 tr|Q4W985|Q4W985_ASPFU ------------------------------------------------------------MHCMQL---LWL--L---
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU ----------------------------------------------------MIVKASM--KT-----VLP---LCLT
132 tr|W6QCW3|W6QCW3_PENRF ------------------------------------------------------MRVFNYAS---QP---SWM--GLLA
133 tr|W3WMD3|W3WMD3_PESFW ------------------------------------------------------------MV---RLSS----------
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH ------------------------------------------------------MQFR-----K-------------A
135 tr|I2GYH1|I2GYH1_TETBL ----------------------------------------------------MI-------------LNE---LILS
```

```
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH ----------------------------------------MIM---LLMN
137 tr|Q752P3|Q752P3_ASHGO         ----------------------------------------------MT
138 tr|B8MYP3|B8MYP3_ASPFN         ------------------------------MPPDWL----------Y-------
139 tr|Q2TXL6|Q2TXL6_ASPOR         ------------------------------MWSYF----------W-------
140 tr|Q0CN17|Q0CN17_ASPTN         ------------------------------MQPNRL----------W-------
141 tr|J3NHD5|J3NHD5_GAGT3         ----------------------------------MF---APLVRVG------
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI ----------------------------------MKTG---SFFA--LAT
143 tr|A1CVB0|A1CVB0_NEOFI         ----------------------------------MHCMQL---LWL--L---
144 tr|B8NVI3|B8NVI3_ASPFN         ------------------------------MWSYF----------W-------
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH ----------------------------------M----------------RL
146 tr|A1DJ54|A1DJ54_NEOFI         ------------------------------MRMLSLRQL----------CG-------
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS ------------------------------MRTT--SSP---RGATWLTA------
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA ----------------------------------------------------MKP
149 tr|A0A177A618|A0A177A618_9PEZI ------------------------------MRL-----------------WSV-AR--
150 tr|Q75CW6|Q75CW6_ASHGO         ----------------------------MTVAFVRVGAE----------------PS
151 tr|Q7S4K4|Q7S4K4_NEUCR         ------------------------------MRTT--SSP---RGATWLTA------
152 tr|G2QB99|G2QB99_MYCTT         ------------------------------MR-I-----------------SG
153 tr|G0S3F2|G0S3F2_CHATD         ----------------------------MTTWS--SSK---RPLRWLVA------
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO ----------------------------MK----L-----------------
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB ----------------------------------------MK---LLFS
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM MFTHPAPQPSNSDALLNRLISENIAMRLQSLPRL----------GW-------
157 tr|G2QKJ0|G2QKJ0_MYCTT         ----------------------------MAP----TTTVLLG------
158 tr|Q4WFX5|Q4WFX5_ASPFU         ----------------------------------MGRQCPCHWPL-SLRL
159 tr|Q5BGD7|Q5BGD7_EMENI         ------------------------------MRWGSLLSL----------GG-------
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH ----------------------------------M----------------KV
161 tr|G8BYN9|G8BYN9_TETPH         ----------------------------------M------FYL---TLL-
162 tr|F7VVT4|F7VVT4_SORMK         ----------------------------MLTTKIAG------
163 tr|G8YB80|G8YB80_PICSO         ------------------------------MKTL----P-----------------
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA ----------------------------------------------------
165 tr|G2WRS7|G2WRS7_VERDV         ----------------------------MKGSFVARA------
166 tr|A0A124BVB5|A0A124BVB5_ASPNG ----------------------------------MHIP----PL-RLWR
167 tr|A7EIG7|A7EIG7_SCLS1         ----------------------------MKLSTS---SSAL---AT
168 tr|C5M5U7|C5M5U7_CANTT         ----------------------------MKFSCS----F-----------------
169 tr|B9WAA8|B9WAA8_CANDC         ------------------------------MISL----Q-----------------
170 tr|F8MRU1|F8MRU1_NEUT8         ------------------------------MRTT--SSP---RGAKWLTA------
171 tr|A1DJS0|A1DJS0_NEOFI         ------------------------------MRLQSLPRL----------GW-------
172 tr|B6HLM8|B6HLM8_PENRW         ------------------------------MRFKGAGH----------M-------
173 tr|A0A117DWS0|A0A117DWS0_ASPNG ------------------------------MVVPGWI----------W-------
174 tr|A0A124BXM1|A0A124BXM1_ASPNG ------------------------------MRLVRDC-L----------YG-------
175 tr|J3NZQ6|J3NZQ6_GAGT3         ----------------------------------MARRWGATALLL-
176 tr|W0THH8|W0THH8_KLUMD         ----------------------------------------------MR
177 tr|B2AEF2|B2AEF2_PODAN         ----------------------------MATTP--APK---SPSRWLTA------
178 tr|Q96TX1|Q96TX1_NEUCS         ----------------------------MKL---SPLPRLA------
179 tr|E9DYA6|E9DYA6_METAQ         ----------------------------MK----------LSSS------
180 tr|B9WAI2|B9WAI2_CANDC         ------------------------------MK----F-----------------
181 tr|Q6C171|Q6C171_YARLI         ------------------------------MRFG----L----------------IP
182 tr|A6ZMV1|A6ZMV1_YEAS7         ----------------------------MIVNISA--KM-----ILS---ICFT
183 tr|B6K7X7|B6K7X7_SCHJY         ----------------------------------------------------MK
184 tr|I1RSA2|I1RSA2_GIBZE         ------------------------------MRSL----------------LQ
185 tr|F7VWZ9|F7VWZ9_SORMK         ------------------------------MRWN----------------VA
186 tr|A0A1B2J789|A0A1B2J789_PICPA ----------------------------------------M----WRP
187 tr|G0V8N4|G0V8N4_NAUCC         ----------------------------------M----------------N-
188 tr|H8WXN8|H8WXN8_CANO9         ------------------------------MKLF----L-----------------
189 tr|A6ZZS0|A6ZZS0_YEAS7         ----------------------------MLV----N----------------KV
190 tr|G9MQD1|G9MQD1_HYPVG         ----------------------------------------MLP------
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH ----------------------------------M----------------KI
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM ------------------------------MRMLSLRQL----------CG-------
193 tr|A0A254UET1|A0A254UET1_ASPNG ----------------------------------MSRIP----QF-RLWR
194 tr|A1D587|A1D587_NEOFI         ------------------------------MKLSS-RT----------W-------
195 tr|Q1K7I4|Q1K7I4_NEUCR         ----------------------------MKL---SPLPRLA------
196 tr|G8JS88|G8JS88_ERECY         ----------------------------------------M----RSV
197 tr|A0A136J7D3|A0A136J7D3_9PEZI ------------------------------MRFS----------------LP
198 tr|E9E4X8|E9E4X8_METAQ         ----------------------------MK---PFISPRPT------
199 tr|Q6CP42|Q6CP42_KLULA         ----------------------------------M----------------RV
200 tr|G0WHL4|G0WHL4_NAUDC         ----------------------------MVRLLRR---PS-----LYF---HFFQ
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU ------------------------------MLV----N----------------KF
202 tr|F9WYX6|F9WYX6_ZYMTI         ------------------------------MRGISGFG----------------
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 ----------------------------MKL---APLARVG------
204 tr|J3P147|J3P147_GAGT3         ----------------------------MFGSSK--LSP---AGVLLLAAASILP
```

```
205 tr|I2H842|I2H842_TETBL
206 tr|S6E3R3|S6E3R3_ZYGB2
207 tr|E5AD94|E5AD94_LEPMJ
208 tr|Q6FJM7|Q6FJM7_CANGA
209 tr|A5DUW0|A5DUW0_LODEL
210 tr|G8BF46|G8BF46_CANPC
211 tr|W3XAF6|W3XAF6_PESFW
212 tr|B6H7I2|B6H7I2_PENRW
213 tr|Q5ATF9|Q5ATF9_EMENI
214 tr|A0A167CDD3|A0A167CDD3_9ASCO
215 tr|G0RXS3|G0RXS3_CHATD
216 tr|B8NX26|B8NX26_ASPFN
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM
219 tr|A1CMM3|A1CMM3_ASPCL
220 tr|B6HU84|B6HU84_PENRW
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS
222 tr|G2WHY6|G2WHY6_YEASK
223 tr|G0W7G8|G0W7G8_NAUDC
224 tr|G4NGV8|G4NGV8_MAGO7
225 tr|W3X855|W3X855_PESFW
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA
227 tr|J8Q2K7|J8Q2K7_SACAR
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO
229 tr|J6EGN8|J6EGN8_SACK1
230 tr|A0A100ISD9|A0A100ISD9_ASPNG
231 tr|G8YQC0|G8YQC0_PICSO
232 tr|F7VVP8|F7VVP8_SORMK
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH
234 tr|F8MBP4|F8MBP4_NEUT8
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG
236 tr|A1CD27|A1CD27_ASPCL
237 tr|A5DHF3|A5DHF3_PICGU
238 tr|D4ATT7|D4ATT7_ARTBC
239 tr|E2PT42|E2PT42_ASPNC
240 tr|C7ZPE5|C7ZPE5_NECH7
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA
242 tr|Q4WKP7|Q4WKP7_ASPFU
243 tr|F8MXT4|F8MXT4_NEUT8
244 tr|A0A177AJ74|A0A177AJ74_9PEZI
245 tr|J7RQR5|J7RQR5_KAZNA
246 tr|Q75CW7|Q75CW7_ASHGO
247 tr|G3AKK4|G3AKK4_SPAPN
248 tr|B2W762|B2W762_PYRTR
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH
250 tr|C4XWK1|C4XWK1_CLAL4
```

```
001 tr|A0A254U2J9|A0A254U2J9_ASPNG
002 sp|Q6FLP9|DCW1_CANGA
003 tr|A1CCM5|A1CCM5_ASPCL
004 tr|C1H190|C1H190_PARBA
005 tr|Q2UR85|Q2UR85_ASPOR
006 tr|A5DV30|A5DV30_LODEL
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO
008 tr|J3PGN0|J3PGN0_GAGT3
009 tr|A0A124BXU6|A0A124BXU6_ASPNG
010 tr|C5P4A1|C5P4A1_COCP7
011 tr|Q0CG55|Q0CG55_ASPTN
012 tr|C4Y2X5|C4Y2X5_CLAL4
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG
014 tr|C7Z068|C7Z068_NECH7
015 tr|G8BXX2|G8BXX2_TETPH
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR
017 tr|A1CSC2|A1CSC2_ASPCL
018 sp|Q05031|DFG5_YEAST
019 tr|H2AQJ6|H2AQJ6_KAZAF
020 tr|G2XFH9|G2XFH9_VERDV
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH
```

```
022 sp|Q5AD78|DCW1_CANAL          -----SIYLIISLFSSFSHAIW----------LDTN-------NET--
023 tr|B6H7E8|B6H7E8_PENRW        ----TL-LTF-LLL--QNALAIE----------VNPE-------DET--
024 tr|B6GZU2|B6GZU2_PENRW       A---------ALLGTSTVQAVE----------LDVE-------DGD--
025 tr|E9E4W7|E9E4W7_METAQ       ----LVAIASSLLFSGVAEAQYKVDT--------------------RD--
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO I---------GLAGLGYVSAFD----------MDIE-------DPD--
027 tr|G2R7G8|G2R7G8_THITE       -------V------LLAAT-----GAIAAGEAKLQVDLS-------SPD--
028 tr|G3JBY1|G3JBY1_CORMM      A----VSLVSLLLA--APSNALEID---LS-SK--------------A--
029 tr|A5DNV5|A5DNV5_PICGU       -----SCGFLVVSALIWVVSALD----------LQTD-------DFD--
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI A---------LLSICGISSAIH----------VDLS-------SSD--
031 tr|Q2TWC1|Q2TWC1_ASPOR       ----AVL---LLLQGQVALSALS----------VQVN-------SKD--
032 tr|J5RXY9|J5RXY9_SACK1       I----GL--LSVLFFTKFSNAVE----------LDLD-------NYE--
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA A----QTLLGLSAT--ASAAQLN----------LDLT-------SES--
034 tr|A0A254U5V1|A0A254U5V1_ASPNG ----AV-FLTIFVSFQRTVTALQ----------LDIN-------DEQ--
035 tr|A0A061B6H7|A0A061B6H7_CYBFA R----LTSLLAVSVLNSVAQAIW----------LTPD-------DLD--
036 tr|G8ZQ93|G8ZQ93_TORDC       L----A----ALASTFACANAVW----------IDLN-------NTE--
037 tr|Q9C2J1|Q9C2J1_NEUCS       -----VCGLMGLLA--QSATAIT----------MDID-------DTQ--
038 tr|C5NZK5|C5NZK5_COCP7       V---------LLGVQAVHAAIP----------LDLG-------SPD--
039 tr|Q6C0T7|Q6C0T7_YARLI       -----LVAALCGLITTVTC-ALN----------LDLN-------SPV--
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB V----TL--LSGLVSLVSVFGLE----------LDLD-------DYA--
041 tr|F7VZ72|F7VZ72_SORMK       -----AVALV--GVDVA-NAAFSVAS--------------------VA--
042 tr|A3LN37|A3LN37_PICST       ----------LTASVSLFMACASAMW----------LDTN-------NDT--
043 tr|Q6CAI2|Q6CAI2_YARLI      I----------SRLLFATTASALT----------VDLD-------DPE--
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO -------VLS-AVIGLVSTVQGIW----------LDTN-------NVT--
045 tr|G9MJS5|G9MJS5_HYPVG       ----AA-LLSVTDMAVAQQSPFSIAS--------------------NS--
046 tr|G9MHI4|G9MHI4_HYPVG      A----ATAALCLLG--QGANAIA----------LTVT-------DSN--
047 tr|D4AP27|D4AP27_ARTBC      L---------LLLQSWLASAIK----------LDLE-------SDD--
048 tr|B6GZT8|B6GZT8_PENRW       ----AA-IFALLLSS--IARAID----------LDIS-------DEQ--
049 sp|O74556|YCZ2_SCHPO        L----ATILFSFA------EAIS----------VDLN-------DTS--
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 ----AAALLVAASLQTANAQYYKIDT--------------------VD--
051 tr|B8MHG0|B8MHG0_TALSN       ----SI-LLGQLLLTMPFATALT----------VNAN-------DPN--
052 tr|A3LMV8|A3LMV8_PICST       -------STFVLTLSAILSVSAVS----------LDVD-------SKE--
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU -------AALMVAGALASSATAIE----------VELG-------NEE--
054 sp|Q9P6I3|YHG7_SCHPO        ----------FFEFFFLFSFAFAFD----------FDVT-------SDD--
055 tr|Q2US57|Q2US57_ASPOR       ----ALVYLGVQLLR--TVKAFS----------VDVS-------SEQ--
056 tr|W7N622|W7N622_GIBM7      K----TVAVTSALA--AVVHGID----------VSWD-------DDK--
057 sp|Q9P6I4|YHG6_SCHPO        ----------LFYFFLVFIGAQAMD----------LDIT-------DYQ--
058 tr|C7GRB3|C7GRB3_YEAS2      F----L-------SFFKATHAMD----------LDTT-------SKT--
059 tr|C7GP28|C7GP28_YEAS2      I----GL--LGVLFATRFTNAVE----------LDLD-------NYE--
060 tr|G3JGZ5|G3JGZ5_CORMM       ----AASLVLSGLYDSVQAQYYKVDT--------------------KE--
061 tr|C4QXV4|C4QXV4_KOMPG      L-------VLVLALLRVLSPAKGLE----------LELS-------DLS--
062 tr|I1S0Z5|I1S0Z5_GIBZE       -----IALLASTGLVSAQVSPYSISS--------------------TD--
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM ------------TLSPAYSIP----------LDPN-------DPT--
064 tr|G0W5X4|G0W5X4_NAUDC      R----VFNILFLLTFTHRANALS----------LDLN-------NYQ--
065 tr|G8ZPC1|G8ZPC1_TORDC      ----------------MGLD----------LDPS-------SKD--
066 tr|A0A167CDC5|A0A167CDC5_9ASCO V----SLFALTLLQCVYNTVAIT----------VDFQ-------DDT--
067 tr|R9XAT7|R9XAT7_ASHAC      T----FTVAATLSLLAGCGRTLQ----------LDVD-------DLQ--
068 tr|B2B747|B2B747_PODAN       -----AVLLS--GAAVANAATYSIAS--------------------AA--
069 tr|G0SE99|G0SE99_CHATD       -----TAVLL--GARVT-YAAFKLDT--------------------IE--
070 tr|J8Q6F0|J8Q6F0_SACAR      F----L-------TFFKASHAMD----------LDTT-------SNT--
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH I----GLLNSLFLAVT-SVLAID----------LDTS-------SKD--
072 tr|J3K6E2|J3K6E2_COCIM      A---------LVGMLSIVSAID----------LNLD-------DHE--
073 tr|G4N3E1|G4N3E1_MAGO7       ----T-ATLSLAAVPIANAQFYKISS--------------------SQ--
074 tr|C5M5J2|C5M5J2_CANTT       ----SLIVYVLLLFTSVVYSVD----------MDTS-------SKE--
075 tr|A0A254U0X0|A0A254U0X0_ASPNG ----FLL---LLLHSQRAYSDLG----------IQVN-------DAN--
076 sp|Q75DG6|DCW1_ASHGO        T----FTAAAVLSLLAASGRTLN----------LDVD-------DLQ--
077 tr|Q6BZF0|Q6BZF0_DEBHA       -------VLS-AVIGLVSTVQGIW----------LDTN-------NVT--
078 tr|E9E3Q1|E9E3Q1_METAQ       ----WLAFAGTLLIGTTGAQYYKIDT--------------------ED--
079 sp|P36091|DCW1_YEAST        I----GL--LGVLFATRFTNAVE----------LDLD-------NYE--
080 tr|W0TAH6|W0TAH6_KLUMD      W----KLMLTTLTAFLTVASGID----------LDVD-------DLN--
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA S----LRALMVPMFLAQQALSID----------LDVD-------SET--
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI T----IAGMAGLLL--QPAAAID----------VNFD-------DTQ--
083 tr|Q2TYU3|Q2TYU3_ASPOR       ----AIVC-LAGIGR--VVEALS----------IDIN-------DAD--
084 tr|G2XGF6|G2XGF6_VERDV       -----LSAVCLM-G--NAAHAIQ----------LDID-------SEQ--
085 tr|G3AMT4|G3AMT4_SPAPN       -------TLVLTTLCITNVVHAIW----------LDTN-------NDT--
086 tr|A0A1C1D213|A0A1C1D213_9EURO S----------ALLFTQNALAID----------LDLS-------SPD--
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS ----------------------------------------------------
088 tr|A0A100I5A6|A0A100I5A6_ASPNG ----SLL---LLLHSQRAYSDLG----------IQVN-------DAD--
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH G----LLW----QLMVLPVRALD----------LDTT-------SKD--
090 tr|G2Q8A7|G2Q8A7_MYCTT       -----AAALA--GAGVA-SAAYSIAT--------------------TD--
091 tr|B8NPZ1|B8NPZ1_ASPFN       ----AVL---VFLSGQGALADLG----------IDVN-------NVD--
```

```
092 tr|Q7SAB2|Q7SAB2_NEUCR
093 tr|A2R8R5|A2R8R5_ASPNC
094 tr|C5DHG0|C5DHG0_LACTC
095 tr|G0SFA3|G0SFA3_CHATD
096 tr|K1WJG5|K1WJG5_MARBU
097 tr|H2B005|H2B005_KAZAF
098 tr|C5DYD7|C5DYD7_ZYGRC
099 tr|G8YM23|G8YM23_PICSO
100 tr|W3X8E3|W3X8E3_PESFW
101 tr|A0A254U3K7|A0A254U3K7_ASPNG
102 tr|Q6CER8|Q6CER8_YARLI
103 tr|G2QT05|G2QT05_THITE
104 sp|Q5ACZ2|DFG5_CANAL
105 tr|G2WKU6|G2WKU6_YEASK
106 tr|A0A0C7N752|A0A0C7N752_9SACH
107 tr|W3WN68|W3WN68_PESFW
108 tr|G2QFS1|G2QFS1_MYCTT
109 tr|G3J9G4|G3J9G4_CORMM
110 tr|Q2UJ03|Q2UJ03_ASPOR
111 tr|G8YRT2|G8YRT2_PICSO
112 tr|W3X554|W3X554_PESFW
113 tr|Q1K7A8|Q1K7A8_NEUCR
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM
115 tr|A7F3P0|A7F3P0_SCLS1
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA
117 tr|Q6CIP9|Q6CIP9_KLULA
118 tr|H2AMQ5|H2AMQ5_KAZAF
119 tr|G8JMI3|G8JMI3_ERECY
120 tr|C5DGT1|C5DGT1_LACTC
121 tr|K1XJR4|K1XJR4_MARBU
122 tr|A0A124BYC5|A0A124BYC5_ASPNG
123 tr|A7TLL2|A7TLL2_VANPO
124 tr|Q0CB86|Q0CB86_ASPTN
125 tr|K1XLH0|K1XLH0_MARBU
126 tr|J7S438|J7S438_KAZNA
127 tr|Q4WG09|Q4WG09_ASPFU
128 tr|H8WZ09|H8WZ09_CANO9
129 tr|I1RL09|I1RL09_GIBZE
130 tr|Q4W985|Q4W985_ASPFU
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU
132 tr|W6QCW3|W6QCW3_PENRF
133 tr|W3WMD3|W3WMD3_PESFW
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH
135 tr|I2GYH1|I2GYH1_TETBL
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH
137 tr|Q752P3|Q752P3_ASHGO
138 tr|B8MYP3|B8MYP3_ASPFN
139 tr|Q2TXL6|Q2TXL6_ASPOR
140 tr|Q0CN17|Q0CN17_ASPTN
141 tr|J3NHD5|J3NHD5_GAGT3
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI
143 tr|A1CVB0|A1CVB0_NEOFI
144 tr|B8NVI3|B8NVI3_ASPFN
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH
146 tr|A1DJ54|A1DJ54_NEOFI
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA
149 tr|A0A177A618|A0A177A618_9PEZI
150 tr|Q75CW6|Q75CW6_ASHGO
151 tr|Q7S4K4|Q7S4K4_NEUCR
152 tr|G2QB99|G2QB99_MYCTT
153 tr|G0S3F2|G0S3F2_CHATD
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM
157 tr|G2QKJ0|G2QKJ0_MYCTT
158 tr|Q4WFX5|Q4WFX5_ASPFU
159 tr|Q5BGD7|Q5BGD7_EMENI
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH
161 tr|G8BYN9|G8BYN9_TETPH
```

```
162 tr|F7VVT4|F7VVT4_SORMK          -----VAP---LLLAAASSVY-AATTTGQSTLQVNVN-------DTA--
163 tr|G8YB80|G8YB80_PICSO          -------VFSALLGILVTVTGLE----------LDVT-------SKS--
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA M---ILLPILTAFLVTRSGAID----------LDLS-------SRD--
165 tr|G2WRS7|G2WRS7_VERDV          -----ATTLL--SAHTAHAAYYSIQS------------------DD--
166 tr|A0A124BVB5|A0A124BVB5_ASPNG S---LLVCVIWSLPCFQVLAID----------LDLQ-------NAD--
167 tr|A7EIG7|A7EIG7_SCLS1          A----------LIAGSHLVNAIE----------VDLT-------SPD--
168 tr|C5M5U7|C5M5U7_CANTT          -----TAYVIISIFSTLANAIW----------LDTN-------NET--
169 tr|B9WAA8|B9WAA8_CANDC          -----QLIISILLFFTSSVQSVD----------LTVD-------DKD--
170 tr|F8MRU1|F8MRU1_NEUT8          -----LFAAAACLLP-AANAQGYAIDT------------------TD--
171 tr|A1DJS0|A1DJS0_NEOFI          -----NALWTTLIFSS--LAQAID----------VNIN-------DDQ--
172 tr|B6HLM8|B6HLM8_PENRW          -----AI-VALRILPFIGHVYALD----------LDVN-------SQE--
173 tr|A0A117DWS0|A0A117DWS0_ASPNG -----AV-FLTIFVFFQRTVTALQ----------LDIN-------DEQ--
174 tr|A0A124BXM1|A0A124BXM1_ASPNG -----GLFLNTALFSQ--LSHAID----------IDIS-------STS--
175 tr|J3NZQ6|J3NZQ6_GAGT3         LH--ATASIQGVAAQANGPNFYKIST------------------KE--
176 tr|W0THH8|W0THH8_KLUMD         L----GIWSYLILSISSLVSAID----------LDTS-------SKD--
177 tr|B2AEF2|B2AEF2_PODAN          -----LLATTSCLLP-GARATVYKLGT------------------KA--
178 tr|Q96TX1|Q96TX1_NEUCS          -----AVALV--GVDVA-NAAYSIAT------------------VA--
179 tr|E9DYA6|E9DYA6_METAQ          -----SWTLLA-AGRLSAAASVFSLDS------------------DD--
180 tr|B9WAI2|B9WAI2_CANDC          -------STYIIISLFSSLSHAIW----------LDTN-------NET--
181 tr|Q6C171|Q6C171_YARLI          A----ILSALAFLLGTARADLSF----------MDLS-------KPD--
182 tr|A6ZMV1|A6ZMV1_YEAS7          F----L-------SFFKATHAMD----------LDTT-------SKT--
183 tr|B6K7X7|B6K7X7_SCHJY          I----TTVLLGFLTCIRLSFAFS----------LDVT-------DEN--
184 tr|I1RSA2|I1RSA2_GIBZE          T----TTAVLGWTA--VAANAYE----------LT---------ITD--
185 tr|F7VWZ9|F7VWZ9_SORMK          -----AAGLMGLFA--QSATAIK----------MDID-------NDQ--
186 tr|A0A1B2J789|A0A1B2J789_PICPA L-------VLVLAFLRVLSPAKGLE----------LELS-------DLS--
187 tr|G0V8N4|G0V8N4_NAUCC          I----IYLLFTVSSFSFPIAAVD----------LNLN-------SLQ--
188 tr|H8WXN8|H8WXN8_CANO9          -------SFAILLTSLLTFSQAIS----------LDTN-------NVT--
189 tr|A6ZZS0|A6ZZS0_YEAS7          I----GL--LGVLFATRFTNAVE----------LDLD-------NYE--
190 tr|G9MQD1|G9MQD1_HYPVG          -----SFVIIATCFFYATVEAIYRIDN------------------AD--
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH S----VAWILALLASFHNCAGVY----------LDVN-------NTA--
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM -----ATT---LLLSGQVALADLG----------IQAD-------NVD--
193 tr|A0A254UET1|A0A254UET1_ASPNG S----IILCVVWSLPCFHVLAID----------LDLQ-------NPD--
194 tr|A1D587|A1D587_NEOFI          -----TS-LVTVVLAGQGAVTALE----------LDIN-------DVQ--
195 tr|Q1K7I4|Q1K7I4_NEUCR          -----AVALV--GVDVA-NAAYSIAT------------------VA--
196 tr|G8JS88|G8JS88_ERECY          W----VILGVFLNLLGTDVSALE----------LDVD-------DFS--
197 tr|A0A136J7D3|A0A136J7D3_9PEZI -------FAGLGLLA--GSVGAIQ----------VNFD-------SDD--
198 tr|E9E4X8|E9E4X8_METAQ          -----GVMLLAAAAFAQAQTSPFSIDSTGKATPIPTPNT-------VRA--
199 tr|Q6CP42|Q6CP42_KLULA          F----KLVLSLVTLFSSVSFGLD----------LDVN-------DLN--
200 tr|G0WHL4|G0WHL4_NAUDC          L----L-------LTVHNTQAID----------LDIT-------SKE--
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU V----GL--LSVLFATRFTNAVE----------LDLD-------NYQ--
202 tr|F9WYX6|F9WYX6_ZYMTI          ----------LTSLALSAGLANAID----------LDVN-------NRD--
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 -----ATALL--AGGLV-DAAYSIAN------------------TA--
204 tr|J3P147|J3P147_GAGT3          VTAQAPAAPAPAAPAGDPHPFYKIGT------------------RD--
205 tr|I2H842|I2H842_TETBL          F----HSLLLKISVITTITNAVS----------LNLN-------DSS--
206 tr|S6E3R3|S6E3R3_ZYGB2          I----IAW-LLH--AIIPVWALD----------LEVE-------SKD--
207 tr|E5AD94|E5AD94_LEPMJ          A----LSLLSS--HVVPQVQALVLD---PTSPGKILTA-------TLE--
208 tr|Q6FJM7|Q6FJM7_CANGA          L----QLW-VLIGTLFPDAGAIN----------L-QD-------NKD--
209 tr|A5DUW0|A5DUW0_LODEL          -------VSCFITFCFLTLTQAIW----------LDTN-------NDT--
210 tr|G8BF46|G8BF46_CANPC          -------SLALLLANLLPFLQALP----------LDTN-------NVT--
211 tr|W3XAF6|W3XAF6_PESFW          -----VFGLLQTLA--YQVSAVD----------LDIT-------QPD--
212 tr|B6H7I2|B6H7I2_PENRW          ----GL-IATLLFST--LTPAIE----------LDLN-------DEQ--
213 tr|Q5ATF9|Q5ATF9_EMENI          ----------ALLASLGRISALE----------IQLN-------DPQ--
214 tr|A0A167CDD3|A0A167CDD3_9ASCO -------ITASVALLVQQATSLT----------LDET-------SVE--
215 tr|G0RXS3|G0RXS3_CHATD          -------ALL---FLTAAC-----VHAQDGRAKLQVDLD-------DYD--
216 tr|B8NX26|B8NX26_ASPFN          -----TLL--LLFLHAQTVWSDLG----------IVAN-------NED--
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM V----------LLGVQAVHAAIP----------LDLG-------SPG--
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM S----LL--LSLSLLPLRATTLD----------LNVT-------DPD--
219 tr|A1CMM3|A1CMM3_ASPCL          A----------LLAGQMAQAAID----------LDLT-------S----
220 tr|B6HU84|B6HU84_PENRW          -----TL-LGGAFAV--QGALAIE----------LNID-------DEMRL
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS -----AVALV--GVDVA-NAAYSIAT------------------VA--
222 tr|G2WHY6|G2WHY6_YEASK          I----GL--LGVLFATRFTNAVE----------LDLD-------NYE--
223 tr|G0W7G8|G0W7G8_NAUDC          L----SIVTTTTLLLSHSTKALE----------IDLD-------DID--
224 tr|G4NGV8|G4NGV8_MAGO7          -----AAAACLLLQQLPGAVSLN----------VNFN-------DPN--
225 tr|W3X855|W3X855_PESFW          -------YLSICVASVLPLVSAVD----------IDWN-------DDS--
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA -----ATALL--S-LASGQNTYSIDS------------------DD--
227 tr|J8Q2K7|J8Q2K7_SACAR          I----GV--LGVLFAARSVAAVE----------LDLD-------NYE--
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO -------ILS--LSTIISFTQAVF----------LDTN-------NVT--
229 tr|J6EGN8|J6EGN8_SACK1          F----L-------SFVGISHAMD----------LDTT-------SKT--
230 tr|A0A100ISD9|A0A100ISD9_ASPNG ----------TFIVVVQAID----------YDVD-------DPD--
231 tr|G8YQC0|G8YQC0_PICSO          -------VLLLLFLFNSLCFAVD----------INPG-------DVD--
```

```
232 tr|F7VVP8|F7VVP8_SORMK          ----LLVATGCLLP-AANAQGYAVDT--------------------PD--
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH F---L-------SFFKATHAMD----------LDTT-------SKT--
234 tr|F8MBP4|F8MBP4_NEUT8          -----VCGLMGLLA--QSATAIT----------MDID-------DTQ--
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG ----------AFTVATLAGTPAAID----------LDIT-------NEQ--
236 tr|A1CD27|A1CD27_ASPCL          ----ATL---LL-LQQVALADLG----------IQAN-------NVE--
237 tr|A5DHF3|A5DHF3_PICGU          ----------FVSALLATVSAVS----------IDVN-------DKD--
238 tr|D4ATT7|D4ATT7_ARTBC          P---------LLGAQVALGSID----------LNLD-------DEN--
239 tr|E2PT42|E2PT42_ASPNC          ----------TFIVVVQAID----------YDVD-------DPD--
240 tr|C7ZPE5|C7ZPE5_NECH7          -----IAALLASAGLVGAQSQYSIAS--------------------ND--
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA L---------LLLAALPIHAIE----------LDLE-------SDE--
242 tr|Q4WKP7|Q4WKP7_ASPFU          ----TS-LVTVVLAGQGAVTALE----------LDIN-------DVQ--
243 tr|F8MXT4|F8MXT4_NEUT8          -----AVALV--GVDLA-NAAYSIAT--------------------VA--
244 tr|A0A177AJ74|A0A177AJ74_9PEZI A---------AAGFLGHAVAIE----------VDFT-------SPA--
245 tr|J7RQR5|J7RQR5_KAZNA          T----LAS-LTLFTLFTGA----------------SAL-------SLD--
246 tr|Q75CW7|Q75CW7_ASHGO          -----MYVLLVILNIVLSNALD----------VDFN-------SSE--
247 tr|G3AKK4|G3AKK4_SPAPN          I----TTYISVFLTFTSLVSAVP----------LTVG-------DKD--
248 tr|B2W762|B2W762_PYRTR          S---------LL--LSGVGAVE----------LDIN-------SPE--
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH I----GL--LGVLFATRFTNAVE----------LDLD-------NYE--
250 tr|C4XWK1|C4XWK1_CLAL4          -----IQALSALSLCFLRVFAVT----------LDVE-------SED--

001 tr|A0A254U2J9|A0A254U2J9_ASPNG --SIKDAASKTAYGSMTWYHG--NET------GQ--IPGA---FP-----
002 sp|Q6FLP9|DCW1_CANGA           --SLQNATALVAYGLMDYYTG--DQY------GK--TVGM---FS-----
003 tr|A1CCM5|A1CCM5_ASPCL         --SIKDAASTAAFGAMSYYHG--NES------GQ--IPGA---FP-----
004 tr|C1H190|C1H190_PARBA         --SIKGAAKIVAKKLVSYYTG--WRP------GD--VPGN---LP-----
005 tr|Q2UR85|Q2UR85_ASPOR         --SIKSAAATAAYNMMSYYHG--NES------GQ--TPGK---LP-----
006 tr|A5DV30|A5DV30_LODEL         --SICNAAMKVQNGTWNYYEG--TKY------GG--TVGM---FA-----
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO --SVCSAATELVDGIMDYYLG--TRY------GG--TVGM---FQ-----
008 tr|J3PGN0|J3PGN0_GAGT3         --AIKQSAKTLAYDMMLMYPG--NKT------GQ--EPGILPGPPTENK
009 tr|A0A124BXU6|A0A124BXU6_ASPNG --SIKNAASTAAYGMMKHYKG--NES------GE--IPGK---IP-----
010 tr|C5P4A1|C5P4A1_COCP7         --SIKLAAKTAATGMTSFYTG--HHP------GG--IPGT---LP-----
011 tr|Q0CG55|Q0CG55_ASPTN         --SIKKAAQVVAKNMLSFYTG--DRP------GD--NPGN---LP-----
012 tr|C4Y2X5|C4Y2X5_CLAL4         --TIKQATALIAEGLLDYYEG--TKY------GG--TIGM---FS-----
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG --SIKAACHSVARQMLTHYTG--NQP------GD--NPGN---LP-----
014 tr|C7Z068|C7Z068_NECH7         --EIKDTAAKLAYDLMLQYDG--NTT------GM--IPGILPGPPTEFK
015 tr|G8BXX2|G8BXX2_TETPH         --SLQNATALVAYGLMDYYTG--LQY------GK--TIGM---FS-----
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR --TIKDATKLIAKGLLDYYQG--NKY------GG--TIGM---FS-----
017 tr|A1CSC2|A1CSC2_ASPCL         --SIKNAAATTAFNMMSDYHG--NET------GQ--IPGK---LP-----
018 sp|Q05031|DFG5_YEAST           --SICDATALIQGGMLDYYEG--TRY------GG--TVGM---FQ-----
019 tr|H2AQJ6|H2AQJ6_KAZAF         --SICDATALIQDGMLDYYEG--IRY------GG--TVGM---FQ-----
020 tr|G2XFH9|G2XFH9_VERDV         --AIRETARTLAYDMMLFYEG--NKT------GM--IPGILPGPPADGK
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH --TLQNATALVAKGLMDYYTG--NQY------GQ--TIGM---FS-----
022 sp|Q5AD78|DCW1_CANAL           --TIREDCNIIAKGLLDYYEG--TKY------GG--VIGM---FS-----
023 tr|B6H7E8|B6H7E8_PENRW         --SLKNAAKTVATTMVDYYNA--RDS------KA--IPGK---FD-----
024 tr|B6GZU2|B6GZU2_PENRW         --SIRKAAKNVATNMMSYYTG--MNP------GD--TPGN---LP-----
025 tr|E9E4W7|E9E4W7_METAQ         --NIIASSKSLAYDLMKFYHG--NES------GQ--IPGLLPGPPASGT
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO --SIKKAAKDIVTNMMTYYTG--MNP------GD--NPGN---LP-----
027 tr|G2R7G8|G2R7G8_THITE         --SIRAAAKIVAANLFSYYHG--NQP------GQ--TPGILPGPPP---G
028 tr|G3JBY1|G3JBY1_CORMM         --SIKDTASTIAYGLVKYYNG--NLT------GG--IPGL---LP-----
029 tr|A5DNV5|A5DNV5_PICGU         --SIKQDTALIAKGLLDYYDG--YKK------GG--VIGM---FT-----
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI --AIKLSAKVLAGDMLSHYHG--NEP------GQ--IPGN---LP-----
031 tr|Q2TWC1|Q2TWC1_ASPOR         --SLANAGKEIADPMMDFYAQ--NQT------EG--IPGK---LT-----
032 tr|J5RXY9|J5RXY9_SACK1         --SLENATSLIAYGLMDYYTG--NQY------GK--TIGM---FS-----
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA --SIKDTASTLAYGLMKYYNG--NQT------GG--IPGL---LP-----
034 tr|A0A254U5V1|A0A254U5V1_ASPNG --SIKDAARTTAFNMMSYYHG--NES------GQ--TPGK---LP-----
035 tr|A0A061B6H7|A0A061B6H7_CYBFA --SLKSVTGVIADGLMDYYTG--DQY------GG--TVGM---FS-----
036 tr|G8ZQ93|G8ZQ93_TORDC         --SIHNASSLVAVGLMDYYTG--LQY------GE--TIGM---FS-----
037 tr|Q9C2J1|Q9C2J1_NEUCS         --SVKDAAATIAYGMLKYYTG--NNT------GD--TPGN---LP-----
038 tr|C5NZK5|C5NZK5_COCP7         --SIKSAAKQIAGGMVKYYTG--YKP------GD--VPGN---LP-----
039 tr|Q6C0T7|Q6C0T7_YARLI         --SVNNAVALVAQGMLDYYDG--FRF------GG--TIGM---FV-----
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB --SLQNATALVAYGLMDYYTG--DQY------GK--TVGM---FS-----
041 tr|F7VZ72|F7VZ72_SORMK         --DIKKTAADIAFDMMQYYKG--NLT------GQ--TPGILPGPPPA---
042 tr|A3LN37|A3LN37_PICST         --TIVEATNLIVDGVLDYYDG--KNY------GG--VVGM---FV-----
043 tr|Q6CAI2|Q6CAI2_YARLI         --SISANAALIAQGLMNYYDG--HRY------GG--VVGM---FV-----
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO --NLKEVSALFAEGLLDYYDG--NKY------GG--TIGM---FS-----
045 tr|G9MJS5|G9MJS5_HYPVG         --DIKKTAATVAWDMLQYYHG--NES------GQ--TPGILPGPPPA---
046 tr|G9MHI4|G9MHI4_HYPVG         --SAKSAAGTIAYGLMKYYTG--NNT------GD--NPGN---LP-----
047 tr|D4AP27|D4AP27_ARTBC         --SIKLAAKTAAKAMMKYYTG--HQP------GG--VPGN---LP-----
048 tr|B6GZT8|B6GZT8_PENRW         --SIKDAASTSTYAMMKYYYG--NET------GQ--IPGA---FP-----
```

```
049 sp|O74556|YCZ2_SCHPO          --SVDLATSLVADGLLNYYAG--QHK------GG--TIGM---FL----
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 --AIKQSAKTLAYDMMLMYPG--NKT------GQ--EPGILPGPPTEN-K
051 tr|B8MHG0|B8MHG0_TALSN        --SLKSAASTAAASAVNYYNN--RES------KL--IPGK---FD----
052 tr|A3LMV8|A3LMV8_PICST        --SVCDAAKYVLDGTLNYYEG--LKP------GG--TVGM---FA----
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU --SVCNAATDLVDGIMDYYLG--TRY------AG--TVGM---FQ----
054 sp|Q9P6I3|YHG7_SCHPO          --SINSALTTVTDGMLNYYQS--TSH------------TF----T----
055 tr|Q2US57|Q2US57_ASPOR        --SIKDAASTAAYGMMHYYHG--NES------GN--IPGK---LP----
056 tr|W7N622|W7N622_GIBM7        --SIKQAASTVAYGLVKYYTG--NNT------GD--TPGN---LP----
057 sp|Q9P6I4|YHG6_SCHPO          --SIDNTVNIMMKDLMNYWNA--SSQ------------AF---VA----
058 tr|C7GRB3|C7GRB3_YEAS2        --SICDATALIQGGMLDYYEG--TRY------GG--TVGM---FQ----
059 tr|C7GP28|C7GP28_YEAS2        --SLQNATSLIAYGLMDYYTG--NQY------GK--TVGM---FS----
060 tr|G3JGZ5|G3JGZ5_CORMM        --AIKTTASTLAYDLMLLYEG--NKT------GQ--IPGILPGPPSEN-K
061 tr|C4QXV4|C4QXV4_KOMPG        --SLQHATSLVADGLMDYYEG--FHL------GG--TIGM---FT----
062 tr|I1S0Z5|I1S0Z5_GIBZE        --AIKKSAKQLAADLIEYYHG--NDP------GG--IPGILPGPPPN---
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM --SIKQAAHHVAANMLSHYTG--MKP------GG--NPGN---LP----
064 tr|G0W5X4|G0W5X4_NAUDC        --SLENDTALVAYGLMDYYNG--DQY------GQ--TVGM---FS----
065 tr|G8ZPC1|G8ZPC1_TORDC        --SICYATALIQKGMLDYYEG--TRI------GG--TVGM---FQ----
066 tr|A0A167CDC5|A0A167CDC5_9ASCO --SLNNALALVADGLMDYYNG--DQY------GG--TPGM---FV----
067 tr|R9XAT7|R9XAT7_ASHAC        --SIREATSLVATGLMDYYHG--HDY------GQ--TVGK---FT----
068 tr|B2B747|B2B747_PODAN        --DIKQTSSLLAWDLLQYYKG--NLT------GQ--TPGILPGPPPA---
069 tr|G0SE99|G0SE99_CHATD        --DIKETAATVAWDLMQYYHG--NES------GQ--TPGILPGPPPA---
070 tr|J8Q6F0|J8Q6F0_SACAR        --SICDATALIQDGMLDYYEG--TRY------GG--AVGM---FQ----
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH --SICDATSLIQQGILDYYAG--DDY------GG--AVGM---FV----
072 tr|J3K6E2|J3K6E2_COCIM        --SIKLAAKIAATGMTSFYTG--HHP------GG--IPGT---LP----
073 tr|G4N3E1|G4N3E1_MAGO7        --EIKDSTTTLAYDLMTLYPG--NQT------GK--EPGILPGPPSEN-K
074 tr|C5M5J2|C5M5J2_CANTT        --SICEAAKVVADGMWNYYEG--FKY------GG--VVGM---FA----
075 tr|A0A254U0X0|A0A254U0X0_ASPNG --SLANAGSTIADPLMEFYKQ--NQT------EG--IPGK---LT----
076 sp|Q75DG6|DCW1_ASHGO          --SIREATSLLATGLMDYYHG--HDY------GE--TVGK---FS----
077 tr|Q6BZF0|Q6BZF0_DEBHA        --NLKEVSAIFAEGLLDYYDG--NEY------GG--TIGM---FT----
078 tr|E9E3Q1|E9E3Q1_METAQ        --AIKESARTLAYDLMLLYKG--NQS------GE--IPGILPGPPADG-K
079 sp|P36091|DCW1_YEAST          --SLQNATSLIAYGLMDYYTG--NQY------GK--TVGM---FS----
080 tr|W0TAH6|W0TAH6_KLUMD        --SLQKATSLVSSGLMDYYTG--LQK------GQ--TIGM---FA----
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA --SICDASALVIGGVMDYYQG--TRY------GG--TVGM---FQ----
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI --SVLDAASTISYGLLKFYTG--NNT------GD--VDGN---LP----
083 tr|Q2TYU3|Q2TYU3_ASPOR        --SIKSAASQTAYGSMLWYSG--NET------GQ--IPGA---FP----
084 tr|G2XGF6|G2XGF6_VERDV        --SVKDAASTIAFGLVSFYTG--NNT------GD--VPGN---LP----
085 tr|G3AMT4|G3AMT4_SPAPN        --TIREDCNLIAKGLLDYYEG--TKY------GG--TIGM---FT----
086 tr|A0A1C1D213|A0A1C1D213_9EURO --SIKSAAKIVADEMVTYYTG--YRP------GD--VPGN---LP----
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS ----------------------------------------------
088 tr|A0A100I5A6|A0A100I5A6_ASPNG --SLANAGSTIADPLMVFYKQ--NQT------EG--IPGK---LT----
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH --SICDATDLIQGGIMDYYAG--SQY------GG--TVGT---FQ----
090 tr|G2Q8A7|G2Q8A7_MYCTT        --GIKKTAADVAWDLLQYYHG--NES------GQ--TPGILPGPPPN---
091 tr|B8NPZ1|B8NPZ1_ASPFN        --SLKQAGKAVAAPMMDFYKK--NET------EG--IPGK---LT----
092 tr|Q7SAB2|Q7SAB2_NEUCR        ----------------------------------------------
093 tr|A2R8R5|A2R8R5_ASPNC        --SIKNAASIAAYGMMKHYKG--NES------GE--IPGK---IP----
094 tr|C5DHG0|C5DHG0_LACTC        --SICDATSLIQGGIMDYYAG--TRI------GG--TVGM---FQ----
095 tr|G0SFA3|G0SFA3_CHATD        --SVKAAAKTIAFGMTEYYTG--DNP------GD--VPGN---LP----
096 tr|K1WJG5|K1WJG5_MARBU        --SIKSAASTIAYGMMKYYTG--NNT------GD--NPGN---LP----
097 tr|H2B005|H2B005_KAZAF        --SLQNATALVAYGLMDYYTG--NQY------GK--TVGM---FS----
098 tr|C5DYD7|C5DYD7_ZYGRC        --SVCSATALIQQGMLDYYAG--DKY------GG--AIGM---FQ----
099 tr|G8YM23|G8YM23_PICSO        --SIQKATSLIAQGLLDYYEG--TKY------GG--TIGM---FS----
100 tr|W3X8E3|W3X8E3_PESFW        --DIKATTRTLAADLMTYYKG--NQS------GQ--TPGILPGPPPA---
101 tr|A0A254U3K7|A0A254U3K7_ASPNG --SIKNAASIAAYGMMKHYKG--NES------GE--IPGK---IP----
102 tr|Q6CER8|Q6CER8_YARLI        --TVDKALALIAEGLLDYYQG--FTY------GG--TIGM---FV----
103 tr|G2QT05|G2QT05_THITE        --DIKRTAADVAWDLLQYYKG--NQT------GQ--TPGILPGPPPA---
104 sp|Q5ACZ2|DFG5_CANAL          --SICSAAKYVVQGIWNYYEG--LKY------GG--TVGM---FA----
105 tr|G2WKU6|G2WKU6_YEASK        --SICDATALIQGGMLDYYEG--TRY------GG--TVGM---FQ----
106 tr|A0A0C7N752|A0A0C7N752_9SACH --SICAATSLIQGGIMDYYAG--SQY------GG--TVGM---FQ----
107 tr|W3WN68|W3WN68_PESFW        --SIKEGAATIAWGLVKYYTG--NNS------GD--VAGN---LP----
108 tr|G2QFS1|G2QFS1_MYCTT        --EIRESARTLAYDLMLFYKG--NQS------GE--IPGILPGPPTEH-K
109 tr|G3J9G4|G3J9G4_CORMM        --SIKSTASTLVYGLMKYYNG--NVT------GG--IPGL---LP----
110 tr|Q2UJ03|Q2UJ03_ASPOR        --SVKRASYAVANNMLSHYTG--MNP------GD--TPGV---LP----
111 tr|G8YRT2|G8YRT2_PICSO        --SVCSAAKKVADGEFNYYEG--LKK------GG--TVGM---FS----
112 tr|W3X554|W3X554_PESFW        --SIKSASSTLAFGLLKYYTG--NNT------GD--VPGN---LP----
113 tr|Q1K7A8|Q1K7A8_NEUCR        --SVKDAAATIAYGMLKYYTG--NNT------GD--TPGN---LP----
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM --SIKDAAATTAYNMMSNYTG--NQT------GQ--IPGK---LP----
115 tr|A7F3P0|A7F3P0_SCLS1        --SIKTAASDVAYDMMKYYTG--NRT------GD--VPGN---LP----
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA --SIKDAASTIAYGLVKYYNG--NLT------GG--IPGL---LP----
117 tr|Q6CIP9|Q6CIP9_KLULA        --SICTATALIQQGISDYYEG--DVY------GG--TVGM---FV----
118 tr|H2AMQ5|H2AMQ5_KAZAF        --SLQNATALVAWGLMDYYEG--TKY------GG--TVGM---FS----
```

```
119 tr|G8JMI3|G8JMI3_ERECY          --SICAATQKIQRGIMDYYWG--STP------GG--IVGM---FI-----
120 tr|C5DGT1|C5DGT1_LACTC          --SIKNATALVAGGLLDYYTG--QQY------GQ--TIGM---FS-----
121 tr|K1XJR4|K1XJR4_MARBU          --SIKDGAATLAYDMMSFYKG--NQS------GG--IIGVLPGPPPNP-P
122 tr|A0A124BYC5|A0A124BYC5_ASPNG --SIKDAAATAAFNAMSYYTG--NQT------GQ--IPGY---IN-----
123 tr|A7TLL2|A7TLL2_VANPO          --SLRNATSLVAYGLMDYYNG--LQY------GQ--TIGM---FS-----
124 tr|Q0CB86|Q0CB86_ASPTN          --SVKSAASEAVYGAMKWYSG--NDT------GG--IPGA---FP-----
125 tr|K1XLH0|K1XLH0_MARBU          --SIIDAAGTVAYDMMTYYTG--NRT------GD--TPGN---LP-----
126 tr|J7S438|J7S438_KAZNA          --SIRNATSLIAYGLMDYYTG--EQY------GK--TVGM---FA-----
127 tr|Q4WG09|Q4WG09_ASPFU          --SLKSAAKTVAAPMMDFYDK--NQT------EG--IPGK---LT-----
128 tr|H8WZ09|H8WZ09_CANO9          --SICSAAKIVSDGMWNYYEG--FKH------GG--TVGM---FA-----
129 tr|I1RL09|I1RL09_GIBZE          --EIKQSARTLAYDLMLQYDG--NTT------GM--IPGILPGPPTEY-K
130 tr|Q4W985|Q4W985_ASPFU          --SIKQAAHHVAANMLSHYTG--MKP------GD--NPGN---LP-----
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU --SICDATALIQDGMLDYYEG--TRY------GG--AVGM---FQ-----
132 tr|W6QCW3|W6QCW3_PENRF          --SIKKAAKQVANNMMSYYTG--MNP------GD--NPGN---LP-----
133 tr|W3WMD3|W3WMD3_PESFW          --SIKQAASQVAEDLLTFYRG--DEP------GW--VPGILPGPPPD---
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH --TLQNATALVAKGLMDYYTG--NQY------GQ--TIGM---FS-----
135 tr|I2GYH1|I2GYH1_TETBL          --SICGATSLIIKGMLDYYWG--TQY------GG--TVGM---FQ-----
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH --SVCAATALIQQGMLDYYAG--DKY------GG--AVGM---FQ-----
137 tr|Q752P3|Q752P3_ASHGO          --DVCHFTNEIQKGIMDYYWG--SKY------GG--IVGM---FQ-----
138 tr|B8MYP3|B8MYP3_ASPFN          --SIKSAAATAAYNMMSYYHG--NES------GQ--TPGK---LP-----
139 tr|Q2TXL6|Q2TXL6_ASPOR          --SIKDAAATAAYGMMTYYHG--NES------GQ--IPGK---LP-----
140 tr|Q0CN17|Q0CN17_ASPTN          --SIKTAAATTAFNMMSQYHG--NES------GQ--IPGK---LP-----
141 tr|J3NHD5|J3NHD5_GAGT3          --DIKSTASLVAADLMTYYKG--NLP------GQ--VPGILPGPPPA---
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI --SVKSAASSIAFDLMSYYTG--NQT------GQYNIPGLLEQPP-----
143 tr|A1CVB0|A1CVB0_NEOFI          --SIKQAAHDVAANMLSHYTG--MKP------GD--NPGN---LP-----
144 tr|B8NVI3|B8NVI3_ASPFN          --SIKDAAATAAYGMMTYYHG--NES------GQ--IPGK---LP-----
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH --ATYNATALVAQGLLDYYTG--LQY------GK--PVGM---FS-----
146 tr|A1DJ54|A1DJ54_NEOFI          --SLKSAAKTVAAPMMEFYDK--NQT------EG--IPGK---LT-----
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS --NIRASAKTLAFDLMKFYNG--NQS------GQ--IPGILPGPPSDG-K
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA --SLKNVTSLIADGLMDYYTG--DQY------GQ--TIGM---FS-----
149 tr|A0A177A618|A0A177A618_9PEZI --SIKSATGEVAKKLMSFYTG--NKP------GD--TPGY---LP-----
150 tr|Q75CW6|Q75CW6_ASHGO          --SICDATKLIQVGILDYYEG--LRP------GG--RVGF---FQ-----
151 tr|Q7S4K4|Q7S4K4_NEUCR          --NIRASAKTLAFDLMKFYNG--NQS------GQ--IPGILPGPPSDG-K
152 tr|G2QB99|G2QB99_MYCTT          --SVKSAASTIAFGLVKYYTG--NNT------GD--TPGN---LP-----
153 tr|G0S3F2|G0S3F2_CHATD          --EILESARTLAYDMMLFYKG--NQS------GE--IPGILPGPPTEH-K
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO --NIKQNAALVADGLMNYYDG--YRR------GG--TIGK---FI-----
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB --SICAATALIQGGMLDYYEG--TRY------GG--TVGM---FQ-----
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM --SIKNAASTAAYGAMSYYHG--NES------GQ--IPGA---FP-----
157 tr|G2QKJ0|G2QKJ0_MYCTT          --SIKSAAKIVAKNLYSYYHG--NEP------GQ--TPGILPGPPP---G
158 tr|Q4WFX5|Q4WFX5_ASPFU          --NIKEVASQLAWDLVSFYTG--NNT------GD--VPGN---LP-----
159 tr|Q5BGD7|Q5BGD7_EMENI          --SLKEAGKTITGPMWEYYLA--NQT------EG--IPGK---LT-----
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH --SLQNATALVAYGLMDYYTG--LQY------GE--AVGM---FA-----
161 tr|G8BYN9|G8BYN9_TETPH          --SVCEATALIQKGMLDYYQG--TRY------GG--TVGM---FQ-----
162 tr|F7VVT4|F7VVT4_SORMK          --SIKTAAKTVAKNLISYYSG--DKP------GQ--TIGILPGPPP---A
163 tr|G8YB80|G8YB80_PICSO          --SIQNATALIAKGLLDYYEG--TKY------GG--TIGM---FT-----
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA --SICDAAALIQDGVLDYYEG--TRY------GG--AVGM---FV-----
165 tr|G2WRS7|G2WRS7_VERDV          --AIRSAAKTLAHDLVAYYHG--NET------GQ--TPGILPGPPPA---
166 tr|A0A124BVB5|A0A124BVB5_ASPNG --SIKSVASDLAWDLVSFYTG--NNT------GD--VPGN---LP-----
167 tr|A7EIG7|A7EIG7_SCLS1          --SIKSAASTIAYDMMTYYTG--NQT------GG--IPGN---LP-----
168 tr|C5M5U7|C5M5U7_CANTT          --TIREDCNIIAKGLLDYYEG--TTY------GG--TIGM---FT-----
169 tr|B9WAA8|B9WAA8_CANDC          --SVCSAAKVVVQGVWNYYEG--LKN------GG--TVGM---FA-----
170 tr|F8MRU1|F8MRU1_NEUT8          --NIRASAKTLAFDLMKFYNG--NQS------GQ--IPGILPGPPSDG-K
171 tr|A1DJS0|A1DJS0_NEOFI          --SIKNAASTAAYGAMSYYHG--NES------GQ--IPGA---FP-----
172 tr|B6HLM8|B6HLM8_PENRW          --SIKDAGSTATYNMMTWYKQ--NGT------D--NPGF---IS-----
173 tr|A0A117DWS0|A0A117DWS0_ASPNG --SIKDAARTTAFNMMSYYHG--NES------GQ--TPGK---LP-----
174 tr|A0A124BXM1|A0A124BXM1_ASPNG --SIKDAASKTAYGSMTWYHG--NET------GQ--IPGA---FP-----
175 tr|J3NZQ6|J3NZQ6_GAGT3          --EVKQSAKDLAWHVMTLYPG--NQT------GK--EVGIFPGPPSEG-K
176 tr|W0THH8|W0THH8_KLUMD          --SICSATSLIQQGIVDYYDG--YKY------GG--SVGM---FT-----
177 tr|B2AEF2|B2AEF2_PODAN          --EIKESAATLAYDLMLYYKG--NQS------GE--IPGILPGPPTEH-K
178 tr|Q96TX1|Q96TX1_NEUCS          --DIKKTSADIAFDMMQYYKG--NLT------GQ--TPGILPGPPPA---
179 tr|E9DYA6|E9DYA6_METAQ          --AIKESASILAWDMLQHYHG--NES------GG--TPGILPGPPPA---
180 tr|B9WAI2|B9WAI2_CANDC          --TIREDCNIIAKGLLDYYDG--NKY------GG--VIGM---FS-----
181 tr|Q6C171|Q6C171_YARLI          --TVHEALALVAGGLMDYYDG--TRP------GG--TVGM---FV-----
182 tr|A6ZMV1|A6ZMV1_YEAS7          --SICDATALIQGGMLDYYEG--TRY------GG--TVGM---FQ-----
183 tr|B6K7X7|B6K7X7_SCHJY          --SILDGLNIVKSGLMNYYDS--ST-------------T---VF-----
184 tr|I1RSA2|I1RSA2_GIBZE          --TIESAAAKAAKGLTSWYTG--MNP------GD--TPGN---LP-----
185 tr|F7VWZ9|F7VWZ9_SORMK          --SVKDAAATIAYGMMSYYTG--NQT------GD--TPGN---LP-----
186 tr|A0A1B2J789|A0A1B2J789_PICPA --SLQHATSLVADGLMDYYEG--FHL------GG--TIGM---FT-----

187 tr|G0V8N4|G0V8N4_NAUCC          --SLQNATSLVAYGLMDYYTG--DQY------GQ--TVGM---FS-----
188 tr|H8WXN8|H8WXN8_CANO9          --SVREGCNLIAQGLLDYYEG--TKY------GG--TIGE---FS-----
```

```
189 tr|A6ZZS0|A6ZZS0_YEAS7          --SLQNATSLIAYGLMDYYTG--NQY------GK--TVGM---FS------
190 tr|G9MQD1|G9MQD1_HYPVG          --NIRDSARGLAYDLMLQYEG--NKT------GQ--IPGILPGPPSEN-K
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH --SIYNATALVAEGLLDYYTG--LQY------GQ--PIGM---FS------
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM --SLKSAAKTVAAPMMDFYDK--NQT------EG--IPGK---LT------
193 tr|A0A254UET1|A0A254UET1_ASPNG --SVKSVASELAWDLVSFYTG--NNT------GD--VPGN---LP------
194 tr|A1D587|A1D587_NEOFI          --SIKDAAATTAYNMMSNYTG--NQT------GQ--IPGK---LP------
195 tr|Q1K7I4|Q1K7I4_NEUCR        --DIKKTSADIAFDMMQYYKG--NLT------GQ--TPGILPGPPPA---
196 tr|G8JS88|G8JS88_ERECY        --SMQSATSLISTGVLDYYTG--LDP------GN--TIGM---FS------
197 tr|A0A136J7D3|A0A136J7D3_9PEZI --SIKAAASTVAFGLARFYTG--NNT------GD--VPGN---LP------
198 tr|E9E4X8|E9E4X8_METAQ        --AIKQSSSTLAWDMLQYYKG--NLT------GQ--TPGILPGPPPA---
199 tr|Q6CP42|Q6CP42_KLULA        --SLKNATSLVAYGLMDYYTG--LQY------GK--TIGM---FA------
200 tr|G0WHL4|G0WHL4_NAUDC        --SICDATSLIQGGMMDYYWG--TQY------GG--TVGM---FQ------
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU --SLQNATSLVAYGLMDYYTG--NQY------GK--TVGM---FS------
202 tr|F9WYX6|F9WYX6_ZYMTI        --SVLKASKIVVDNILSVYNNYTESP------GG--IPGL---LP------
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 --DIKTTAASIAADLMAYYKG--NLP------GQ--VPGILPGPPPA---
204 tr|J3P147|J3P147_GAGT3        --QIIQSSRDLAWEVMALYPG--NQT------GQ--ETGILPGPPDDG-K
205 tr|I2H842|I2H842_TETBL        --SLDNAAALVAYGLMDYYNG--NQY------GQ--TIGM---FS------
206 tr|S6E3R3|S6E3R3_ZYGB2        --SVCAATALIQQGMLDYYAG--DKY------GG--AVGM---FQ------
207 tr|E5AD94|E5AD94_LEPMJ        --SIKDVSSKVAKQLVAQYAK--IDDKGVHVLGG--YPGI---LY------
208 tr|Q6FJM7|Q6FJM7_CANGA        --SICAATALIQGGMLDYYEG--TRY------GG--TVGM---FQ------
209 tr|A5DUW0|A5DUW0_LODEL        --TIREDCALIADGILDYYDG--YNY------GG--TIGE---FV------
210 tr|G8BF46|G8BF46_CANPC        --SVREACNLIAKGLLDYYEG--TKY------GG--TIGE---FT------
211 tr|W3XAF6|W3XAF6_PESFW        --SVKEAAATIAWGLVKYYTG--NNT------GD--VPGN---LP------
212 tr|B6H7I2|B6H7I2_PENRW        --SIKDAASTSTYSMMGWYYG--NET------GQ--IPGS---FP------
213 tr|Q5ATF9|Q5ATF9_EMENI        --SIKDAASKTAYGSLLWYSG--NET------GG--NPGA---FP------
214 tr|A0A167CDD3|A0A167CDD3_9ASCO --SIYGAQALVAQGLMDYYNG--NQT------GQ--TPGM---FS------
215 tr|G0RXS3|G0RXS3_CHATD        --SIRRAAKIVAGNLWTYYHG--DKP------GQ--TPGILPGPHPAVPS
216 tr|B8NX26|B8NX26_ASPFN        --SLKKGAQEIIPPMMEFYKE--NQT------EG--IPGK---LT------
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM --SIKSAAKQIAGGMVKYYTG--YKP------GD--VPGN---LP------
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM --NIKEVASQLAWDLVSFYTG--NNT------GD--VPGN---LP------
219 tr|A1CMM3|A1CMM3_ASPCL          ----------------------Q--GRP------GD--VPGN---LP------
220 tr|B6HU84|B6HU84_PENRW        SDSLKSAAKTVATTMMEYYDA--RES------KD--IPGK---FD------
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS --DIKKTSADIAFDMMQYYKG--NLT------GQ--TPGILPGPPPA---
222 tr|G2WHY6|G2WHY6_YEASK        --SLQNATSLIAYGLMDYYTG--NQY------GK--TVGM---FS------
223 tr|G0W7G8|G0W7G8_NAUDC        --SLRNATSLVAYGLMDYYTG--NQY------GK--TIGM---FS------
224 tr|G4NGV8|G4NGV8_MAGO7        --SVKSAASTVAFGMMKYYTG--NNT------GD--TPGN---LP------
225 tr|W3X855|W3X855_PESFW        --SIKDGAAQIAYGLLQYYTG--NNT------GD--VPGN---LP------
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA --AIKKTASQMAEDLLKHYHG--DEP------GQ--VPGILPGPPPA---
227 tr|J8Q2K7|J8Q2K7_SACAR        --SLQNATSLVAYGLMDYYTG--NQY------GE--TVGM---FS------
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO --NIEQNTALIAKGLLDYYEG--EKY------GG--TIGM---FS------
229 tr|J6EGN8|J6EGN8_SACK1        --SICDATALIQDGMLDYYDG--TRY------GG--AVGM---FQ------
230 tr|A0A100ISD9|A0A100ISD9_ASPNG --SIKAACHSIARQMLTHYTG--NQP------GD--NPGN---LP------
231 tr|G8YQC0|G8YQC0_PICSO        --SVCSAAKKVADGEFNYYEG--IKK------GG--TVGM---FT------
232 tr|F7VVP8|F7VVP8_SORMK        --NIRASAKTLAYDTMLFYKG--NQS------GE--IPGILPGPPSDG-K
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH --SICDATALIQGGMLDYYEG--TRY------GG--TVGM---FQ------
234 tr|F8MBP4|F8MBP4_NEUT8        --SVKDAAATIAYGMLKYYTG--NNT------GD--TPGN---LP------
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG --SIKNAAATAAFNAMSYYTG--NQT------GQ--IPGK---IN------
236 tr|A1CD27|A1CD27_ASPCL        --TLKSAGKDLAAPMMDFYKK--NQT------EG--IPGK---LT------
237 tr|A5DHF3|A5DHF3_PICGU        --SICEAASYVANGELNYYEG--LKY------GG--TVGM---FA------
238 tr|D4ATT7|D4ATT7_ARTBC        --SIKKAASHIAHGMVSYYTG--NHT------GD--VPGN---LP------
239 tr|E2PT42|E2PT42_ASPNC        --SIKAACHSVARQMLTHYTG--NQP------GD--NPGN---LP------
240 tr|C7ZPE5|C7ZPE5_NECH7        --DIKKTAKQLAADLIEYYHG--LDP------GG--IPGILPGPPPN---
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA --SIKSAAKTLARGCKSYYTG--DNP------GD--VPGN---LP------
242 tr|Q4WKP7|Q4WKP7_ASPFU        --SIKDAAATTAYNMMSNYTG--NQT------GQ--IPGK---LP------
243 tr|F8MXT4|F8MXT4_NEUT8        --DIKKTSADIAFDMMQYYKG--NLT------GQ--TPGILPGPPPA---
244 tr|A0A177AJ74|A0A177AJ74_9PEZI --SVKTAASSIAFDLMSYYTG--NQS------GQYNIPGLLEQPP---
245 tr|J7RQR5|J7RQR5_KAZNA        --PVCSATALVTDGLLDYYSG--TRY------GG--TVGM---FQ------
246 tr|Q75CW7|Q75CW7_ASHGO        --SICDATKLIQVGILDYYEG--AKP------GG--RAGF---LV------
247 tr|G3AKK4|G3AKK4_SPAPN        --STCAAAKVVADGIWNYYEG--LKY------GG--VVGM---FA------
248 tr|B2W762|B2W762_PYRTR        --SIKKAAKDISINLRQYYTG--DRP------GD--TPGN---LP------
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH --SLQNATSLIAYGLMDYYTG--NQY------GK--TVGM---FS------
250 tr|C4XWK1|C4XWK1_CLAL4        --SICEAAWEVIVGELNYYEG--TKY------GG--TVGM---FS------

001 tr|A0A254U2J9|A0A254U2J9_ASPNG S--KWWEGSALFTSLLLYWYYT--GDSTYNDEVRQGMQWQAG--D-CDYM
002 sp|Q6FLP9|DCW1_CANGA          DPYYWWQAGGAWGCMLDYWYFM--QNDTYNDKIMAALLHQTG-DN-NDYV
003 tr|A1CCM5|A1CCM5_ASPCL        T--KWWEGSALFMAMLQYWYFT--GDSTYNSAVSEGLEWQAG--P-GDYM
004 tr|C1H190|C1H190_PARBA        DPYYWWEAGAMFGALIDYWYYT--GDDQYNEIVMQAMLHQTG-PD-FDYQ
005 tr|Q2UR85|Q2UR85_ASPOR        D--TWWEGGAMFMTLIQYWFWT--GDTSYNEVTTQGMLWQKG-HD--DYF
```

```
006 tr|A5DV30|A5DV30_LODEL          SPNYWWNAGEAFGGLLDYYIFCDPENEQLESRIFDGMYHQAG-EN-YNYI
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO QPYYWWESALVFGGMIDTWKIC--NNYTYVSTIQSAISHQKG-DS-NDFY
008 tr|J3PGN0|J3PGN0_GAGT3         GDYYWWQGGALMGTMIDYWHLT--GDTEYNKIISEGILNQVG-EG-RDFQ
009 tr|A0A124BXU6|A0A124BXU6_ASPNG N--TWWEGGAMFMTLMQYYHYT--GDSTYNQEVIQGMQWQSG--D-CDFM
010 tr|C5P4A1|C5P4A1_COCP7          LPYYWWEAGAMFGALIDYWFYT--GDDTWNEMTTTALLHQAS-ST-KNFM
011 tr|Q0CG55|Q0CG55_ASPTN         DPYYWWEAGAMFNSLVNYWYYT--GDDTWNDITTQAILWQAG-DT-GTFM
012 tr|C4Y2X5|C4Y2X5_CLAL4         APYYWWESGGAWGSMLEYTIYM--ENDTYTDLIKEALLYQVG-DD-FNYI
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG DPYYWWEAGAMFTALVDYWYLT--SDDTWNNITTQGITWQAG-PS-GSFM
014 tr|C7Z068|C7Z068_NECH7          GDYYWWEGGAMMGTYIDYWKLT--GDSSYNHVIMEGMLHQTG-DG-HDYM
015 tr|G8BXX2|G8BXX2_TETPH         DPYYWWEAGGAWGCMLDYWFFM--ENDTYNDIIKQALLYQVG-TA-NDYV
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR WPYYWWHAGGAFGSLLDYSYYT--QDTQYDDLIEQALTYQVG-EN-YNYI
017 tr|A1CSC2|A1CSC2_ASPCL          D--TWWEGGAMFMTLIQYWYWT--GDPSYNDVVTQGMLWQKG-ND--DYF
018 sp|Q05031|DFG5_YEAST           SPYYWWHAGEAFGGMLENWFLC--ENDTYQELLYDALLAQTG-SN-YDYI
019 tr|H2AQJ6|H2AQJ6_KAZAF         SPYYWWEAGEAFGGMLENWFLC--ENDTFKEIIYDAILAQTG-SD-YDFM
020 tr|G2XFH9|G2XFH9_VERDV         GDYYWWQGGAMMGTYVDYWHLT--GDTSYNDVVREGMVHQVG-DN-RDYM
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH WPYYWWEAGGAWNTLIDFWFYM--KNDTYNSIVQQALVYQSG-EN-HDYV
022 sp|Q5AD78|DCW1_CANAL            WPYYWWEAGGAWGSLIDYTFYF--DNDTLVPLITDALLYQTG-DD-DNYI
023 tr|B6H7E8|B6H7E8_PENRW         G--TWWEGGSMFMTLIQYWYLT--GDSQFNDAIQEGMYWQKG-DD-NDFF
024 tr|B6GZU2|B6GZU2_PENRW         DPYYWWEAGAMFNSLIDYWFYT--GDDTWNDITIQAMLWQAG-EN-KAFM
025 tr|E9E4W7|E9E4W7_METAQ         GDYYWWEAGAMMGTYIDYWKLT--GDATYNDVVTQGMLFQVG-PQ-RDYM
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO DPYYWWEAGAMFNVLIDYWYFT--GDDTYNDITMQGMLWQAG-DN-GAFM
027 tr|G2R7G8|G2R7G8_THITE          GPYYWWQAGAMWGTYIDYWFYT--GDSTYNDETVRAMEFQSE---NSYQ
028 tr|G3JBY1|G3JBY1_CORMM         GPYYWWEAGGMFGHLIDYWYYT--GDTTYNEIVMQGMMHQINAPT-GDFL
029 tr|A5DNV5|A5DNV5_PICGU         WPYYWWEAGGAWGSLIEYSYYM--DNDTLVPLIKEALGYQVG-EN-YNYV
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI YPYYWWECGAMFGSLIDYWLYT--GDSAHNDLVTQGMQFQVG-PN-NDFM
031 tr|Q2TWC1|Q2TWC1_ASPOR          D--TWYIAGAMFMTLIQYWATS--GVEQYNKVVSHDLMFQSG-EN-YDYF
032 tr|J5RXY9|J5RXY9_SACK1         DPYYWWEAGGAWGCMLDYWFFM--DNDTYNDEIIAAMVHQSG-DD-NDYI
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA GPYYWWEAGGMFGHLIDYWYYT--GDASYNAVTMEGMMHQIT-PK-GDFL
034 tr|A0A254U5V1|A0A254U5V1_ASPNG D--TWWEGGAMFMTLIQYWYWT--GDTSYNEVTTQGMLWQKG-DN--DYF
035 tr|A0A061B6H7|A0A061B6H7_CYBFA NPYYWWEAGGAWGSMIDFWYYM--ENDTYLNQTYDALLAQVG-DN-WDYV
036 tr|G8ZQ93|G8ZQ93_TORDC         DPYYWWEAGGAWGCMLDYWWYM--QNDTYNDIIMQALLYQVG-TH-NDYV
037 tr|Q9C2J1|Q9C2J1_NEUCS         DPYYWWEAGAMFGAMVDYWWVT--GDTSYVEVTTQAIVHQAG-DA-RDFN
038 tr|C5NZK5|C5NZK5_COCP7         DPYYWWEAGAMFSALIDYWYYT--GDDQYNDITTQAMLHQVG-KE-NNYM
039 tr|Q6C0T7|Q6C0T7_YARLI         QPYYWWHAGAAWNAMIEYWHIT--GNSSLNDITYEALLHQRG-ER-FDLM
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB DPYYWWQAGGAWGCMLDYWYFM--QNDTYNDKIMAALLHQTG-DN-NDYV
041 tr|F7VZ72|F7VZ72_SORMK          GDYYWWEAGAMWGTLIDYWRFT--GDESYNEVTTQALLWQVG-PG-QDYM
042 tr|A3LN37|A3LN37_PICST         WPYYWWEAGGVWGSLIDYTYFT--QNDTLVPLIKEALLYQTG-DD-NNYV
043 tr|Q6CAI2|Q6CAI2_YARLI         FPYYWWEAGAAWNSMLDFWYYT--GNDTYNDVVKEALLYQVG-KN-NDYL
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO WPYYWWEAGGAWGSLIDYTYYM--ENDTFVPMIKEALEYQTG-DD-LNYI
045 tr|G9MJS5|G9MJS5_HYPVG          GDYYWWEGGAMWGTLIDYWYLT--NDSTYNDLIMQAIQWQTG-PD-DDFQ
046 tr|G9MHI4|G9MHI4_HYPVG         DPYYWWEAGAMFGTMIDYWRMT--KDSSYNTATMQAMLWQAG-TD-GAFL
047 tr|D4AP27|D4AP27_ARTBC         DPYFWWETGAMFGGLIDYWFYT--GDSQFNDIVTQGMLWQVG-PD-SNFM
048 tr|B6GZT8|B6GZT8_PENRW         D--KWWEGAPLFMALLEYWHFT--GDTTYNEELSVGLQWQGG-TD-GDYM
049 sp|O74556|YCZ2_SCHPO           PPAYWWEAGAAWNGLLNRYIAT--GNSTYNELVKTSMLYQSG-ED-SDYM
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 GDYYWWQGGALMGTMIDYWHLT--GDTEYNKIITEGILNQVG-EG-RDFQ
051 tr|B8MHG0|B8MHG0_TALSN         G--TWWEGGAFFTFLINYWHWT--GDDQYNNLVAEGLSWQGG-ED-NDFF
052 tr|A3LMV8|A3LMV8_PICST         PPNYWWNAGEAFGGIVDFYTFCNSTNETLRELIIDAMYHQAG-EN-FNYI
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU QPYYWWEAGLVFGGMIDTWKFC--NNDTYVGIIQEAIVHQKG-PD-NDFF
054 sp|Q9P6I3|YHG7_SCHPO           A--YWWMTGAGLNSMTDTYAAT--GNTTHLDMLISALVANKG-DN-NDYA
055 tr|Q2US57|Q2US57_ASPOR         D--TWWEGGAMFMTLIEYWHFT--GDATYNDEVSEGMQWQAG--D-GDYM
056 tr|W7N622|W7N622_GIBM7         DPYYWWTAGGMFGTLIDYWWLT--GDESYNKITTQAMLHQVG-TN-DDYM
057 sp|Q9P6I4|YHG6_SCHPO           S--YWWVTGATMGALLYNYELF--NNDTYVDLISSSLLYNAG-SG-FDYQ
058 tr|C7GRB3|C7GRB3_YEAS2         SPYYWWHAGEAFGGMLENWFLC--ENDTYQELLYDALLAQTG-SN-YDYI
059 tr|C7GP28|C7GP28_YEAS2         DPYYWWEAGGAWGCMLDYWFFM--DNDTYNDEIIAAMIHQAG-DD-NDYI
060 tr|G3JGZ5|G3JGZ5_CORMM         GNYYWWQGGAMMGTYIDYWHLT--GDTSYNDVVMQGILHQVG-EN-RDFM
061 tr|C4QXV4|C4QXV4_KOMPG         NPYYWWQSGAAFGSMLDYWWYM--ENDTYHDAIMQAIVYQAG-DN-ADFM
062 tr|I1S0Z5|I1S0Z5_GIBZE         GDYYWWEGGAMWGTYMDYWRCT--GDTKYNDLVMQGMQFQVG-DD-QDYQ
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM PPYYWWEAGAMFNALIDYWYLT--GDSTWNAITTQALTWQAG-HT-GTFM
064 tr|G0W5X4|G0W5X4_NAUDC         NPYYWWEAGGAWGTMLDYWFFM--ENDTYNDLIMSALLHQTG-EN-NDYI
065 tr|G8ZPC1|G8ZPC1_TORDC         SPYYWWQAGEAFGGMIDNWYFC--ENTTFESLIYDGMLAQTG-PH-YDYM
066 tr|A0A167CDC5|A0A167CDC5_9ASCO NPYYWWEAGAAFGSMLDYWFYT--GNTTYNDVIKAGMLYQTG-KN-NDYM
067 tr|R9XAT7|R9XAT7_ASHAC         DPYYWWEAGGAWGSMLDYWYYM--ENSTYNNLLTDSLLHQAG-ED-LAYT
068 tr|B2B747|B2B747_PODAN         GDYYWWEGGAMWGTLIDYWKFT--GDDSYNDLITQAMLWQVG-PD-KDYM
069 tr|G0SE99|G0SE99_CHATD         GDYYWWEAGAMWGTLIDYWHWT--GDDSYNDVITQAMLWQVG-PN-RDYM
070 tr|J8Q6F0|J8Q6F0_SACAR         SPYYWWHAGEAFGGMIENWFLC--ENDTYKELLYDALLAQTG-SD-YDYI
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH SPYYWWEAGLVFGGMIENWYLC--ENTTYEDMLYEALMAQTG-PD-YDYM
072 tr|J3K6E2|J3K6E2_COCIM         LPYYWWEAGAMFGALIDYWFYT--GDDTWNEMTTTALLHQAS-ST-KNFM
073 tr|G4N3E1|G4N3E1_MAGO7         GDYYWWEGGALMGTMIDYWHFT--GHTEYNDEVTQGILYQVG-EN-ADFQ
074 tr|C5M5J2|C5M5J2_CANTT         APNYWWNAGEAFGGLVDYYTFCDPDNSTLEGWIYDAMYHQAG-EN-YNYI
075 tr|A0A254U0X0|A0A254U0X0_ASPNG D--TWYIAGAMFMTLIQFWKAS--GVDTYNSVVQHDLMFQAG-EN-YDYF
```

```
076 sp|Q75DG6|DCW1_ASHGO                    DPYYWWEAGGAWGSILDYWYYM--ENSTYNDLLTDSLLHQAG-ED-LSYT
077 tr|Q6BZF0|Q6BZF0_DEBHA                  WPYYWWEAGGAWGSLIDYTYYM--ENDTFVPLIKQALEYQTG-DD-LNYI
078 tr|E9E3Q1|E9E3Q1_METAQ                  GPYYWWEGGAMMGTYIDYWKLT--GDSSYNKVVMEGMLHQVG-EN-KNYM
079 sp|P36091|DCW1_YEAST                    DPYYWWEAGGAWGCMLDYWFFM--DNDTYNDEIIAAMIHQAG-DD-NDYI
080 tr|W0TAH6|W0TAH6_KLUMD                  SPSYWWQAGGAWGSILDYWYYM--QNDTYNSILTQALLYQTG-DN-NDYM
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA          DPYYWWEAGEAMGSMLGFWYFC--NNDTYEDVIYDALMAQRG-SN-NDYI
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI          SPYYWWEAGAMFGTLVDYWWLT--NDTSNNDAVKRALMHQVG-GD-NNYM
083 tr|Q2TYU3|Q2TYU3_ASPOR                  D--KWWEGSALFLSLLYYWHYT--GDTTYNAEVSQGMEWQAG--N-GDYM
084 tr|G2XGF6|G2XGF6_VERDV                  DPYFWWTAGAMFGTLVDYWFLT--GDDTYNEITTQAMVHQAG-DE-ADFM
085 tr|G3AMT4|G3AMT4_SPAPN                  PPYYWWEAGGAWGGLIDYTFFF--ENDTLVPLIKAALEYQTG-DD-NNYV
086 tr|A0A1C1D213|A0A1C1D213_9EURO          APYYWWQAGAMFGSLIDYWYFT--GDASYNEITTQAMLHQVG-PD-NDFM
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS          ----------------MDYWHYT--GDSSYNELITSSMLFHTG-PPLNAYM
088 tr|A0A100I5A6|A0A100I5A6_ASPNG          D--TWYIAGAMFMTLIQFWQAS--GVDTYNSVVQHDLMFQAG-EN-YDYF
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH          EPYYWWEAGAVFGGMLENWFLC--QNDTYEQVIYDSMIAQAG-PD-YDFM
090 tr|G2Q8A7|G2Q8A7_MYCTT                  GDYYWWEAGAMWGTFIDYWKLT--GDSTYNDLVTQAMLFQVG-PD-RDYR
091 tr|B8NPZ1|B8NPZ1_ASPFN                  D--TWYVAGSMFMTLIQYWQAS--GDDTYNAVVSNDLMFQAG-EN-YDYY
092 tr|Q7SAB2|Q7SAB2_NEUCR                  ------------------MDYWHYT--GDSSYNELITSSMLFHTG-PPLNAYM
093 tr|A2R8R5|A2R8R5_ASPNC                  N--TWWEGGAMFMTLMQYYHYT--GDSTYNKEVIQGMQWQAG--D-CDYM
094 tr|C5DHG0|C5DHG0_LACTC                  QPYYWWEAGEVFGGMLENWFLC--QNDTYHDIIYDAMIHQAG-PD-YDFM
095 tr|G0SFA3|G0SFA3_CHATD                  DPYFWWEAGAMFGMLVEYWALT--KDDTYNVMTMQALLHQAT-EN-GDFM
096 tr|K1WJG5|K1WJG5_MARBU                  APYYWWEAGATFGTMVEYWYYT--NDTSYNPTVTAAILSQVG-DD-KDFM
097 tr|H2B005|H2B005_KAZAF                  DPYYWWEAGGAWGSMIDFWYYM--DNNTYNDDIMAALLHQTG-DD-WDYI
098 tr|C5DYD7|C5DYD7_ZYGRC                  PPYYWWQAGEAFGGMIDNWHFC--KNNTFEELIREAMLTQAG-PH-YDYM
099 tr|G8YM23|G8YM23_PICSO                  WPYYWWEAGGAWGSLIDYSYYT--ENDTLVPIIKQALTYQTG-ED-DNYI
100 tr|W3X8E3|W3X8E3_PESFW                  GDYYWWEGGALWGTMIDYWHLT--GDSTYNDVVTQALLWQVG-PN-QDYM
101 tr|A0A254U3K7|A0A254U3K7_ASPNG          N--TWWEGGAMFMTLMQYYHYT--GDSTYNKEVIQGMQWQAG--D-CDYM
102 tr|Q6CER8|Q6CER8_YARLI                  PPYYWWHAGAAWNAMIEYWHIT--GNSTFNNITFEAMKAQRG-DK-YNLM
103 tr|G2QT05|G2QT05_THITE                  GDYYWWEAGAMWGTLIDYWKLT--GDSTYNDLVTQAMLWQVG-PD-RDYM
104 sp|Q5ACZ2|DFG5_CANAL                    PPNYWWNAGEAFGGLVDFYTYCQSDNSTLEKLIYNGMYHQAG-EN-YNYI
105 tr|G2WKU6|G2WKU6_YEASK                  SPYYWWHAGEAFGGMLENWFLC--ENDTYQELLYDALLAQTG-SN-YDYI
106 tr|A0A0C7N752|A0A0C7N752_9SACH          APYYWWEAGEVFGGMLENWYLC--QNDSYSQVLYDSMLAQAG-PE-YDYM
107 tr|W3WN68|W3WN68_PESFW                  DPYYWWEAGAMFGALVDYWAFT--GDDSYNNITFQALQHQVG-DD-ADFM
108 tr|G2QFS1|G2QFS1_MYCTT                  GDYYWWEGGAMMGTYVDYWFLT--GDPSYNHVVTEGMLHQVG-PN-ADYM
109 tr|G3J9G4|G3J9G4_CORMM                  GKYYWWEAGGMFGHLIDYWYYT--GDASYNDVTMQGMMHQIT-PK-GDYL
110 tr|Q2UJ03|Q2UJ03_ASPOR                  PPYYWWEAGAMFNALIDYWFYT--GDDRWNDIIMQAMTWQAG-DD-GTFM
111 tr|G8YRT2|G8YRT2_PICSO                  YPYYWWHAGAAFGGLLDYYTFCDKDNSTLKKYIYNGMYHQAG-DD-YNYI
112 tr|W3X554|W3X554_PESFW                  DPYFWWEAGAMFGTMVDYWFFT--GDDTYNAITSQALLHQVG-DD-KDFM
113 tr|Q1K7A8|Q1K7A8_NEUCR                  DPYYWWEAGAMFGAMVDYWWVT--GDTSYVEVTTQAIVHQAG-DA-RDFN
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM          D--TWWEGGAMFMTLIQYWFWT--GDTSYNEVTTQGMLWQKG-NN--DYF
115 tr|A7F3P0|A7F3P0_SCLS1                  DPYYWWEAGAMFGIMVDYWYYT--GDETYNEEVKQALLHQAG-ND-YDFM
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA          GPYYWWEAGGMFGHLIDYWYYT--GDATYNDIVMQGMMHQINAPA-GDYL
117 tr|Q6CIP9|Q6CIP9_KLULA                  SPYYWWEAGLVFGGMIENWYLC--QNTSYEDMLYDALMAQTG-SD-YDYM
118 tr|H2AMQ5|H2AMQ5_KAZAF                  SPYYWWEAGGAFGTLLDFWYFM--DNDTYNDEILAAMVYQAG-ND-WDYI
119 tr|G8JMI3|G8JMI3_ERECY                  QPYYWWEAGLVFAGMLENWYLC--DNMEYNSTLYESMIAQTG-AN-YDYI
120 tr|C5DGT1|C5DGT1_LACTC                  NPYYWWEAGGAWGSILDFWWYT--QNDTYNAMLEEALLYQAG-EN-HDYI
121 tr|K1XJR4|K1XJR4_MARBU                  DGYYWWESGAMWGAMIDYWHYT--GDESYNDVVSTGIQAQVG-EN-LDLM
122 tr|A0A124BYC5|A0A124BYC5_ASPNG          H--TWWEGGVLFDTMIRYWYFT--NDASNNAAVSQGMYHQRG-AG-NDYM
123 tr|A7TLL2|A7TLL2_VANPO                  NPYYWWEAGGAWGTMIDYWYYM--QNDTYNDIIMQALQYQTG-VN-NDYV
124 tr|Q0CB86|Q0CB86_ASPTN                  D--KWWEGSALFLSTLLYWHYT--GDDTYNSVTSQGMEWQAG--N-GDYM
125 tr|K1XLH0|K1XLH0_MARBU                  DPYYWWEAGAMFGTMINYWYYT--GDTTYNPTTKQALLHQIG-DN-EDFM
126 tr|J7S438|J7S438_KAZNA                  DPYYWWEAGGAWGSMLDYSWFM--DNDTYDSQIMSAMLHQTG-EH-NNYI
127 tr|Q4WG09|Q4WG09_ASPFU                  G--TWYVAGAMFMTLIQYWQSS--GDDTYNSIVSHDLMFQSG-EN-YDFF
128 tr|H8WZ09|H8WZ09_CANO9                  SPNYWWNAGEAFGGLVDYFTYCDSDNSTLEKWIYDGMYHQAG-DN-YNYI
129 tr|I1RL09|I1RL09_GIBZE                  GDYYWWEGGAMMGTYIDYWKLT--GDSSYNKVVMEGMLHQTG-EG-RDYM
130 tr|Q4W985|Q4W985_ASPFU                  PPYYWWEAGAMFNALIDYWYLT--GDSTWNAITTQALTWQAG-HT-GTFM
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU          SPYYWWHAGEAFGGMVENWFLC--KNDTYQELLYNALLAQTG-AD-YDYI
132 tr|W6QCW3|W6QCW3_PENRF                  DPYYWWEAGAMFNALIDYWFYT--GDDTWNNITIQGMLWQAG-EN-KAFM
133 tr|W3WMD3|W3WMD3_PESFW                  GDYYWWQGGAMWGTLLDYRHHT--GDKTYDDMTSTAILFQVG-DD-RDFM
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH          WPYYWWEAGGAWNTLIDFWFYM--KNDTYNSIVQQALVYQSG-EN-HDYV
135 tr|I2GYH1|I2GYH1_TETBL                  PPYYWWEAGEAFGGMLENWYLC--QNDTYADLLFDALTAQTG-PD-YDYI
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH          SPYYWWQAGEAFGGMLDNWYFC--QNNTFENLIRKAMLAQTG-TK-YDYM
137 tr|Q752P3|Q752P3_ASHGO                  PPYYWWEAGEVFGGMLENWYLC--DNQEYESVLRESMIAQTG-DK-FDYI
138 tr|B8MYP3|B8MYP3_ASPFN                  D--TWWEGGAMFMTLIQYWFWT--GDTSYNEVTTQGMLWQKG-HD--DYF
139 tr|Q2TXL6|Q2TXL6_ASPOR                  G--TWWTGGEVFMALVQYWYWT--GDTSYNDVTKQALIWQKG-HN--DYL
140 tr|Q0CN17|Q0CN17_ASPTN                  D--TWWEGGAMFMTLIQYWYWT--GDTSYNDVTTQGMLWQKG-SN--DYF
141 tr|J3NHD5|J3NHD5_GAGT3                  GDYYWWEGGALMGQMVEYWHLT--GDTTYNDEVSQGMLHQVG-QD-RDYQ
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI          AGYYWWQAGAMWNTMVDYWYIT--GDTSYNEATTQALTFQVG-DD-QDYM
143 tr|A1CVB0|A1CVB0_NEOFI                  PPYYWWEAGAMFNALIDYWYLT--GDSTWNAITTQALTWQAG-HT-GTFM
144 tr|B8NVI3|B8NVI3_ASPFN                  G--TWWTGGEVFMALVQYWYWT--GDTSYNDVTKQALIWQKG-HN--DYL
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH          DPYYWWEAGGAWGSILDYWWYT--GNTTYNGILKQALLAQTG-DN-NDYV
```

```
146 tr|A1DJ54|A1DJ54_NEOFI          G--TWYVAGAMFMTLIQYWQAS--GDDTYNSVVSHDLMFQSG-EN-YDFF
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS GDYYWWEGGALMGTMIDYWHLT--GDTTYNDVITQGILHQVG-DN-RDFQ
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA NPYYWWEAGGAWGSMIDYWYYM--ENDTYLNQTYEALLAQVG-DD-WNYV
149 tr|A0A177A618|A0A177A618_9PEZI DPYYWWEAGAMFGAMIDYWYYT--NDTTYNDIVEQALVHQAG-DD-RNYM
150 tr|Q75CW6|Q75CW6_ASHGO         KPYYWWMAGEAFGGMVENWYLC--NNTEYQDLLTDAMVSQIG-EN-KDYV
151 tr|Q7S4K4|Q7S4K4_NEUCR         GDYYWWEGGALMGTMIDYWHLT--GDTTYNDVITQGILHQVG-DN-RDFQ
152 tr|G2QB99|G2QB99_MYCTT         DPYFWWEAGAMFGTLVDYWALT--GDDSYNAITTQAILHQAT-EK-GDFM
153 tr|G0S3F2|G0S3F2_CHATD         GDYYWWEGGAMMGTYVDYWHLT--GDPSYNHVIMEGMLHQVG-PN-ADYQ
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO PPYYWWEAGGAWGSLIDYSYYM--ENNTYHDVIQQAMEYQTG-DD-YNYI
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB PPYYWWQAGVAFGGMLENWFLC--QNDTYKDLLMNALVAQTG-PN-YDYI
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM T--KWWEGSALFMAMLQYQYFT--GDHTYNSDVSLGLQWQAG--D-GDYM
157 tr|G2QKJ0|G2QKJ0_MYCTT         GPYYWWQAGAMWGTYIDYWFYT--GDDTYNAETTRSLLFQAE-PPANAYM
158 tr|Q4WFX5|Q4WFX5_ASPFU         APYYWWEAGAFFGTLVNYWRYT--GDDTYNNITMQAILHQAG-T--GDFM
159 tr|Q5BGD7|Q5BGD7_EMENI         D--TWYVAGAMFMTMIQFWHVS--GYNEHNSVVSHDLMFQAG-RN-YDYF
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH SPNYWWQAGGAWGSILDYWFYM--QNDTYNDMLKQALLAQTG-DN-YDYM
161 tr|G8BYN9|G8BYN9_TETPH         PPYYWWEAGEAFGGMLENWYLC--QNDTFESLISDALIYQAG-SN-YDYI
162 tr|F7VVT4|F7VVT4_SORMK         GPYYWWEAGAMWGTIIDYWHYT--GDDTYNTLAKQALVFQAD-PPQNSFM
163 tr|G8YB80|G8YB80_PICSO         WPYYWWEAGGAWGSLVDYIYYM--ENDTLTPLVKQALTYQTG-DD-DNYI
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA YPYYWWQAGHAFGALIDTWYLC--GNDTYEQLIHDALMHQTG-ED-NDYM
165 tr|G2WRS7|G2WRS7_VERDV         GDYYWWEAGAMWGALIDYWHLT--GDDTYNGIITQAMLWQVG-EG-RDYM
166 tr|A0A124BVB5|A0A124BVB5_ASPNG DPYYWWEAGAFFGTLVNYWAYT--GDTTYNELTAQAILHQVG-PN-RDFM
167 tr|A7EIG7|A7EIG7_SCLS1         APYYWWEAGAMFGSLIDYWYYT--GDSSYNDVTTEAMLWQTG-PD-NDYM
168 tr|C5M5U7|C5M5U7_CANTT         NPYYWWEAGGAWGSLIDYTFFF--DNDTLVPLIKEALLYQTG-DD-DNYI
169 tr|B9WAA8|B9WAA8_CANDC         PPNYWWNAGEAFGGLVDFYTFCESDNSTLESLIYNGMYHQAG-EN-YNYI
170 tr|F8MRU1|F8MRU1_NEUT8         GDYYWWEGGALMGTMIDYWHLT--GDTTYNDVITQGILHQVG-DN-RDFQ
171 tr|A1DJS0|A1DJS0_NEOFI         T--KWWEGSALFMAMLQYQYFT--GDHTYNSEVSLGLQWQAG--E-GDYM
172 tr|B6HLM8|B6HLM8_PENRW         R--SWWTGAALFLACLNYWHAT--NDTTYNEDVSIGLQHQGG-AG-GDYM
173 tr|A0A117DWS0|A0A117DWS0_ASPNG D--TWWEGGAMFMTLIQYWYWT--GDSSYNEVTTQGMLWQKG-DN--DYF
174 tr|A0A124BXM1|A0A124BXM1_ASPNG S--KWWEGSALFTSLLLYWYYT--GDSTYNDEVRQGMQWQAG--D-CDYM
175 tr|J3NZQ6|J3NZQ6_GAGT3         GPYYWWQAGAVWGAMIDYWHYT--GDSTYNKIVTEGIMNQIG-ED-QDFE
176 tr|W0THH8|W0THH8_KLUMD         QPYYWWEAGLVFGGMIENWYLC--QNNSYQNLLYDAMIAQTG-SN-YDFM
177 tr|B2AEF2|B2AEF2_PODAN         GDYYWWEGGAMMGTYVDYWFLT--GDESYNKVVTEGMIHQVG-PN-EDYM
178 tr|Q96TX1|Q96TX1_NEUCS         GDYYWWEAGAMWGALIDYWRFT--GDASYNDVITQALLWQVG-PG-QDYM
179 tr|E9DYA6|E9DYA6_METAQ         GDYYWWEGGAMWGTLIQYWAFT--GDSTYNDNIMQAMQFQVG-EG-QDYM
180 tr|B9WAI2|B9WAI2_CANDC         WPYYWWEAGGAWGSLIDYTFYF--DNDTLVPLITEALLYQTG-DD-DNYI
181 tr|Q6C171|Q6C171_YARLI         APYYWWHAGAAWNAFIDYWHIT--GNDTYNNLTWAAMKHQRG-EK-YDLM
182 tr|A6ZMV1|A6ZMV1_YEAS7         SPYYWWHAGEAFGGMLENWFLC--ENDTYQELLYDALLAQTG-SN-YDYI
183 tr|B6K7X7|B6K7X7_SCHJY         LTVYWWMTGAGMNALLNNYYLT--GNSTYNDLVTKVLLSNTG-DE-NDFA
184 tr|I1RSA2|I1RSA2_GIBZE         DPYYWWEAGAMFGTLVDYWWIT--GDDTYNAITKQAMIHQAG-SD-GDYM
185 tr|F7VWZ9|F7VWZ9_SORMK         DPYYWWEAGAMFGALVDYWWIT--GDTSYVEATSQAIVHQAS-ET-RDFM
186 tr|A0A1B2J789|A0A1B2J789_PICPA NPYYWWQSGAAFGSMLDYWWYM--ENDTYHDAIMQAIIYQAG-DN-ADFM
187 tr|G0V8N4|G0V8N4_NAUCC         DPYYWWEAGGAWGTMLDYWYFM--QNDTYNDIITAALLHQVG-TD-NNYI
188 tr|H8WXN8|H8WXN8_CANO9         SPYYWWEAGGAWGGMIDYTFYF--NNDTLVPLIKQALEYQVG-DD-DNYI
189 tr|A6ZZS0|A6ZZS0_YEAS7         DPYYWWEAGGAWGCMLDYWFFM--DNDTYNDEIIAAMIHQAG-DD-NDYI
190 tr|G9MQD1|G9MQD1_HYPVG         GDYYWWEGGAMMGTYVDYWHLT--GDTSYNAVVMEGMLHQVG-DH-ENYM
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH NPYYWWEAGGAWGSILDYWWYT--GNDTFNGILKQALLAQTG-EN-NDYV
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM G--TWYVAGAMFMTLIQYWQSS--GDDTYNSIVSHDLMFQSG-EN-YDFF
193 tr|A0A254UET1|A0A254UET1_ASPNG DPYYWWEAGAFFGTLVNYWAYT--GDTTYNDLTAQAILHQVG-PN-RDFM
194 tr|A1D587|A1D587_NEOFI          D--TWWEGGAMFMTLIQYWYWT--GDTSYNEVTTQGMLFQKG-NN--DYF
195 tr|Q1K7I4|Q1K7I4_NEUCR         GDYYWWEAGAMWGALIDYWRFT--GDASYNDVITQALLWQVG-PG-QDYM
196 tr|G8JS88|G8JS88_ERECY         SPYYWWEAGAAWGALIDYWFYM--ENDTYNDIIKQALLYQVG-KY-EDYV
197 tr|A0A136J7D3|A0A136J7D3_9PEZI DPYFWWEAGAMFGTLVDYWFHT--GDDTYNAITKQALLHQVG-KD-RDYM
198 tr|E9E4X8|E9E4X8_METAQ         GDYYWWEGGAMWGTLIDYWYWT--GDSTYNDEVMQSMLFQVG-EN-MDYM
199 tr|Q6CP42|Q6CP42_KLULA         SPSYWWQAGGAWGSILDYWFYM--ENDTYNDILTQALLYQTG-DN-HDYM
200 tr|G0WHL4|G0WHL4_NAUDC         PPYYWWEAGEAFGGMIENWYLC--QNDTYEDMLKSALLAQTG-SN-YDYI
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU DPYYWWEAGGAWGCMLDYWFYM--ENDTYNDEITAAMIHQAG-DD-NDYI
202 tr|F9WYX6|F9WYX6_ZYMTI         QPYYWYNAGNMFNSLIKYWALS--GDQSIVPTLQSALVFQLG-PD-FNYM
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 GDYYWWEGGALMGQMIEYWHLT--GDTTYNDQVMQGMLHQVG-QD-RDYQ
204 tr|J3P147|J3P147_GAGT3         GPYYWWQAGAMWAAMIDYWHLT--GDTTYNKLVTDGILAQHG-PD-RDFQ
205 tr|I2H842|I2H842_TETBL         DPYYWWEAGGAWGCMLDYWYYM--QNDTYNDVIYEALQYQTG-DD-NDYI
206 tr|S6E3R3|S6E3R3_ZYGB2         SPYYWWQAGEAFGGMLDNWYFC--QNNTFENLIRKAMLAQTG-TK-YDYM
207 tr|E5AD94|E5AD94_LEPMJ         PPYYWWQAGAMFGTLLDYWHYT--GDDQYNDMIREGLIHQFG-EH-LDLM
208 tr|Q6FJM7|Q6FJM7_CANGA         PPYYWWQAGVAFGGMLENWFLC--QNDTYKDLLMNALVAQTG-PN-YDYI
209 tr|A5DUW0|A5DUW0_LODEL         APYYWWEAGGAWGSLIDYTFFF--ENDTYVPLIKQALLYQVG-DD-NNYI
210 tr|G8BF46|G8BF46_CANPC         SPYYWWEAGGAWGGMIDYTFYF--QNDTLVPLIKQALEYQVG-DD-NNYI
211 tr|W3XAF6|W3XAF6_PESFW         DPYYWWQCGAMFGSLLDYWSYT--GDDSYNNITYQAMWHQRG-DD-YDYM
212 tr|B6H7I2|B6H7I2_PENRW         E--KWWEGSALFLALLQYWHFT--GDTTYNSLMSEGMQWQSG-EE-GDYM
213 tr|Q5ATF9|Q5ATF9_EMENI         E--KWWEGSALFMSLMLYWYYT--GDSQYNSLITQGMQHQAG--N-GDYL
214 tr|A0A167CDD3|A0A167CDD3_9ASCO NPYYWWEAGVAWGTMLDYWYYT--QNDTYVDIIKSSMLFQTG-KD-WNYM
215 tr|G0RXS3|G0RXS3_CHATD         GPYYWWQAGAMWGTLIDYWYYT--GDDTYVNETRRSLVFQAE-GPGYSFE
```

```
216 tr|B8NX26|B8NX26_ASPFN          D--TWYIAGSMFMTLIQYWQAS--GDDSQNAVVSHDLMFQAG-EN-YDFF
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM DPYYWWEAGAMFSALIDYWYYT--GDDQYNDITTQAMLHQVG-KE-NNYM
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM APYYWWEAGAFFGTLVNYWRYT--GDDAYNNITMQAILHQAG-T--GDFM
219 tr|A1CMM3|A1CMM3_ASPCL         DPYYWWEAGAMFGSLVEYWYYT--GDSQWNEIITQAILHQTG-PS-FNFM
220 tr|B6HU84|B6HU84_PENRW        G--TWWEGGSMFMTLILYWYVT--GDSQFNDAIQEGMYFQKG-DD--NFF
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS GDYYWWEAGAMWGALIDYWRFT--GDASYNDVITQALLWQVG-PG-QDYM
222 tr|G2WHY6|G2WHY6_YEASK         DPYYWWEAGGAWGCMLDYWFFM--DNDTYNDEIIAAMIHQAG-DD-NDYI
223 tr|G0W7G8|G0W7G8_NAUDC        DPYYWWEAGGAWGSMLDYWYYM--DNDTYNDEIMAALLHQTG-DN-NDYV
224 tr|G4NGV8|G4NGV8_MAGO7        DPYFWWEAGAMFGTMVDYWLLT--GDATYNPVTTQALLHQVG-DN-KDYM
225 tr|W3X855|W3X855_PESFW        DPYYWWEAGAYFGTLVDYWAFT--GDEQYVDLTFQALQHQVG-DD-GDFM
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA GPYYWWIGGAMWGTLIDYWHLT--GDTTYNDRIMQAMQFQVG-EN-KAYM
227 tr|J8Q2K7|J8Q2K7_SACAR         DPYYWWEAGGAWGCMLDYWYFM--DNDSYNDEIIAAMIHQTG-DD-NDYI
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO WPYYWWEAGGAWGSLIDYTYYM--KNDTLVPLIKQALEYQVG-DD-YNYI
229 tr|J6EGN8|J6EGN8_SACK1         SPYYWWHAGEAFGGMLENWFLC--ENETYQELLYDALLAQTG-SN-YDYI
230 tr|A0A100ISD9|A0A100ISD9_ASPNG DPYYWWEAGAMFTALVEYWYLT--SDDTWNNITTQGITWQAG-PS-GSFM
231 tr|G8YQC0|G8YQC0_PICSO         YPYYWWHAGCAFGGLLDYFTFCDKDNSTLKEIIYDGMYHQAG-DN-YDYI
232 tr|F7VVP8|F7VVP8_SORMK        GDYYWWEGGALMGTVVDYWHLT--GDTTYNDVVIQGILHQVG-DR-RDFQ
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH SPYYWWHAGEAFGGMLENWFLC--ENDTYQELLYDALLAQTG-SN-YDYI
234 tr|F8MBP4|F8MBP4_NEUT8         DPYYWWEAGAMFGAMVDYWWLT--GDTSYVEVTTQAIVHQAG-DA-RDFN
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG H--TWWEGGVLFDTMIRYWYFT--DDASNNAAVSQGMYHQRG-AG-NDYM
236 tr|A1CD27|A1CD27_ASPCL         G--TWYVAGAMFMTLIQYWQAS--GDDTYNSIVSHDLMFQSG-EN-YDFF
237 tr|A5DHF3|A5DHF3_PICGU        SPNYWWNAGEAFGGLLDYFTFCDSGNKSLEKLIFNGMYHQAG-DQ-YNYI
238 tr|D4ATT7|D4ATT7_ARTBC        DPYYWWEAGAMFGALIDYWWYT--GDDAYNKITMQAMLHQVG-TE-KNYE
239 tr|E2PT42|E2PT42_ASPNC        DPYYWWEAGAMFTALVDYWYLT--SDDTWNNITTQGITWQAG-PS-GSFM
240 tr|C7ZPE5|C7ZPE5_NECH7        GDYYWWEGGAMWGTYMDYWKCT--GDATYNAKVIQAMQFQTG-DD-NDYQ
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA SPYYWWEAGALFGAFIDYWFYT--GDDTYNDITMQAMIHQAS-ED-RDFM
242 tr|Q4WKP7|Q4WKP7_ASPFU        D--TWWEGGAMFMTLIQYWFWT--GDTSYNEVTTQGMLWQKG-NN--DYF
243 tr|F8MXT4|F8MXT4_NEUT8        GDYYWWEAGAMWGALIDYWRFT--GDASYNDVITQALLWQVG-PG-QDYM
244 tr|A0A177AJ74|A0A177AJ74_9PEZI SGYFWWQAGAMWNTMLDYWHIT--GDNSYNEAISQALVFQVG-ND-RDYM
245 tr|J7RQR5|J7RQR5_KAZNA         PPYYWWHAGEAFGGMIENWFLC--QNSTYETLLMDAFIAQTG-NN-YDYI
246 tr|Q75CW7|Q75CW7_ASHGO        KPYYWWMAGEAFGGMVENWYLC--NNTEYQDLLTDAMVSQIG-EN-KDYM
247 tr|G3AKK4|G3AKK4_SPAPN        PPNYWWNAGEAFGGLVDYYTFCDPNNDTLKDLIFNGMYHQAG-EN-YNYV
248 tr|B2W762|B2W762_PYRTR        DPYYWWECGAMFNAFIDYWYYT--GDDTYNGITTQALEAQIG-DY-NAFM
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH DPYYWWEAGGAWGCMLDYWFFM--DNDTYNDEIIAAMIHQAG-DD-NDYI
250 tr|C4XWK1|C4XWK1_CLAL4         PPYYWWNAGEVFGGWVDYWAFCAQDNDTFTELLSNAMYHQAG-DD-FNYM

001 tr|A0A254U2J9|A0A254U2J9_ASPNG -PANYSSYLGNDDQMFWGLAAITAAEIEFADS---SSGYS-WLALAQGV
002 sp|Q6FLP9|DCW1_CANGA          -PLNQSTTEGNDDQAFWGIAVMQAAERKFPNPP---DDKPQ-WLYLTQAV
003 tr|A1CCM5|A1CCM5_ASPCL       -PSNYSSYLGNDDQMFWGLAAMLAAETKFPDR----ANGYS-WLSLAQGV
004 tr|C1H190|C1H190_PARBA       -PHNQSRTSGNDDQAFWGMAAMTAAETKFQDPAE---DKPQ-WLALAQGV
005 tr|Q2UR85|Q2UR85_ASPOR       -PANYSNYLGNDDQVFWGLAAMTAAELNFPEK----DDDSS-WLSLAQGV
006 tr|A5DV30|A5DV30_LODEL       -PSNQSMTEGNDDQGVWGMAIMEAVERNFTNPD---L--HS-WLEMVQAI
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO GVPNQTHVEANDDQVFWGFTVLEAAERDFPAYSN-NSGDPS-YADLALNV
008 tr|J3PGN0|J3PGN0_GAGT3        -PAKHRASLGNDDQGFWGMTAMLAAENKFPDPPK---DKPQ-WLALAQAV
009 tr|A0A124BXU6|A0A124BXU6_ASPNG -PANWSSYIGNDDQMFWGLAAMTGAELNFPDS----SDGYS-WLSLAQGV
010 tr|C5P4A1|C5P4A1_COCP7        -PLNQTSTLGNDDQAFWGLAALAAAERNFPNPPA---HEPQ-WLALAQGV
011 tr|Q0CG55|Q0CG55_ASPTN       -PTNQTRTEGNDDQAFWAFAAMSAAERNFPNPPE---TSPG-WLAMAQAV
012 tr|C4Y2X5|C4Y2X5_CLAL4       -PLNQSTTEGNDDQAFWGLAVMSAAEKNFSNPT---DHRAA-WLTLAQAV
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG -PANQTRTEGNDDQSFWAFAAMSAAERNFPDPPP-SSGSPG-WLAMAQAV
014 tr|C7Z068|C7Z068_NECH7        -PENHTASLGNDDQGFWGMSAMLAAENKFPNPPE---DKAQ-WLALAQAV
015 tr|G8BXX2|G8BXX2_TETPH       -PLNQSTTEGNDDQAFWGIAAMMAAERNFTNPA---EDEPQ-WLYLAQAV
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR -PLNQSTTEGNDDQAFWGFAVMAAAEHNFTNPE---DPKLA-WLALAQAV
017 tr|A1CSC2|A1CSC2_ASPCL       -PSNYSNYLGNDDQVFWGLAAMTAAELNYPEQ----NGQPS-WVSLAQGV
018 sp|Q05031|DFG5_YEAST         -PSNQTMVEGNDDQGIWGITVMGAVERNFTDPG---DGKPG-WLAMVQAV
019 tr|H2AQJ6|H2AQJ6_KAZAF       -PENQTMVEGNDDQCVWGLTIMDAVERNFTDPT---DGSPG-WLYMTQAI
020 tr|G2XFH9|G2XFH9_VERDV       -PLNHTASLGNDDQGFWGMTAMLAAENKFPDPPE---DQPQ-WLALAQAV
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH -PLNQSTTEGNDDQAFWGIAVLNAAERKFPNPP---SDEPQ-WLYLAQAV
022 sp|Q5AD78|DCW1_CANAL          -PLNQSTTEGNDDQAFWGIAVMAAAERNFTNPK---DPTKA-WLTLAQAV
023 tr|B6H7E8|B6H7E8_PENRW       -PSNYSQYLGNDDQVFWGLAAITAGELNFPER----SGEPS-WISLAEGV
024 tr|B6GZU2|B6GZU2_PENRW       -PANQSKTEGNDDQVFWAFAAMTAAERNFPNPPE---DKPQ-WLEMAQAV
025 tr|E9E4W7|E9E4W7_METAQ       -PPNQTLSLGNDDQGFWGLSALLAAENKFPDPPP---DQPQ-WLELAQAV
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO -PNNQSKTEGNDDQAFWGFAAMSAAERNFPNPPD---GKPQ-WLEMAQAV
027 tr|G2R7G8|G2R7G8_THITE        -PANWTLSLGNDDQGFWGMAAMLAAETNFQNPPD---GQTP-WLTLAQAV
028 tr|G3JBY1|G3JBY1_CORMM       -PRNQSNAMGNDDQGFWAMTTMMAAEAAFPNPPS---DTPQ-WLAGGQAV
029 tr|A5DNV5|A5DNV5_PICGU       -PLNQSTTEGNDDQGFWGIAVMAAAEKNMSNPK---DERKA-YLALAQAV
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI -PPNQTKTLGNDDQCFWALAAISAAESNFPNPPK---DKPQ-WLALAQAV
031 tr|Q2TWC1|Q2TWC1_ASPOR       -SSNYSQWLGNDDQMFWGLASITASETGFPEV----SGKPT-WTSLARTV
032 tr|J5RXY9|J5RXY9_SACK1       -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---ADEPQ-WLYLAQAV
```

```
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA  -PVNQTNAMGNDDQGFWALTSMMAAEAVFPNPPS---DTPQ-WLAGVQAV
034 tr|A0A254U5V1|A0A254U5V1_ASPNG  -PANDSNYLGNDDQVFWGLAAMTAAELNYPEQ----DGQPS-WLSLAQGV
035 tr|A0A061B6H7|A0A061B6H7_CYBFA  -PLNQSTTEGNDDQGFWGIAVMAAAERDFPNPP---EDGPQ-WVALAQAV
036 tr|G8ZQ93|G8ZQ93_TORDC          -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---SDQPQ-WLYLAQAV
037 tr|Q9C2J1|Q9C2J1_NEUCS         -PANQSRTSSNDDVGFWTITAMMAAEDAFPDPPP---DQPQ-WLALVQAV
038 tr|C5NZK5|C5NZK5_COCP7          -PLNQTSTLGNDDQAFWGMAAMSAAENKFPNPPD---DKPQ-WLELAQAV
039 tr|Q6C0T7|Q6C0T7_YARLI         -VFNFSSSEGNDDQGVWGMTMMTAAERNFTAPT----DEYG-WLYVAQGA
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB -PLNQSTTEGNDDQAFWGIAVMQAAERKFPNPP---DDKPQ-WLYLAQAV
041 tr|F7VZ72|F7VZ72_SORMK          -PPNVTASLGNDDQGFWGMTAMTAAENNFPDPPP---DQPQ-WLGLAQAV
042 tr|A3LN37|A3LN37_PICST          -PLNQSTTEGNDDQGFWGIAVMGAAERNFSNPD---DDSKA-WLTLAQAV
043 tr|Q6CAI2|Q6CAI2_YARLI         -PVNQSTTEGNDDQGFWGITAMAAAEKNFSNPG---KDEPQ-WLYLAQAV
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO -PLNQSTTEGNDDQGFWGIAVMAAAEKNFSNPN---DEKKQ-WLALTQAV
045 tr|G9MJS5|G9MJS5_HYPVG          -PPNVTLSLGNDDQGFWGMTAMLAAEEKFPNPPA---DKPG-WLALAQGV
046 tr|G9MHI4|G9MHI4_HYPVG          -PTNQSLTEGNDDQGFWAMAAMSAAENVFPNPPS---DQPQ-WLAMAQAV
047 tr|D4AP27|D4AP27_ARTBC          -PPNQTLTEGNDDQAFWAIAAMSAAERKFPNPPS---DKPQ-WLSLVEAV
048 tr|B6GZT8|B6GZT8_PENRW          -PGNYSSYLGNDDQMFWGLAAMTAAEFGFPDR----STGFS-WLSLAQGV
049 sp|O74556|YCZ2_SCHPO            -PSNYTTSEGNDDQAFWGLTVISAAEANFSNPAA---DEPQ-WLELAQAV
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 -PAKHRASLGNDDQGFWGMTAMLAAENKFPDPPK---DQPQ-WLALAQAV
051 tr|B8MHG0|B8MHG0_TALSN          -PSNYSSYLGNDDQEFWALAALTAIEQKFPDN----PGHAS-WLSLAQGA
052 tr|A3LMV8|A3LMV8_PICST          -PSNQSMTEGNDDQGVWVMAIMQAVERNFTNPA---D--HS-WLSMTIAA
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU NVENQTDVEANDDQVFWGFTVMEAAERGFPMYSN-DENDVN-YAQLAMNV
054 sp|Q9P6I3|YHG7_SCHPO            -PNSEKFDLGNDDQGIWGLSAMSAAEVNMTTGD----SSAS-FTELAQAV
055 tr|Q2US57|Q2US57_ASPOR          -PSNWSSYLGNDDQMFWGLAAITAAELQFPEV----SDGYS-WLSLAQGV
056 tr|W7N622|W7N622_GIBM7          -PDNQTMTEGNDDQGFWAVAAMSAAEHKYPDPPA---DQPQ-WLALVQAV
057 sp|Q9P6I4|YHG6_SCHPO            -PSFEYFNLGNDDQGMWAAAAMDAAEANFSPPN---STEHS-WLELTQAA
058 tr|C7GRB3|C7GRB3_YEAS2          -PSNQTMVEGNDDQGIWGITVMGAVERNFTDPG---DGKPG-WLAMVQAV
059 tr|C7GP28|C7GP28_YEAS2          -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---ENEPQ-WLYLAQAV
060 tr|G3JGZ5|G3JGZ5_CORMM          -PRNHTASLGNDDQGFWGMTAMLAAENKFPDPPA---DKPQ-WLALAQAV
061 tr|C4QXV4|C4QXV4_KOMPG          -PLNQTTTEGNDDQGFWGITAMAAVERNFTNPP---EDEPQ-WLYLVQAT
062 tr|I1S0Z5|I1S0Z5_GIBZE          -PSNVTLSLGNDDQGFWGMSAMLAAELAFPNPPS---DQPG-WLALAQAV
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM -PTNQTKTEGNDDQAFWAFAAMSAAERNFPDPDP-D-HGPG-WLAMAQAV
064 tr|G0W5X4|G0W5X4_NAUDC          -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---ESSPQ-WLYLAQAV
065 tr|G8ZPC1|G8ZPC1_TORDC          -PENQTMVEGNDDQGVWGLTLMGAVERNFTNPT---NGAPG-WLAMAQAI
066 tr|A0A167CDC5|A0A167CDC5_9ASCO -PSNQTTTEGNDDQGFWGILVMDAAEKNFSNPD---PDQAQ-WLSLAQAV
067 tr|R9XAT7|R9XAT7_ASHAC          -PWNETTTEGNDDQAFWGMAVMAAAERNYPTPP---PDQPQ-WLALAQAV
068 tr|B2B747|B2B747_PODAN          -PPNVTASLGNDDQGFWGMSAMLAAENNFPDPPE---DEPQ-WLALAQAV
069 tr|G0SE99|G0SE99_CHATD          -PPNVTASLGNDDQGFWGMSAMLAAENGFPDPPE---DEPQ-WLALAQAV
070 tr|J8Q6F0|J8Q6F0_SACAR          -PSNQTMVEGNDDQGIWGITVMGAVERNFTDPT---NGAPG-WLAMVQAV
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH -PSNQTMVEGNDDQGVWALTVMSAVERNFTNPK---DGAPG-WLAMAQAV
072 tr|J3K6E2|J3K6E2_COCIM          -PLNQTSTLGNDDQAFWGLAALAAAERNFPNPPA---HEPQ-WLALAQGV
073 tr|G4N3E1|G4N3E1_MAGO7          -PLSKRASLGNDDQGFWGMTAMLAAETNFPNPPK---DKPQ-WLALAQAV
074 tr|C5M5J2|C5M5J2_CANTT          -PSNQSMTEGNDDQGVWGMAIMQAVERNFTEPE---S--HS-WLEMTQAV
075 tr|A0A254U0X0|A0A254U0X0_ASPNG -SSNYSQWLGNDDQMFWGLAAITASETGFPEV----SGKPS-WTSLARAV
076 sp|Q75DG6|DCW1_ASHGO            -PWNETTTEGNDDQFFWGMAVMAAAERNFPNPP---ADQPQ-WLALAQAV
077 tr|Q6BZF0|Q6BZF0_DEBHA          -PLNQSTTEGNDDQGFWGLAVMAATEKNFSNPD---NKDKH-WLALTQAV
078 tr|E9E3Q1|E9E3Q1_METAQ          -PNNHTMSLGNDDQGFWGLSAMLAAENKFPDPPA---DQPQ-WLALAQAV
079 sp|P36091|DCW1_YEAST            -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---ENEPQ-WLYLAQAV
080 tr|W0TAH6|W0TAH6_KLUMD          -PLNQTVTEGNDDQAFWGFVVMAAAERNYPNPP---KDQPQ-WLYLAQAV
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA -PSNQSMTEGNDDQGFWGLAVMEATERNFTNPP---SDQPG-WLALTQAV
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI -PENQTRSLGNDDQGFWALAAMSAAESGFENPPS---DKPQ-WLALVQAV
083 tr|Q2TYU3|Q2TYU3_ASPOR          -PANYSSYLGNDDQGFWGIAAMTAAEIGFPDV----DDGYS-WLSLAQGV
084 tr|G2XGF6|G2XGF6_VERDV          -PLNQTRTLGNDDQGFWTLSAMMATEAGYPDPPA---DQPQ-WLALVQAC
085 tr|G3AMT4|G3AMT4_SPAPN          -PLNQSTTEGNDDQGFWGIAVMGAAERNFSNPE---DPDKA-WLTLAQSV
086 tr|A0A1C1D213|A0A1C1D213_9EURO -PPNQTKTEGNDDQAFWGMAALSAAENVFPNPPS---DQPQ-WLALAQAV
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS -PENYTISMGNDDQGFWGLSAMLAAEIGFPDPPA---DKPQ-WLQLAQAV
088 tr|A0A100I5A6|A0A100I5A6_ASPNG -SSNYSQWLGNDDQMFWGLAAITASETGFPEV----SGKPS-WTSLARAV
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH -PSNQSMVEGNDDQGVWGITVMGAVERNFTNPS--TKGAPG-WLAMVQGV
090 tr|G2Q8A7|G2Q8A7_MYCTT          -PPNYTASLGNDDQAFWGMSAMTAAENNFPNPPA---DQPQ-WLALAQAV
091 tr|B8NPZ1|B8NPZ1_ASPFN          -SKNVSDWLGNDDQMFWGLATITASEAGFPEI----SGKPS-WTSLARVV
092 tr|Q7SAB2|Q7SAB2_NEUCR          -PENYTISMGNDDQGFWGLSAMLAAEIGFPDPPA---DKPQ-WLQLAQAV
093 tr|A2R8R5|A2R8R5_ASPNC          -PANWSSYIGNDDQMFWGLAAMTGAELNFPDS----SDGYS-WLSLAQGV
094 tr|C5DHG0|C5DHG0_LACTC          -PSNQSMVEGNDDQGVWGLTVMSAAERNFTNPT--EDGVPG-WLAMAQGV
095 tr|G0SFA3|G0SFA3_CHATD          -PPNQTRTLGNDDQGFWGMAAMSAAEYNFPNPPQ---DKPQ-WLALAQSL
096 tr|K1WJG5|K1WJG5_MARBU          -PKNQTKSEGNDDQGFWGMTAMSAAEMNYPNPPD---TEPQ-WLALAQAV
097 tr|H2B005|H2B005_KAZAF          -PLNQSTTEGNDDQAFWGIAVMTAAERNFTNPP---DDQPQ-WLYLAQAV
098 tr|C5DYD7|C5DYD7_ZYGRC          -PQNQTMVEGNDDQGVWGLTVMEAVERNFTDPG---DGKPG-WLAMAQAV
099 tr|G8YM23|G8YM23_PICSO          -PLNQSTTEGNDDQGFWGIAVMAAAEKNFSNPS---DPKKA-WLTLAQAV
100 tr|W3X8E3|W3X8E3_PESFW          -PPNVTASLGNDDQGFWGMTAMLAAEENFPDPPK---DQPQ-WLALAQAV
101 tr|A0A254U3K7|A0A254U3K7_ASPNG -PANWSSYIGNDDQMFWGLAAMTGAELNFPDS----SDGYS-WLSLAQGV
102 tr|Q6CER8|Q6CER8_YARLI          -VRNYSSSEGNDDQGVWAMTMMTAAERNFTNPD----DDYG-WLYVAQGA
```

```
103 tr|G2QT05|G2QT05_THITE          -PPNVTASLGNDDQGFWGMSAMLAAENNFPNPPA---DQPQ-WLALAQAV
104 sp|Q5ACZ2|DFG5_CANAL           -PSNQSMTEGNDDQGVWGMAIMEAVERNFTEPE---S--HS-WLEMVQAV
105 tr|G2WKU6|G2WKU6_YEASK          -PSNQTMVEGNDDQGIWGITVMGAVERNFTDPG---DGKPG-WLAMVQAV
106 tr|A0A0C7N752|A0A0C7N752_9SACH -PSNQSMVEGNDDQGVWGLTVMGAVERNFTNPH--DKNAPG-WLAMVQGV
107 tr|W3WN68|W3WN68_PESFW          -PENQTRSLGNDDQGFWAMASMTAAEMNFTNPSS---DEPQ-WLALSQAV
108 tr|G2QFS1|G2QFS1_MYCTT          -PPNHTASLGNDDQGFWGMSAMLAAENKFPNPPE---DKPQ-WLALAQAV
109 tr|G3J9G4|G3J9G4_CORMM          -PVNQTNAMGNDDQGFWALTSMMAAEAVFPNPPS---DTPQ-WLAGVQAV
110 tr|Q2UJ03|Q2UJ03_ASPOR          -PPNQTHAEGNDDQAFWAFAAMSAAERNFPNPPD---GQPG-WLAMAQAV
111 tr|G8YRT2|G8YRT2_PICSO          -PSNQSMTEGNDDQGIWGLTTMEAAERNFTDPP------HS-WLSMTQAI
112 tr|W3X554|W3X554_PESFW          -PQNQTRSEGNDDQGFWAMAAMSAAENKFPDPPA---DQPQ-WLALAQAV
113 tr|Q1K7A8|Q1K7A8_NEUCR          -PANQSRTSSNDDVGFWTITAMMAAEDAFPDPPP---DQPQ-WLALVQAV
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM -PSNYSNYLGNDDQVFWGLAAMTAAELNFPEE----DGQPS-WVSLAQGV
115 tr|A7F3P0|A7F3P0_SCLS1          -PLNQTKTEGNDDQGFWGMAAMTAAETNFTNPDN---GTPG-WVAMAQAV
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA -PRNQSNAMGNDDQGFWAVTSMMATEAAFPNPPS---DTPQ-WLAGAQAV
117 tr|Q6CIP9|Q6CIP9_KLULA          -PSNQTMVEGNDDQGVWALTVMTAVERNFTNPK--EDGTPG-WLAMVQAV
118 tr|H2AMQ5|H2AMQ5_KAZAF          -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPA---SDEPQ-WLYLAQAV
119 tr|G8JMI3|G8JMI3_ERECY          -PSNQSMTEGNDDQGIWGLFIMGAVERNFPDAQ--MGNVPG-WLTMAQAV
120 tr|C5DGT1|C5DGT1_LACTC          -PLNQSTTEGNDDQAFWGLVVMAAAERNFTNPP---ETEPQ-WLYLAQAV
121 tr|K1XJR4|K1XJR4_MARBU          -PRNWSQSMGNDDQGFWGLTAMSAAETNFTNPPS---GSPG-WLALAQAV
122 tr|A0A124BYC5|A0A124BYC5_ASPNG -PSNWSSYMGNDDQVAWGLAAMTAAELNYPED----AGMPS-WVDLAERV
123 tr|A7TLL2|A7TLL2_VANPO          -PLNQSTTEGNDDQAFWGIAVMSAAERNFTNPP---ADRPQ-WLYLAQAV
124 tr|Q0CB86|Q0CB86_ASPTN          -PANYSSYLGNDDQVFWGIAAMTAAEINFPDV----PDKYS-WLSLAQGV
125 tr|K1XLH0|K1XLH0_MARBU          -PLNQTRTLGNDDQCFWGMSAMTAAETGFEDPPE---GQPG-WLALVQGV
126 tr|J7S438|J7S438_KAZNA          -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---EDKPQ-WLYLAQAV
127 tr|Q4WG09|Q4WG09_ASPFU          -SGNYSQWLGNDDQMFWGLASITASETGFPEI----SGKPT-WTSLARVV
128 tr|H8WZ09|H8WZ09_CANO9          -PSNQSTTEGNDDQGVWGMAIMEAVERNFTNPE---S--HS-WLEMAQAI
129 tr|I1RL09|I1RL09_GIBZE          -PSNHSASLGNDDQGFWGMSAMLAAENKFPNPPE---DQAQ-WLALAQAV
130 tr|Q4W985|Q4W985_ASPFU          -PTNQTKTEGNDDQAFWAFAAMSAAERNFPDPDP-D-HGPG-WLAMAQAV
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU -PANQTMVEGNDDQGIWGITVMGAVERNFTDPD---NGAPG-WLAMVQAV
132 tr|W6QCW3|W6QCW3_PENRF          -PANQSKTEGNDDQVFWGFAAMTAAERNFPNPPA---DQPQ-WLEMAQAV
133 tr|W3WMD3|W3WMD3_PESFW          -PANWSASMGNDDQAFWALSSLVAAETGFTDPPE---DQPQ-WLSLAQAV
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH -PLNQSTTEGNDDQAFWGIAVLNAAERKFPNPP---SDEPQ-WLYLAQAV
135 tr|I2GYH1|I2GYH1_TETBL          -PQNQTMVEGNDDQCIWGLTVMDAVERNFTNPT---EG-PG-WLAMGQAV
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH -PDNQTMVEGNDDQGVWGLTVMGAVERNFTDPG---DGKPG-WLAMAQAI
137 tr|Q752P3|Q752P3_ASHGO          -PANQSTTEGNDDQGVWALFIMGAVERNFKDPE--KEGMPG-WLTMVQTV
138 tr|B8MYP3|B8MYP3_ASPFN          -PANYSNYLGNDDQVFWGLAAMTAAELNFPEK----DDDSS-WLSLAQGV
139 tr|Q2TXL6|Q2TXL6_ASPOR          -PDNYTQDLGNDDQVFWGLAAMTAAELNFPED-----EEVS-WLALAQGV
140 tr|Q0CN17|Q0CN17_ASPTN          -PTNYSQWLGNDDQVFWGLAAMTAAELNFPEQ----DGQPS-WLSLAQGV
141 tr|J3NHD5|J3NHD5_GAGT3          -PSNYTASLGNDDQGFWGMSAMTAAELNFPNPPK---DQPQ-WLALAQAV
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI -PTNQTKALGNDDQGMWGMAAMTAAEYNFPNPPK---GQPQ-WLALAQAV
143 tr|A1CVB0|A1CVB0_NEOFI          -PTNQTKTEGNDDQAFWAFAAMSAAERNFPDPDP-D-HGPG-WLAMAQAV
144 tr|B8NVI3|B8NVI3_ASPFN          -PDNYTQDLGNDDQVFWGLAAMTAAELNFPED-----EEVS-WLALAQGV
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH -PLNQSVTEGNDDQAFWGLVVMAAAERNFENPA---ENEPQ-WLYLAQAV
146 tr|A1DJ54|A1DJ54_NEOFI          -SGNYSQWLGNDDQMFWGLASITASETGFPEI----SDKPT-WTSLARVV
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS -PLNFTASLGNDDQGFWGMTAMLAAENKFPNPPA---DQPQ-WLALAQAV
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA -PLNQSTTEGNDDQGFWGIAVMAAAERNFPNPP---DDQPQ-WLYLAQAV
149 tr|A0A177A618|A0A177A618_9PEZI -PINASKSMGNDDQGFWGMTAMMAAERNFQNPPK---DQPQ-WLALAQAV
150 tr|Q75CW6|Q75CW6_ASHGO          -PSHQKMIEGNDDQGIWGLLVMGAAERNFPDAP--PERAPG-WLALAQAV
151 tr|Q7S4K4|Q7S4K4_NEUCR          -PLNFTASLGNDDQGFWGMTAMLAAENKFPNPPA---DQPQ-WLALAQAV
152 tr|G2QB99|G2QB99_MYCTT          -PPNQTRTLGNDDQGFWGMAAMSAAENNFPNPPE---DQPQ-WLALAQSL
153 tr|G0S3F2|G0S3F2_CHATD          -PPNHTASLGNDDQGFWGMSAMLAAENKFPNPPD---DKPQ-WLALAQAV
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO -PLNESTTEGNDDQGFWGVAVIAAAEKNFTNPK---DPKKQ-WLALAQAV
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB -PANQTTVEGNDDQGVWGLTILDAVERNFSAPI---DGKPG-WLAMSQGI
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM -PSNYSSYLGNDDQMFWGLAAMLAAELQFPDR----SEGYS-WLSLAQGV
157 tr|G2QKJ0|G2QKJ0_MYCTT          -PKNWTASLGNDDQGFWGMAAMLAAEVNFPNPPK---DQPQ-WLALAQAV
158 tr|Q4WFX5|Q4WFX5_ASPFU          -PSNQTRTEGNDDQAFWAFTALMAAGHNFPNPPE---DSPS-WLAMAQAV
159 tr|Q5BGD7|Q5BGD7_EMENI          -DSNYSQWLGNDDQMFWGLATITASETGFPEV----SNKPT-WTSLARAV
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH -PLNQTTTEGNDDQAFWGFAVMAAAERNFTNPT---DEEPQ-WLYLAQAV
161 tr|G8BYN9|G8BYN9_TETPH          -PSNQTLVEGNDDQGVWGLTVLGAVERNLP--D--RDGVPG-WLAISQAV
162 tr|F7VVT4|F7VVT4_SORMK          -SPNWTASLGNDDQGFWGLSAMLAAEVNLPLA-P---GEAD-WLEMAQAV
163 tr|G8YB80|G8YB80_PICSO          -PLNQSTTEGNDDQGFWGIAVMAAAEKNFSNPS---DSKKA-WLTLTQAV
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA -PKNQTMTEGNDDQGFWGLTVMAAAERNFTAP----QDGPD-WLALTQAV
165 tr|G2WRS7|G2WRS7_VERDV          -PSNVTMSLGNDDQGFWGMSAMLAAEVGFPDPPE---DQPQ-WLSLAQAV
166 tr|A0A124BVB5|A0A124BVB5_ASPNG -PANQTRTEGNDDQAFWAFAALRAAERNFPNPPA---DQPS-WLALAQAV
167 tr|A7EIG7|A7EIG7_SCLS1          -TPNQTKTEGNDDQGFWGLAAMSAAEVNFPNPPS---NKPQ-WLALAQAV
168 tr|C5M5U7|C5M5U7_CANTT          -PLNQSTTEGNDDQAFWGIAVMAAAERNFSNPE---DDSKA-WLTLAQAV
169 tr|B9WAA8|B9WAA8_CANDC          -PSNQSMTEGNDDQGVWGMAIMQAVERNFTEPE---S--HS-WLAMVQAV
170 tr|F8MRU1|F8MRU1_NEUT8          -PLNFTASLGNDDQGFWGMTAMLAAENKFPNPPA---DQPQ-WLALAQAV
171 tr|A1DJS0|A1DJS0_NEOFI          -PSNYSSYLGNDDQMFWGLAAMLAAELQFPDR----SEGYS-WLSLAQGV
```

```
172 tr|B6HLM8|B6HLM8_PENRW          -PS-WAIGVGNDDQMFWGLAAITAAEYKFPNR---PSGDT-WLTLAIGV
173 tr|A0A117DWS0|A0A117DWS0_ASPNG -PANDSNYLGNDDQVFWGLAAMTAAELNYPEQ---DGQPS-WLSLAQGV
174 tr|A0A124BXM1|A0A124BXM1_ASPNG -PANYSSYLGNDDQMFWGLAAITAAEIEFADS---SSGYS-WLALAQGV
175 tr|J3NZQ6|J3NZQ6_GAGT3         -PRNHRASLGNDDQGFWGMTAMLAAENKFPDAPGSGSTVPG-WLGLAQAV
176 tr|W0THH8|W0THH8_KLUMD         -PNNQTMVEGNDDQGVWALTIMSAVERNFTNPK--KAGAPD-WLAMVQAV
177 tr|B2AEF2|B2AEF2_PODAN         -PPNHTASLGNDDQGFWGMSAMLAAENKFPNPPE---DQPQ-WLALAQAV
178 tr|Q96TX1|Q96TX1_NEUCS         -PLNVTASLGNDDQGFWGMTAMTAAENNFPNPPA---DQPQ-WLGLAQAV
179 tr|E9DYA6|E9DYA6_METAQ         -PRNVTASLGNDDQAFWGMAAMLAAEIKFPDPPA---DRPG-WLGLAQAV
180 tr|B9WAI2|B9WAI2_CANDC         -PLNQSTTEGNDDQAFWGIAVMAAAERNFTNPK---DSTKA-WLTLAQAV
181 tr|Q6C171|Q6C171_YARLI         -VSNFSSSEGNDDQGVWAMTMMTAAEKNFPEPG----DEYG-WLYVAQGA
182 tr|A6ZMV1|A6ZMV1_YEAS7         -PSNQTMVEGNDDQGIWGITVMGAVERNFTDPG---DGKPG-WLAMVQAV
183 tr|B6K7X7|B6K7X7_SCHJY         -PSSQTLDLGNDDQAFWALTAMSAAELNFTESSS---QSSS-WLEMAQAV
184 tr|I1RSA2|I1RSA2_GIBZE         -PDNQTMTEGNDDQGFWAMAAMSAAEHQFPNPPE---DSPG-WLAQAQAV
185 tr|F7VWZ9|F7VWZ9_SORMK         -PANQSRTSSNDDVGFWTLTAMSAAEDAFPDPPE---DQPQ-WLALVQAV
186 tr|A0A1B2J789|A0A1B2J789_PICPA -PLNQTTTEGNDDQGFWGITAMAAVERNFTNPP---EDEPQ-WLYLAQAT
187 tr|G0V8N4|G0V8N4_NAUCC         -PLNQSTTEGNDDQAFWGITVMTAAERNFTNPP---EDQPQ-WLYLAQAV
188 tr|H8WXN8|H8WXN8_CANO9         -PLNQSTTEGNDDQGFWGIAVMGAAERNFSNPD---DEKLS-WLGLAQAV
189 tr|A6ZZS0|A6ZZS0_YEAS7         -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---ENEPQ-WLYLAQAV
190 tr|G9MQD1|G9MQD1_HYPVG         -PLNHTASLGNDDQGFWGMTAMLAAENKFPNPPA---DKPQ-WLALAQAV
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH -PLNQSVTEGNDDQAFWGLVVMAAAERNFENPS---SDEPQ-WLYLAQAV
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM -SGNYSQWLGNDDQMFWGLASITASETGFPEI----SGKPT-WTSLARVV
193 tr|A0A254UET1|A0A254UET1_ASPNG -PANQTRTEGNDDQAFWAFAALRAAERNFPNPQP---TNRRGYLSR----
194 tr|A1D587|A1D587_NEOFI          -PSNYSNYLGNDDQVFWGLAAMTAAELNYPEE----DGQPS-WVSLAQGV
195 tr|Q1K7I4|Q1K7I4_NEUCR        -PLNVTASLGNDDQGFWGMTAMTAAENNFPNPPA---DQPQ-WLGLAQAV
196 tr|G8JS88|G8JS88_ERECY         -PLNQSTTEGNDDQAFWGIAVMSAAEKGFPNPD---PDQPQ-WLALAQAV
197 tr|A0A136J7D3|A0A136J7D3_9PEZI -PVNQTRTLGNDDQGFWAMAAMSAAEAKFPDPSP---DEPQ-WLALAQAV
198 tr|E9E4X8|E9E4X8_METAQ         -PRNVTASLGNDDQGFWGMAAMSAAENGFPDPPS---DKPQ-WLELAQAV
199 tr|Q6CP42|Q6CP42_KLULA         -PLNQTVTEGNDDQAFWGFAVMAAAERNYPNPP---DDEPQ-WLYLAQAV
200 tr|G0WHL4|G0WHL4_NAUDC         -PANQSMVEGNDDQGVWGLTLMGAAERNFTDPD---DGTPG-WLAMTQAV
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---SDQPQ-WLYLAQAV
202 tr|F9WYX6|F9WYX6_ZYMTI         -PPNQSKSLGNDDQAAWALAAMSAAEYDLPVPNDLLSNNIT-WASIADTV
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 -PSNYTASLGNDDQGFWGMSAMTAAELNFPNPPK---DQPQ-WLSLAQAV
204 tr|J3P147|J3P147_GAGT3         -PPAHRASLGNDDQGFWGLAVMQAAEMRFPDSPKE--GDAH-WLALAQGV
205 tr|I2H842|I2H842_TETBL         -PLNQSTTEGNDDQAFWGIAAITAAERNFTNPP---PDKPQ-WLYLAQAV
206 tr|S6E3R3|S6E3R3_ZYGB2         -PDNQTMVEGNDDQGVWGLTVMGAVERNFTDPG---DGKPG-WLAMAQAI
207 tr|E5AD94|E5AD94_LEPMJ         -PSNQSKNEGNDDQVFWAFSMIAAAEYKLPDPPA---DQPG-WLAMTQSV
208 tr|Q6FJM7|Q6FJM7_CANGA         -PANQTTVEGNDDQGVWGLTILDAVERNFSAPI---DGKPG-WLAMSQGI
209 tr|A5DUW0|A5DUW0_LODEL         -PLNQSTTEGNDDQGFWGIAVMGAAERNFSNPD---SDEIS-WLGLAQAV
210 tr|G8BF46|G8BF46_CANPC         -PLNQSTTEGNDDQGFWGIAVMGAAERNFSNPN---NKELS-WLGLAQAV
211 tr|W3XAF6|W3XAF6_PESFW         -PDNQTRSLGNDDQGFWALTAMSAAELNFTNPPD---DEPG-WLGLTQAV
212 tr|B6H7I2|B6H7I2_PENRW         -PSNYSSYLGNDDQMFWGLAAMMAAELKFPDV----EDGFS-WLSLAQGV
213 tr|Q5ATF9|Q5ATF9_EMENI         -PSNYSSYLGYDDQFFWGATAMLAAEIGFPED----DVEYS-WLSLAQGV
214 tr|A0A167CDD3|A0A167CDD3_9ASCO -PSNQTTTEGNDDQGYWGVTVMAAAEKNFSDPG---PGNPS-WLYLAQGV
215 tr|G0RXS3|G0RXS3_CHATD         -PRNWTLSLGNDDQGFWGMAAMLAAETNFPNPAP---DEPQ-WLALAQAV
216 tr|B8NX26|B8NX26_ASPFN         -SSNYSQWLGNDDQMFWGLAAITASEAGFPER----DGKPS-WTSLARAV
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM -PLNQTSTLGNDDQAFWGMAAMSAAENKFPNPPD---DKPQ-WLELAQAV
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM -PSNQTRTEGNDDQAFWAFTALMAAEHNFPNPPE---DSPS-WLAMARAV
219 tr|A1CMM3|A1CMM3_ASPCL         -PPNQSRTLGNDDQAFWGLAAMTAAEVRFPNPPE---DQPQ-WLALAQAV
220 tr|B6HU84|B6HU84_PENRW         -PSNYSQYLGNDDQVFWGLAAITAAEFNFPER----EGEPS-WVSLAQGV
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS -PLNVTASLGNDDQGFWGMTAMTAAENNFPNPPA---DQPQ-WLGLAQAV
222 tr|G2WHY6|G2WHY6_YEASK         -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---ENEPQ-WLYLAQAV
223 tr|G0W7G8|G0W7G8_NAUDC         -PLNQSTTEGNDDQAFWGIAVMTAAERNFTNPP---EDSPQ-WLYLAQAV
224 tr|G4NGV8|G4NGV8_MAGO7         -PVNQTRTLGNDDQGFWALAAMSAAESRYPDPPP---GQPQ-WLALAQAV
225 tr|W3X855|W3X855_PESFW         -PENQTLTEGNDDQGFWALSAMTAAEMGFQNPGA---DGVQ-WLAAAQAV
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA -PNNYTASLGNDDQGFWGMSAMLAAEVKFPDPPS---DKPS-WLALAQAV
227 tr|J8Q2K7|J8Q2K7_SACAR         -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---ESDPQ-WLYLAQAV
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO -PLNQSTTEGNDDQGFWGIAVMAAAEKNFSNPE---NPKKA-WLSLAQAV
229 tr|J6EGN8|J6EGN8_SACK1         -PANQTMVEGNDDQGIWGITVMGAVERNFTDPD---EGKPG-WLAMVQAV
230 tr|A0A100ISD9|A0A100ISD9_ASPNG -PANQTRTEGNDDQSFWAFAAMSAAERNFPDPPP-SSGSPG-WLAMAQAV
231 tr|G8YQC0|G8YQC0_PICSO         -PSNQSMTEGNDDQGIWGLTTMEAAERNFTDPP------HS-WLSMTQAI
232 tr|F7VVP8|F7VVP8_SORMK         -PLNYTASLGNDDQGFWGMTAMLAAENKFPNPPD---DQPQ-WLALAQAV
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH -PSNQTMVEGNDDQGIWGITVMGAVERNFTDPG---DGKPG-WLAMVQAV
234 tr|F8MBP4|F8MBP4_NEUT8         -PANQSRTSSNDDVGFWTITAMMAAEDAYPDPPE---DQPQ-WLALVQAV
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG -PSNWSSYIANDDQMAWGLAAMTAAELNYPED----AGMPS-WVDLAERV
236 tr|A1CD27|A1CD27_ASPCL         -SKNYSQWLGNDDQMFWGLASITASETGFPDI----SGKPT-WTSLARVV
237 tr|A5DHF3|A5DHF3_PICGU         -PSNQSLTEGNDDQGVWGMAIIAAVERNFTDPE------HS-WLSMAQAI
238 tr|D4ATT7|D4ATT7_ARTBC         -PRNQSRSLGNDDQAFWAISAMSAVENKFPDPPP---DQPQ-WLYLVQAV
239 tr|E2PT42|E2PT42_ASPNC         -PANQTRTEGNDDQSFWAFAAMSAAERNFPDPPP-SSGSPG-WLAMAQAV
240 tr|C7ZPE5|C7ZPE5_NECH7         -PGNVTLSLGNDDQGFWGMSAMLAAELKFPNPPE---DRPG-WLALAQAV
```

```
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA  -PTNQTKTEGNDDQAFWAMSAMSAAERNFPNPSS---DQPQ-WLALVQAV
242 tr|Q4WKP7|Q4WKP7_ASPFU         -PSNYSNYLGNDDQVFWGLAAMTAAELNFPEE----DGQPS-WVSLAQGV
243 tr|F8MXT4|F8MXT4_NEUT8         -PLNVTASLGNDDQGFWGMTAMTAAENNFPNPPA---DQPQ-WLGLAQAV
244 tr|A0A177AJ74|A0A177AJ74_9PEZI -PTNQTKALGNDDQGFWGMAAMTAAEYNFPNPPK---DQPQ-WLALAQAV
245 tr|J7RQR5|J7RQR5_KAZNA         -PENQTMVEGNDDQGIWGLTIMDAVERNFTKPT--KDGVPG-WLAMSQAV
246 tr|Q75CW7|Q75CW7_ASHGO         -PSHQKMVEGNDDQGIWGLLVMGAAERNFPDAP--PERAPG-WLALAQAV
247 tr|G3AKK4|G3AKK4_SPAPN         -PSNQSMTEGNDDQGVWGMGIMQAAERNFTNPE---D--HS-WISLTQAI
248 tr|B2W762|B2W762_PYRTR         -PANQTKTLGNDDQAFWGMAAMTAAENKLPDLPD---GKPS-WLSLAQAV
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH -PLNQSTTEGNDDQAFWGIAAMTAAERNFTNPP---ENEPQ-WLYLAQAV
250 tr|C4XWK1|C4XWK1_CLAL4         -PSNQSMTEGNDDQGVWAMAIIQAVERNFTNTG---D--HS-WLYMTQAI
```

```
001 tr|A0A254U2J9|A0A254U2J9_ASPNG  YNTQI----DRWD--D----STCGGGLRWQIYPYE--AGYAMKNSISNGG
002 sp|Q6FLP9|DCW1_CANGA            FNTMA----LRWD--S----ETCGGGLRWQIFVWN--SGYDYKNTVSNGA
003 tr|A1CCM5|A1CCM5_ASPCL          YNTQA----GRWD--P----ATCGGGLRWQIWPYQ--AGYTMKNAISNGG
004 tr|C1H190|C1H190_PARBA          FNTQA----PRWA--N----HTCNGGLKWSINTFG--SGWTYKNTISNGC
005 tr|Q2UR85|Q2UR85_ASPOR          FNTQV----PRWD--T----SSCDGGLRWQIWPYQ--AGYTTKNAISNGG
006 tr|A5DV30|A5DV30_LODEL          FNTMN----ARWD--T----AHCNGGLRWQIFTWN--SGYNYKNSISNGC
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO  YNSMA----DRWD--A----DNCGGGLRWQIMSNM--SGWTYKSTVSNAG
008 tr|J3PGN0|J3PGN0_GAGT3          WTTQA----APER--H---DNTCNGGMRWQVPPTN--AGYDYKNTIANGC
009 tr|A0A124BXU6|A0A124BXU6_ASPNG  FNTQV----ARWD--N----STCDGGMRWQIYTYE--SGYTMKNAISNGG
010 tr|C5P4A1|C5P4A1_COCP7          FNSQA----ARWN--T----ATCGGGLNWQIFQFN--RGFNYKNSISNGC
011 tr|Q0CG55|Q0CG55_ASPTN          FNTQA----ARWD--T----ETCGGGLRWQIFTFN--NGFNYKNTISNGC
012 tr|C4Y2X5|C4Y2X5_CLAL4          FNTMS----ARWD--L----GECNGGLRWQIFQWN--SGYDYKNSVSNGA
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG  FNTQA----ARWD--K----STCNGGLRWQIFTFN--NGWTYKNTISNGC
014 tr|C7Z068|C7Z068_NECH7          WTTQA----HPNR--H---DKECNGGLRWQIPFTN--AGYNYKNTIANGC
015 tr|G8BXX2|G8BXX2_TETPH          FNTMA----LRWD--T----DSCGGGLRWQIFQWN--SGYDYKNSVSNGA
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR  FNTMS----SRWD--M----DHCNGGLRWQIFQWN--SGYDYKNSVSNGA
017 tr|A1CSC2|A1CSC2_ASPCL          FNTQV----PRWD--T----STCHGGLRWQIWPYQ--DGYRTKNAISNGG
018 sp|Q05031|DFG5_YEAST            FNTMY----SRWD--S----EHCGGGLRWQIFTWN--SGYNYKNTVSNAC
019 tr|H2AQJ6|H2AQJ6_KAZAF          FNEMW----SRWD--S----ANCGGGLRWQIFTWN--SGYDYKSTISNVC
020 tr|G2XFH9|G2XFH9_VERDV          WSTQN----D--R--F---DDHCNGGLRWQAVSTN--AGYNYKNTIANGC
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH  FNTMA----ARWD--S----ATCGGGLRWQIFTWN--SGYDYKNTVSNGA
022 sp|Q5AD78|DCW1_CANAL            FNTMQ----ARWD--T----ETCNGGLRWQIFQWN--SGYDYKNSVSNGA
023 tr|B6H7E8|B6H7E8_PENRW          FNGQI----PRWD--M----TACNGGLRWQIWPYQ--DGYNLKNAISNGG
024 tr|B6GZU2|B6GZU2_PENRW          FNTQV----PRWD--T----SSCGGGLRWQIFTWN--NGYNYKNTISSGG
025 tr|E9E4W7|E9E4W7_METAQ          WNTQA----DPSR--Y---DETCNGGLRWQIPRTN--AGYDYKNTIANGI
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO  FNTQA----PRWD--P----ATCGGGLRWQIFTWN--NGYNYKNTISTGG
027 tr|G2R7G8|G2R7G8_THITE          FNTQA----WR-W--D---MEYCNGGLHWQIPTTN--GGYDYKNSIANVV
028 tr|G3JBY1|G3JBY1_CORMM          FNEYV----VRWN--EE--NGLCGGGLRWQVYSFL--KGYDYKNSIANGC
029 tr|A5DNV5|A5DNV5_PICGU          FNTMQ----ARWD--H----NDCHGGLRWQIFQWN--SGYDYKNSVSNGC
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI  FNTQA----QRWD--K----STCGGGLKWQIYTFN--NGYTYKNTISNGC
031 tr|Q2TWC1|Q2TWC1_ASPOR          FNMQI----ERWD--E----TACNGGLRWQIWPYQ--AGYTMKNSISNGG
032 tr|J5RXY9|J5RXY9_SACK1          FNTMA----LRWD--A----DSCGGGLRWQIFVWN--SGYDYKNTVSNGA
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA  FNEYA----VRWE--EE--NGLCGGGLRWQVYSFL--NGWAYKNSISNGC
034 tr|A0A254U5V1|A0A254U5V1_ASPNG  FNT-----------------------------Q--AGYTTKNAISNGG
035 tr|A0A061B6H7|A0A061B6H7_CYBFA  FNTMS----SRWD--H----DNCGGGLRWQIFEWN--SGYDYKNSVSNGC
036 tr|G8ZQ93|G8ZQ93_TORDC          FNTMA----LRWD--T----TSCNGGLRWQIFMWN--TGYDYKNTVSNGA
037 tr|Q9C2J1|Q9C2J1_NEUCS          FNQMA----SRWD--D----LNCGGGLRWAINDFQ--TGKDYKNSISNGI
038 tr|C5NZK5|C5NZK5_COCP7          LHSQV----PRWD--D----ETCGGGLRWQIFSFN--KGYNYKNSISNGC
039 tr|Q6C0T7|Q6C0T7_YARLI          FNTMA----SRWD--D-----ICGGGLHWQIYQWN--SGYSYKNAIASGA
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB  FNTMA----LRWD--S----ETCGGGLRWQIFVWN--SGYDYKNTVSNGA
041 tr|F7VZ72|F7VZ72_SORMK          FNTQA----HPDR--H---DETCNGGLRWQIPPLN--NGYNYKNSIANGC
042 tr|A3LN37|A3LN37_PICST          FNTMT----ARWD--S----AECNGGLRWQIFQWN--SGYDYKNSVSNGC
043 tr|Q6CAI2|Q6CAI2_YARLI          FNTMS----ARWD--T----DNCGGGLRWQIYTWN--NGYDYKNTVSNGC
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO  FNTMT----ARWD--H----KECHGGLRWQIFQWN--SGYDYKNSVSNGC
045 tr|G9MJS5|G9MJS5_HYPVG          FNTQA----SPER--H---DGTCGGGIRWQIPPTN--PGYSYKNSIANGC
046 tr|G9MHI4|G9MHI4_HYPVG          FNQYV----GRWD--------SHCGGGLRWQIFSWN--KGYDYKNSVANGC
047 tr|D4AP27|D4AP27_ARTBC          FNSQI----PRWD--T----ATCGGGLRWQIFRFN--RGFDYKNTISNGA
048 tr|B6GZT8|B6GZT8_PENRW          FNTQI----KRWD--T----SSCGGGLRWQIWPYQ--SGYTMKNSISNGG
049 sp|O74556|YCZ2_SCHPO            FNQQV----TRWD--T----DHCNGGLRWQITEFN--SGYNYKNTVSNGA
050 tr|A0A0C4E991|A0A0C4E991_MAGP6  WTTQA----APER--H---DNTCNGGMRWQVPPTN--AGYDYKNTIANGC
051 tr|B8MHG0|B8MHG0_TALSN          FNDFV----ARWEVDK----ASCGGGLRWQLYPTL--AGYDSKNAISNGA
052 tr|A3LMV8|A3LMV8_PICST          FNTMN----ARWD--T----DHCGGGLRWQIFTWN--SGYSYKNSVSNGC
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU  YNNMA----PRWD--D----QNCGGGLRWQILDTM--NGWDYKSMISTGG
054 sp|Q9P6I3|YHG7_SCHPO            FNEIM----SRWD--T----SSCGGGVRWQIYSFN--NGYSYKNSISNGI
055 tr|Q2US57|Q2US57_ASPOR          FNTQV----KRWD--T----SSCAGGMRWQIWTYE--AGYRMKNSISNGG
056 tr|W7N622|W7N622_GIBM7          FNEYV----SRWD--T----QHCDGGMRWQIFTWN--AGYDYKNSISNGC
057 sp|Q9P6I4|YHG6_SCHPO            FNRMS----GRWD--S----STCGGGLRWQAFAWL--NGYSYKASVSNAL
```

```
058 tr|C7GRB3|C7GRB3_YEAS2                FNTMY----SRWD--S----EHCGGGLRWQIFTWN--SGYNYKNTVSNAC
059 tr|C7GP28|C7GP28_YEAS2                FNTMA----LRWD--A----DSCGGGLRWQIFVWN--SGYDYKNTVSNGA
060 tr|G3JGZ5|G3JGZ5_CORMM                WATQA----DPSR--H---DSYCNGGMRWQIPFSN--AGYDYKNTIANGC
061 tr|C4QXV4|C4QXV4_KOMPG                VNTMW----ERWD--L----EHCNGGLRWQIFQWN--AGYNYKNTVSNAC
062 tr|I1S0Z5|I1S0Z5_GIBZE                FNTQA----NPDR--H---DDTCGGGLRWQIPRTN--NGYDYKNSIANGC
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM        FNTQA----ARWD--E----DTCGGGLRWQIFSFN--NGWNYKNTISNGC
064 tr|G0W5X4|G0W5X4_NAUDC                FNTMA----LRWD--A----QNCGGGLRWQIFVWN--SGYDYKNSVSNGA
065 tr|G8ZPC1|G8ZPC1_TORDC                FNQLN----SRWD--T----ANCGGGLRWQIFTWN--GGYNYKNTISNAC
066 tr|A0A167CDC5|A0A167CDC5_9ASCO        FNTMA----SRWD--T----DTCHGGLRWQIFTWN--NGYNYKNSVANGC
067 tr|R9XAT7|R9XAT7_ASHAC                FNTMA----LRWD--T----DTCNGGLRWQIFRWN--AGYHYKNSVSNGA
068 tr|B2B747|B2B747_PODAN                FNTQA----SPDR--H---DETCNGGLRWQIPWSN--NGYNYKNSIANGC
069 tr|G0SE99|G0SE99_CHATD                FNTQA----SPTR--H---DDTCGGGLRWQIPWSN--NGYNYKNSIANGC
070 tr|J8Q6F0|J8Q6F0_SACAR                FNTMY----SRWD--S----ENCGGGLRWQIFTWN--SGYNYKNTVSNAC
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH        FNEMY----SRWD--T----ANCGGGLRWQIFTWN--SGYNYKNTISNGC
072 tr|J3K6E2|J3K6E2_COCIM                FNSQA----ARWN--T----ATCGGGLNWQIFQFN--RGFNYKNSISNGC
073 tr|G4N3E1|G4N3E1_MAGO7                WTTQN----SPER--H---DDKCGGGMRWQIPPLN--AGYDYKNTIANAI
074 tr|C5M5J2|C5M5J2_CANTT                FNTMY----ARWD--D----EHCGGGLRWQIFTWN--SGYDYKNSISNGC
075 tr|A0A254U0X0|A0A254U0X0_ASPNG        FNMQV----NRWD--D----TTCDGGMRWQIWPYQ--AGYTMKNAISNGG
076 sp|Q75DG6|DCW1_ASHGO                  FNTMA----LRWD--M----ETCNGGLRWQIFRWN--DGYHYKNSVSNGA
077 tr|Q6BZF0|Q6BZF0_DEBHA                FNTMT----ARWD--H----KECHGGLRWQIFQWN--SGYDYKNSVSNGC
078 tr|E9E3Q1|E9E3Q1_METAQ                WNTQA----DPSR--Y---DETCNGGLRWQIPFSN--QGYGYKNTIANGI
079 sp|P36091|DCW1_YEAST                  FNTMA----LRWD--A----DSCGGGLRWQIFVWN--SGYDYKNTVSNGA
080 tr|W0TAH6|W0TAH6_KLUMD                FNTMA----SRWD--S----DTCGGGLRWQIFTWN--SGYDYKNSVSNAA
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA        FNTMW----ARWD--D----QYCGGGLRWQIFTWN--SGYTYKNTISTAC
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI        FNDYA----SRWD--E----ASCSGGLRWQIYTFN--AGYNYKNTISNGC
083 tr|Q2TYU3|Q2TYU3_ASPOR                FNTQV----ARWD--S----SNCGGGLRWQIFPYQ--AGYAMKNSISNGL
084 tr|G2XGF6|G2XGF6_VERDV                FNQWA----QRWI--EA--SDSCGGGLRWQIFSFN--NGFNYKNSISNGC
085 tr|G3AMT4|G3AMT4_SPAPN                FNTMT----ARWD--E----NECNGGLRWQIFQWN--SGYDYKNSVSNGC
086 tr|A0A1C1D213|A0A1C1D213_9EURO        FNTQA----ARWD--T----MSCGGGLKWQIFAFN--NGYNYKNTISNGC
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS        FHTQA----DPGR--H---DDVCGGGLRWQIPMAN--IGYDYKNSISNGI
088 tr|A0A100I5A6|A0A100I5A6_ASPNG        FNMQV----NRWD--D----TACDGGMRWQIWPYQ--AGYTMKNAISNGG
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH        FNLMY----SRWD--D----QHCGGGLRWQIFTWN--SGYNYKNTVSNGC
090 tr|G2Q8A7|G2Q8A7_MYCTT                FNTQA----SPEE--H---DDTCNGGLRWQIYLSN--NGYDYKNSIANGC
091 tr|B8NPZ1|B8NPZ1_ASPFN                FNMEV----ERWD--K----SACNGGMRWQLWPYQ--EGYTMKNAISNGG
092 tr|Q7SAB2|Q7SAB2_NEUCR                FHTQA----DPGR--H---DDVCGGGLRWQIPMAN--IGYDYKNSISNGI
093 tr|A2R8R5|A2R8R5_ASPNC                FNTQV----ARWD--N----STCNGGMRWQIYTYE--SGYTMKNAISNGG
094 tr|C5DHG0|C5DHG0_LACTC                FNTMW----MRWD--S----EHCGGGVRWQIFTWN--SGYNYKNTVSNAC
095 tr|G0SFA3|G0SFA3_CHATD                FNQWA----TRWE--E----STCQGGLRWQIFSFN--AGYNYKNSISNGC
096 tr|K1WJG5|K1WJG5_MARBU                FNEMV----ARWD--T----STCGGGMRWQIFTFN--SGYTYKNSIANGC
097 tr|H2B005|H2B005_KAZAF                FNTMA----LRWD--D----ATCGGGLRWQIFTWN--SGYDYKNTVSNGA
098 tr|C5DYD7|C5DYD7_ZYGRC                FNEIN----ERWD--D----KSCAGGVRWQIFTWN--SGYNYKNTISNGC
099 tr|G8YM23|G8YM23_PICSO                FNTMS----ARWD--T----KDCGGGLRWQIFQWN--SGYDYKNTVSTGC
100 tr|W3X8E3|W3X8E3_PESFW                FNTQA----DPGR--H---DDTCGGGLRWQIPFSN--NGYNYKNSIANGC
101 tr|A0A254U3K7|A0A254U3K7_ASPNG        FNTQV----ARWD--N----STCNGGMRWQIYTYE--SGYTMKNAISNGG
102 tr|Q6CER8|Q6CER8_YARLI                FNSMA----DRWD--T----QVCDGGLHWQIFQWN--SGYDYKNAIANGA
103 tr|G2QT05|G2QT05_THITE                WTTQA----SPDR--H---DNTCGGGLRWQIPVVN--NGYDYKNSIANGC
104 sp|Q5ACZ2|DFG5_CANAL                  FNTMN----ARWD--A----DNCGGGLRWQIFTWN--SGYDYKNSISNGC
105 tr|G2WKU6|G2WKU6_YEASK                FNTMY----SRWD--S----EHCGGGLRWQIFTWN--SGYNYKNTVSNAC
106 tr|A0A0C7N752|A0A0C7N752_9SACH        FNTMY----SRWD--D----QHCGGGVRWQIFTWN--SGYNYKNTVSNAC
107 tr|W3WN68|W3WN68_PESFW                FNEYV----TRWE--DA--EDTCGGGLRWQIYTFN--NGYNYKNSISNGC
108 tr|G2QFS1|G2QFS1_MYCTT                FHTQA----APDR--H---DGTCNGGLRWQVPPTN--AGYNYKNTIANAC
109 tr|G3J9G4|G3J9G4_CORMM                FNEYA----VRWE--EE--NGLCGGGLRWQVYSFL--NGWAYKNSISNGC
110 tr|Q2UJ03|Q2UJ03_ASPOR                FNTQA----PRWN--T----ETCGGGLRWQIFSFN--NGYHYKNTIANGC
111 tr|G8YRT2|G8YRT2_PICSO                YNTMN----DRWD--------SHCGGGLRWQIFTWN--SGYDYKNSIANGC
112 tr|W3X554|W3X554_PESFW                FNEYV----DRWD--P----DTCGGGLRWQIFTFN--NGFNYKNSISNGC
113 tr|Q1K7A8|Q1K7A8_NEUCR                FNQMA----SRWD--D----LNCGGGLRWAINDFQ--TGKDYKNSISNGI
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM        FNTQV----PRWD--T----STCHGGLRWQISTYQ--DGYRTKNAISNGG
115 tr|A7F3P0|A7F3P0_SCLS1                FNDMA----ARWD--E----STCGGGLRWQIFTFN--SGYNYKNSIANGC
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA        FNEYV----IRWQ--EE--NGLCGGGLRWQVYSFL--KGYDYKNSIANGC
117 tr|Q6CIP9|Q6CIP9_KLULA                FNQMY----SRWD--S----DHCNGGLRWQIFTWN--SGYNYKNTISNGC
118 tr|H2AMQ5|H2AMQ5_KAZAF                FNTMA----LRWD--T----DSCGGGLRWQIFTWN--SGYDYKNSVSNGA
119 tr|G8JMI3|G8JMI3_ERECY                FNTMY----ARWD--S----QHCGGGLRWQIFTWN--LGYDYKNTIANGC
120 tr|C5DGT1|C5DGT1_LACTC                FNTMA----LRWD--D----ESCGGGLRWQIFTWN--SGYDYKNTVSNGA
121 tr|K1XJR4|K1XJR4_MARBU                FNSQS------RR--T---DGECGGGLRWQVFPYL--IGYDYKNSIANGC
122 tr|A0A124BYC5|A0A124BYC5_ASPNG        FDVQS----ARWD--N----SSCEGGLRWQIWPYE--VGYTMKNSMSNGG
123 tr|A7TLL2|A7TLL2_VANPO                FNTMA----LRWD--D----QTCNGGLRWQIFVWN--SGYDYKNSVSNGA
124 tr|Q0CB86|Q0CB86_ASPTN                YNTQV----ARWD--T----SNCGGGLRWQIFPYQ--AGYNMKNSISNGG
125 tr|K1XLH0|K1XLH0_MARBU                FNTQI----PRWD--T----TTCGGGLRWQIFPFN--AGYTYKNSISNGC
126 tr|J7S438|J7S438_KAZNA                FNTMA----LRWD--T----ESCGGGLRWQIFNWN--SGYDYKNTVSNGA
127 tr|Q4WG09|Q4WG09_ASPFU                FNMQV----ARWD--K----VACDGGMRWQIWPYQ--AGYTMKNAISNGG
```

```
128 tr|H8WZ09|H8WZ09_CANO9          FNTMD----ARWD--T----EHCGGGLRWQIFTWN--SGYNYKNTISNGC
129 tr|I1RL09|I1RL09_GIBZE          WTTQA----NPER--H---DDECNGGMRWQIPFTN--SGYDYKNTIANGC
130 tr|Q4W985|Q4W985_ASPFU          FNTQA----ARWD--E----DTCGGGLRWQIFSFN--NGWNYKNTISNGC
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU FNTMY----SRWD--S----EHCGGGLRWQIFTWN--SGYNYKNTVSNAC
132 tr|W6QCW3|W6QCW3_PENRF          FNTQV----PRWD--T----TSCGGGLRWQIFTWN--NGYNYKNTISTGG
133 tr|W3WMD3|W3WMD3_PESFW          FNEQT----HEER--RVPAGSNCEWGLRWQVYRTN--NGFDYINTIANAC
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH FNTMA----ARWD--S----ATCGGGLRWQIFTWN--SGYDYKNTVSNGA
135 tr|I2GYH1|I2GYH1_TETBL          FNTMW----SRWD--A----ADCGGGLRWQIFTWN--SGYDYKNTISNAC
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH FNEVW----SRWD--D----EHCGGGVRWQIFTWN--SGYNYKNTISNAC
137 tr|Q752P3|Q752P3_ASHGO          FNKME----ARWD--A----ENCGGGLRWQIFQWN--NGYTYKNTIANGC
138 tr|B8MYP3|B8MYP3_ASPFN          FNTQV----PRWD--T----SSCDGGLRWQIWPYQ--AGYTTKNAISNGG
139 tr|Q2TXL6|Q2TXL6_ASPOR          FNTQA----EKWD--P----DTCHGGLRWQRNSWN--GGYDLKNSVSNGG
140 tr|Q0CN17|Q0CN17_ASPTN          FNTQV----PRWD--K----TSCEGGLRWQIWPYQ--AGYTTKNSVSNGG
141 tr|J3NHD5|J3NHD5_GAGT3          FNTQA----APGR--H---DNTCGGGLRWQIPPTN--NGYDYKNSIANGC
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI FATQA----SRWD--D----AHCSGGLKWQIFRFN--VGYDYKNAISNGC
143 tr|A1CVB0|A1CVB0_NEOFI          FNTQA----ARWD--E----DTCGGGLRWQIFSFN--NGWNYKNTISNGC
144 tr|B8NVI3|B8NVI3_ASPFN          FNTQA----EKWD--P----DTCHGGLRWQRNSWN--GGYDLKNSVSNGG
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH FNTMA----LRWD--D----DTCGGGLRWQIFTWN--SGYDYKNTVSNGA
146 tr|A1DJ54|A1DJ54_NEOFI          FNMQV----ARWD--K----VACDGGMRWQIWPYQ--AGYTMKNSISNGG
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS WATQA----APDR--H---DDTCNGGLRWQIPPTN--NGYDYKNTIANAI
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA FNTMA----LRWD--P----DHCGGGLRWQIFEWN--SGYDYKNSVSNGC
149 tr|A0A177A618|A0A177A618_9PEZI FSLQS----GRWD--T----ETCGGGLKWQVYVFN--KGYDYKNSISNGA
150 tr|Q75CW6|Q75CW6_ASHGO          FNTMW----ARWD--P----ENCGGGLRWQIFPWN--NGYDYKNTISNAC
151 tr|Q7S4K4|Q7S4K4_NEUCR          WATQA----APDR--H---DDTCNGGLRWQIPPTN--NGYDYKNTIANAI
152 tr|G2QB99|G2QB99_MYCTT          FNQYA----SRWE--E----ATCGGGLRWQIFTFN--NGFNYKNSISNGC
153 tr|G0S3F2|G0S3F2_CHATD          WTTQA----SPER--H---DGTCNGGLRWQIPPTN--AGYNYKNTIANAC
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO FNTMS----ARWD--P----EHCGGGLRWQIFQWN--SGYNYKNSVSNGC
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB FNTMY----ARWD--M----QSCNGGLRWQIFTWN--SGYNYKNTISNAC
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM YNTQV----DRWD--T----STCGGGLRWQIYTYQ--AGYTMKNAISNGG
157 tr|G2QKJ0|G2QKJ0_MYCTT          FNTQA----PR-W--E---TEYCAGGLRWQVVSTN--GGYDYKNTIANAV
158 tr|Q4WFX5|Q4WFX5_ASPFU          FNEQI----ARWD--D----QACGGGLNWQIFPFN--NGYTYRNSISNGG
159 tr|Q5BGD7|Q5BGD7_EMENI          FNMQA----NRWD--E----RACDGGITWQIHPWQ--AGYTLRNSISNGG
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH FNTMS----ARWD--E----DSCGGGLRWQIFTWN--SGYDYKNSVSNAA
161 tr|G8BYN9|G8BYN9_TETPH          FNTMY----ARWD--T----AYCGGGLRWQIYTWN--SGYNYKNTISNAC
162 tr|F7VVT4|F7VVT4_SORMK          FNTQA----ADDR--H---DDECGGGLRWQIFHSN--AGYDYKNTIANAC
163 tr|G8YB80|G8YB80_PICSO          FNTMS----ARWD--T----TDCGGGLRWQIFQWN--SGYDYKNTVSTGC
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA YNTMW----ERWD--M----EHCGGGLRWQIFTWN--SGYNYKNTISTAC
165 tr|G2WRS7|G2WRS7_VERDV          FHTQA----SPDR--H---DETCGGGLRWQIPFAN--NGYNYKNSIANGC
166 tr|A0A124BVB5|A0A124BVB5_ASPNG FNEQA----QRWN--T----QKCGGGLKWQIYTFN--SGYTYRNSISNGC
167 tr|A7EIG7|A7EIG7_SCLS1          FNTQA----PRWD--D----STCGGGLRWQIFTFN--NGYNYKNSISNGC
168 tr|C5M5U7|C5M5U7_CANTT          FNTMQ----ARWD--E----EDCNGGLRWQIFQWN--SGYDYKNSVSNGC
169 tr|B9WAA8|B9WAA8_CANDC          FNTMN----ARWD--A----DNCGGGLRWQIFTWN--SGYDYKNSISNGC
170 tr|F8MRU1|F8MRU1_NEUT8          WATQA----APDR--H---DDTCNGGLRWQIPPTN--NGYDYKNTIANAI
171 tr|A1DJS0|A1DJS0_NEOFI          YNTQV----DRWD--T----STCGGGLRWQIYTYQ--AGYTMKNAISNGG
172 tr|B6HLM8|B6HLM8_PENRW          FNDQKSPDGKGWD--T----TICGGGLRWQKEIWQ--SGYTMKNAVSNGG
173 tr|A0A117DWS0|A0A117DWS0_ASPNG FNTQV----PRWD--T----SSCQGGLRWQIWPYQ--AGYTTKNAISNGG
174 tr|A0A124BXM1|A0A124BXM1_ASPNG YNTQV----DRWD--D----STCGGGLRWQIYPYE--AGYAMKNSISNGG
175 tr|J3NZQ6|J3NZQ6_GAGT3          WVRQT----LPER--Q---DSECGGGLRWQVPGAGWGPGQTYKNTIANAC
176 tr|W0THH8|W0THH8_KLUMD          FNEMY----SRWD--P----AHCGGGLRWQIFTWN--SGYNYKNTISNGC
177 tr|B2AEF2|B2AEF2_PODAN          FHTQA----APER--H---DNTCNGGLRWQVPPTN--AGYNYKNTIANGC
178 tr|Q96TX1|Q96TX1_NEUCS          FNTQA----NPDR--H---DKTCNGGLRWQIPPLN--NGYNYKNSIANGC
179 tr|E9DYA6|E9DYA6_METAQ          FNTQA----SPQR--H---DGTCGGGLRWQIPFTN--NGYGYKNSIANGC
180 tr|B9WAI2|B9WAI2_CANDC          FNTMQ----ARWD--T----EECNGGLRWQIFQWN--SGYDYKNSVSNGA
181 tr|Q6C171|Q6C171_YARLI          FNTMA----ARWD--T----QSCYGGLHWQIFQWN--SGYDYKNAIANGA
182 tr|A6ZMV1|A6ZMV1_YEAS7          FNTMY----SRWD--S----EHCGGGLRWQIFTWN--SGYNYKNTVSNAC
183 tr|B6K7X7|B6K7X7_SCHJY          FNEQV----ARWD--T----TLCDGGLRWQIYPFN--NGYTYKNSITNGL
184 tr|I1RSA2|I1RSA2_GIBZE          FNEYV----SRWD--D----EYCGGGMRWQIFKWN--TGYDYKNAISNGC
185 tr|F7VWZ9|F7VWZ9_SORMK          FNQMA----SRWD--D----INCGGGLRWAINDFQ--TGKDYKNSISNGL
186 tr|A0A1B2J789|A0A1B2J789_PICPA VNTMW----ERWD--L----EHCNGGLRWQIFQWN--AGYNYKNTVSNAC
187 tr|G0V8N4|G0V8N4_NAUCC          FNTMA----LRWD--A----DSCGGGLRWQIFVWN--SGYDYKNSVSNGA
188 tr|H8WXN8|H8WXN8_CANO9          FNTMT----ARWD--S----DDCNGGLRWQIFQWN--SGYDYKNSVSNGC
189 tr|A6ZZS0|A6ZZS0_YEAS7          FNTMA----LRWD--A----DSCGGGLRWQIFVWN--SGYDYKNTVSNGA
190 tr|G9MQD1|G9MQD1_HYPVG          WNTQA----DPSR--H---DETCNGGLRWQIPFSN--AGYNYKNTIANGC
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH FNTMA----LRWD--D----DTCGGGLRWQIFTWN--SGYDYKNTVANGA
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM FNMQV----ARWD--K----VACDGGMRWQIWPYQ--AGYTMKNAISNGG
193 tr|A0A254UET1|A0A254UET1_ASPNG ----------RPCS--T----NKRSDGIRRSIYTFN--SGYTYRNSISNGC
194 tr|A1D587|A1D587_NEOFI          FNTQV----PRWD--T----STCHGGLRWQISTYQ--DGYRTKNAISNGG
195 tr|Q1K7I4|Q1K7I4_NEUCR          FNTQA----NPDR--H---DKTCNGGLRWQIPPLN--NGYNYKNSIANGC

196 tr|G8JS88|G8JS88_ERECY          FNTMA----YRWD--S----SSCNGGLRWQIFTWN--TGYDYKNSVSNGA
```

```
197 tr|A0A136J7D3|A0A136J7D3_9PEZI  FNAYV----SRWD--E----QHCDGGLRWQIFQFN--AGFEYKNSISNGC
198 tr|E9E4X8|E9E4X8_METAQ          FNTQA----SPDR--H---DSTCGGGLRWQIPFAN--NGYDYKNSIANGC
199 tr|Q6CP42|Q6CP42_KLULA          FNTMS----SRWD--M----DSCGGGLRWQIFTWN--SGYDYKNSVSNAA
200 tr|G0WHL4|G0WHL4_NAUDC          FNTMN----ARWD--S----AHCGGGLRWQIFTWN--SGYNYKNTISNAC
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU FNTMA----LRWD--A----DSCGGGLRWQIFVWN--SGYDYKNTVSNGA
202 tr|F9WYX6|F9WYX6_ZYMTI          FKEQV----ARWD--T----ESCGGGLRWQIFTFN--NGYNYKNSISQAS
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 FNTQA----APGR--H---DNTCGGGLRWQIPPTN--NGYDYKNSIANGC
204 tr|J3P147|J3P147_GAGT3          WVTQA----SKER--H---DELCGGGMRWQVPPTN--TGYNYKNTIANGI
205 tr|I2H842|I2H842_TETBL          FNTMA----LRWD--S----DTCNGGLRWQIFVWN--SGYSYKNSVSNGA
206 tr|S6E3R3|S6E3R3_ZYGB2          FNEVW----SRWD--D----EHCGGGVRWQIFTWN--SGYNYKNTISNAC
207 tr|E5AD94|E5AD94_LEPMJ          FNQFV----TRYS--KEVDEGTCGGGMRWQIFPWL--NGWNYKNTASNGG
208 tr|Q6FJM7|Q6FJM7_CANGA          FNTMY----ARWD--M----QSCNGGLRWQIFTWN--SGYNYKNTISNAC
209 tr|A5DUW0|A5DUW0_LODEL          FNTMT----ARWD--S----ADCNGGLRWQIFQWN--SGYDYKNSVSNGC
210 tr|G8BF46|G8BF46_CANPC          FNTMT----ARWD--V----DDCNGGLRWQIFQWN--SGYDYKNSVSNGC
211 tr|W3XAF6|W3XAF6_PESFW          FNEYV----ERWN--EA--TDTCGGGLRWQIFTFN--NGYNYKNSISNGC
212 tr|B6H7I2|B6H7I2_PENRW          FNTQT----ARWD--T----TTCGGGLRWQLFPYQ--SGYTMKNSVSNGG
213 tr|Q5ATF9|Q5ATF9_EMENI          YNTQI----RAWD--T----SNCGGGLRWQMFPYQ--AGYAMKNSISNAG
214 tr|A0A167CDD3|A0A167CDD3_9ASCO FNTMA----SRWD--T----QTCNGGLRWQIFTWN--GGYDYKNSVANGC
215 tr|G0RXS3|G0RXS3_CHATD          FNRQA----RL-W--D---MEFCNGGLRWQIPHTN--AGYNYKNTIAQVV
216 tr|B8NX26|B8NX26_ASPFN          FNEQV----ERWD--G----QTCGGGMRWQVWPYQ--AGYQLKNAISNGG
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM LHSQV----PRWD--D----ETCGGGLRWQIFSFN--KGYNYKNSISNGC
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM FNEQI----ARWD--D----QACGGGLKWQIFPFN--NGYTYRNSISNGG
219 tr|A1CMM3|A1CMM3_ASPCL          FNTQA----PRWD--N----RTCGGGLKWQIFTFN--NGYNYKNTISNGC
220 tr|B6HU84|B6HU84_PENRW          FNTQY----PRWD--T----SSCHGGMRWQIWPYQ--DGYLTKNAISNGG
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS FNTQA----NPDR--H---DKTCNGGLRWQIPPLN--NGYNYKNSIANGC
222 tr|G2WHY6|G2WHY6_YEASK          FNTMA----LRWD--A----DSCGGGLRWQIFVWN--SGYDYKNTVSNGA
223 tr|G0W7G8|G0W7G8_NAUDC          FNTMA----LRWD--G----DTCGGGLRWQIFVWN--SGYDYKNTVSNGA
224 tr|G4NGV8|G4NGV8_MAGO7          FNEYT----ARWD--M----KTCGGGLRWQIFQFN--NGFNYKNSISNGA
225 tr|W3X855|W3X855_PESFW          FNEYV----WRWD--N----STCGGGLRWQIYTFN--NGYNYKNSVSNGC
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA WNTQA----SPSR--H---DTECNGGLRWQIPFAN--NGFDYKNSIANGI
227 tr|J8Q2K7|J8Q2K7_SACAR          FNTMA----LRWD--A----AYCGGGLRWQIFVWN--SGYDYKNTVSNGA
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO FNTMT----ARWD--H----KECDGGLRWQIFQWN--SGYDYKNSVSNGC
229 tr|J6EGN8|J6EGN8_SACK1          FNTMY----SRWD--S----EHCGGGLRWQIFTWN--SGYNYKNTVSNAC
230 tr|A0A100ISD9|A0A100ISD9_ASPNG FNTQA----ARWD--T----STCNGGLRWQIFTFN--NGWTYKNTISNGC
231 tr|G8YQC0|G8YQC0_PICSO          YNTMN----DRWD--------SHCNGGLRWQIFTWN--SGYNYKNSIANGC
232 tr|F7VVP8|F7VVP8_SORMK          WATQA----APER--H---DDTCNGGLRWQIPLAN--NGYDYKNTIANAI
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH FNTMY----SRWD--S----EHCGGGLRWQIFTWN--SGYNYKNTVSNAC
234 tr|F8MBP4|F8MBP4_NEUT8          FNQMA----SRWD--D----LNCGGGLRWAINDFQ--TGKDYKNSISNGI
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG FDVQA----ARWD--N----NTCSGGLRWQLWPYQ--AGYAMKNAMSNGG
236 tr|A1CD27|A1CD27_ASPCL          FNMQV----ERWD--R----GTCDGGMRWQIWPYQ--AGYTMKNAISNGG
237 tr|A5DHF3|A5DHF3_PICGU          YNTMN----SRWD--T----KHCGGGLRWQIFTWN--SGYNYKNSISNGC
238 tr|D4ATT7|D4ATT7_ARTBC          FNRQS----ARWD--N----ETCGGGLRWQIFSLN--NGYHYKNAISNGA
239 tr|E2PT42|E2PT42_ASPNC          FNTQA----ARWD--K----STCNGGLRWQIFTFN--NGWTYKNTISNGC
240 tr|C7ZPE5|C7ZPE5_NECH7          FNTQA----APDR--H---DKTCNGGLRWQIPPTN--NGYDYKNSIANGC
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA FNLQV----ARWD--K----GTCGGGLKWQIFRFN--NGFNYKNTISNGC
242 tr|Q4WKP7|Q4WKP7_ASPFU          FNTQV----PRWD--T----STCHGGLRWQISTYQ--DGYRTKNAISNGG
243 tr|F8MXT4|F8MXT4_NEUT8          FNTQA----NPDR--H---DKTCNGGLRWQIPPLN--NGYNYKNSIANGC
244 tr|A0A177AJ74|A0A177AJ74_9PEZI FYTQA----SRWD--D----AHCNGGLKWQIFRFN--VGYDYKNAISNGC
245 tr|J7RQR5|J7RQR5_KAZNA          FNTMW----ARWD--T----QHCGGGLRWQIFTWN--SGYNYKNTISNAC
246 tr|Q75CW7|Q75CW7_ASHGO          FNTMW----ARWD--P----ENCGGGLRWQIFPWN--GGYDYKNTISNAC
247 tr|G3AKK4|G3AKK4_SPAPN          FNTMN----ARWD--T----SHCNGGLRWQIFTWN--SGYNYKNSIANGC
248 tr|B2W762|B2W762_PYRTR          FNTQQ----NRWD--M----STCGGGLRWQIFSFN--NGYNYKNAISNGC
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH FNTMA----LRWD--A----DSCGGGLRWQIFVWN--SGYDYKNTVSNGA
250 tr|C4XWK1|C4XWK1_CLAL4          FNTMK----NRWD--T----TTCGGGLRWQIFTWN--SGYNYKNSISNGC

001 tr|A0A254U2J9|A0A254U2J9_ASPNG LFQLAARLARYTSND-T-YADWAEKIFDWCASTPLLNNE---------TW
002 sp|Q6FLP9|DCW1_CANGA           LFHIAARLARYTGNQ-S-YVDWAERVYDWMEDVHLID------NG-T-YR
003 tr|A1CCM5|A1CCM5_ASPCL         FFQLAARLARYTNNH-T-YYEWGEKAWDWSCTTPLLNNK---------TW
004 tr|C1H190|C1H190_PARBA         FFNLASRLALYTGNT-T-YANWAEASWNWMSGVGLISP----------TY
005 tr|Q2UR85|Q2UR85_ASPOR         LFQLAARLGRYTKNQ-T-YIDWAEKIWDWSATTPLLKTA---------DW
006 tr|A5DV30|A5DV30_LODEL         LFHLAARLARYTGNGTV-YVDVAEKVWQWMDDVGFLTEED---N--G-DF
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO VFALGARLARYTGDN-E-LVKSSSRILRWMKKAFFVHAPTDDSDNAGNYY
008 tr|J3PGN0|J3PGN0_GAGT3         FFNIGARLARYTGNK-T-YAEHAAKTFDWLWAVNYIDN----------ENY
009 tr|A0A124BXU6|A0A124BXU6_ASPNG LFQLSARLARYTSNQ-T-YYDWAERIWDWSTTTPLLSNS---------TW
010 tr|C5P4A1|C5P4A1_COCP7         FFNIAARLARYTGND-T-YAEWAVRTWDWTKGVGLLTG----------DY
011 tr|Q0CG55|Q0CG55_ASPTN         FFNLAARLAKYTGNQ-T-YAEWADRVWNWTTSVNLMTN----------DW
012 tr|C4Y2X5|C4Y2X5_CLAL4         LFHMAARLARYTAND-S-YVEWAERVFDWMYGVNLLT--------EGEWW
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG FFNLAARLAKYTGNS-T-YADWADTVWDWTGEVGFMTD----------TY
```

```
014 tr|C7Z068|C7Z068_NECH7          FFNLGARLARYTGNS-T-YAKYAEETWDWLWDVQYIDH----------DDW
015 tr|G8BXX2|G8BXX2_TETPH          LFHIAARLARYTGND-S-YAEWAEKVYDWMVDVDLIT------QD-Q-YY
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR  LFHMSARLARYTGND-S-YVDWAEKVFDWMYGVGLLT--------EGEHW
017 tr|A1CSC2|A1CSC2_ASPCL          LFQLSARLARYTNNA-T-YADWAERIWDWSATTPLLKES---------DW
018 sp|Q05031|DFG5_YEAST            LFQIAARLGRYTGNT-T-YLEVAEQVFDWLVDVGYVVL--------N-DTA
019 tr|H2AQJ6|H2AQJ6_KAZAF          LFQLASRLGRYTGND-T-YLDVAESVYDWLVDVGFINVA------N-GTA
020 tr|G2XFH9|G2XFH9_VERDV          FFNIGARLARYTRND-T-YTQAAEETWDWLWGVNYIDH----------ENW
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH  LFHIAARLAHFTGNN-T-YAEWAEKVYNWMSDVNLIS------SG--DTK
022 sp|Q5AD78|DCW1_CANAL            LFHLAARLARYTGND-S-YVVWAERVWDWMYGVGLLT--------EQNWW
023 tr|B6H7E8|B6H7E8_PENRW          LFQLSARLALYTKNA-T-YAEWAEKIWDWSHSTPLLR-K---------TW
024 tr|B6GZU2|B6GZU2_PENRW          FFNLASRLAKYTNNQ-T-YQDWAEKAWDWTVDVGFMTP----------EY
025 tr|E9E4W7|E9E4W7_METAQ          FFNMGARLARYTGND-T-YADRAEKAWDWLWGVKYIDH----------DNW
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO  FFHLASRLAKYTGNQ-T-YSDWAEKAWDWISDVGFMTK----------DY
027 tr|G2R7G8|G2R7G8_THITE          FLNIGARLARYTKND-T-YAKWVDTAWDWIEGVGYITK----------DY
028 tr|G3JBY1|G3JBY1_CORMM          FFNVAARLARFTRND-T-YAEWAAKIYAWEAGIGLIGT----------NY
029 tr|A5DNV5|A5DNV5_PICGU          LFNIAARLARYTGND-S-YADWAEKVWEWMVGTHLIN------ETQPGYY
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI  YFNLGARLALYTRND-T-YADWATAAWNWMSEVGLIDE----------EY
031 tr|Q2TWC1|Q2TWC1_ASPOR          LFELSARLARFTKNE-T-YAEWAEKIWDWSASSPLILTD---------KW
032 tr|J5RXY9|J5RXY9_SACK1          LFHIAARLARYTGNQ-T-YVDWAEKVYEWMVGVNLIS------NG-T-YK
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA  FFNIAARLARFTGNA-T-YGEWAQRIYDWEVKEGLISD---------DY
034 tr|A0A254U5V1|A0A254U5V1_ASPNG  LFQLAARLGRYTNND-T-YTDWAEKIWDWSATTPLLQTE---------DW
035 tr|A0A061B6H7|A0A061B6H7_CYBFA  LFHIGARLARYTGNT-T-YVEWCDRVHDWMYDIGLIY------ES--DWY
036 tr|G8ZQ93|G8ZQ93_TORDC          LFNLAARLARYTGNQ-S-YVDWADKVYDWMYGVGLIS------NG-N-SR
037 tr|Q9C2J1|Q9C2J1_NEUCS          FFNLGARLARFTGNS-S-YGEWASRTWDWERSINLITD----------EY
038 tr|C5NZK5|C5NZK5_COCP7          FINLAARLALYTKNE-T-YVELAEKHWDWMTAIGLISP----------TS
039 tr|Q6C0T7|Q6C0T7_YARLI          LFNIGARLYRYTGNE-T-YLEWCDKIWNWAEAVDLVT----------KQY
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB  LFHIAARLARYTGNQ-S-YVDWAERVYDWMEDVHLID------NG-T-YR
041 tr|F7VZ72|F7VZ72_SORMK          FFNLGARLARYTGNK-T-YADWAEKTWDWMRGVGLMDD----------DY
042 tr|A3LN37|A3LN37_PICST          LFHIGARLARFTAND-T-YVDWAEKVWDWMYDIGLLDVVN---SNDKDYW
043 tr|Q6CAI2|Q6CAI2_YARLI          LFNLAARLYRYTGND-T-YLHWAEKAWDWCEKSGLLN------KED---H
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO  LFNLGARLARFTSND-T-YMEWAEKVWDWMYGIGLLT--------EGEHW
045 tr|G9MJS5|G9MJS5_HYPVG          FFNLGARLARYTGNT-T-YSDWAEKTWDWMVQIGFLDN----------STY
046 tr|G9MHI4|G9MHI4_HYPVG          FFNIAARLARYTGNQ-T-YADWALKVWQWEKQIGLIGP----------QY
047 tr|D4AP27|D4AP27_ARTBC          FFQMGARLARYTGNE-T-YAKWAEKTWDWSRAIGLINE----------NY
048 tr|B6GZT8|B6GZT8_PENRW          LFQLSARLARYTGDA-I-YLQWANKIWDWSVSSPLLSNT---------TW
049 sp|O74556|YCZ2_SCHPO            FFQLAARLARFTDND-T-YAEWANVAYDWSQRIGFIQE----------DY
050 tr|A0A0C4E991|A0A0C4E991_MAGP6  FFNIGARLARYTGNK-T-YAEHAEKTFDWLWAVNYIDH----------ENY
051 tr|B8MHG0|B8MHG0_TALSN          FFQIAARLARYTNND-T-YATWAETVWDWSSTVPLFDNS---------TW
052 tr|A3LMV8|A3LMV8_PICST          LFHLAARLARYLGND-T-YADVADKVWDWMENTVEFLTDV---D--G-TV
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU  VFALGARLARYTGDD-S-LKRSSKRILQWMKQSKFVVQN-----EGNDFY
054 sp|Q9P6I3|YHG7_SCHPO            LFQLAARLARYTNND-T-YVDLAQKVWDWSTTVGFVDLD---------DY
055 tr|Q2US57|Q2US57_ASPOR          LFQLAARLARYTNNQ-T-YADWAEKIWDWSASTPLLNTN---------TW
056 tr|W7N622|W7N622_GIBM7          FFNIAARLARYTGNT-T-YGDWAEKVWDWETKIGLINS----------DY
057 sp|Q9P6I4|YHG6_SCHPO            LFQLSSRLARFTNES-V-YSDWANKIWDWTTDVGFVNTT---------TY
058 tr|C7GRB3|C7GRB3_YEAS2          LFQIAARLGRYTGNT-T-YLEVAEQVFDWLVDVGYVVL--------N-DTA
059 tr|C7GP28|C7GP28_YEAS2          LFHIAARLARYTGNQ-T-YVDWAEKVYEWMVGVNLIS------NG-T-YK
060 tr|G3JGZ5|G3JGZ5_CORMM          FFNIGARLARYTKNE-T-YAKYADETWDWLWNVKYIDH----------ENW
061 tr|C4QXV4|C4QXV4_KOMPG          LFQLSARLARYTAND-T-YITLAEEAFDWMYGAGFLT------EG--DWW
062 tr|I1S0Z5|I1S0Z5_GIBZE          FFNLGARLARYTGNK-T-YADWAEKTWDWMEGVGFLDP----------TSY
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM  FFHLAARLARYTGNR-T-YAEWAERVWDWTVDVGFITD----------DW
064 tr|G0W5X4|G0W5X4_NAUDC          LFHLAARLARYTGNQ-S-YVDWAEKTYDWMSDVGLIS------NG-T-QK
065 tr|G8ZPC1|G8ZPC1_TORDC          LFQIAARLGRYTGND-T-YLEVAENVFDWLVDVGYVVL--------K-DVG
066 tr|A0A167CDC5|A0A167CDC5_9ASCO  FFHMGARLARYTGNT-T-YSDWSEKVWDWMTTVGFIN------ETS---G
067 tr|R9XAT7|R9XAT7_ASHAC          LFHMAARLARYTGNS-T-YLEWADRVYNWMYGVGLIN------VTNSYWH
068 tr|B2B747|B2B747_PODAN          FFNLGARLARYTGNT-T-YADWAEKTWDWMRGVGFMDE----------KY
069 tr|G0SE99|G0SE99_CHATD          FFNLGARLARYTGNQ-S-YADWAEKTWDWMVGVGYIDK----------DY
070 tr|J8Q6F0|J8Q6F0_SACAR          LFQLAARLGRYTGNT-T-YLDVAEQVFDWLVDVGYVVL--------N-DTA
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH  LFQLAARLGRYTGND-T-YLEVAESVFNWLLEVDFVVL--------K-DEA
072 tr|J3K6E2|J3K6E2_COCIM          FFNIAARLARYTGND-T-YAEWAVRTWDWTKGVGLLTD----------DY
073 tr|G4N3E1|G4N3E1_MAGO7          YMNLGARLARYTGNT-T-YSDAAEKTYDWLWSVNYIDH----------DTY
074 tr|C5M5J2|C5M5J2_CANTT          LFHLAARLARYTGNESN-YVEAAEKVWNWMDDVGFLTEED---N--G-NF
075 tr|A0A254U0X0|A0A254U0X0_ASPNG  LFELSARLARFTKND-T-YAEWAEKIFDWSTTTPLLNTN--------ATIW
076 sp|Q75DG6|DCW1_ASHGO            LFHMAARLTRYTGNA-T-YLEWAERVYDWMYGVGLIS------IVQPNWH
077 tr|Q6BZF0|Q6BZF0_DEBHA          LFNIGARLARYTSND-T-YMEWAEKVWDWMYGIGLLT--------EGEHW
078 tr|E9E3Q1|E9E3Q1_METAQ          FFNMGARLARYTNND-T-YAKRAEKAWDWMVGVGYIDS----------KSW
079 sp|P36091|DCW1_YEAST            LFHIAARLARYTGNQ-T-YVDWAEKVYEWMVGVNLIS------NG-T-YK
080 tr|W0TAH6|W0TAH6_KLUMD          LFHMASRLARYTGNS-T-YADWATKVYDWMQDVGMLT------DKDENSS
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA  MIAIGTRLGRYTGNT-T-YIDIANEAFEWLENIGFVS------ISD-GSA
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI  FFNIAARLARYTGNS-S-YSDWAAKIWNWEKDLLFINT----------DW
083 tr|Q2TYU3|Q2TYU3_ASPOR          LFQLAARLARYTNNQ-T-YTEWAEKVWDWSASSPLLNNQ---------TW
```

```
084 tr|G2XGF6|G2XGF6_VERDV          FFNIAARLARYTGNQ-T-YAEWAERVWDWEVKAGLITD----------EF
085 tr|G3AMT4|G3AMT4_SPAPN          LFHIAARLARYTSND-T-YVEWAERVWDWMYGVGLLT--------EGDWW
086 tr|A0A1C1D213|A0A1C1D213_9EURO  FFNMGARLAVYTGNQ-T-YADWADRMWDWVTAIGLRDE----------QY
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS  FFNLGARLARYTGNE-T-YAHHARKTWDWMMAVGLMTE----------DG
088 tr|A0A100I5A6|A0A100I5A6_ASPNG  LFELSARLARFTKND-T-YAEWAEKIFDWSETTPLLNTN-------ATIW
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH  LFQLAARLGRYTGND-T-YLDVAEKVFDWMVDVGYIVLS------DT--A
090 tr|G2Q8A7|G2Q8A7_MYCTT          FFNLGARLARYTGNK-T-YAEWAEKTWDWVRGVGYMDD----------QY
091 tr|B8NPZ1|B8NPZ1_ASPFN          LFELAARLARFTKNE-T-YSEWADRIWDWSASTPLLQTD---------RW
092 tr|Q7SAB2|Q7SAB2_NEUCR          FFNLGARLARYTGNE-T-YAHHARKTWDWMMAVGLMTE----------DG
093 tr|A2R8R5|A2R8R5_ASPNC          LFQLSARLARYTSNQ-T-YYDWAERIWDWSVTTPLLSNS---------TW
094 tr|C5DHG0|C5DHG0_LACTC          MFQLAARLGRYTGND-T-YLTKAETVWDWLVDVGFVVLN------ENGNA
095 tr|G0SFA3|G0SFA3_CHATD          FFNIAARLARYTGNQ-T-YAEWAARIWAWEEAIGLITS----------DY
096 tr|K1WJG5|K1WJG5_MARBU          LFNIAARLARYTGNQ-T-YADWAEKTWDWMETMGLISD----------DG
097 tr|H2B005|H2B005_KAZAF          LFHLAARLARYTGNT-S-YVDWAEKVYNWMEDTYLIS------NG-T-QM
098 tr|C5DYD7|C5DYD7_ZYGRC          YFQLAARLGRYTNNQ-T-YLEVAEKVFDWLTDVGFVVL-------K-DKG
099 tr|G8YM23|G8YM23_PICSO          LFNIAARLARYTGND-S-YSDWAEKIWEWFDSVGLLA------VSGDS-Y
100 tr|W3X8E3|W3X8E3_PESFW          FFNLGARLARYTNNE-T-YANWAETTWNWVRSVGFMDD----------EY
101 tr|A0A254U3K7|A0A254U3K7_ASPNG  LFQLSARLARYTSNQ-T-YYDWAERIWDWSVTTPLLSNS---------TW
102 tr|Q6CER8|Q6CER8_YARLI          LFNLASRLYRYTGNT-T-YLEWAEKIWTWSEKVKIID----------G-G
103 tr|G2QT05|G2QT05_THITE          FFNLGARLARYTGNS-T-YAEWAEKTWDWVRAVGFMDD----------KY
104 sp|Q5ACZ2|DFG5_CANAL           LFHLAARLARYTGNSSV-YVDTAEKVWKWMEDVGFLTEED---N--G-DV
105 tr|G2WKU6|G2WKU6_YEASK          LFQIAARLGRYTGNT-T-YLEVAEQVFDWLVDVGYVVL-------N-DTA
106 tr|A0A0C7N752|A0A0C7N752_9SACH  LFQLAARLGRYTGNE-T-YIAVAEKIFDWLVDVGYVVLQ------EN--A
107 tr|W3WN68|W3WN68_PESFW          FFNLASRLARYTGND-T-YGDWAERIFAWEQGVDFINS----------DW
108 tr|G2QFS1|G2QFS1_MYCTT          FFNLGARLARYTENQ-T-YADWASKTFQWLWDVGYIDH----------ESW
109 tr|G3J9G4|G3J9G4_CORMM          FFNIAARLARFTGNS-T-YADWAQKIYDWEVKEGLITS----------DY
110 tr|Q2UJ03|Q2UJ03_ASPOR          FFNLAARLARYTGNQ-T-YADWAVRVWDWTESVGFITE----------DY
111 tr|G8YRT2|G8YRT2_PICSO          LFQIAARLARYTGND-T-YAKTATKVWDWMHSVGFIGEDD---N--HENM
112 tr|W3X554|W3X554_PESFW          FFNVASRLARFTGNQ-T-YADWAEKVYDWMTDTGLITA----------EF
113 tr|Q1K7A8|Q1K7A8_NEUCR          FFNLGARLARFTGNS-S-YGEWASRTWDWERSINLITD----------EY
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM  LFQLAARLARYTNNE-T-YSQWAERIWDWSATTPLLKES---------DW
115 tr|A7F3P0|A7F3P0_SCLS1          FFNLGARLARYTGNS-S-YAEWAERIFTWEQSVGLIGD----------GY
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA  FFNVAARLARFTRND-T-YAEWATKIYEWQEGIGLIGK----------NY
117 tr|Q6CIP9|Q6CIP9_KLULA          LFQLAARLGRYTGNT-T-YLEVAEKVFDWLVDIDFVVM-------K-EEA
118 tr|H2AMQ5|H2AMQ5_KAZAF          LFHLAARLARYTGND-S-YVDWAEKVYNWMEDTGFIT------TG-N-TL
119 tr|G8JMI3|G8JMI3_ERECY          LFQLAARLGRYTGNS-T-YLDIANETFSWLVDVKFVVL-------K-DEA
120 tr|C5DGT1|C5DGT1_LACTC          LFHLAARLARYTGNN-T-YVDWAEKVFDWMKGIGLLT------EN--DNS
121 tr|K1XJR4|K1XJR4_MARBU          FFNIAARLARYTDNS-T-YYDHAVQTWDWITSVGFIDK----------DY
122 tr|A0A124BYC5|A0A124BYC5_ASPNG  LFQLAARLARYTNNQ-T-YVDWANTIWDWSTV-ILVNQE---------SW
123 tr|A7TLL2|A7TLL2_VANPO          LFHIGARLARFTGND-S-YVEWAEKVYDWMESVNLIE------DD-GKGA
124 tr|Q0CB86|Q0CB86_ASPTN          LFQLAARLARYTNNQ-T-YADWADKIWDWSASTPLLNNQ---------TW
125 tr|K1XLH0|K1XLH0_MARBU          LFNIAARLALYTGNQ-T-YADWAVKTWDWMSAVGLLNE----------KY
126 tr|J7S438|J7S438_KAZNA          LFHMASRLARYTGNQ-S-YVDWAEKVYDWMWDVHVIS------NS--TYM
127 tr|Q4WG09|Q4WG09_ASPFU          LFELSARLARFTKNE-T-YAEWANKIWDWSASTPLLQTD---------RW
128 tr|H8WZ09|H8WZ09_CANO9          LFHIAARLARYTGNESA-YLPTAERVWNWMEEVNFLTEED---N--G-NL
129 tr|I1RL09|I1RL09_GIBZE          FFNLGARLARYTGNE-T-YAKHAEETWEWLWGVNYIDH----------DRW
130 tr|Q4W985|Q4W985_ASPFU          FFHLAARLARYTGNR-T-YAEWAERVWDWTVDVGFITD----------DW
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU  LFQIAARLGRYTGNT-T-YLDVAEQVFDWLVDVGYVVL-------N-DTA
132 tr|W6QCW3|W6QCW3_PENRF          FFNLASRLAKYTNNQ-T-YHDWAEKAWNWTAEVGFMTP----------EY
133 tr|W3WMD3|W3WMD3_PESFW          YFNIGARLARYTNND-T-YMELAGRTFDIMEKLGYVDA----------DW
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH  LFHIAARLAHFTGNN-T-YAEWAEKVYNWMSDVKLIN------SG--DTK
135 tr|I2GYH1|I2GYH1_TETBL          LFQMAARLGRYTGNN-T-YLDVANDVFDWLTGVGYVNM-------D-NAA
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH  LFQLAARLGRYTNNQ-T-YLHIAEEVFDWLHDVGYVVL-------S-DKG
137 tr|Q752P3|Q752P3_ASHGO          LFQLAARLGRYTGNQ-T-YLDVASKVFQWLVDVKYIVL-------E-DKA
138 tr|B8MYP3|B8MYP3_ASPFN          LFQLAARLGRYTKNQ-T-YIDWAEKIWDWSATTPLLKTA---------DW
139 tr|Q2TXL6|Q2TXL6_ASPOR          FFQLAARLARYTKNE-T-YTEWAEKAFTWATSVPLIIEK---------GW
140 tr|Q0CN17|Q0CN17_ASPTN          LFQLAARLGRYTNNH-T-YTDWAEKIWDWSATTPLLRTT---------DW
141 tr|J3NHD5|J3NHD5_GAGT3          FFNLGARLARYTTNK-T-YADWAEKQWDWVTSVGLMDA----------QY
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI  FFNIGSRLARYTGND-T-YANWAEKTWDWTQKIGLIDA----------QY
143 tr|A1CVB0|A1CVB0_NEOFI          FFHLAARLARYTGNS-T-YAEWAERVWDWTVDVGFITD----------DW
144 tr|B8NVI3|B8NVI3_ASPFN          FFQLAARLARYTKNE-T-YTEWAEKAFTWATSVPLIIEK---------GW
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH  LFHIAARLARYTGNT-S-YVDWAEKVFEWMKGIHLLE------E--DGTN
146 tr|A1DJ54|A1DJ54_NEOFI          LFELSARLARFTKND-T-YAQWADKIWDWSASTPLLQMD---------RW
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS  FFNMGARLARYTRND-T-YATWATKQFQWIYDVNYIDH----------DSW
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA  LFHIGARLARYTGNT-S-YVDWCDKVHDWMYDTGLIH------ES--DWY
149 tr|A0A177A618|A0A177A618_9PEZI  MFNIASRLAVYTGNS-T-YAEWAEKTWDWVESIGLIND----------KG
150 tr|Q75CW6|Q75CW6_ASHGO          LFQIAARLARFTGNS-T-YQDVAETVFDWLVDSKLVQL-------E-NDA
151 tr|Q7S4K4|Q7S4K4_NEUCR          FFNMGARLARYTRND-T-YATWATKQFQWIYDVNYIDH----------DSW
152 tr|G2QB99|G2QB99_MYCTT          FFNIAARLARYTGNE-T-YAEWAEKIFAWEESVGLIDA----------NL
```

```
153 tr|G0S3F2|G0S3F2_CHATD          FFDLGARLARYTKNN-T-YAEWAEKIFDWLYAVGYIDH----------ETW
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO LFHIGARLARYTGND-S-YADWSEKVWDWMIDSDLII--------PGPYW
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB LFQIAARLGRYTGNT-T-YLDVAERVFDWLVGVGYIVL--------S-EKG
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM LFQLAARLARYTNNH-T-YYEWAEKVWDWSCSSPLLNNK---------TW
157 tr|G2QKJ0|G2QKJ0_MYCTT          FMNIAARLARYTNND-T-YALWATKAWDWIEAMGYVTD----------KY
158 tr|Q4WFX5|Q4WFX5_ASPFU          LFNIAARLARYTGDA-T-YAKWANKVWDWVTETGLIGR----------EY
159 tr|Q5BGD7|Q5BGD7_EMENI          LFQLAARLGRFTNNQ-T-YFDWAEKIWDWAAASPLIDTN---------TW
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH LFHLAARLARYTGND-T-YTEWAEKVYDWMQDIDLLI------DEEPNQY
161 tr|G8BYN9|G8BYN9_TETPH          LFQIAARLGRYTGNN-T-YLEVAEKVFDWMVDVEYIVL--------K-DVA
162 tr|F7VVT4|F7VVT4_SORMK          FFNLGARLARYTDNS-T-YSKWAEKTYNWIRDVGYIDK----------NW
163 tr|G8YB80|G8YB80_PICSO          LFNIAARLARYTGND-S-YSDWAEKVWEWFDSAGLLA------DSGDS-Y
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA LFNLAGRLARFTRND-T-YVETAQTAYDWLVDIGFIV------DSN-GMA
165 tr|G2WRS7|G2WRS7_VERDV          FFNIGARLARYTKNS-T-YADWADKTWDWMTNVGFIDS----------EY
166 tr|A0A124BVB5|A0A124BVB5_ASPNG FFNMAARLARYTGNA-T-YADWANKVWDWVTDVGLIGP----------EY
167 tr|A7EIG7|A7EIG7_SCLS1          FFNLGARLAKYTGND-T-YAQWAERTWDWMENVGLMDE----------NY
168 tr|C5M5U7|C5M5U7_CANTT          LFHLAARLARYTGND-S-YVVWAERVWDWMYGVGLLT--------EGNWW
169 tr|B9WAA8|B9WAA8_CANDC          LFHLAARLARYTGNSSV-YVDTAEKVWKWMEDVGFLTEEQ---N--G-DV
170 tr|F8MRU1|F8MRU1_NEUT8          FFNMGARLARYTRNE-T-YATWATKQFQWIYDVNYIDH----------DSW
171 tr|A1DJS0|A1DJS0_NEOFI          LFQLAARLARYTNNH-T-YYEWAEKVWDWSCSSPLLNNK---------TW
172 tr|B6HLM8|B6HLM8_PENRW          FFMLAARLAWYTQDE-E-YATWAEKVWDWSTSVNLVNNE---------TW
173 tr|A0A117DWS0|A0A117DWS0_ASPNG LFQLAARLGRYTNND-T-YTDWAEKIWDWSATTPLLQTE---------DW
174 tr|A0A124BXM1|A0A124BXM1_ASPNG LFQLAARLARYTSND-T-YADWAEKIFDWCASTPLLNNQ---------TW
175 tr|J3NZQ6|J3NZQ6_GAGT3          YFNIGARLTRYTGNS-T-YAEHAEKIWDWMWALGYIDH----------ESW
176 tr|W0THH8|W0THH8_KLUMD          LFQLAARLGRYTGNQ-T-YLDVAEKVFNWLTDVNYVVL--------K-DEA
177 tr|B2AEF2|B2AEF2_PODAN          FFDLGARLAAYTFNQ-T-YADWADKTFQWLWDVGYIDN----------KDW
178 tr|Q96TX1|Q96TX1_NEUCS          FFNLGARLARYTGNK-T-YADWAETTWDWMRGVGLMDD----------NY
179 tr|E9DYA6|E9DYA6_METAQ          FFNMAARLARYTGDA-K-YSDWAEKTWDWVEKIGFIDP----------QNY
180 tr|B9WAI2|B9WAI2_CANDC          LFHLAARLARYTGND-S-YVVWAERVWDWMYGVGLLT--------EQNWW
181 tr|Q6C171|Q6C171_YARLI          LFNLGARLYRYTGNV-T-YLEWAERIWDWSTQIEIVD----------N-G
182 tr|A6ZMV1|A6ZMV1_YEAS7          LFQIAARLGRYTGNT-T-YLEVAEQVFDWLVDVGYVVL--------N-DTA
183 tr|B6K7X7|B6K7X7_SCHJY          LFQLAARIARFTAND-T-CAEFAEKVWDWTTTVGFLNE----------STF
184 tr|I1RSA2|I1RSA2_GIBZE          FFNVASRLARYTGND-T-YADWAGKVFDWHIKAGIITD----------KF
185 tr|F7VWZ9|F7VWZ9_SORMK          FFNLGARLSRFTGNS-S-YGDWATRTWDWERSINLITD----------EY
186 tr|A0A1B2J789|A0A1B2J789_PICPA LFQLSARLARYTAND-T-YITLAEEAFDWMYGAGFLT------EG--DWW
187 tr|G0V8N4|G0V8N4_NAUCC          LFHMAARLARYTGNQ-S-YVDWAERVYDWMSDVGLIS------NG-S-YK
188 tr|H8WXN8|H8WXN8_CANO9          LFHLAARLARYTQND-S-YVEWAQKVWDWMYGVGLLT--------EGDHW
189 tr|A6ZZS0|A6ZZS0_YEAS7          LFHIAARLARYTGNQ-T-YVDWAEKVYEWMVGVNLIS------NG-T-YK
190 tr|G9MQD1|G9MQD1_HYPVG          FFNIGARLARYTRNE-T-YAKRAEDTWNWLWGVGYIDH----------KSW
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH LFHIAARLARYTGNQ-S-YVDWAEKVFDWMKGINLLT------ENTGGNN
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM LFELSARLARFTKNE-T-YAEWANKIWDWSASTPLLQTD---------RW
193 tr|A0A254UET1|A0A254UET1_ASPNG FFNIAARLARYTGNA-T-YADWANKVWDWVTDVGLIGP----------EY
194 tr|A1D587|A1D587_NEOFI          LFQLAARLARYTNNE-T-YSQWAERIWDWSATTPLLKES---------DW
195 tr|Q1K7I4|Q1K7I4_NEUCR          FFNLGARLARYTGNK-T-YADWAETTWDWMRGVGLMDD----------NY
196 tr|G8JS88|G8JS88_ERECY          LFHLAARLARYTGNQ-S-YADWAEKVYDWMASVQLIN------QT-GDWT
197 tr|A0A136J7D3|A0A136J7D3_9PEZI FFNLASRLARYTGNS-T-YADWAGRVWTWLEGHELIDK----------DF
198 tr|E9E4X8|E9E4X8_METAQ          FFNMGARLARYTGNT-T-YSDWADKTWDWMWNIGFIDN----------KNY
199 tr|Q6CP42|Q6CP42_KLULA          LFHLSARLARYTGNS-S-YTEWAEKIYDWMQDIQLLV------DEEPTQY
200 tr|G0WHL4|G0WHL4_NAUDC          LFQLAARLGRYTGNT-T-YLEVAEKVFTWLVDVGYVLL--------D-DTA
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU LFHIAARLARYTGNQ-T-YVDWAEKVYEWMVGVDLIS------NG-T-YK
202 tr|F9WYX6|F9WYX6_ZYMTI          FASLAARLHRFTGNS-TFYGEWATRAVEWTRSIGLMNQN-------GQT-
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 FFNLGARLARYTLNT-T-YADWAEKQWDWITSVGLMDS----------QY
204 tr|J3P147|J3P147_GAGT3          YFNLGARLARYTGNQ-T-YLKEAERTWDWLWAVKYIDH----------DTY
205 tr|I2H842|I2H842_TETBL          LFHIAARLARYTGNQ-S-YVDWSEKVYDWMLGTGLIS------NG-TTYK
206 tr|S6E3R3|S6E3R3_ZYGB2          LFQLAARLGRYTNNQ-T-YLHIAEEVFDWLHDVGYVVL--------S-DKG
207 tr|E5AD94|E5AD94_LEPMJ          LFHLGARLAMYTKND-T-YAQWAEKAFDWLSQSPILPG----------DG
208 tr|Q6FJM7|Q6FJM7_CANGA          LFQIAARLGRYTGNT-T-YLDVAERVFDWLVGVGYIVL--------S-EKG
209 tr|A5DUW0|A5DUW0_LODEL          LFHLAARLARFTGND-S-YVEWAERVWNWMYGAGLLT--------EGNEW
210 tr|G8BF46|G8BF46_CANPC          LFHLAARLARYTQND-S-YVEWAQKTWDWMYGVGLLT--------EGDHW
211 tr|W3XAF6|W3XAF6_PESFW          FFNIASRLARYTGND-T-YREWADKIFDWELNVGFISD----------TW
212 tr|B6H7I2|B6H7I2_PENRW          LFQLAARLARYTDDD-K-YSKWAEKIWDWSTSSPLVNNK---------TW
213 tr|Q5ATF9|Q5ATF9_EMENI          LFQLAARLARYTNND-T-YAEKAQMVWDWVVSSPLVNNK---------TW
214 tr|A0A167CDD3|A0A167CDD3_9ASCO LFHIGARLARYTGNN-S-YVDWAERVWNWAEGVNFLN------QTFVNQT
215 tr|G0RXS3|G0RXS3_CHATD          FMNIASRLARYTRND-T-YARWAETAWDWTQGVGYITK----------EY
216 tr|B8NX26|B8NX26_ASPFN          LFELAARLARFTKNE-T-YTEWAEKIWDWSAKSGLMDVN---------KW
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM FINLAARLALYTKNE-T-YVELAEKHWDWMTAIGLISP----------TS
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM LFNIAARLARYTGDA-T-YAEWANKVWDWVTETGLIGP----------EY
219 tr|A1CMM3|A1CMM3_ASPCL          FFNLASRLAVYTGNA-T-YAEWAEKTWDWVVRIGLTNE----------AY
220 tr|B6HU84|B6HU84_PENRW          LLQLSARLALYTGNK-T-YADWAERIWDWSATTPLLKQK---------NW
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS FFNLGARLARYTGNK-T-YADWAETTWDWMRGVGLMDD----------NY
```

```
222 tr|G2WHY6|G2WHY6_YEASK          LFHIAARLARYTGNQ-T-YVDWAEKVYEWMVGVNLIS------NG-T-YK
223 tr|G0W7G8|G0W7G8_NAUDC          LFHMAARLARYTGNQ-S-YVDWAEKVYDWMYEVGLVS------NG-S-SM
224 tr|G4NGV8|G4NGV8_MAGO7          FFNLAARLARFTGNQ-T-YAEWANKIWDWERSTNLMTE---------KF
225 tr|W3X855|W3X855_PESFW          FFNIAARLARYTGNT-T-YADWAETVYDWMETVGFIDD---------EY
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA FFNMGARLGRYTGNQ-T-YIDWAEKTFDWVTGVGYIDP---------QTH
227 tr|J8Q2K7|J8Q2K7_SACAR          LFHIAARLARYTGNS-T-YVDWAEKVYEWMVSVNLIS------NG-T-YK
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO LFNIAARLARYTSND-S-YVEWAEKVWDWMYGIGLLS--------EGDWW
229 tr|J6EGN8|J6EGN8_SACK1          LFQIAARLGRYTGNS-T-YLDVAEQVFDWLVDVGYVVL-------N-DTA
230 tr|A0A100ISD9|A0A100ISD9_ASPNG FFNLAARLAKYTGNS-T-YADWADTVWDWTGEVGFMTD---------TY
231 tr|G8YQC0|G8YQC0_PICSO          LFQIAARLARYTGND-T-YAKTAEKVWDWMSDVGFIGEDD---N--HENM
232 tr|F7VVP8|F7VVP8_SORMK          FFNLGARLARYTRNE-T-YATWATKQFEWLYNVNYIDH---------ESW
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH LFQIAARLGRYTGNT-T-YLEVAEQVFDWLVDVGYVVL-------N-DTA
234 tr|F8MBP4|F8MBP4_NEUT8          FFNLGARLARFTGNS-S-YGEWASRTWDWERSINLITD---------EY
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG LFQLAARLARYTNNQ-T-YVDWANTIWDWSTS-HLVQEG--------SW
236 tr|A1CD27|A1CD27_ASPCL          LFELSARLARFTKNG-T-YADWAEKIWDWSVTTPLLQTD---------RW
237 tr|A5DHF3|A5DHF3_PICGU          LFHISARLARYLKND-T-YAKTAEKVWDWMEELFF--HED---D--G-EL
238 tr|D4ATT7|D4ATT7_ARTBC          FFHLAARLARYTNDE-S-YSKWADKAWRWMTAIKLISP---------TF
239 tr|E2PT42|E2PT42_ASPNC          FFNLAARLAKYTGNS-T-YADWADTVWDWTGEVGFMTD---------TY
240 tr|C7ZPE5|C7ZPE5_NECH7          FFNLGARLARYTGNT-T-YADWAETTWEWMESVGFLDP---------TSY
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA FFNMAARLATYTGNA-T-YAEWAEIAWDWMESVGFVTE---------NF
242 tr|Q4WKP7|Q4WKP7_ASPFU          LFQLAARLARYTNNE-T-YSQWAERIWDWSATTPLLKES--------DW
243 tr|F8MXT4|F8MXT4_NEUT8          FFNLGARLARYTGNK-T-YADWAETTWDWMRGVGLMDD---------NY
244 tr|A0A177AJ74|A0A177AJ74_9PEZI FFNLGSRLARYTGND-T-YSNWAEKTWEWTQSIGLIDA---------QF
245 tr|J7RQR5|J7RQR5_KAZNA          LFNLAARLGRYTGNS-T-YLDVADQVFDWMVSVGYVVL-------G-DTA
246 tr|Q75CW7|Q75CW7_ASHGO          LFQIAARLARFTGNS-T-YQDVAETVFDWLVDVKLVEL-------E-NDA
247 tr|G3AKK4|G3AKK4_SPAPN          LFHLAARLARYTGNSTV-YLDVAEKVWDWMDEVGFLTEED---N--G-NF
248 tr|B2W762|B2W762_PYRTR          FFNIAARLYKYLGNE-T-YADWAEKAWKWELSVGLMSA---------DY
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH LFHIAARLARYTGNQ-T-YVDWAEKVYEWMVGVNLIS------NG-T-YK
250 tr|C4XWK1|C4XWK1_CLAL4          LFHLAARLYRFTGEK-V-YLETAEKVWNWMWDVGFIR--D---S--P-EF

001 tr|A0A254U2J9|A0A254U2J9_ASPNG NVADSTDV-------DN-SCST-QGNNQWSYNYGVF------LTGAAYMY
002 sp|Q6FLP9|DCW1_CANGA          YVYDGVS----I---ND-NCTT-VTKYQWTYNQGLM------LSGSAYLF
003 tr|A1CCM5|A1CCM5_ASPCL        NVADSTTI-------ED-GCTS-QGNNQWSYNYGTY------LMGAAYMW
004 tr|C1H190|C1H190_PARBA        QFFDGTD----D---TT-NCTR-VNHVQWSYNSGVF------LAGAASMY
005 tr|Q2UR85|Q2UR85_ASPOR        NIADTTTS-------EA-NCKD-HGDLQWTYNYGTY------LSGAAYMY
006 tr|A5DV30|A5DV30_LODEL        RVYDGAE----V---DD-NCTA-VTKLRWSYTYGVF------MAGCAYLY
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO NVYDGAD----I---VNGSCPV-VNGAIWSYNYALM------TMGSAYLY
008 tr|J3PGN0|J3PGN0_GAGT3        RVYDGGH----V---NN-NCTD-IHKAQFSYNIGVL------LQGCAFMY
009 tr|A0A124BXU6|A0A124BXU6_ASPNG NVADSTST-------TN-DCST-QGDNQWSYNYGAY------LGGAAYMY
010 tr|C5P4A1|C5P4A1_COCP7        QFFDGSD----E---KL-NCSE-VNRIQWTYNAGVY------LLGAASMY
011 tr|Q0CG55|Q0CG55_ASPTN        LFYDGTD----D---QS-NCTS-YNQIEWTYNSGVY------LLGAAHMY
012 tr|C4Y2X5|C4Y2X5_CLAL4        FVYDGVK----I---AN-NCSQ-VTKLQWTYNQGLM------LAGCAYLY
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG RFWDGAD----V---SS-GCGD-WNYIEWTYNTGVY------LLGAAVMY
014 tr|C7Z068|C7Z068_NECH7        RVYDGGH----V---EK-NCTD-INKATFSYNAAIM------LQGAAFMY
015 tr|G8BXX2|G8BXX2_TETPH        FVYDGVD----S---ND-NCTT-ITKYQWTYNQGLI------LSGCAYLY
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR FVYDGTH----A---PD-NCSN-VTKWQWTYNQGLM------LAGCAYLW
017 tr|A1CSC2|A1CSC2_ASPCL        TIADTTSP-------DK-QCTD-HGDLQWTYNYGTY------ISGAAYMF
018 sp|Q05031|DFG5_YEAST          NVFDGAE----I---DT-NCTD-ITKIEWTYNHGIV------LGGLAYMY
019 tr|H2AQJ6|H2AQJ6_KAZAF        YVYDGAD----I---SE-NCTD-ITTIDWSYNHGIL------LGGAAYAY
020 tr|G2XFH9|G2XFH9_VERDV        RVYDGGH----I---EH-NCTD-INKATFSYNIAVL------MQGAAFMY
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH YVYDGVD----I---TN-NCSGPVHKYQWTYNQGLL------MSGCAYLY
022 sp|Q5AD78|DCW1_CANAL          FVYDGVK----I---AN-NCSN-ITKYQWSYNQGLM------LAGCAYLY
023 tr|B6H7E8|B6H7E8_PENRW        VIADTTTN-------AA-NCKD-HGDHQWTYNYATY------IMGAGYMH
024 tr|B6GZU2|B6GZU2_PENRW        HFLDGAS----C---LT-ECKP-INQIEWTYNAGVY------LLGAANMY
025 tr|E9E4W7|E9E4W7_METAQ        AVYDGAS----V---ND-NCTD-IHKTQYSYNAGVL------IQGVAFMY
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO HFWDGAS----D---LT-DCKE-INQIEWTYNNGVF------LLGAANMY
027 tr|G2R7G8|G2R7G8_THITE        NVLDGGH----T---EE-NCTD-INPVQFSANMAVL------LHGSAIMY
028 tr|G3JBY1|G3JBY1_CORMM        EVYDGYN----VIDATQ-TCDR-LHDVQFSYNAGIW------LAGAAAMY
029 tr|A5DNV5|A5DNV5_PICGU        FVYDGVT----V---ND-NCSE-VVKLQWSYNQGLF------LSGAAVLY
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI AVFDGSD----E---LK-HCAS-INRVQWSYNVGIL------LNGAAAMY
031 tr|Q2TWC1|Q2TWC1_ASPOR        YVADSTSN-------EN-NCKD-SGNNQWTYNYGTF------LSGAAFMY
032 tr|J5RXY9|J5RXY9_SACK1        YVYDGVS----I---GD-NCTK-VTSYQWTYNQGLL------LAGSAYLY
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA IIYDGIE----VDDATK-SCKN-IHKVQFTYNAGVW------LQGAAAMY
034 tr|A0A254U5V1|A0A254U5V1_ASPNG YIADTTTT-------EA-NCKD-HGDIQWTYNYGTY------LSGAAYMY
035 tr|A0A061B6H7|A0A061B6H7_CYBFA FVYDGIS----I---TE-NCSD-ITKLQWTYNQGLM------MSGSAYLY
036 tr|G8ZQ93|G8ZQ93_TORDC        FVYDGVN----I---AH-NCAN-ITGYQWSYNHGLM------LSGCAYLY
037 tr|Q9C2J1|Q9C2J1_NEUCS        DVKDGAH----FDVTTH-VCRNDSGPHVWSYNIGVF------LQGAAFMY
038 tr|C5NZK5|C5NZK5_COCP7        QVFDGSD----V---VK-NCSE-LSHIQWTYNNGVL------MGGAAALY
```

```
039 tr|Q6C0T7|Q6C0T7_YARLI          EVYDGIETGSGN---VP-NCTN-RTPYIWTYNSGIF------MHGAAALY
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB YVYDGVS----I---ND-NCTT-VTKYQWTYNQGLM------LSGSAYLF
041 tr|F7VZ72|F7VZ72_SORMK          NIHDGAH----V---EK-NCTD-VNGAQFSYNFGVF------TLGAAHMY
042 tr|A3LN37|A3LN37_PICST          FVYDGLT----I---DN-NCSN-ITKYQWTYNQGLM------LSGCAYLY
043 tr|Q6CAI2|Q6CAI2_YARLI          YHIRDGV----D---VN-NCTN-ITPYVWTYNAGLF------IGGSAFLY
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO YVYDGVT----V---GG-NCSD-ITKLQWTYNQGLM------MSGCAFLY
045 tr|G9MJS5|G9MJS5_HYPVG          AIYDGAD----V---SN-NCTQ-INKAEFSYNSAVW------VLGTAYMY
046 tr|G9MHI4|G9MHI4_HYPVG          QIYDGLS----INTD---NTCGHMDQNQWSYNAGIY------MHGAAAMY
047 tr|D4AP27|D4AP27_ARTBC          QFFDGSS----D---TL-NCTE-LNRLQWTYNAGVY------LLGAAAMY
048 tr|B6GZT8|B6GZT8_PENRW          NVADSAVV--------GD-NCAT-QGNAQWTYNYGTY------LMGAAYMY
049 sp|O74556|YCZ2_SCHPO            TVFDGSS----I---KD-NCSS-IEITQWTYNIGLY------MAGAAYMY
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 KVYDGGH----V---NN-NCTD-IHKAQFSYNIAVL------LQGCAFMY
051 tr|B8MHG0|B8MHG0_TALSN          KIYDMTRI--------SE-NCKD-QDVMQWSYNYGSY------VTGAAYMY
052 tr|A3LMV8|A3LMV8_PICST          TLYDGAN----I---DT-NCTD-LTKKQWTYTYGIF------MAGAAYMY
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU NVNDGAE----I---VHGACPV-VNGALWSYNYALM------LMGSAYLY
054 sp|Q9P6I3|YHG7_SCHPO            TVYDGASV--------TS-NCSS-ITNEQWSYNVGVY------LAGTAFLY
055 tr|Q2US57|Q2US57_ASPOR          NVADSTDV--------DD-NCTS-QGNTQWTYNYGAY------IGGAAYMY
056 tr|W7N622|W7N622_GIBM7          QVRDGVH----FEGKCP---SSMDTNQWTYNSGVY------LYGAAAMY
057 sp|Q9P6I4|YHG6_SCHPO            AVYDGADT--------ST-NCTT-LDPSQWSYNIGIF------MVGAAYMY
058 tr|C7GRB3|C7GRB3_YEAS2          NVFDGAE----I---DT-NCTD-ITKIEWTYNHGIV------LGGLAYMY
059 tr|C7GP28|C7GP28_YEAS2          YVYDGVS----I---DD-NCTK-VTSYQWTYNQGLL------LAGSAYLY
060 tr|G3JGZ5|G3JGZ5_CORMM          LVYDGGR----V---PF-NCTD-INKATFSYNAAIL------TQGAAFMY
061 tr|C4QXV4|C4QXV4_KOMPG          FVYDGAF----V---DD-NCTE-IVMLQWTYNAGLM------VSGCAYLA
062 tr|I1S0Z5|I1S0Z5_GIBZE          KIYDGGH----V---GK-NCTD-INKAQFSYNSGVF------LQGAAFMY
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM LFYDGAD----V---LL-NCSD-LNRIEWTYNSGVY------LLGAANMY
064 tr|G0W5X4|G0W5X4_NAUDC          YVYDGTS----I---ET-NCTT-ITSYQWSYNQGLM------LAGAAYLY
065 tr|G8ZPC1|G8ZPC1_TORDC          NVFDGAE----I---DN-NCTS-INKLEWSYNHGVV------LGGLAFMY
066 tr|A0A167CDC5|A0A167CDC5_9ASCO QLIDGAD----I---ST-NCTT-HTPYEWTYNYGLF------TAGAAYLY
067 tr|R9XAT7|R9XAT7_ASHAC          FVFDGTA----I---DH-NCTV-IDRLQWSYNHGLI------MAGCAFIY
068 tr|B2B747|B2B747_PODAN          NIYDGGH----V---EH-NCTD-INRAQFSYNNGVF------LLGAAYMY
069 tr|G0SE99|G0SE99_CHATD          NIYDGGH----V---EQ-NCTD-INRAQFSYNNAIF------LLGAAYMY
070 tr|J8Q6F0|J8Q6F0_SACAR          NVFDGAE----I---DT-NCTD-ITKIEWTYNHGIV------LGGIAYMY
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH NVYDGAT----I---ED-NCTD-IVKYEWSYNHGVV------LGGCAYMY
072 tr|J3K6E2|J3K6E2_COCIM          QFFDGSD----E---KL-NCSE-VNRIQWTYNAGVY------LLGAASMY
073 tr|G4N3E1|G4N3E1_MAGO7          KVYDGGH----V---EH-NCTD-INKAQFSYSIAVL------TQATAFLW
074 tr|C5M5J2|C5M5J2_CANTT          RMYDGAK----I---EN-NCSS-VTDLRWSYTYGVF------MAGCAYLY
075 tr|A0A254U0X0|A0A254U0X0_ASPNG NVADSTSN--------EA-NCKD-VGNNQWTYNYGTY------LSGAAFMY
076 sp|Q75DG6|DCW1_ASHGO            VVYDGTD----I---ND-NCTN-LNKLQWTYNHGLI------MAGCAFIY
077 tr|Q6BZF0|Q6BZF0_DEBHA          YVYDGVT----V---GG-NCTD-ITKLQWSYNQALM------LSGCAFLY
078 tr|E9E3Q1|E9E3Q1_METAQ          HAYDGAH----V---AQ-NCTD-INKATFSYNAGVL------VQGAAFLY
079 sp|P36091|DCW1_YEAST            YVYDGVS----I---DD-NCTK-VTSYQWTYNQGLL------LAGSAYLY
080 tr|W0TAH6|W0TAH6_KLUMD          FVYDGAS----I---EN-NCTK-IVKYQWTYNQGLI------LSGSAYMY
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA RVYDGAS----I---TD-NCTS-ITGLEWTYNIGLL------MAGSAYAY
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI EVEDGAH----T---DD-ECKD-MTKVLWSYNFGIH------MHAAANMY
083 tr|Q2TYU3|Q2TYU3_ASPOR          NVADSTDI--------EG-GCKS-QGNNQWSYNYGTY------LIGAAYMY
084 tr|G2XGF6|G2XGF6_VERDV          KVYDGVE----VVENSP-DCPK-IDKNQWSYNAGIY------LHGAAVMH
085 tr|G3AMT4|G3AMT4_SPAPN          FVYDGAK----V---AN-NCSN-ITKYQWTYNQGLM------LAGCAYLY
086 tr|A0A1C1D213|A0A1C1D213_9EURO NFFDGSD----D---NL-NCTE-LDHIQWTYNAGTF------LVGAANMY
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS NVYDGAH----T---GV-NCTD-VFKAQFSYNAAIF------LQGAAFMY
088 tr|A0A100I5A6|A0A100I5A6_ASPNG NVADSTSN--------EA-NCKD-IGNNQWTYNYGTY------LSGAAFMY
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH NVYDGAT----I---DE-NCTD-ITKLEWSYNQGIV------LGGAAYMY
090 tr|G2Q8A7|G2Q8A7_MYCTT          NIYDGGH----V---QH-NCTD-INRAQFSYNNAVY------LLGAAFMY
091 tr|B8NPZ1|B8NPZ1_ASPFN          YIADSTSN--------LN-NCSD-AGDQQWSYNYGTY------LAGAAFMY
092 tr|Q7SAB2|Q7SAB2_NEUCR          NVYDGAH----T---GV-NCTD-VFKAQFSYNAAIF------LQGAAFMY
093 tr|A2R8R5|A2R8R5_ASPNC          NVADSTST--------TN-DCST-QGDNQWSYNYGAY------LGGAAYMY
094 tr|C5DHG0|C5DHG0_LACTC          NVYDGAN----I---DD-NCTD-IVKNEWSYNQGIV------LGGAAYMY
095 tr|G0SFA3|G0SFA3_CHATD          QVHDGVH----INTNDN-TCSQ-TDTNQWTYNAGIF------LHGAAVMY
096 tr|K1WJG5|K1WJG5_MARBU          KVYDGSS----D---TE-NCTS-IDHLQFSYNHGIF------LFGAAVMY
097 tr|H2B005|H2B005_KAZAF          FVYDGAS----V---TD-NCTT-VTKYQWTYNQGLL------LSGSAYLY
098 tr|C5DYD7|C5DYD7_ZYGRC          NVFDGAV----T---DD-NCTQ-VTELEWTYNHGII------LGGSAYMY
099 tr|G8YM23|G8YM23_PICSO          YIYDGVE----A---K-ACEN-ITKYQWTYNQGLL------LSGCAFLY
100 tr|W3X8E3|W3X8E3_PESFW          NIYDGAH----I---ET-NCTD-INKAQFSYNNGVY------LLGAAYMY
101 tr|A0A254U3K7|A0A254U3K7_ASPNG NVADSTST--------TN-DCST-QGDNQWSYNYGAY------LGGAAYMY
102 tr|Q6CER8|Q6CER8_YARLI          HVYDGI-----H---LP-NCSD-RSPYLWTYNSGIY------MSGAASIY
103 tr|G2QT05|G2QT05_THITE          NIYDGGH----V---EQ-NCTD-INRAQFSYNNAVY------LLGAAYMY
104 sp|Q5ACZ2|DFG5_CANAL            RIYDGAK----I---TN-NCSS-VTDLRWSYTYGVF------MAGCAYLY
105 tr|G2WKU6|G2WKU6_YEASK          NVFDGAE----I---DT-NCTD-ITKIEWTYNHGIV------LGGLAYMY
106 tr|A0A0C7N752|A0A0C7N752_9SACH NVYDGAN----I---DD-NCTD-ITKLEWSYNQGIV------LGGLAYLY
107 tr|W3WN68|W3WN68_PESFW          NVLDGAG----N-ADDA-NCTE-INAAQFTYNAGIY------LIGAAHMY
108 tr|G2QFS1|G2QFS1_MYCTT          KVYDGGH----V---EH-NCTD-INKAQFSYNAALL------LHGSAFMY
```

```
109 tr|G3J9G4|G3J9G4_CORMM          VIYDGIE----VDDATQ-SCKN-IHQVQFTYNAGVW------LQGAAAMY
110 tr|Q2UJ03|Q2UJ03_ASPOR          LFLDGAD----E---LK-NCSE-FNYLQWTYNSGVY------LFGAASMY
111 tr|G8YRT2|G8YRT2_PICSO          YIYDGAN----I---GK-NCSD-VTKTKWSYIYGVF------IAGCAYMH
112 tr|W3X554|W3X554_PESFW          QVFDGAQ----V---TS-NCSD-IDKAQWTYNAGIF------LHGAAVMY
113 tr|Q1K7A8|Q1K7A8_NEUCR          DVKDGAH----FDVTTH-VCRNDSGPHVWSYNIGVF------LQGAAFMY
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM TIADTTSP--------ET-GCTD-HGDLQWTYNYGTY------ISGAAYMY
115 tr|A7F3P0|A7F3P0_SCLS1          EVYDGTS----D---TD-NCSS-IDHIQWTYNAGIY------LLGSAMMY
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA EVYDGYN----VLDDTK-TCGH-LHDVQFSYNAGIW------LAGTAAMY
117 tr|Q6CIP9|Q6CIP9_KLULA          NVFDGAT----I---DD-NCTS-IVKYEWSYNHGVV------LGGCAYMY
118 tr|H2AMQ5|H2AMQ5_KAZAF          YVYDGAS----T---ED-NCSS-VTKYQWTYNQGLI------LSGAAYLY
119 tr|G8JMI3|G8JMI3_ERECY          NVFDGAH----I---GD-NCQD-IKGIEWTYNHGVI------LGGCAYMY
120 tr|C5DGT1|C5DGT1_LACTC          YVYDGAK----V---DG-NCST-ITKLIWTYNQGLL------MSGCAYLY
121 tr|K1XJR4|K1XJR4_MARBU          NVFDGGH----I---GH-NCTD-INRVQFSYNIAIW------LLGAANMY
122 tr|A0A124BYC5|A0A124BYC5_ASPNG NVADSTDA--------AN-GCTN-LGNNQWSYNYAVY------LSGAAYMY
123 tr|A7TLL2|A7TLL2_VANPO          FVYDGVN----I---GD-NCTK-VTNLQWTYNHGLI------LSGCAYLY
124 tr|Q0CB86|Q0CB86_ASPTN          NVADSTDI--------ET-GCKT-QGNNQWSYNYGTY------LTGAAYMY
125 tr|K1XLH0|K1XLH0_MARBU          EVFDGSQ----N---TD-NCTE-KDHNKWTYNTGIF------IMGAATMY
126 tr|J7S438|J7S438_KAZNA          FVYDGVD----A---ND-NCST-VTQYQWTYNHGLL------LAGSAYLY
127 tr|Q4WG09|Q4WG09_ASPFU          YIADSTSN--------EA-NCKD-AGNTQWSYNYGTY------LSGASFMY
128 tr|H8WZ09|H8WZ09_CANO9          RIYDGAN----I---EE-NCTD-VTDLRWSYTYGVF------MTGCAYLY
129 tr|I1RL09|I1RL09_GIBZE          LVYDGGH----V---GK-NCTD-INKATFSYNAAIL------IQGAAFMY
130 tr|Q4W985|Q4W985_ASPFU          LFYDGAD----V---LL-NCSD-LNRIEWTYNSGVY------LLGAANMY
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU NVFDGAE----I---DT-NCTD-ITKIEWTYNHGIV------LGGMAYMY
132 tr|W6QCW3|W6QCW3_PENRF          RFWDGAS----D---LT-ECKP-INQIEWTYNAGVY------LLGAANMY
133 tr|W3WMD3|W3WMD3_PESFW          NVYDGAH----L---P--DCTD-INKAQFSYNSAML------MQGAAFLY
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH YVYDGVD----I---TN-NCSGAVHKYQWTYNQGLL------MSGCAYLY
135 tr|I2GYH1|I2GYH1_TETBL          EVYDGAD----I---SE-NCTD-ITKYQWSYNNGCM------LGGAAYMY
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH NVYDGAE----I---DS-NCTE-VTRLEWTYNHGIV------LGGLAYMY
137 tr|Q752P3|Q752P3_ASHGO          RVFDGAH----I---HE-NCSD-IKGIEWSYNHGII------LAGAAYMY
138 tr|B8MYP3|B8MYP3_ASPFN          NIADTTTS--------EA-NCKD-HGDLQWTYNYGTY------LSGAAYMY
139 tr|Q2TXL6|Q2TXL6_ASPOR          TINDLVTV--------ES-NCQA-PNQMQWSYNYGIY------FNGAAYMY
140 tr|Q0CN17|Q0CN17_ASPTN          TIADTTTT--------QS-NCKA-HGDLQWTYNYGMY------LSGAAYMY
141 tr|J3NHD5|J3NHD5_GAGT3          NVYDGGH----I---GK-NCTD-INRAQFSYNNAVW------LLGAAHMY
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI NIYDGTD----D---LS-NCTT-INHVQYSYNAGVW------ILGAAHMY
143 tr|A1CVB0|A1CVB0_NEOFI          LFYDGAD----V---LL-NCSD-FNRIEWTYNSGVY------LLGAANMY
144 tr|B8NVI3|B8NVI3_ASPFN          TINDLVTV--------ES-NCQA-PNQMQWSYNYGIY------FNGAAYMY
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH LVFDGAK----V---EG-NCSN-ITKLVWTYNQGLL------MSGCAYLW
146 tr|A1DJ54|A1DJ54_NEOFI          YIADSTSN--------EA-NCKD-AGNTQWSYNYGTY------LSGASFMY
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS KVYDGGH----V---EH-NCTD-INKAQFSYSAAIL------VQGAAFMY
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA FVYDGAN----I---DD-NCTE-ITKLQWTYNQGLM------MSGSAYLY
149 tr|A0A177A618|A0A177A618_9PEZI DVFDGSS----D---LL-NCTE-LDHTQWSYNAGIW------LHGAANMY
150 tr|Q75CW6|Q75CW6_ASHGO          KVLDGAH----I---SE-RCED-HSKLEWTYNHGVV------LAGCAYMY
151 tr|Q7S4K4|Q7S4K4_NEUCR          KVYDGGH----V---EH-NCTD-INKAQFSYSAAIL------VQGAAFMY
152 tr|G2QB99|G2QB99_MYCTT          TVRDGVH----VSLEDG-SCNS-RDENQWTYNAGIF------LHGAAVMY
153 tr|G0S3F2|G0S3F2_CHATD          AVYDGGH----V---EH-NCTD-INRAQFSYNAALL------LHGAAFMW
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO KIVDGLE----V----D-NCSD-ISPYQWTYNHGLF------LAGCAYLY
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB NVYDGAK----V---ED-NCTD-ITAIEWTYNHGVV------LGGLAYMY
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM SVADSTNI--------ED-GCES-QGNNQWSYNYGTY------LMGAAYMW
157 tr|G2QKJ0|G2QKJ0_MYCTT          DVLDGAH----I---GH-NCTD-LNPVQFSANAAML------IHATAVMY
158 tr|Q4WFX5|Q4WFX5_ASPFU          YVFDGTY----E---SD-NCSA-LNRVEWTYNNGVF------LHGAAHMW
159 tr|Q5BGD7|Q5BGD7_EMENI          FVADSTSG--------SN-DCVD-ADRMQWSYNYGTF------IAGAAYMY
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH YVYDGAS----I---DE-NCSD-IVKYQWTYNQGLM------LGGSAYLY
161 tr|G8BYN9|G8BYN9_TETPH          NVYDGAT----T---NY-NCSD-IVQYEWSYNHGVV------LGGLAYMY
162 tr|F7VVT4|F7VVT4_SORMK          NIYDGGH----V---PH-NCTD-INKVQWSANAAIM------IHGVAIMY
163 tr|G8YB80|G8YB80_PICSO          YIYDGVE----A----K-ACEN-ITKYQWTYNQGLL------LAGCAFLY
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA RVYDGAD----I---AE-NCTD-IVHIEWTYNIGIM------IGGAAYLY
165 tr|G2WRS7|G2WRS7_VERDV          NIYDGGH----V---EH-NCTD-INKAQFSYNNAVF------LLGAAHMY
166 tr|A0A124BVB5|A0A124BVB5_ASPNG HVFDGTS----E---EN-NCTD-LNHIEWTYNNGVF------LLGAAHMW
167 tr|A7EIG7|A7EIG7_SCLS1          YVYDGSD----D---TI-NCTR-LNHIQWTYNAGAF------LLGAANMY
168 tr|C5M5U7|C5M5U7_CANTT          FVYDGVK----I---YN-NCSN-ITKYQWTYNQGLM------LAGCAYLY
169 tr|B9WAA8|B9WAA8_CANDC          RIYDGAK----I---AN-NCSS-VTDLRWSYTYGVF------MTGCAYLY
170 tr|F8MRU1|F8MRU1_NEUT8          KVYDGGH----V---EH-NCTD-INKAQFSYSAAIL------VQGAAFMY
171 tr|A1DJS0|A1DJS0_NEOFI          SVADSTNI--------ED-GCES-QGNNQWSYNYGAY------LMGAAYMW
172 tr|B6HLM8|B6HLM8_PENRW          AVADSVRQGTGG---PN-GCTL-PDHTRWTYNYGTY------LSGAGYMY
173 tr|A0A117DWS0|A0A117DWS0_ASPNG YIADTTTT--------EA-NCKD-HGDIQWTYNYGTY------LSGAAYMY
174 tr|A0A124BXM1|A0A124BXM1_ASPNG NVADSTDV--------DN-NCKT-QGNNQWSYNYGVF------LTGAAYMY
175 tr|J3NZQ6|J3NZQ6_GAGT3          RVYDGAN----S---NH-NCTN-INHVQYMYNPGIL------MQGAAAMY
176 tr|W0THH8|W0THH8_KLUMD          NVYDGAN----I---EE-NCTN-IIKYEWSYNHGVV------LGGCAYMY

177 tr|B2AEF2|B2AEF2_PODAN          RVYDGGH----V---EH-NCTD-INKAQFSYNAALL------LHGSAFMY
```

```
178 tr|Q96TX1|Q96TX1_NEUCS          SIHDGAH----V---ET-NCTD-INNAQFSYNSGVF------VLGAAHMY
179 tr|E9DYA6|E9DYA6_METAQ          AIYDGAN----V---AN-QCKD-INKVQYSAFALRVPDCAKRRLTPRIPT
180 tr|B9WAI2|B9WAI2_CANDC          FVYDGVK----V---AN-NCSN-ITKYQWSYNQGLM------LAGCAYLY
181 tr|Q6C171|Q6C171_YARLI          RVYDGI-----H---LP-NCSD-RSPYLWTYNAGIY------LSGAAALY
182 tr|A6ZMV1|A6ZMV1_YEAS7          NVFDGAE----I---DT-NCTD-ITKIEWTYNHGIV------LGGLAYMY
183 tr|B6K7X7|B6K7X7_SCHJY          TVYDGAT----T---AD-NCST-LTETQWTYNVGVF------LGGAAYLY
184 tr|I1RSA2|I1RSA2_GIBZE          EIYDGVH----IGKEKS-SVCDDIDKTQWSYNSGVF------LHGAANMY
185 tr|F7VWZ9|F7VWZ9_SORMK          DVKDGAH----FDVTTH-VCRNDSGPHVWSYNVGVF------LQGAAFMY
186 tr|A0A1B2J789|A0A1B2J789_PICPA  FVYDGAF----V---ED-NCTE-IVMLQWTYNAGLM------VSGCAYLA
187 tr|G0V8N4|G0V8N4_NAUCC          FVYDGAS----T---TD-NCTV-LTTYQWTYNQGLL------LSGAAYLY
188 tr|H8WXN8|H8WXN8_CANO9          FVYDGVK----I---AN-NCSN-ITKYQWTYNQGLM------LAGCAYLY
189 tr|A6ZZS0|A6ZZS0_YEAS7          YVYDGVS----I---DD-NCTK-VTSYQWTYNQGLL------LAGSAYLY
190 tr|G9MQD1|G9MQD1_HYPVG          LVYDGGH----V---EK-NCTD-INLATFSYNAAVL------THGAAFMY
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH  FVYDGAK----V---EG-NCSN-ITRLVWTYNQGLL------MSGSAYLY
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM  YIADSTSN-------EA-NCKD-AGNTQWSYNYGTY------LSGASFMY
193 tr|A0A254UET1|A0A254UET1_ASPNG  HVFDGTS----E---EN-NCTD-LNHIEWTYNNGVF------LLGAAHMW
194 tr|A1D587|A1D587_NEOFI          TIADTTTP-------DT-ECTD-HGDLQWTYNYGTY------LSGAAYMY
195 tr|Q1K7I4|Q1K7I4_NEUCR          SIHDGAH----V---ET-NCTD-INNAQFSYNSGVF------VLGAAHMY
196 tr|G8JS88|G8JS88_ERECY          FVYDGAF----I---HD-NCSR-PSVLQWSYNHGLI------LGGCAYLT
197 tr|A0A136J7D3|A0A136J7D3_9PEZI  NVFDGAG----I---VD-QCAT-HDKAQWTYNAGIY------LHGAAALY
198 tr|E9E4X8|E9E4X8_METAQ          AIYDGAK----V---AN-GCKD-INKAEFSYNNAVF------AEGVAFMY
199 tr|Q6CP42|Q6CP42_KLULA          YVYDGAS----I---EE-NCTD-TVKYQWTYNQGLM------LSGSAYLY
200 tr|G0WHL4|G0WHL4_NAUDC          NVFDGAE----I---DS-NCTD-MTQIEWTYNHGVV------LGGAAYMY
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU  YVYDGVS----I---GD-NCTK-VTSYQWTYNQGLL------LAGSAYLY
202 tr|F9WYX6|F9WYX6_ZYMTI          --FDGTDA-------TT-SCSE-INRLQWTNNAGLF------LTTGAYSA
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6  AVYDGGH----I---AK-NCTD-INKAQFSYNNAVW------LLGAAHMY
204 tr|J3P147|J3P147_GAGT3          KVFDGGH----V---EH-NCTD-INKQQFSYNAGVL------VQGAAFMY
205 tr|I2H842|I2H842_TETBL          FVHDGVN----I---ET-NCTD-ITPYQWTYNQGLL------LAGCAYLY
206 tr|S6E3R3|S6E3R3_ZYGB2          NVYDGAE----I---DS-NCTE-VTRLEWTYNHGIV------LGGLAYMY
207 tr|E5AD94|E5AD94_LEPMJ          QVFDGTS----I-TTNP-PCKD-ADMTPWTYNYGIM------IAGAAYMY
208 tr|Q6FJM7|Q6FJM7_CANGA          NVYDGAK----V---ED-NCTD-ITAIEWTYNHGVV------LGGLAYMY
209 tr|A5DUW0|A5DUW0_LODEL          FVYDGVK----I---AN-NCSN-ITKYQWTYNQGLM------LAGCAYLY
210 tr|G8BF46|G8BF46_CANPC          FVYDGVK----I---AN-NCSN-ITKYQWTYNQGLM------LAGCAYLY
211 tr|W3XAF6|W3XAF6_PESFW          EVRDGAG----N-AGEA-NCTE-INGALFSYNAGIF------LHGAAYMY
212 tr|B6H7I2|B6H7I2_PENRW          NVADSTQM-------AN-DCAD-SGNYQWTYNYGAY------LMGAVYMY
213 tr|Q5ATF9|Q5ATF9_EMENI          NVADSTDI-------ND-GCTS-QGNNQWSYNYGAF------LMGAAYMY
214 tr|A0A167CDD3|A0A167CDD3_9ASCO  FVYDGAY----V---EA-NCTN-VRNLEWTYNYGLF------ISGCAYLY
215 tr|G0RXS3|G0RXS3_CHATD          NVLDGGH----V---EH-NCTD-INPVQFSTNAALL------IHAAAIMY
216 tr|B8NX26|B8NX26_ASPFN          IIFDSVNN-------DD-QCKS-PDNLQWTYNYGTY------LSGAAFLY
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM  QVFDGSD----V---VK-NCSE-LSHIQWTYNNGVL------MGGAAALY
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM  YVFDGTY----E---SD-NCSA-LNRVEWTYNNGVF------LHGAAHMW
219 tr|A1CMM3|A1CMM3_ASPCL          HFFDGTD----D---NY-NCTS-LNHIQWTYNAGIF------LLGAANMY
220 tr|B6HU84|B6HU84_PENRW          NIADTTTC-------GT-QCTD-HGDFQWTYNYATY------ISGAAYMH
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS  SIHDGAH----V---ET-NCTD-INNAQFSYNSGVF------VLGAAHMY
222 tr|G2WHY6|G2WHY6_YEASK          YVYDGVS----I---DD-NCTK-VTSYQWTYNQGLL------LAGSAYLY
223 tr|G0W7G8|G0W7G8_NAUDC          FVYDGVN----I---AN-NCST-ITPYQWTYNHGLV------MAGAAYLY
224 tr|G4NGV8|G4NGV8_MAGO7          EIYDGAV----A-DGGQ-DCKK-PDMVQWTYNAGIF------MHGAAFMY
225 tr|W3X855|W3X855_PESFW          NIYDGAS----D---DE-GCVS-IDQAQWTYNAGIF------MEGSAAMW
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA  AIYDGGH----V---EH-NCTD-INKVEFSYNNAVF------LQGCAYMY
227 tr|J8Q2K7|J8Q2K7_SACAR          YVYDGVS----I---TD-NCTK-VTSYQWTYNQGLI------LAGSAYLY
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO  FVYDGVK----V---DN-NCSK-VTKYQWSYNQGLL------LSGCAFLY
229 tr|J6EGN8|J6EGN8_SACK1          NVFDGAE----I---DT-NCTD-ITKIEWTYNHGIV------LGGLAYMY
230 tr|A0A100ISD9|A0A100ISD9_ASPNG  RFWDGAD----V---SN-GCGD-WNYIEWTYNTGVY------LLGAAVMY
231 tr|G8YQC0|G8YQC0_PICSO          YIYDGAD----I---SN-NCSD-VTKTKWTYIYGVF------IAGCAYMH
232 tr|F7VVP8|F7VVP8_SORMK          KVYDGGH----V---GH-NCTD-INKAQFSYNAAIL------AQGAAFMY
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH  NVFDGAE----I---DT-NCTD-ITKIEWTYNHGIV------LGGLAYMY
234 tr|F8MBP4|F8MBP4_NEUT8          DVKDGAH----FDVTTH-VCRNDSGPHVWSYNIGVF------LQGAAFMY
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG  NVADSTNT-------EN-DCTT-LGNNQWSYNYAVF------LSGAAYMY
236 tr|A1CD27|A1CD27_ASPCL          YIADSTSN-------EA-NCKD-AGNTQWSYNYGTY------LSGASFMY
237 tr|A5DHF3|A5DHF3_PICGU          VIYDGAP----I---EE-NCTD-HTTLKWSYTYGIF------MSGAAYMY
238 tr|D4ATT7|D4ATT7_ARTBC          QIFDGTD----S---AD-NCSG-INHIQWSYNNGIL------LLGTAHMY
239 tr|E2PT42|E2PT42_ASPNC          RFWDGAD----V---SS-GCGD-WNYIEWTYNTGVY------LLGAAVMY
240 tr|C7ZPE5|C7ZPE5_NECH7          KIYDGAH----V---GT-NCTD-INKAQFSYNAGVF------LQGAAFMY
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA  HFLDGAD----D---TK-NCTQ-INRIQWTYNAGVC------LLGAAHMF
242 tr|Q4WKP7|Q4WKP7_ASPFU          TIADTTSP-------ET-GCTD-HGDLQWTYNYGTY------ISGAAYMY
243 tr|F8MXT4|F8MXT4_NEUT8          SIHDGAH----V---ET-NCTD-INNAQFSYNSGVF------VLGAAHMY
244 tr|A0A177AJ74|A0A177AJ74_9PEZI  NIYDGTD----D---LS-NCTS-INHVQYSYNAGVW------ILGAATMY
245 tr|J7RQR5|J7RQR5_KAZNA          NVFDGAN----I---ED-NCTT-ITTFEWSYNHGVV------LGGTAFMY
246 tr|Q75CW7|Q75CW7_ASHGO          RVLDGAH----I---SE-RCED-HSQLEWTYNHGVV------LGGCAYMY
247 tr|G3AKK4|G3AKK4_SPAPN          RIFDGAN----V---ET-NCTD-VTQLRWSYTYGIF------MAGCAYLY
```

```
248 tr|B2W762|B2W762_PYRTR          HFYDGTD----D---LQ-NCTT-INHIQWTYNAGVH------MAGAAALW
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH YVYDGVS----I---DD-NCTK-VTSYQWTYNQGLL------LAGSAYLY
250 tr|C4XWK1|C4XWK1_CLAL4          IIYDGAD----D---TE-NCTD-LTIHKWSYTYGVY------LAGCAYLY

001 tr|A0A254U2J9|A0A254U2J9_ASPNG NYT----------G---K-AKWKTAAEGLLNVT---LDTFFPAKY----G
002 sp|Q6FLP9|DCW1_CANGA           NM--T----------G--SDLWHERTHAFLNAS---RVFFN---------
003 tr|A1CCM5|A1CCM5_ASPCL         NYT----------N-GTN-TKWESAVNGLLNVT---LNQFFPTAY----G
004 tr|C1H190|C1H190_PARBA         KY--T---------N-G--DKKWEERVRGILKGI---EIFFQN---------
005 tr|Q2UR85|Q2UR85_ASPOR         NLT----------D-GG--EKWKEAIDGLLGTT---IAKFFPHEY----G
006 tr|A5DV30|A5DV30_LODEL         NF--T----------G--DEVWKTRTEEIVDAS---LSYFFT-----S--
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO NA--T----------G--DDAWGSELDKFLGGI---EHYFLS-----P-N
008 tr|J3PGN0|J3PGN0_GAGT3         NYT-----DG---------GAIWKERVEKLTESI---FRDFFA---------
009 tr|A0A124BXU6|A0A124BXU6_ASPNG NYT----------N-GTS-TKWMNAVDGLLNRT---LNKFFPASH----G
010 tr|C5P4A1|C5P4A1_COCP7         NY--T---------D-G--EQIWRERVQGIIDGL---RP-FFP--------E
011 tr|Q0CG55|Q0CG55_ASPTN         NL--T---------G-G--SPVWKERTENILNAT---DVFFQN--------Q
012 tr|C4Y2X5|C4Y2X5_CLAL4         NY--T----------L--DEKWLNRTMNLLHAT---QVFVYN-----KSA
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG NL--T----------E--SPVWKARTEGILNAS---FVFFQD--------D
014 tr|C7Z068|C7Z068_NECH7         NYT-----NG---------SEVWETRVNNLTDSL---LKNFFP---------
015 tr|G8BXX2|G8BXX2_TETPH         NY--T----------G--SELWHTRTKNFLSSS---SVFFN---------
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR NY--T----------E--DEKWHNRTLNFLKSA---EVFFYK-----KLE
017 tr|A1CSC2|A1CSC2_ASPCL         NYT----------N-GG--SKWKKGLDGLLGTT---FSLFFPFQN-----
018 sp|Q05031|DFG5_YEAST           NA--T---------N-G--TGEWETSLTKILNGA---KSYFFK---------
019 tr|H2AQJ6|H2AQJ6_KAZAF         NA--T---------N-G--SSVWESRVTDIVEGA---KSIFFS---------
020 tr|G2XFH9|G2XFH9_VERDV         NYT-----E---------DAKWLDRVEKLLDRC---LTDFFP---------
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH KT--N----------G--QEKWYNRTEDFVKSS---GVFYN---------
022 sp|Q5AD78|DCW1_CANAL           NY--T----------E--EEKWYNYTIKLLESA---QVFFKN-----I-S
023 tr|B6H7E8|B6H7E8_PENRW         NYT----------N-GTE-GKWLEGVTGLLKTS---EQFFPKSG----D
024 tr|B6GZU2|B6GZU2_PENRW         NL----------T-E--DPKWKERTQKIIDAS---GVFFSH--------D
025 tr|E9E4W7|E9E4W7_METAQ         NHT-----S---------DDKWKTRLDSLLDAC---LKSFFP---------
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO NV----------T-E--DPKWKERVQKILDSS---EIFFAK--------T
027 tr|G2R7G8|G2R7G8_THITE         NYT-----NG--S-------EKWRARVAGLLNHT---IDFFFP---------
028 tr|G3JBY1|G3JBY1_CORMM         DY--T---------K----QDIWKTHVDGLLNFT---TRMFVS--------H
029 tr|A5DNV5|A5DNV5_PICGU         NY--T----------Q--DEVWHERTKKLLAGS---VVFFNS-----T--
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI NY--T---------G-G--TDTWRRHTSGLLKTT---SSTFFR--------D
031 tr|Q2TWC1|Q2TWC1_ASPOR         NYT----------N-GD--EKWLKRVNGLLEST---FATFFPSTY-----
032 tr|J5RXY9|J5RXY9_SACK1         NF--T----------G--SDLWHTRTKEFLNSS---QVFFQ---------
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA DY--N---------K----TDVWKGHVNGILNYT---ERTFFQ--------N
034 tr|A0A254U5V1|A0A254U5V1_ASPNG NLT----------N-GA--DKWKKGLDGLLAST---FSRFFPKEY----G
035 tr|A0A061B6H7|A0A061B6H7_CYBFA NF--T----------E--EEKWRNRTLRYLQSS---SVFFN---------
036 tr|G8ZQ93|G8ZQ93_TORDC         NM--T----------K--DEKWGNRTFEYLNSA---SIFYN---------
037 tr|Q9C2J1|Q9C2J1_NEUCS         NV--S---------T-GAEQETWKTRVDGLLGAV---EAKFLT--------N
038 tr|C5NZK5|C5NZK5_COCP7         NF--T---------K-G--ADIWEKRLNGLIDAA---GIFFSK--------D
039 tr|Q6C0T7|Q6C0T7_YARLI         NG--TLLRNGTGS------EKWGQRVRGIFNTAVSPLHFFGG--------P
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB NM--T----------G--SDLWHERTHAFLNAS---RVFFN---------
041 tr|F7VZ72|F7VZ72_SORMK         NYT-----DG---------SPVWKDRLDKLVNAT---LVRFFP---------
042 tr|A3LN37|A3LN37_PICST         NY--T----------E--DQKWLDRTLNLLNAS---QVFFMK-----I-G
043 tr|Q6CAI2|Q6CAI2_YARLI         NA--T----------S--DEQWATRAKNIWTGC---EDLFFQ---------
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO NY--T----------E--DELWYNRTMRLLSGA---QVFFKD--------
045 tr|G9MJS5|G9MJS5_HYPVG         NHT-----N---------SAVWKTRLQKLVDHG---LETFFP---------
046 tr|G9MHI4|G9MHI4_HYPVG         NF--T---------G----NATWKANLDGIISQT---VSKFV---------
047 tr|D4AP27|D4AP27_ARTBC         NY--T---------D-G--SPKWAERVQGILDGL---RP-FFH--------P
048 tr|B6GZT8|B6GZT8_PENRW         NYT----------N-GTIQAQWLTAVNGLLNAT---FNNFFLT-------
049 sp|O74556|YCZ2_SCHPO           NY--T---------N----STVWKTRVEGFANKT---AKTFFF---------
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 NFT-----DG---------GAIWKDRVEKLTDSI---FRDFFA---------
051 tr|B8MHG0|B8MHG0_TALSN         NYT----------N-GG--TKWKEGVDGLLNTT---WNTFFPTEY-----
052 tr|A3LMV8|A3LMV8_PICST         DY--T----------K--DEKWLNRVNDLMETS---ASTFFP-----K-S
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU NA--T----------G--EQFWDNELGLFLGGL---EHYMVN-----T-T
054 sp|Q9P6I3|YHG7_SCHPO           NYT----------N-GS--SVWQTHMEGLMNKA---LDYYFTSD------
055 tr|Q2US57|Q2US57_ASPOR         NYT----------N-GS--TTWLTAVNGLLNAT---IDGFFPTKY----G
056 tr|W7N622|W7N622_GIBM7         NI--T---------G----KASWKTRVDGLLEDI---KTVFV---------
057 sp|Q9P6I4|YHG6_SCHPO           NYT----------GE--TVWRERLDGLISHA---TSYFFT-D------
058 tr|C7GRB3|C7GRB3_YEAS2         NA--T---------N-G--TGEWETSLTKILNGA---KSYFFK---------
059 tr|C7GP28|C7GP28_YEAS2         NF--T----------G--SDLWHTRTKEFLNAS---QVFFH---------
060 tr|G3JGZ5|G3JGZ5_CORMM         NYT-----DG---------DKKWGDRVDKLLDSM---LKNFIR---------
061 tr|C4QXV4|C4QXV4_KOMPG         NY--T----------G--DEMWLDRTENFLHGI---QVFTN---------
062 tr|I1S0Z5|I1S0Z5_GIBZE         NHT-----DG---------EQKWEQRLDKLVDAT---ISNFFP---------
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM NCYKT---------E-G--DSRWEARTKHILQAT---DAFFAE--------D
064 tr|G0W5X4|G0W5X4_NAUDC         NF--T----------G--STIWLNRTMDLLDAT---SVFFN---------
```

```
065 tr|G8ZPC1|G8ZPC1_TORDC          NA--T---------N-G--SSEWETRVTQILGGA---ESYFFK---------
066 tr|A0A167CDC5|A0A167CDC5_9ASCO  DF--T----------N--DTKWLTRAQLLMERG---SSIYFD---------
067 tr|R9XAT7|R9XAT7_ASHAC          NH--T----------Q--DEVWHQRTLRFLESA---AVFIK---------
068 tr|B2B747|B2B747_PODAN          NYT-----NG---------SDVWRERLDGLTDAT---IRVFFP---------
069 tr|G0SE99|G0SE99_CHATD          NYT-----NG---------SDVWAARVNGLVNST---LRTFFR---------
070 tr|J8Q6F0|J8Q6F0_SACAR          NA--T---------N-G--TSEWKTRVTQVLNGA---KSYFFK---------
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH NA--T---------N-G--SSEWQTRTEQILGGA---ESFFFD---------
072 tr|J3K6E2|J3K6E2_COCIM          NY--T---------D-G--EQIWRERVQGMIDGL---RP-FFP--------E
073 tr|G4N3E1|G4N3E1_MAGO7          NIT-----S---------NEKWKTATTKLTESL---LRDFFA---------
074 tr|C5M5J2|C5M5J2_CANTT          NF--T----------G--DTVWQTRALEIVEAS---LSYFFT-----S--
075 tr|A0A254U0X0|A0A254U0X0_ASPNG NYT----------N-GS--TKWEERVNNLLTSL---TTKFFPENFN----
076 sp|Q75DG6|DCW1_ASHGO            NH--T----------Q--DELWHQRTLRFLDSA---RIFLS---------
077 tr|Q6BZF0|Q6BZF0_DEBHA          NY--T----------E--DEMWYNRTMRLLSGA---QVFFKD---------
078 tr|E9E3Q1|E9E3Q1_METAQ          NYT-----DG---------AQKWKTALDGLLDAT---LKTFFP---------
079 sp|P36091|DCW1_YEAST            NF--T----------G--SDLWHTRTKEFLNAS---QVFFH----------
080 tr|W0TAH6|W0TAH6_KLUMD          NH--T----------G--SELWHKRTKDFLNSA---SVFFNS-----T--
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA NA--T----------E--DDVWLERTEALVQGA---SVFFE----------
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI NV--T---------G----DQTWLDATNGILKKT---TELFF---------
083 tr|Q2TYU3|Q2TYU3_ASPOR          NMT----------E---K-ETWKTAVDGLLGVT---LNTFFPQDF----N
084 tr|G2XGF6|G2XGF6_VERDV          NV--T---------G----DDKWRTRTVGLLDET---EKTF-F---------
085 tr|G3AMT4|G3AMT4_SPAPN          SY--T----------Q--DKKWYNYTHRLLDGS---RVFFKN-----I-S
086 tr|A0A1C1D213|A0A1C1D213_9EURO NY--T---------D-G--SQVWADRISGIISR----LNVFLE--------G
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS SFTKG---PE---------QVLWKNRVQALLDRT---IAFFFH---------
088 tr|A0A100I5A6|A0A100I5A6_ASPNG NYT----------N-GS--TKWEQRVNNLLTSI---TTKFFPQDYS----
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH NA--T---------N-G--SSTWETKLSQLVGGS---TKYFFR---------
090 tr|G2Q8A7|G2Q8A7_MYCTT          NYT-----NG---------SDKWREAIEGLADAT---IKTFFP---------
091 tr|B8NPZ1|B8NPZ1_ASPFN          NHT----------N-GS--DKWLKRVDGLLNSV---LTTFFPKEG-------
092 tr|Q7SAB2|Q7SAB2_NEUCR          SFTKG---PE---------QVLWKNRVQALLDRT---IAFFFH---------
093 tr|A2R8R5|A2R8R5_ASPNC          NYT----------N-GTS-TKWMNAVDGLLNRT---LNKFFPASH----G
094 tr|C5DHG0|C5DHG0_LACTC          NA--T---------QND--TSVWKSRLDQIVGGS---TEFFFK---------
095 tr|G0SFA3|G0SFA3_CHATD          NL--T---------S----NPIWRERVEGLLRTT---VSTFFTLPENNTNS
096 tr|K1WJG5|K1WJG5_MARBU          EF--K---------N-Q--SSVWQGRLESLIKN----IDIFFK--------N
097 tr|H2B005|H2B005_KAZAF          NF--T----------E--SELWHERTKGFLNAS---VVFFN----------
098 tr|C5DYD7|C5DYD7_ZYGRC          NA--T---------N-G--SSKWEERVNQILNGA---SDYFFQ---------
099 tr|G8YM23|G8YM23_PICSO          NH--T----------E--DEVWHNRTKALLKGS---RVFFQ----------
100 tr|W3X8E3|W3X8E3_PESFW          NYT-----NG---------NETWKERLDGLIDGT---FRVFFP---------
101 tr|A0A254U3K7|A0A254U3K7_ASPNG NYT----------N-GTS-TKWMNAVDGLLNRT---LNKFFPASH----G
102 tr|Q6CER8|Q6CER8_YARLI          NA--TMLRNGTNATLE--ASVWMNRTYQLWNADIMKGIFFGG-------N
103 tr|G2QT05|G2QT05_THITE          NYT-----NG---------SDTWKERLEGLVDAT---IATFFP---------
104 sp|Q5ACZ2|DFG5_CANAL            NF--T----------G--DDVWLTRTNEIVQAS---LSYFFA---------
105 tr|G2WKU6|G2WKU6_YEASK          NA--T---------N-G--TGEWETSLTKILNGA---KSYFFK---------
106 tr|A0A0C7N752|A0A0C7N752_9SACH NA--T---------N-G--SSTWETKLSQVLGGS---TTFFFK---------
107 tr|W3WN68|W3WN68_PESFW          NY--T---------E----SDTWKERVNGLVTSI---SSVFF----------
108 tr|G2QFS1|G2QFS1_MYCTT          NYT-----NG---------SEIWRERVDKLLEGI---LRDFFP---------
109 tr|G3J9G4|G3J9G4_CORMM          DY--N---------K----TDVWKGHVNGILNYT---ERVFFK--------D
110 tr|Q2UJ03|Q2UJ03_ASPOR          NL--T---------D-G--DPVWKERTQRILDAT---KVYFQN--------D
111 tr|G8YRT2|G8YRT2_PICSO          SY--S----------N--DSVWLDRVHEILKAS---TSAFFK---------
112 tr|W3X554|W3X554_PESFW          NF--T---------G-G--DDKWKQRTQGLLDQT---TG-FHF---------
113 tr|Q1K7A8|Q1K7A8_NEUCR          NV--S---------T-GAEQETWKTRVDGLLGAV---EAKFLT--------N
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM NFT----------N-GG--DKWKKGLDGLLNTT---FQRFFPFQN-----
115 tr|A7F3P0|A7F3P0_SCLS1          NY--T---------N-G--SSIWQNRTQGLLNA----SSVFFQ--------N
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA DY--T---------K----KDIWRTHVDGLLNFT---KGMFIS--------Q
117 tr|Q6CIP9|Q6CIP9_KLULA          NA--T---------N-G--SDVWQTRTEQILGGA---TSFFFE---------
118 tr|H2AMQ5|H2AMQ5_KAZAF          NY--S----------E--SSLWYNRTMTALSSS---GVFLN----------
119 tr|G8JMI3|G8JMI3_ERECY          NA--T---------N-G--DPVWETRLRQILGGA---RAFFFR---------
120 tr|C5DGT1|C5DGT1_LACTC          NY--T----------G--SETWHTRTFEFLNSS---DIFFN----------
121 tr|K1XJR4|K1XJR4_MARBU          NYT-----NG---------SSIWEDRVNLLLNST---FNTFFP---------
122 tr|A0A124BYC5|A0A124BYC5_ASPNG NYT----------N-GET-TQWKTALDGLLNST---FETFFPAKF-----
123 tr|A7TLL2|A7TLL2_VANPO          NY--T----------G--DAKWYNRTERFLEYA---YVFFN----------
124 tr|Q0CB86|Q0CB86_ASPTN          NLT----------E---K-SYWKDRLDGLLGVT---LDKFFPQKY----G
125 tr|K1XLH0|K1XLH0_MARBU          NF--T---------N-G--SALWKERIDGLKGAS---KLFYN----------
126 tr|J7S438|J7S438_KAZNA          NF--T----------G--DEVWHNRTKKLVTAA---QVFFNS-----T--
127 tr|Q4WG09|Q4WG09_ASPFU          NYT----------N-GE--DKWLKRVNGLLDSL---IGTFCPKDK-----
128 tr|H8WZ09|H8WZ09_CANO9          NV--T----------G--DETWKFRTQEIVEAA---TSYFFN-----S--
129 tr|I1RL09|I1RL09_GIBZE          NFT-----NG---------STVWEDRITKLLDAT---LKNFFP---------
130 tr|Q4W985|Q4W985_ASPFU          NCYKT---------E-G--DSRWEARTKHILQAT---DAFFAE--------D
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU NA--T---------N-G--TSEWKTSLQQILNGA---ASYFFK---------
132 tr|W6QCW3|W6QCW3_PENRF          NL----------T-G--DPKWKDRTERIINSS---SVFFSH--------D
133 tr|W3WMD3|W3WMD3_PESFW          NYT-----NG---------EQVWQDRVQGLLTRT---IEVFFP---------
```

```
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH KT--N----------G--QEKWYNRTEDFVKSS---NVFYN---------
135 tr|I2GYH1|I2GYH1_TETBL         NA--T----------N-G--SSIWETRVTQLLNGA---KAYFFQ---------
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH NA--T----------N-G--SSTWEERVSQILGGA---TDYFFE---------
137 tr|Q752P3|Q752P3_ASHGO         NA--T----------G-G--SPEWESHVTKILAGA---ADFYFR---------
138 tr|B8MYP3|B8MYP3_ASPFN         NLT----------D-GG--EKWKEAIDGLLGTT---IAKFFPHEY----G
139 tr|Q2TXL6|Q2TXL6_ASPOR         NLT----------N-GD--TKWKNVVEGLLNTT---WRNFFPQEY----G
140 tr|Q0CN17|Q0CN17_ASPTN         NLT----------N-GD--TKWKKGIDGLLQTT---FSQFFPKQY----G
141 tr|J3NHD5|J3NHD5_GAGT3         NYT-----DG---------SEVWKGRVQSLLDAT---LKTFFP---------
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI NI--T----------K--SDVWKARVQGLLDQS---IKIFSD-------N
143 tr|A1CVB0|A1CVB0_NEOFI         NF--T---------E-G--DPRWEARTKRILQAT---DVFFVK-------D
144 tr|B8NVI3|B8NVI3_ASPFN         NLT----------N-GD--TKWKNVVEGLLNTT---WRNFFPQEY----G
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH NM--T----------Q--DEKWHVRTWEFLNAS---SIFFN---------
146 tr|A1DJ54|A1DJ54_NEOFI         NYT----------N-GE--DKWLKRVNGLLDSL---IGTFCPKDK-----
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS NYT-----EGDAAT----QDMWKTRIEKLTEGL---FRDFFP---------
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA NF--T----------E--EEKWRTRTLDYLRSS---TVFFN---------
149 tr|A0A177A618|A0A177A618_9PEZI KF--T---------N-N--SEKWLERSTRILNKA---STQFFW-------K
150 tr|Q75CW6|Q75CW6_ASHGO         NI--T---------N-G--SAVWGDRLNKVLNGS---LLF-FP---------
151 tr|Q7S4K4|Q7S4K4_NEUCR         NYT-----EGDAAT----QDMWKTRIEKLTEGL---FRDFFP---------
152 tr|G2QB99|G2QB99_MYCTT         NH--T---------N-A--SAIWRERVDGLLAST---TATFVD-------P
153 tr|G0S3F2|G0S3F2_CHATD         NYT-----ED---------QKWKDRVDNLLTGI---LRDFFK---------
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO NY--T----------E--DEKWRNRTLELLKGA---VVFFGT-----E--
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB NA--T---------N-G--SSVWQSRLTSVLGGA---TAYFFQ---------
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM NYT----------N-GTN-SKWATAVDGLLNNT---LDIFFPAKY----G
157 tr|G2QKJ0|G2QKJ0_MYCTT         NYT-----TG--ET----RAKWRRHVAGLLNHT---IDHFFP---------
158 tr|Q4WFX5|Q4WFX5_ASPFU         NY--T---------Q-G--NSSWKARIDGLLDAQ---RTFLSP-----NKS
159 tr|Q5BGD7|Q5BGD7_EMENI         NAT----------K--D--VKWRNRTEGLVDHV---FKHFFPTKYTTAIG
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH NY--T----------E--SEVWHERTKNFLNSS---GVFFN---------
161 tr|G8BYN9|G8BYN9_TETPH         NA--T---------N-G--TSKWETALTQVLSGA---TTYFFK---------
162 tr|F7VVT4|F7VVT4_SORMK         NMT-----ED--A------EKWGKPLTGLVNRT---LEFFFP---------
163 tr|G8YB80|G8YB80_PICSO         NY--T----------E--DEVWHNRTKLLLNGS---RVFFQ----------
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA NF--T----------E--EQYWLDVTAHLVEGA---KVFFR---------
165 tr|G2WRS7|G2WRS7_VERDV         NFT-----DG---------DAKWKTRIDGLLQRT---IEVFFP---------
166 tr|A0A124BVB5|A0A124BVB5_ASPNG NY--T---------N-G--DEKWRERVNGLLQAQ---NSFLST-----NES
167 tr|A7EIG7|A7EIG7_SCLS1         NY--T---------N-G--SDIWRGRVEGLLNG----THVFFQ-------N
168 tr|C5M5U7|C5M5U7_CANTT         NY--T----------E--EEKWYNYTIKLLESS---QVFFMN-----I-T
169 tr|B9WAA8|B9WAA8_CANDC         NF--T----------G--DDVWLTRTNEIVQAS---LSYFFA---------
170 tr|F8MRU1|F8MRU1_NEUT8         NYT-----EGDAAT----QDMWKTRIEKLTEGL---FRDFFP---------
171 tr|A1DJS0|A1DJS0_NEOFI         NYT----------N-GTN-TKWATAVDGLLNNT---LDLFFPTKY----G
172 tr|B6HLM8|B6HLM8_PENRW         AYT----------N-D---TKWLDITNNLLDSL---FTTFFLPEH-----
173 tr|A0A117DWS0|A0A117DWS0_ASPNG NLT----------N-GA--DKWKKGLDGLLAST---FSRFFPKEY----G
174 tr|A0A124BXM1|A0A124BXM1_ASPNG NYT----------G---K-AKWKNAAEGLLNVT---LDTFFPAKY----G
175 tr|J3NZQ6|J3NZQ6_GAGT3         DHT-----KK---------DKWKNIVTNLAKAA---IESFFA---------
176 tr|W0THH8|W0THH8_KLUMD         NA--T---------N-G--SEEWKSRTSQILGGA---TSFFFD---------
177 tr|B2AEF2|B2AEF2_PODAN         NYT-----NG---------SEIWKTRVDKLWEGM---HRDFFE---------
178 tr|Q96TX1|Q96TX1_NEUCS         NYTNTTQTDG---------SPVWKDRLDKLLNAT---LVRFFP---------
179 tr|E9DYA6|E9DYA6_METAQ         AQT-----NG---------DQKWKDRVDKLINYG---LKTFFP---------
180 tr|B9WAI2|B9WAI2_CANDC         NY--T----------E--EEKWYNYTIKLLESA---QVFFKN-----I-T
181 tr|Q6C171|Q6C171_YARLI         NA--TVLRNHTDP--N--HGKWLDRANELWNSTKGPNLFFGG-------P
182 tr|A6ZMV1|A6ZMV1_YEAS7         NA--T---------N-G--TGEWETSLTKILNGA---KSYFFK---------
183 tr|B6K7X7|B6K7X7_SCHJY         NY--T---------N-G--STVWQQRLDGLITKA---SNHFFS---------
184 tr|I1RSA2|I1RSA2_GIBZE         NI--T---------S----SDKWKKRADGLLDDV---FNKFV----------
185 tr|F7VWZ9|F7VWZ9_SORMK         NV--T---------T-GEEQNTWKTRVDGLLGSI---ESKFLT-------N
186 tr|A0A1B2J789|A0A1B2J789_PICPA NY--T----------G--DDLWLDRTENFLHGI---QVFFTN----------
187 tr|G0V8N4|G0V8N4_NAUCC         NY--T----------G--SDTWLLRTNQLLNAS---SVFFN----------
188 tr|H8WXN8|H8WXN8_CANO9         NY--T----------E--DQKWYNYTIQLLTSA---QVFFKN-----M-S
189 tr|A6ZZS0|A6ZZS0_YEAS7         NF--T----------G--SDLWHTRTKEFLNAS---QVFFH----------
190 tr|G9MQD1|G9MQD1_HYPVG         NYT-----DG---------SEIWKNRVDRLLDSL---LARFFP---------
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH NM--T----------E--DEKWHTRTWEFLNSS---NIFFN----------
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM NYT----------N-GE--DKWLKRVNGLLDSL---IGTFCPKDK-----
193 tr|A0A254UET1|A0A254UET1_ASPNG NY--T---------N-G--DENWRKRVNGLLQAQ---SSFLST-----NES
194 tr|A1D587|A1D587_NEOFI         NFT----------N-GG--DKWKKGLDGLLNTT---FKRFFPFQN-----
195 tr|Q1K7I4|Q1K7I4_NEUCR         NYT-----DG---------SPVWKDRLDKLLNAT---LVRFFP---------
196 tr|G8JS88|G8JS88_ERECY         DF--T----------G--SDVWHNRTLRLLRSA---SIFFP----------
197 tr|A0A136J7D3|A0A136J7D3_9PEZI DY--T---------N-G--SALWEGRVNGLIKRT---REHFFN---------
198 tr|E9E4X8|E9E4X8_METAQ         NYT-----NG---------NATWKARLDGLIKHG---METFLP---------
199 tr|Q6CP42|Q6CP42_KLULA         NF--T----------E--SETWHERTKNFLNSS---GIFFN----------
200 tr|G0WHL4|G0WHL4_NAUDC         NA--T---------N-G--SSDWESRVTKLLKGA---TTYFFD-----S--
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU NY--T----------G--SDLWHTRTKEFLNAS---QVFFH----------
202 tr|F9WYX6|F9WYX6_ZYMTI         NAV----------Y-SATGRYWSELNQLVTNSS---LRIFVYEDR-----
```

```
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 NYT-----GG---------SEKWRARVEGLLDAT---LKTFFP---------
204 tr|J3P147|J3P147_GAGT3          NQT-----KA---------QKWKTAVEKLTEAI---FRDFFE---------
205 tr|I2H842|I2H842_TETBL          NF--T----------Q--DEKWLIRTNQFLNGS---GVFFN---------
206 tr|S6E3R3|S6E3R3_ZYGB2          NA--T---------N-G--SSTWEERVSQILGGA---TDYFFE---------
207 tr|E5AD94|E5AD94_LEPMJ          NY--T---------N-G--AEKWKKALTQFIAKT---EMFYPP--------D
208 tr|Q6FJM7|Q6FJM7_CANGA          NA--T---------N-G--SSVWQSRLTSVLGGA---TAYFFQ---------
209 tr|A5DUW0|A5DUW0_LODEL          NY--T----------E--EEKWLNYTKNLLHSS---RVFFMN-----K-T
210 tr|G8BF46|G8BF46_CANPC          NY--T----------E--DQKWYNYTVQLLNSA---QVFFKN-----M-S
211 tr|W3XAF6|W3XAF6_PESFW          NL--T---------E----TDTWKTRVESLTNSI---ATTFF---------
212 tr|B6H7I2|B6H7I2_PENRW          NYT----------N-GD--EKWKKPVDGLLGKT---LKTFFPN--------
213 tr|Q5ATF9|Q5ATF9_EMENI          NYT----------E---K-AEWKTVVDGLLGKL---LDEFFPEQY----G
214 tr|A0A167CDD3|A0A167CDD3_9ASCO NY--T----------N--QTVWLDRLEMLWSRG---KVFFNE---------
215 tr|G0RXS3|G0RXS3_CHATD          NYT-----LE---------EKWRSRVLGLLNRT---IDHFFP---------
216 tr|B8NX26|B8NX26_ASPFN          NYT----------N-GD--EKWLNRVNGLLGTT---LKTFFPEKY-----
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM NF--T---------K-G--ADIWEKRLNGLIDAA---GIFFSK-------D
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM NY--T---------Q-G--NSSWKARIDGLLDAQ---RTFLSP-----NKS
219 tr|A1CMM3|A1CMM3_ASPCL          HF--S---------N-R--SDVWKTRLEGIIKGL---DVFFPV---------
220 tr|B6HU84|B6HU84_PENRW          NYT----------N-GTE-TKWKEGVEGLIKTS---QQFYPTTGFNGAG
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS NYT-----DG---------SPVWKDRLDKLLNAT---LVRFFP---------
222 tr|G2WHY6|G2WHY6_YEASK          NF--T----------G--SDLWHTRTKEFLNAS---QVFFH---------
223 tr|G0W7G8|G0W7G8_NAUDC          NF--T----------E--NEVWHNRTKMLVDAS---TVFFN---------
224 tr|G4NGV8|G4NGV8_MAGO7          NL--T---------E----SPKWKERVSGLLDRT---MISF-V---------
225 tr|W3X855|W3X855_PESFW          NY--T---------N-GT-SDKWEQRTTGIVNRT---ATYFF---------
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA NHT-----EG---------AQKWKNCVDGLVKHG---LEHFFP---------
227 tr|J8Q2K7|J8Q2K7_SACAR          NH--T----------G--SDLWHTRTKQFLNAS---QVFFQ---------
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO NY--T----------E--DVLWLNRSLNLLNGA---QVFFAN-----N--
229 tr|J6EGN8|J6EGN8_SACK1          NA--T---------N-G--TSEWETRLTQILNGA---KAYFFK---------
230 tr|A0A100ISD9|A0A100ISD9_ASPNG NL--T----------E--SPVWKARTEGILNAS---FVFFQD--------D
231 tr|G8YQC0|G8YQC0_PICSO          SY--N----------N--DSVWLDRVNEILKAS---TSAFFK---------
232 tr|F7VVP8|F7VVP8_SORMK          NYT-----DGDS------DNVWKTRVEKLTEGL---LRDFFP---------
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH NA--T---------N-G--TGEWETSLTKILNGA---KSYFFK---------
234 tr|F8MBP4|F8MBP4_NEUT8          NI--S---------T-GTEQETWKTRVDGLLGAV---EAKFLT--------N
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG NYT----------N-GE--TKWKSGLDGLLNTT---FGEFFPHKY-----
236 tr|A1CD27|A1CD27_ASPCL          NYT----------N-GD--QKWLNRVNGLLDSL---LATFSPKQL-----
237 tr|A5DHF3|A5DHF3_PICGU          NY--T----------E--EAIWKQRVEEIWQVA---SSFFFD------D--
238 tr|D4ATT7|D4ATT7_ARTBC          NH--T---------N-K--DPVWKDRLFGLIKGF---EVFFSQ-------KE
239 tr|E2PT42|E2PT42_ASPNC          NL--T----------E--SPVWKARTEGILNAS---FVFFQD--------D
240 tr|C7ZPE5|C7ZPE5_NECH7          NYT-----DG---------EAKWEKRLTKLVDST---LDNFFP---------
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA QF--T---------N-N--SDKWKKRVQGIKNGL---DV-FFF--------K
242 tr|Q4WKP7|Q4WKP7_ASPFU          NFT----------N-GG--DKWKKGLDGLLNTT---FQRFFPFQN-----
243 tr|F8MXT4|F8MXT4_NEUT8          NYT-----DG---------SPVWKDRLDKLLNAT---LVRFFP---------
244 tr|A0A177AJ74|A0A177AJ74_9PEZI NI--T----------K--SDIWKARVQGLLDRS---IKVFSD--------A
245 tr|J7RQR5|J7RQR5_KAZNA          NA--T---------N-G--SSVWEDRVNRILKGA---TTYFFR---------
246 tr|Q75CW7|Q75CW7_ASHGO          NI--T---------N-G--SAVWGDRLNKVLNGS---LLF-FE---------
247 tr|G3AKK4|G3AKK4_SPAPN          NF--T----------G--DTKWQTRANEIVEAS---LSYFFQ---------
248 tr|B2W762|B2W762_PYRTR          NA--T----------G--KDYWKTRVQGLIDGS---KVFFKD---------
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH NF--T----------G--SDLWHTRTKEFLNAS---QVFFH---------
250 tr|C4XWK1|C4XWK1_CLAL4          NA--T----------G--NETWETGVAQILDAS---SYFFN---------

001 tr|A0A254U2J9|A0A254U2J9_ASPNG G-NIMSEILCEP-------TEVC--NDNEILFKGLVTGWLGLVALVMP-S
002 sp|Q6FLP9|DCW1_CANGA          -NSILYEAACQGP-------NTC--NTDQRSFKAYFARFLGSTAELVP-E
003 tr|A1CCM5|A1CCM5_ASPCL        G-NIMSEVLCEP-------SELC--NNNEVLFKGLVSSWLAFTALLVP-S
004 tr|C1H190|C1H190_PARBA        --DVMTEVACERN-------GKC--NVDQRSFKAYLSRWMAMTVKVAP-F
005 tr|Q2UR85|Q2UR85_ASPOR        G-DIMSEISCEQ-------SMMF--DRNQDCFKGFLSSWLTFTTTIAP-F
006 tr|A5DV30|A5DV30_LODEL        -DKIMQETTCQP-------HNLC--NNDQRSFRSLFSRCLGLTKLIIP-E
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO ATDVLYEYQCLQW-------GRC--NNDQRSFRAIVARALGDVVVLAP-S
008 tr|J3PGN0|J3PGN0_GAGT3        D-GAAYELPCEQK-------KGGC--TADMLSFKGYVHRWLAVASQVAP-F
009 tr|A0A124BXU6|A0A124BXU6_ASPNG G-EILTEILCEP-------TEVC--NDNEIIFKGLVSAWLAYTALLVP-S
010 tr|C5P4A1|C5P4A1_COCP7        QADIMVESSCEPH-------DNC--LTDQRSFKAFLSRWMAETTQLAP-F
011 tr|Q0CG55|Q0CG55_ASPTN        ---VLYERACEPI-------NTC--KVDQRSFKGYLASWMAQTTQMAP-F
012 tr|C4Y2X5|C4Y2X5_CLAL4        DAHIIYEAACSSP-TEN--KFTC--NNDQRSFKAYFCRCLGATSVLVP-Q
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG ---VMYERACEPV-------STC--QVDQRSFKGYLARWMAATTQMAP-F
014 tr|C7Z068|C7Z068_NECH7        K-GIAWEVPCEGR-------KGAC--STDMLSFKGYVHRWLSVVTQIVP-H
015 tr|G8BXX2|G8BXX2_TETPH        -DTILYEAACQNQ-------ETC--NTDQRSFKAYFSRFLGVTAQLVP-E
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR DAHIMYEAACQSP-TSN--QYSC--NQDQRSFKAYFSRFLGLTSILVP-E
017 tr|A1CSC2|A1CSC2_ASPCL        G-TVMSEVACEP-------NMMC--DRNQDCFKGFLSSWLTFTTTIAP-Y
018 sp|Q05031|DFG5_YEAST          -DSIMYESACQD-------YGTC--NTDQRTFKSIFSRMLGLTSVMAP-F
019 tr|H2AQJ6|H2AQJ6_KAZAF        -DGIMYESACQD-------YDTC--NNDQRSFKSIFSRMLGFTSVLAP-F
```

```
020 tr|G2XFH9|G2XFH9_VERDV          R-GVAYEVPCERD-------PGRC--SADMLTFKGYMHRWLAVVTQLVP-S
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH -NSILYEATCQPY-------NSC--NTDQRSFKAYFTTMLAATAQLIP-D
022 sp|Q5AD78|DCW1_CANAL            GSMVMYEAACQPS-------NSC--NNDQRSFKAYFSRFLGLTSVLVP-Q
023 tr|B6H7E8|B6H7E8_PENRW          H-QIISDITCEP-------IDMC--DRNQKTFKTYFTSWIGFMSLIVP-T
024 tr|B6GZU2|B6GZU2_PENRW          PPNVMYERACESV-------NTC--MVDQRSFKGYFSRWMAQTAQMAP-F
025 tr|E9E4W7|E9E4W7_METAQ          R-QVAYELSCEFS----LGGSVC--KTDMLSYKGYLVRWLAVVTQLAP-H
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO PQNVMYERACETV-------NTC--MVDQRSFKGYLARWMAATTQMAP-F
027 tr|G2R7G8|G2R7G8_THITE          D-GIMVERACELP-----DRVQC--NTDQHSFKGYMHRALATTAVVAP-F
028 tr|G3JBY1|G3JBY1_CORMM         NGSHLWEPPCETQ-------PRGC--DQNQVSFKGYLLRCLAYTSKMAP-W
029 tr|A5DNV5|A5DNV5_PICGU         -NDIMYEAACQNG-KGGQGAGSC--NQDQRSFKAYFSRFLGLTAIMAP-E
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI --KVMFEPACEL-------DDTC--NTDQLSFKAYLSRWMAATTKLAP-F
031 tr|Q2TWC1|Q2TWC1_ASPOR          GGNVLSEVACEP-------IMSC--DRNQLGFKGYTAMWLAHTAILVP-S
032 tr|J5RXY9|J5RXY9_SACK1         -DGIVYEAACQGP-------NSC--NTDQRSFKAYFARFLGVTAQLVP-E
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA --DIMYEPACET-------VHTC--DQNMVSFKGYLIRFLAATAKMAP-W
034 tr|A0A254U5V1|A0A254U5V1_ASPNG S-NIMSEISCEP-------NMMC--DRNQDCFKGFLSSWLTFTTTIAP-Y
035 tr|A0A061B6H7|A0A061B6H7_CYBFA -NSVLYEAACQGA-------GNC--NNDQRSFKAYFSRFLGLTAYMVP-E
036 tr|G8ZQ93|G8ZQ93_TORDC          -NSVLYEVACQTV-------NKC--NTDQRSFKAYFTRFLGVTAELIP-Q
037 tr|Q9C2J1|Q9C2J1_NEUCS         DTKIIKEWYCESGFSDRGHPYQC--NIDQQTFKGYLLRWLSSTSQVAP-Y
038 tr|C5NZK5|C5NZK5_COCP7         PPDVMTEVACEGN-------GKC--NIDQRSFKAYLSRWMAMTIKLAP-Y
039 tr|Q6C0T7|Q6C0T7_YARLI         EGNIMREIACQQA-------TITC--DADQRTFKGIYSSLLGQTAQLVP-D
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB -NSILYEAACQGP-------NTC--NTDQRSFKAYFARFLGSTAELVP-E
041 tr|F7VZ72|F7VZ72_SORMK          E-KIAYEPACED-------GMTC--TTDMLSFKGYVHRWLSTTTQIAP-H
042 tr|A3LN37|A3LN37_PICST          --SVMYEAACQPS-------NNC--NNDQRSFKAYFSRFLGMTAVMVP-Q
043 tr|Q6CAI2|Q6CAI2_YARLI         -DDKMVEISCQQT-------RITC--NNDQRCFKAIFSRFIGYAALFLP-D
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO --NVMYEAACQDS-------NNC--NQDQRSFKAYFSRFLGLTSVMVP-D
045 tr|G9MJS5|G9MJS5_HYPVG          N-GIAYEPSCEG-------VNTC--TTDMVSFKGYLHRWYATTTQLAP-F
046 tr|G9MHI4|G9MHI4_HYPVG          QNGVIYEQFCEPR-------GFC--SVDQSTFKGYLVRYMASTMQLAS-H
047 tr|D4AP27|D4AP27_ARTBC          ETYIMSEISCEEQ-------GNC--ETDQRSFKAYLSRWMAASTQFAP-F
048 tr|B6GZT8|B6GZT8_PENRW          G-GIIEDYYCEP-------TETC--NNNEILFKGLTSSWLALTALLVP-S
049 sp|O74556|YCZ2_SCHPO            -KDIMFEPVCEIA-------LSC--NYDQTSFKGFLTRFMVYTAQMAP-F
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 D-GAAYELPCEQK-------KGGC--TADMLSFKGYVHRWLAVASQVAP-F
051 tr|B8MHG0|B8MHG0_TALSN          GGNILSEVACDP-------ILTC--NRDQVCFKGLMANWLSTIALIVP-Y
052 tr|A3LMV8|A3LMV8_PICST          NGGYMTEIQCFF-------SNTC--NNDQRSFRSLFSRCIGLTMKLVN-S
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU GGNTLYEYQCEKW-------QRC--NNDQRAFRAVVARTLGEIYQLAP-Q
054 sp|Q9P6I3|YHG7_SCHPO            --KIIYEPSCEP-------TESC--NSDQTAFKGMLARFLGYTMQLAP-Y
055 tr|Q2US57|Q2US57_ASPOR          G-NIMSDYTCET-------TEVC--NNNEIIFKGLLSMWLAFTALLVP-S
056 tr|W7N622|W7N622_GIBM7          KNGVIYEQFCEEH-------KLC--NLDQQTFKGYLARWMAATALVAP-H
057 sp|Q9P6I4|YHG6_SCHPO            --DIAWDPQCEY-------FDDC--NSDQTAFKGIFMQSFGNTIRLAP-Y
058 tr|C7GRB3|C7GRB3_YEAS2          -DSIMYESACQD-------YGTC--NTDQRTFKSIFSRMLGLTSVMAP-F
059 tr|C7GP28|C7GP28_YEAS2          -DGIVYEAACQGP-------NSC--NTDQRSFKAYFARFLGVTAQLVP-E
060 tr|G3JGZ5|G3JGZ5_CORMM         D-GAMYELPCEGT-------PGTC--TADMLTFKGYVHRWLSVVAQIVP-H
061 tr|C4QXV4|C4QXV4_KOMPG         -QSVFFEAACQGS-------GNC--NTDQRSFKAYLARFLGLTAQMVP-S
062 tr|I1S0Z5|I1S0Z5_GIBZE         K-DIAVEIACEN-------HDTC--TTDMYSFKGYVHRWMSQATHLAP-F
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM PAMVMYERACELV-------DTC--QVDQRAFKGFLARWMAAATQVAP-F
064 tr|G0W5X4|G0W5X4_NAUDC          -NSIMYEVACQPT-------GTC--NNDQRSFKSYLARFLGLTAQLVT-I
065 tr|G8ZPC1|G8ZPC1_TORDC          -DSIMYESLCQGD-----GTSTC--NSDQRSFKAIFSRMLGLTSVLVP-S
066 tr|A0A167CDC5|A0A167CDC5_9ASCO -DGVMFEAACQNT-------GRC--NNDQRSFKAIYSRFLGLTAQLAP-P
067 tr|R9XAT7|R9XAT7_ASHAC          -NDTLYEAACQDT-------NTC--NVDQRSFKAYFSRFLGLTAQLVP-E
068 tr|B2B747|B2B747_PODAN          D-NIAYEVACEE-------HMSC--TTDMLSFKGYVARWMATATQVAP-F
069 tr|G0SE99|G0SE99_CHATD          D-NVAYEVACEE-------HMSC--TTDMLSFKGYLHRWLATATEVAP-F
070 tr|J8Q6F0|J8Q6F0_SACAR          -DGIMYESACQD-------YSTC--NTDQRTFKSIFSRMLGLTSVMAP-F
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH -NSIMYESTCQGS-------GKC--NTDQRVFKAIFSRMLGYTAVLAP-F
072 tr|J3K6E2|J3K6E2_COCIM          QADIMVEYSCEPH-------DNC--LTDQRSFKAFLSRWMAETTQLAP-F
073 tr|G4N3E1|G4N3E1_MAGO7          D-GAAFELPCETT-------PGGC--TADMLSFKGYVHRWMSVVTQVAP-F
074 tr|C5M5J2|C5M5J2_CANTT          -DKIMQETTCQP-------YNLC--NNDQRSFRSLFSRCLGLTMLLMP-E
075 tr|A0A254U0X0|A0A254U0X0_ASPNG NGTVLSEVTCEP-------ILSC--DRNQLGFKGYVAMWLAFTALLVP-S
076 sp|Q75DG6|DCW1_ASHGO            -NDTLYEAGCQGG-------DNC--NIDQRSFKAYFSRFLGLTAQLVP-E
077 tr|Q6BZF0|Q6BZF0_DEBHA          --DIMYEAACQGS-------NNC--NQDQRSFKAYFSRLLGLTSVMVL-D
078 tr|E9E3Q1|E9E3Q1_METAQ          K-GIAFEVACERD----NGSGTC--TPDMLSYKGFLHRWLAVTSQIAP-Y
079 sp|P36091|DCW1_YEAST            -DGIVYEAACQGP-------NSC--NTDQRSFKAYFARFLGVTAQLVP-E
080 tr|W0TAH6|W0TAH6_KLUMD          -NNIMYEAACQGS-------GYC--NNDQRSFKAYFSRFLGLTAQLVP-E
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA -DDIMYERACQGS-------SSGC--NTDQRSFKSIFSRCLGQTAVMAP-T
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI RENIVYEPACEPF-------GTC--NNDQRSFKTYLLRWMAQTAQLVS-S
083 tr|Q2TYU3|Q2TYU3_ASPOR          --YIMSEVLCEP-------NEVC--NDNEILFKGLVSGWLAFVALLVP-S
084 tr|G2XGF6|G2XGF6_VERDV          QNGVMFEQRCEPF-------KQC--NIDQSSFKGYLARWMAGTAQVVP-D
085 tr|G3AMT4|G3AMT4_SPAPN          GSMVMYEAACQPS-------GTC--NNDQRSFKAYFSRFLGLTSVLVP-D
086 tr|A0A1C1D213|A0A1C1D213_9EURO --NIMKETACEAV--NADGSDTC--NVDQRSFKAYLARWIAATIVRAP-F
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS A-GPMIELSCETP-----TVILC--KTDMLSFKGYTHRWLATTTQLAP-F
088 tr|A0A100I5A6|A0A100I5A6_ASPNG NGTVLSEVTCEP-------ILSC--DRNQLGFKGYVAMWLAFTALLVP-S
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH -NKIMYEATCQD-------SNSC--NNDQRSFKSIFSRMLSLTKILAP-F
```

```
090 tr|G2Q8A7|G2Q8A7_MYCTT          D-GVAYEIACEQ-------GMTC--TADMLSFKGYLHRWLATATQVAP-I
091 tr|B8NPZ1|B8NPZ1_ASPFN          NGVVLSEVACEP-------ILTC--DRNQLCFKGYVAMWLAFTAILVP-S
092 tr|Q7SAB2|Q7SAB2_NEUCR          A-GPMIELSCETP-----TVILC--KTDMLSFKGYTHRWLATTTQLAP-F
093 tr|A2R8R5|A2R8R5_ASPNC          G-EILTEILCEP-------TEVC--NDNEIIFKGLVSAWLAYTALLVP-S
094 tr|C5DHG0|C5DHG0_LACTC          -DKVMYESTCQD-------SKTC--NQDQRSFKSIFSRMLSLTRSLAP-Y
095 tr|G0SFA3|G0SFA3_CHATD          RTPVMHEQLCEPL-------SLC--NIDQRSFKGYLTRWLANTAQLAP-F
096 tr|K1WJG5|K1WJG5_MARBU          --GVMFEAACETV--NNN-AGAC--NTDQLSFKAYFSRWLAATAKLIP-A
097 tr|H2B005|H2B005_KAZAF          -NSIMYEAACQGP-------QTC--NTDQRSFKAYFSRFLGATAQLVP-E
098 tr|C5DYD7|C5DYD7_ZYGRC          -DNVMYESGCQHG-----DSSNC--NNDQRSFKSIFSRMLGLTSVLVP-S
099 tr|G8YM23|G8YM23_PICSO          -NDTMYEAACQP-------ANTC--SQDQRSFKAYFSRFLGLTAVMAP-E
100 tr|W3X8E3|W3X8E3_PESFW          N-DIAFEVACEE-------HMTC--TTDMLSFKGYVHRWMSTITQIAP-Y
101 tr|A0A254U3K7|A0A254U3K7_ASPNG  G-EILTEILCEP-------TEVC--NDNEIIFKGLVSAWLAYTALLVP-S
102 tr|Q6CER8|Q6CER8_YARLI          QSNIMVEIACQQV------TITC--NQDQRTFKGIYSSLLGQTAQMMP-S
103 tr|G2QT05|G2QT05_THITE          N-NIAYEIACEE-------QMSC--TTDMLSFKGYLHRWMATATQIAP-F
104 sp|Q5ACZ2|DFG5_CANAL           -NKIMQETTCQP-------QNKC--NNDQRSFRCLFSRCLGLTTQLAP-E
105 tr|G2WKU6|G2WKU6_YEASK         -DSIMYESACQD-------YGTC--NTDQRTFKSIFSRMLGLTSVMAP-F
106 tr|A0A0C7N752|A0A0C7N752_9SACH -DKIMYEVTCQD-------SNSC--NNDQRSFKSIFSRMLSLTSILAP-Y
107 tr|W3WN68|W3WN68_PESFW          ENGIMYEPPCES-------TSC--NTDQQAFKGHLARWMAYTAKLAP-F
108 tr|G2QFS1|G2QFS1_MYCTT          D-GVAFELPCEGR------KGAC--TADMLSFKGYVHRWMAVVTKLVP-D
109 tr|G3J9G4|G3J9G4_CORMM          --GIMYEPACET-------VHTC--DQNMVSFKGYLIRFLAYTSKVAP-W
110 tr|Q2UJ03|Q2UJ03_ASPOR          ---VLYERACEPI-------NTC--GVDQRSFKGYLARWMAASAQVAP-F
111 tr|G8YRT2|G8YRT2_PICSO          -NGVMYEASCEP-------DRC--NNDQTAFKALFARSLGQTAVLIP-E
112 tr|W3X554|W3X554_PESFW          DNGIMVEKPCEEG-------GFC--DIDQQSFKAYLARWLAGTSQLAP-F
113 tr|Q1K7A8|Q1K7A8_NEUCR          DTKIIKEWYCESGFSDRGHPYQC--NIDQQTFKGYLLRWLSSTSQVAP-Y
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM G-MVMSEIACEP-------NMKC--DRNQDCFKGFLSSWLTFMTTIVP-Y
115 tr|A7F3P0|A7F3P0_SCLS1          --NVMLEVACETT------TVGC--DTDEQSFKAYLSRWMAATTKIAP-F
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA EGGHLWEPPCEKT-------ARGC--DQNQVSFKGYVIRFLAFTAQMAP-W
117 tr|Q6CIP9|Q6CIP9_KLULA          -DDIMYESTCQAS-------NKC--NTDQRVFKAIFSRMLGYTSVLAP-F
118 tr|H2AMQ5|H2AMQ5_KAZAF          -NSIMCEAACQPA-------GTC--DTDQRSFKAYFARFLGVTAQLVP-S
119 tr|G8JMI3|G8JMI3_ERECY          -NNIMYERACQD-------NKIC--NNDQRSFKSIFSRMLAVTSVLAP-F
120 tr|C5DGT1|C5DGT1_LACTC          -NSILYERTCQEA-------NTC--NTDQRSFKAYFSRFLGLTAQLVP-A
121 tr|K1XJR4|K1XJR4_MARBU          N-DIAYEVACEP-------KLTC--SIDMFSFKSYVMRWLAATSMVAP-F
122 tr|A0A124BYC5|A0A124BYC5_ASPNG GGEIMSEVLCEP-------ANVC--NDNEITFKGLLAESLTLTSMLAP-Y
123 tr|A7TLL2|A7TLL2_VANPO          -RTVMFEAACQGV-------HTC--TTDQRTFKAYFSRFLGMTAQLVP-D
124 tr|Q0CB86|Q0CB86_ASPTN          G-NVMSEILCEP-------QELC--NNNEILFKGLVTSWLGFVAMIVP-S
125 tr|K1XLH0|K1XLH0_MARBU          PEGIMYE-PCESV-------KSC--NVDQRSFKAYFSRQLAATAILAP-Y
126 tr|J7S438|J7S438_KAZNA          -TGIMYEAACQAA-------LTC--NNDQRSFKAYFSRFLGLTAQLVP-E
127 tr|Q4WG09|Q4WG09_ASPFU          GGNVLSEVACEP-------IMTC--DRNQIGFKGYTAMWLAHTAILVP-S
128 tr|H8WZ09|H8WZ09_CANO9          TNKVMQETTCQP-------SKKC--NNDQRSFRSLFSRCLGLTMVIIP-D
129 tr|I1RL09|I1RL09_GIBZE          K-NIMWEVPCEGR------KGAC--STDMLSFKGYVHRWLAVTTQVAP-F
130 tr|Q4W985|Q4W985_ASPFU          PAMVMYERACELV-------DTC--QVDQRAFKGFLARWMAAATQVAP-F
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU -DGIMYESACQD-------YGTC--NTDQRTFKSIFSRMLGLTSVMAP-F
132 tr|W6QCW3|W6QCW3_PENRF          PPDVMYERACETV-------NTC--MVDQRSFKGYLSRWMAQTAQMAP-F
133 tr|W3WMD3|W3WMD3_PESFW          D-GIAFEPACEP-------G-NC--NADMRSFKGFLHRWMASTAMMAP-F
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH -NSILYEATCQPY-------NSC--NTDQRSFKAYFTTMLAATAQLIP-D
135 tr|I2GYH1|I2GYH1_TETBL          -DDIMYESACQT-------AKTC--NNDQRSFKSLFSRMLGFTSVLVP-S
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH -NDIMYESGCQHN-----DSSTC--NNDQRSFRSIFSRMLGLTSVLVP-S
137 tr|Q752P3|Q752P3_ASHGO          -DKIMYERACQD-------YNTC--DNDQRSFKSIFSRMLAITSVVAP-F
138 tr|B8MYP3|B8MYP3_ASPFN          G-DIMSEISCEQ-------SMMC--DRNQDCFKGFLSSWLTFTTTIAP-F
139 tr|Q2TXL6|Q2TXL6_ASPOR          G-NIMVEP-CEPQ---KQGVVPC--DGNQSTFKSLVTAWLAFTTTIMP-E
140 tr|Q0CN17|Q0CN17_ASPTN          G-NVMSEISCEP-------NMKC--DRNQDCFKGFLSSWLTFMTTIVP-Y
141 tr|J3NHD5|J3NHD5_GAGT3          D-DVAYEIACET-------GDTC--TTDMLSFKGYLHRWLATATRIAP-F
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI ATGIIIERDCELT--E---PASC--NTDQKSFKAYLTRWMAATSKVAP-F
143 tr|A1CVB0|A1CVB0_NEOFI          PAMVMYERACELV-------NTC--QVDQRAFKGFLARWMAAATQVAP-F
144 tr|B8NVI3|B8NVI3_ASPFN          G-NIMVEP-CEPQ---KQGVVPC--DGNQSTFKSLVTAWLAFTTTIMP-E
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH -DSILYEVTCQGA-------NSC--NTDQRSFKAYFSRFLGLTAQLVP-Q
146 tr|A1DJ54|A1DJ54_NEOFI          GGNILSEVACEP-------IMTC--DRNQIGFKGYTAMWLAHTAILVP-S
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS K-GIAFELACEGR------QGAC--TPDMVSFKGYVHRWMAMVTQIAP-F
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA -DSVLYEAACQGS-------NNC--NNDQRSFKAYFSRFLGLTAYMVP-E
149 tr|A0A177A618|A0A177A618_9PEZI DTGIMYE-PCEAT----PGEPLC--NIDQMSFKAYLSRWLGATSRLIP-S
150 tr|Q75CW6|Q75CW6_ASHGO          -NGILYERACQD-------VQKC--NDDQRSFRSIFSRMLALTSVLAP-H
151 tr|Q7S4K4|Q7S4K4_NEUCR          K-GIAFELACEGR------QGAC--TPDMVSFKGYVHRWMAMVTQIAP-F
152 tr|G2QB99|G2QB99_MYCTT          ATGVLVEQLCEPS-------GFC--NTDQRSFKGYLTRWLAGTAQMAP-H
153 tr|G0S3F2|G0S3F2_CHATD          D-GVVFEIPCEGR------QGAC--TADMLTFKGYVHRWMAVVTQIAP-H
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO -YDIMYEAACQPS-------NTC--NQDQRSFKAYFSRFLGLTMTLVP-E
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB -DNIMYESACQP-------YKTC--NNDQRCFKSIFSRMLGFTSVLAP-F
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM G-NIMSEVLCEP-------NEVC--NDNEILFKGLVTSWLAFTALLVP-S
157 tr|G2QKJ0|G2QKJ0_MYCTT          E-GIMVERACELE------DRVQC--NTDQHSFKGYMHRALATVAVLAP-F
158 tr|Q4WFX5|Q4WFX5_ASPFU          SENVLYEYACETI-------NTC--RTDQYSFKAYLARWMGDVAQLAP-W
```

```
159 tr|Q5BGD7|Q5BGD7_EMENI          PGTIFSDVACEP-------LQTC--DRNMLNFKGWSSMWLAMAAIMVP-E
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH -NSVMYEAACQGG-------NYC--NNDQRSFKAYFSRFLGLTAQLVP-E
161 tr|G8BYN9|G8BYN9_TETPH          -NNIMYESACQE-------AKTC--NTDQRSFKSIFSRMLGYTRVLAP-Y
162 tr|F7VVT4|F7VVT4_SORMK          K-GVMVERACELS-----DRLLC--NVDQHSFKGYLLRSLATAALMAPDL
163 tr|G8YB80|G8YB80_PICSO          -NDTMYEAACQP-------ANSC--SQDQRSFKAYFARFLGLTAVMAP-E
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA -DGIMYERACQDG-------DNC--NNDQRFFKGVFARCLGLTALLVP-D
165 tr|G2WRS7|G2WRS7_VERDV          N-NIAYEVACEK-------KMSC--TTDMLSFKGYVARWLSTMTQVAP-F
166 tr|A0A124BVB5|A0A124BVB5_ASPNG SKNVLYEAACEMV-------GTC--QTDQFSFKAYLVRWMADTARLAP-F
167 tr|A7EIG7|A7EIG7_SCLS1          --NIMMEVACEN-------NGKC--NIDQQSFKAYLARWLAATTKMAP-F
168 tr|C5M5U7|C5M5U7_CANTT          GSMVMYEAACQPS-------GTC--NNDQRSFKAYFSRFLGLTSVLVP-Q
169 tr|B9WAA8|B9WAA8_CANDC          -NKIMQETTCQP-------QNKC--NNDQRSFRCLFSRCLGLTTQLVP-E
170 tr|F8MRU1|F8MRU1_NEUT8          K-GIAFELACEGR-------QGAC--TPDMVSFKGYVHRWMAMVTQIAP-F
171 tr|A1DJS0|A1DJS0_NEOFI          G-NIMSEVLCEP-------NEVC--NNNEILFKGLVASWLAFTALLVP-S
172 tr|B6HLM8|B6HLM8_PENRW          G-GVVTDWRCEA-------VGQCYKDANGPLFKGLTVSWLSDIALIIP-S
173 tr|A0A117DWS0|A0A117DWS0_ASPNG S-NVMSEVSCEP-------NMMC--DRNQDCFKGFLSSWLTFTTTIAP-Y
174 tr|A0A124BXM1|A0A124BXM1_ASPNG G-NIMSEILCEP-------AEVC--NDNEILFKGLVTGWLGLVALIMP-S
175 tr|J3NZQ6|J3NZQ6_GAGT3          D-GAALEASCEPY-----GPTNC--TSDMVMFKGFVHRWMAQTTLLAP-F
176 tr|W0THH8|W0THH8_KLUMD          -NNIMYESTCQSE-----KQMTC--NTDQRVFKAIFSRMLGYTAVLAP-F
177 tr|B2AEF2|B2AEF2_PODAN          D-DIAYEIPCEGR-------KGAC--TADMLSFKGYVHRWLSVVTKVAP-H
178 tr|Q96TX1|Q96TX1_NEUCS          D-NIAFEPACED-------QMSC--TTDMLSFKGYVHRWLSITTQIAP-Y
179 tr|E9DYA6|E9DYA6_METAQ          H-DVAVEISCEL-------NDGC--KTDMFTYKGFVHRWYATTTQIAP-F
180 tr|B9WAI2|B9WAI2_CANDC          GSMVMYEAACQPS-------NSC--NNDQRSFKAYFSRFLGLTSVLVP-Q
181 tr|Q6C171|Q6C171_YARLI          QKNIMVEIACQQK-------TITC--NSDQRTFKGIFSSLLGQTAQMVP-T
182 tr|A6ZMV1|A6ZMV1_YEAS7          -DSIMYESACQD-------YGTC--NTDQRTFKSIFSRMLGLTSVMAP-F
183 tr|B6K7X7|B6K7X7_SCHJY          -GGVIYDPQCETS-------NSC--NTDQPSFKGYLARFMMYAMQLAP-Y
184 tr|I1RSA2|I1RSA2_GIBZE          KNEIIYEQFCEPH-------KQC--SQDQQSFKGYLARWLAATTQLYP-E
185 tr|F7VWZ9|F7VWZ9_SORMK          DTKIIKEWYCESGFTDRGHPYQC--NIDQQTFKGYLLRWLSSTSQVAP-F
186 tr|A0A1B2J789|A0A1B2J789_PICPA -QSVFFEAACQGS-------GNC--NTDQRSFKAYLARFLGLTAQMVP-S
187 tr|G0V8N4|G0V8N4_NAUCC          -NSILYEAACQKS-------GTC--NNDQRSFKAYFLRFLGQTAQLVS-E
188 tr|H8WXN8|H8WXN8_CANO9          GSSVMYEAACQPS-------NTC--NNDQRSFKAYFSRFLALTAQLVP-E
189 tr|A6ZZS0|A6ZZS0_YEAS7          -DGIVYEAACQGP-------NSC--NTDQRSFKAYFARFLGVTAQLVP-E
190 tr|G9MQD1|G9MQD1_HYPVG          E-GIAYEVACEPR-------QGAC--STDMLSFKGYVHRWLSVVAQIAP-H
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH -NSVLYEVTCQAG-------KTC--NTDQRSFKAYYSRFLGLTAQLVP-A
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM GGNVLSEVACEP-------IMTC--DRNQIGFKGYTAMWLAHTAILVP-S
193 tr|A0A254UET1|A0A254UET1_ASPNG SKNVLYEAACETV-------GTC--QTDQFSFKAYLVRWMADTARLAP-F
194 tr|A1D587|A1D587_NEOFI          G-MVMSEIACEP-------NMMC--DRNQDCFKGFLSSWLTFMTTIVP-Y
195 tr|Q1K7I4|Q1K7I4_NEUCR          D-NIAFEPACED-------QMSC--TTDMLSFKGYVHRWLSITTQIAP-Y
196 tr|G8JS88|G8JS88_ERECY          -NGVVTEVACQEA-------GTC--NTDQRSFKAYFMRFLGLTAQLIP-E
197 tr|A0A136J7D3|A0A136J7D3_9PEZI EDGILVERACETF-------GTC--NNDQQSFRGYLMRWFGATMQLAP-F
198 tr|E9E4X8|E9E4X8_METAQ          K-GIAVEISCEN-------AGTC--TTDMLTFKGFLHRWYSTITQLAP-Y
199 tr|Q6CP42|Q6CP42_KLULA          -NSIMYEAACQGS-------GYC--NNDQRSFKAYFSRFLGLTAQLVP-E
200 tr|G0WHL4|G0WHL4_NAUDC          -NGIMFENACQE-------YNTC--NNDQRSFKSIFSRMLGLTSVLAP-F
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU -DGIVYEAACQGP-------NSC--NTDQRSFKAYFARFLGATAQLVP-E
202 tr|F9WYX6|F9WYX6_ZYMTI          --NIIYEVACEP-------SDNC--NIDQTGFRAQFVRSLATYRDMAP-A
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 D-DVAYEIACET-------GNTC--TTDMLSFKGYLHRWLSTATRIAP-F
204 tr|J3P147|J3P147_GAGT3          K-GAAFEISCEHL-------DNQC--TQDMVMFKGFVARWMPVVVQLVP-E
205 tr|I2H842|I2H842_TETBL          -NSVIYEAACQPN-------KNC--NNDQRSFKAYFARFLGVTAQLLP-S
206 tr|S6E3R3|S6E3R3_ZYGB2          -NDIMYESGCQHN-----DSSTC--NNDQRSFRSIFSRMLGLTSVLVP-S
207 tr|E5AD94|E5AD94_LEPMJ          KGGVMVE-VCEAK-------TLC--NADQESFKAYLARWLGNAIQMAP-F
208 tr|Q6FJM7|Q6FJM7_CANGA          -DDIMYESACQP-------YKTC--NNDQRCFKSIFSRMLGFTSVLAP-F
209 tr|A5DUW0|A5DUW0_LODEL          GSMVMYEAACQPG-------YTC--NNDQRSFKAYFSRFLGLTSVLVP-E
210 tr|G8BF46|G8BF46_CANPC          GSSVMYEAACQPS-------NTC--NNDQRSFKAYFSRFLALTAQLVP-E
211 tr|W3XAF6|W3XAF6_PESFW          TDGIMWELPCEAT-------ATSC--NNDQRMFKGFLSRWMAGAAKLAP-T
212 tr|B6H7I2|B6H7I2_PENRW          G-DVIEDVTCEP-------IKMC--NFNEILFKGLTSSWLAFTTLLAP-D
213 tr|Q5ATF9|Q5ATF9_EMENI          GGKIFSEYLCEP-------KALC--NYNEILFKGIVSTWITFVGLIVP-E
214 tr|A0A167CDD3|A0A167CDD3_9ASCO -DGIMYEAACQPS-------NQC--NTDQRCFKGIYSRFLGLTMLLVP-S
215 tr|G0RXS3|G0RXS3_CHATD          E-GIMVERACELP-----DRMQC--NVDQHSFKGYMHRALATVAVVAP-F
216 tr|B8NX26|B8NX26_ASPFN          GGNTMSEVACEP-------IMSC--DRNQIGFKGYLSMWLAFTAILVP-S
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM PPDVMTEVACEGN-------GKC--NIDQRSFKAYLSRWMAMTIKLAP-Y
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM SENVLYEYACETI-------NTC--RTDQYSFKAYLARWMGDVAQLAP-W
219 tr|A1CMM3|A1CMM3_ASPCL          -GNIMSEVACEST-------GKC--NVDQRSFKAYLSRWMAATVQIAP-F
220 tr|B6HU84|B6HU84_PENRW          G-QILSDITCEG-------SGNC--DRNQITFKAYFTNWLGMLTTIVP-G
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS D-NIAFEPACED-------QMSC--TTDMLSFKGYVHRWLSITTQIAP-Y
222 tr|G2WHY6|G2WHY6_YEASK          -DGIVYEAACQGP-------NSC--NTDQRSFKAYFARFLGVTAQLVP-E
223 tr|G0W7G8|G0W7G8_NAUDC          -NSIMYEAACQGP-------NSC--NTDQRSFKAYFSRFLGATAQLVP-E
224 tr|G4NGV8|G4NGV8_MAGO7          QNQVLVEPGCEFV-------GTC--NNDQRSFKGYLTRWLGGTSQLAP-F
225 tr|W3X855|W3X855_PESFW          NESVMYEPPCEPQ-------SNC--ATDSFSFKAYLVRWMAKTTQLMS-S
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA KEDIAYEPACEG-------VKTC--TTDMLSFKGYLHRWWSASAQMAP-Y
227 tr|J8Q2K7|J8Q2K7_SACAR          -DGIVYEAACQGP-------NSC--NTDQRSFKAYFARFLGVTAQLVP-E
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO --SVMYEAACQNS-------NNC--NQDQRSFKAFFSRFLGLTSVLIP-Q
```

```
229 tr|J6EGN8|J6EGN8_SACK1          -DTIMYESACQD-------YSTC--NTDQRTFKSIFSRMLGLTSVMAP-F
230 tr|A0A100ISD9|A0A100ISD9_ASPNG  ---VMYERACEPV-------NTC--QVDQRSFKGYLARWMAATTQMAP-F
231 tr|G8YQC0|G8YQC0_PICSO          -NDTMYEASCEP-------DRC--NNDQTSFKSLFARSLAQTAVLIP-E
232 tr|F7VVP8|F7VVP8_SORMK          K-GIAFELPCETR-------HGAC--TPDMLSFKGYVHRWMATVAQVAP-F
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH -DSIMYESACQD-------YGTC--NTDQRTFKSIFSRMLGLTSVMAP-F
234 tr|F8MBP4|F8MBP4_NEUT8          DTKIIKEWYCESGFSNNGHPYQC--NIDQQTFKGYLLRWLSSTSQVAP-Y
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG GGEIMSEILCEP-------TEVC--NDDEIIFKGLLAESLTFTSMVAP-Y
236 tr|A1CD27|A1CD27_ASPCL          GGNVLSEVACEP-------IMTC--DRNQLGFKGYVAMWLAQTSILVP-S
237 tr|A5DHF3|A5DHF3_PICGU          -NKIMQETQCWNN-----GDPKC--NNDQRSFRSLFSRSIQLTSVLVP-D
238 tr|D4ATT7|D4ATT7_ARTBC          PKGVMIEVACEAQ-------GNC--NNDQRSFKAYMARWMAATMIIAP-D
239 tr|E2PT42|E2PT42_ASPNC          ---VMYERACEPV-------STC--QVDQRSFKGYLARWMAATTQMAP-F
240 tr|C7ZPE5|C7ZPE5_NECH7          D-GIAVEIACEN-------HGTC--TTDMLSFKGYVHRWMSQVTHLAP-I
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA ETNIMFEVPWEPN-------GKW--EKDQRSFKAYLARWMAATTQVAP-F
242 tr|Q4WKP7|Q4WKP7_ASPFU          G-MVMSEIACEP-------NMKC--DRNQDCFKGFLSSWLTFMTTIVP-Y
243 tr|F8MXT4|F8MXT4_NEUT8          D-NIAFEPACED-------QMSC--TTDMLSFKGYVHRWLSITTQIAP-H
244 tr|A0A177AJ74|A0A177AJ74_9PEZI ATGIIIERDCELT--E---PASC--NTDQKSFKAYITRWMAATSKVAP-F
245 tr|J7RQR5|J7RQR5_KAZNA          -DSIMFESACQD-------TGKC--NNDQRTFKAIFSRMLSFTAVLVP-S
246 tr|Q75CW7|Q75CW7_ASHGO          -NGILYERACQD-------VNKC--NDDQRSFRSIFSRMLALTSVLAP-H
247 tr|G3AKK4|G3AKK4_SPAPN          -NKIMTETTCAP-------FNKC--NNDQRSFRSLFARCLGLTTLLMP-E
248 tr|B2W762|B2W762_PYRTR          --NVMQEVACENN-------GKC--DVDQRSFKAYLARWMAYTAIVAP-W
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH -DGIVYEAACQGP-------NSC--NTDQRSFKAYFARFLGVTAQLVP-E
250 tr|C4XWK1|C4XWK1_CLAL4          -NSIMTETTCAQ-------SNSC--NNDQRSFRSLFARCLSLTAGLMP-E
```

```
001 tr|A0A254U2J9|A0A254U2J9_ASPNG TYD-----SI-LPKLQGSAEAA-AASCSGMS--------------------
002 sp|Q6FLP9|DCW1_CANGA           TRQ-----QI-MTWLNTSALAA-AKSCSGGT--------------------
003 tr|A1CCM5|A1CCM5_ASPCL         TYS-----TI-LPKLQGSAVAA-AQSCTGNG--------------------
004 tr|C1H190|C1H190_PARBA         SRP-----LI-MPKLRASAEAA-AKQCSGP---------------------
005 tr|Q2UR85|Q2UR85_ASPOR         TQD-----QI-LPKIQASAQAA-AKQCSGGD--------------------
006 tr|A5DV30|A5DV30_LODEL         FKE-----KI-EPYLEASAQAA-AQSCSGGT--------------------
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO FAD-----RA-NKIIDASAKGA-ASACSGGS--------------------
008 tr|J3PGN0|J3PGN0_GAGT3         MAD-----KI-RPVLRKSAEMA-VKQCTGGA-------------------T
009 tr|A0A124BXU6|A0A124BXU6_ASPNG TYN-----RI-LPKLQGSAQGA-AATCTGYG--------------------
010 tr|C5P4A1|C5P4A1_COCP7         TRE-----LI-MPKLRASAMAA-AKACTGGN--------------------
011 tr|Q0CG55|Q0CG55_ASPTN         TFD-----WV-MPRLRVSAKAA-AETCTGGS--------------------
012 tr|C4Y2X5|C4Y2X5_CLAL4         TYD-----RI-RRWLVDSANAA-AYSCSGGS--------------------
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG TYD-----LV-MPKLRASAKAA-AETCTGGE--------------------
014 tr|C7Z068|C7Z068_NECH7         LKE-----KI-LPVLQTSTEAA-VKQCTGGK-------------------S
015 tr|G8BXX2|G8BXX2_TETPH         TRD-----NI-MKWIDASAYGA-AESCSGGT--------------------
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR TYD-----HI-RRWLVDSANAA-AWSCVGGS--------------------
017 tr|A1CSC2|A1CSC2_ASPCL         TAD-----EI-VPRIQQSALAA-AKQCSGGD--------------------
018 sp|Q05031|DFG5_YEAST           TRD-----TI-DDLIKTSAEAA-AKSCNGGT--------------------
019 tr|H2AQJ6|H2AQJ6_KAZAF         TAS-----TI-DPLIKSSAAAC-ALSCDGGT--------------------
020 tr|G2XFH9|G2XFH9_VERDV         TAS-----KI-LPVLQNSTQAA-IRQCTGGA-------------------S
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH LRD-----QI-MGWINTTAQAA-AQSCVGGY--------------------
022 sp|Q5AD78|DCW1_CANAL           TEP-----VI-TKWLVDSANGA-AGSCSGGS--------------------
023 tr|B6H7E8|B6H7E8_PENRW         NVT----AEV-MGKFKASAVAA-GQQCSGGS--------------------
024 tr|B6GZU2|B6GZU2_PENRW         TYD-----TV-MKRLKASAKAA-AKTCTGGV--------------------
025 tr|E9E4W7|E9E4W7_METAQ         TAG-----NI-LGPLRKSGEAA-AGQCTGGA-------------------S
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO TFD-----QI-MPKLKASGQAA-AKTCTGGE--------------------
027 tr|G2R7G8|G2R7G8_THITE         TRE-----QI-VRVLRSSTEGA-VSSCLADG-------------------T
028 tr|G3JBY1|G3JBY1_CORMM         TAD-----GI-RALIQQAARDA-AAACTGPVGGTGP------VGGTFQGI
029 tr|A5DNV5|A5DNV5_PICGU         TSD-----TI-MNWLEKSAIAA-AQSCSGGT--------------------
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI TYG-----TI-KPLLRASAEAA-MSHCNGGN--------------------
031 tr|Q2TWC1|Q2TWC1_ASPOR         TAE-----RI-TPKLQGSAEAI-AKQCSGES--------------------
032 tr|J5RXY9|J5RXY9_SACK1         TRN-----QI-MTWLNTSAIAA-AKSCSGGT--------------------
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA TAD-----SI-TKLISTAAKDA-VKVCNGAN----------GPKFPGP
034 tr|A0A254U5V1|A0A254U5V1_ASPNG TAD-----QI-LPKIQQSALAA-AKQCSGGD--------------------
035 tr|A0A061B6H7|A0A061B6H7_CYBFA TYE-----TI-YSLIEASARAA-SNSCSGGS--------------------
036 tr|G8ZQ93|G8ZQ93_TORDC         ARD-----QA-MKWINASALGA-AQSCSGGY--------------------
037 tr|Q9C2J1|Q9C2J1_NEUCS         TYE-----RI-NPWIRATAAAA-VATCTGPVGAAAPQVDSGGIQPGFKGI
038 tr|C5NZK5|C5NZK5_COCP7         TRD-----RL-LPKLQASATAA-ALQCSGP---------------------
039 tr|Q6C0T7|Q6C0T7_YARLI         LAP-----EI-MKYLVPSAMAG-ARTCSGGR--------------------
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB TRQ-----QI-MTWLNTSALAA-AKSCSGGT--------------------
041 tr|F7VZ72|F7VZ72_SORMK         TAP-----II-LPVLKVSAEAA-VQTCTGEA-------------------S
042 tr|A3LN37|A3LN37_PICST         TYD-----GI-RQWLVDSANAAAKNSCTGGT--------------------
043 tr|Q6CAI2|Q6CAI2_YARLI         IRD-----DI-MKKLRASATAA-LATCSGGS--------------------
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO TEP-----VI-SEWLITSANAAAKNSCTGGS--------------------
045 tr|G9MJS5|G9MJS5_HYPVG         LAP-----KI-LPILKTSAQAA-VKQCTGGA-------------------L
```

```
046 tr|G9MHI4|G9MHI4_HYPVG          TVQ-----TL-MPLITGSASAA-AAICTGPA----------SAAFKGI
047 tr|D4AP27|D4AP27_ARTBC         STD-----FI-MRRLRACARGA-AKACTGGE-------------------
048 tr|B6GZT8|B6GZT8_PENRW         TFD-----SI-LAKLQKSGQAA-AASCTGHN-------------------
049 sp|O74556|YCZ2_SCHPO           TAP-----LL-EPLLISTAKAA-AGACCGGY-------------------
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 IAP-----KI-RPVLRKSAEMA-VKQCTGGP-------------------T
051 tr|B8MHG0|B8MHG0_TALSN         TYS-----TI-LPKLQGSAVGA-GAQCSGPN-------------------
052 tr|A3LMV8|A3LMV8_PICST         TQE-----VL-GPLIEQSAKGA-AASCSGGA-------------------
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU YSE-----RV-LALIDSSAAGA-AASCSGGS-------------------
054 sp|Q9P6I3|YHG7_SCHPO           TVE-----TI-LPYIQSSAEAA-ALACSGGS-------------------
055 tr|Q2US57|Q2US57_ASPOR         TYS-----TI-IPKLQGSGVAA-AETCTGHN-------------------
056 tr|W7N622|W7N622_GIBM7         TSE-----YI-TATLLSTAKKA-AVSCSGSP----------SSGFAGQ
057 sp|Q9P6I4|YHG6_SCHPO           TYD-----TL-YPLIQTSAAAA-AKQCCGGY-------------------
058 tr|C7GRB3|C7GRB3_YEAS2         TRD-----TI-DDLIKTSAEAA-AKSCNGGT-------------------
059 tr|C7GP28|C7GP28_YEAS2         TRN-----QI-MSWLNTSAIAA-AKSCSGGT-------------------
060 tr|G3JGZ5|G3JGZ5_CORMM         TAD-----KI-LPALAKSAKAA-ANQCTGRE-------------------S
061 tr|C4QXV4|C4QXV4_KOMPG         TAE-----TI-MNWMNTSAVAV-AQSCSGGT-------------------
062 tr|I1S0Z5|I1S0Z5_GIBZE         IRP-----KI-LPVLEKSAQAA-IAQCTGGD-------------------T
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM TYD-----WV-MPRLRASAAAA-ARTCTGGP-------------------
064 tr|G0W5X4|G0W5X4_NAUDC         TQS-----RI-MNWLNTSAIAV-AHSCSGGT-------------------
065 tr|G8ZPC1|G8ZPC1_TORDC         TQS-----KV-DDLLKQSALAA-AASCSGGT-------------------
066 tr|A0A167CDC5|A0A167CDC5_9ASCO MAD-----QI-MNLLSSSAAAA-AISCSGGT-------------------
067 tr|R9XAT7|R9XAT7_ASHAC         SRD-----TV-MRWLRGSAMGA-AASCSGGT-------------------
068 tr|B2B747|B2B747_PODAN         LAP-----KV-LPVLQNSAKAA-IASCVGEK-------------------N
069 tr|G0SE99|G0SE99_CHATD         IRD-----TV-LPVLRTSAEAA-VSTCTGEE-------------------N
070 tr|J8Q6F0|J8Q6F0_SACAR         SSD-----TI-DDLIKTSAEAA-AKSCDGGT-------------------
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH TED-----SI-TPLVDASAAAA-AVSCSGGS-------------------
072 tr|J3K6E2|J3K6E2_COCIM         TRE-----LI-MPKLRASAMAA-AKACTGGN-------------------
073 tr|G4N3E1|G4N3E1_MAGO7         TAA-----TV-IPALKKSAEAC-IKQCTGGP-------------------T
074 tr|C5M5J2|C5M5J2_CANTT         SRD-----TI-LPYMEASAAGA-AESCSGGS-------------------
075 tr|A0A254U0X0|A0A254U0X0_ASPNG THD-----LI-LPKLQGSAAAI-SKQCSGTG-------------------
076 sp|Q75DG6|DCW1_ASHGO           SRE-----TI-VRWIRASAQGA-AASCSGGR-------------------
077 tr|Q6BZF0|Q6BZF0_DEBHA         TEP-----VI-TEWLVASANAAAKNSCTGGS-------------------
078 tr|E9E3Q1|E9E3Q1_METAQ         TKD-----KI-LPVLRKSTEAA-IKQCTGGP-------------------T
079 sp|P36091|DCW1_YEAST           TRN-----QI-MSWLNTSAIAA-AKSCSGGT-------------------
080 tr|W0TAH6|W0TAH6_KLUMD         TRD-----DI-MKLLKASAKGA-AQSCSGGT-------------------
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA TYN-----TI-YTWLVASAEGA-AKSCTGGY-------------------
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI TAS-----TI-QPLLLASGKAA-AATCTGTA----------DGFKGI
083 tr|Q2TYU3|Q2TYU3_ASPOR         TYD-----EI-LPKLQASAQGA-AASCSGMS-------------------
084 tr|G2XGF6|G2XGF6_VERDV         LFD-----RI-MGLLRPNAQSV-ASVCVGDP----------DQ-GFPGI
085 tr|G3AMT4|G3AMT4_SPAPN         TEP-----II-TKWLVDSANSA-AGSCSGGT-------------------
086 tr|A0A1C1D213|A0A1C1D213_9EURO TYP-----LL-KPILEQSAAAA-AATCTGGV-------------------
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS TRD-----AI-EKALRNSTAAA-VDSCRGPL-------------------H
088 tr|A0A100I5A6|A0A100I5A6_ASPNG THD-----II-LPKLQGSAAAI-SKQCSGTG-------------------
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH TAE-----KI-DPLLTASAEGA-AKSCSGGT-------------------
090 tr|G2Q8A7|G2Q8A7_MYCTT         LSE-----KI-LPVLKTSTEAA-VSTCTGEA-------------------N
091 tr|B8NPZ1|B8NPZ1_ASPFN         TRE-----LI-TPKLQGSAAAI-SKQCSGGD-------------------
092 tr|Q7SAB2|Q7SAB2_NEUCR         TRD-----AI-EKALRNSTAAA-VDSCRGPL-------------------H
093 tr|A2R8R5|A2R8R5_ASPNC         TYD-----RI-LPKLQGSAQGA-AATCTGYG-------------------
094 tr|C5DHG0|C5DHG0_LACTC         TAE-----TI-DPLMEASAEGA-AKSCTGGT-------------------
095 tr|G0SFA3|G0SFA3_CHATD         TFQ-----TI-QPLLLADAAAA-AQACTGSVRTLVPNIPGARPEPPFRGE
096 tr|K1WJG5|K1WJG5_MARBU         YHD-----TI-MPLLKTSAAAA-AAQCSGGT-------------------
097 tr|H2B005|H2B005_KAZAF         TRD-----TV-MHWVNTSAVAA-ATSCSGGT-------------------
098 tr|C5DYD7|C5DYD7_ZYGRC         VSD-----KV-DQLLNASAEGA-AKSCSGGI-------------------
099 tr|G8YM23|G8YM23_PICSO         TSD-----TI-MTWLTASAKAA-AQSCSGGT-------------------
100 tr|W3X8E3|W3X8E3_PESFW         TAD-----KI-LPVLKTSATAG-IQQCTGGD-------------------N
101 tr|A0A254U3K7|A0A254U3K7_ASPNG TYD-----RI-LPKLQGSAQGA-AATCTGYG-------------------
102 tr|Q6CER8|Q6CER8_YARLI         LAG-----SI-LSYLRPSAYAA-ARTCSGGF-------------------
103 tr|G2QT05|G2QT05_THITE         ISA-----KV-LPVLKTSTAAA-VDTCTGGT-------------------N
104 sp|Q5ACZ2|DFG5_CANAL           TKD-----RI-REVLEASAEGA-AKSCSGGS-------------------
105 tr|G2WKU6|G2WKU6_YEASK         TRD-----TI-DDLIKTSAEAA-AKSCNGGT-------------------
106 tr|A0A0C7N752|A0A0C7N752_9SACH TAD-----TI-NPLMEASAAAA-AKSCSGGS-------------------
107 tr|W3WN68|W3WN68_PESFW         TYD-----TI-VPLLKSSATAA-AEQCSGSP----------TSGFKGH
108 tr|G2QFS1|G2QFS1_MYCTT         TAG-----TI-LPVLRTSAEAA-ARQCTGGD-------------------T
109 tr|G3J9G4|G3J9G4_CORMM         TSD-----SI-TKLISTAASDA-VKVCNGAN----------GASYPGP
110 tr|Q2UJ03|Q2UJ03_ASPOR         ILD-----QV-MPKLRTSASAA-ARTCTGGS-------------------
111 tr|G8YRT2|G8YRT2_PICSO         TSD-----DI-MKLLDKSASAA-GESCSGGT-------------------
112 tr|W3X554|W3X554_PESFW         TFD-----TI-MPLLQSTATAA-AQQCNGSP----------AAALYKGP
113 tr|Q1K7A8|Q1K7A8_NEUCR         TYE-----RI-NPWIRATAAAA-VATCTGPVGAAAPQVDSGGIQPGFKGI
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM TSS-----EV-VPRIQQSALAA-AKQCSGGQ-------------------
115 tr|A7F3P0|A7F3P0_SCLS1         TEP-----TI-MTYINASAQGA-AAQCSGGS-------------------
```

```
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA  TSD-----GI-KTILQQAAKDA-AAACIGPR----------NEKYQGA
117 tr|Q6CIP9|Q6CIP9_KLULA         TTD-----TI-TPLIDTSAAAA-AKSCSGGT--------------------
118 tr|H2AMQ5|H2AMQ5_KAZAF         TRT-----QI-MTWMNASAYAV-AESCSGGT--------------------
119 tr|G8JMI3|G8JMI3_ERECY         TRE-----TI-DPLLEASAAAA-AASCSGGK--------------------
120 tr|C5DGT1|C5DGT1_LACTC         ARN-----QI-MDWIRASAKGA-AQSCSGGT--------------------
121 tr|K1XJR4|K1XJR4_MARBU         TRD-----RI-MPKLRKSAQAA-AAQCTGGA--------------------T
122 tr|A0A124BYC5|A0A124BYC5_ASPNG TSS-----DI-LPRLQGSAVGA-AKQCSGGS--------------------
123 tr|A7TLL2|A7TLL2_VANPO         TQH-----TI-MTYMNASAYGA-AQSCSGGY--------------------
124 tr|Q0CB86|Q0CB86_ASPTN         TYD-----QI-LPKLQGSAQSA-AQTCTGMG--------------------
125 tr|K1XLH0|K1XLH0_MARBU         THD-----GI-MKEIKTSAMAA-IQTCTAGD--------------------
126 tr|J7S438|J7S438_KAZNA         TRT-----QI-MGWLNTSAVAA-GKSCSGGY--------------------
127 tr|Q4WG09|Q4WG09_ASPFU         TAA-----RI-FPVLQGSALAL-SKQCSKAP--------------------
128 tr|H8WZ09|H8WZ09_CANO9         FED-----MI-RPYLETSAEAA-AQSCSGGS--------------------
129 tr|I1RL09|I1RL09_GIBZE         TAK-----KI-LPVLKTSTEAA-VKQCTGGD--------------------S
130 tr|Q4W985|Q4W985_ASPFU         TYD-----WV-MPRLRASAAAA-ARTCTGGP--------------------
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU TSD-----TI-DALIKTSAEAA-AKSCDGGR--------------------
132 tr|W6QCW3|W6QCW3_PENRF         TYD-----IV-MKRLKASAQAA-AKTCTGGI--------------------
133 tr|W3WMD3|W3WMD3_PESFW         TYD-----TI-MPVLRTSVEAA-VKQCTGGD--------------------N
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH LRD-----QI-MGWINTTAQAA-AQSCVGGY--------------------
135 tr|I2GYH1|I2GYH1_TETBL         TRD-----II-DPLILASAQGA-ARTCNGGT--------------------
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH VSD-----QI-DTLLTASAAAA-AKSCSGGT--------------------
137 tr|Q752P3|Q752P3_ASHGO         TRG-----TI-DELIYASATAA-SKSCEQGT--------------------
138 tr|B8MYP3|B8MYP3_ASPFN         TQD-----QI-LPKIQASAQAA-AKQCSGGD--------------------
139 tr|Q2TXL6|Q2TXL6_ASPOR         TLD-----QI-LPKLQGSAEGA-AKQCSGPS--------------------
140 tr|Q0CN17|Q0CN17_ASPTN         TSG-----EI-LPKIQASAQAA-AKQCSGGS--------------------
141 tr|J3NHD5|J3NHD5_GAGT3         IHD-----KT-LAVLRKSAEAA-VATCTGGS--------------------S
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI TAD-----QI-SKFLLISAKAA-AAQCTGGD--------------------
143 tr|A1CVB0|A1CVB0_NEOFI         TYD-----WV-MPRLRASAAAA-ARTCTGGP--------------------
144 tr|B8NVI3|B8NVI3_ASPFN         TLD-----QI-LPKLQGSAEGA-AKQCSGPS--------------------
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH ARD-----KI-MSWIHASAKGA-AQSCSGGT--------------------
146 tr|A1DJ54|A1DJ54_NEOFI         TAA-----RI-FPVLQGSALAI-SKQCSKAP--------------------
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS TRD-----TI-LPVLKTSAEAA-AKQCTGGA--------------------T
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA TYD-----DI-YSLLDASARAA-SNSCSGGT--------------------
149 tr|A0A177A618|A0A177A618_9PEZI LAP-----KI-MGLLQSSAVAA-ANQCIPGP--------------------
150 tr|Q75CW6|Q75CW6_ASHGO         TKA-----RI-EPNIRKSAEAA-AKSCSGGR--------------------
151 tr|Q7S4K4|Q7S4K4_NEUCR         TRD-----TI-LPVLKTSAEAA-AKQCTGGA--------------------T
152 tr|G2QB99|G2QB99_MYCTT         TLD-----AV-RPLLEADAAAA-ARACTGDP----------APPVFRGH
153 tr|G0S3F2|G0S3F2_CHATD         TKD-----RI-LPVLRTSAEAA-VKQCVGPP--------------------T
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO TEA-----VI-RPWLVNAANAA-ANSCSGGT--------------------
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB TSD-----TI-DPLLKASAMAA-AGSCDGGT--------------------
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM TYD-----RI-LPKLQGSAVAA-GATCTGNG--------------------
157 tr|G2QKJ0|G2QKJ0_MYCTT         TYE-----TI-TKTLRSSTEGC-VSSCLADG--------------------T
158 tr|Q4WFX5|Q4WFX5_ASPFU         TRD-----TI-ATRLRASATAA-AAQCVGGK--------------------
159 tr|Q5BGD7|Q5BGD7_EMENI         LRE-----KI-TPKLQGSALAI-GRSCDGSS--------------------
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH TRD-----NI-LTLLKASALGA-AEACSGGT--------------------
161 tr|G8BYN9|G8BYN9_TETPH         TAA-----TL-DTLIETSAKAA-AGSCDGGT--------------------
162 tr|F7VVT4|F7VVT4_SORMK         VRE-----PI-IKTFKTNIEGV-IDSCLADG--------------------T
163 tr|G8YB80|G8YB80_PICSO         TSD-----TI-MTWLTASAKAA-AASCSGGT--------------------
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA LYD-----DI-MPLLDTSASAA-AQSCVGGI--------------------
165 tr|G2WRS7|G2WRS7_VERDV         TAD-----TI-LPVLRNSAEKA-IAQCTGGP--------------------E
166 tr|A0A124BVB5|A0A124BVB5_ASPNG TYE-----TI-MARLQPTAKAA-AAQCVGGK--------------------
167 tr|A7EIG7|A7EIG7_SCLS1         TYD-----AV-MAYLRPSAAGA-AASCVGGD--------------------
168 tr|C5M5U7|C5M5U7_CANTT         TEP-----VI-RKWLVDSANAA-AGSCSGGS--------------------
169 tr|B9WAA8|B9WAA8_CANDC         TKD-----KI-REVLEASAEGA-AKSCSGGS--------------------
170 tr|F8MRU1|F8MRU1_NEUT8         TRD-----TI-LPVLKTSAEAA-AKQCTGGS--------------------T
171 tr|A1DJS0|A1DJS0_NEOFI         TYD-----RI-LPKLQGSAVAA-GATCTGNG--------------------
172 tr|B6HLM8|B6HLM8_PENRW         LKE-----KI-LPRLQVSAEGA-AKSCTGDG--------------------
173 tr|A0A117DWS0|A0A117DWS0_ASPNG TSD-----QI-IPKIQQSALAA-AKQCSGGD--------------------
174 tr|A0A124BXM1|A0A124BXM1_ASPNG TSS-----LI-LPKLQGSAEAA-AASCNGMG--------------------
175 tr|J3NZQ6|J3NZQ6_GAGT3         VKD-----EI-RPVLRKSAAQA-LKTCTGAP--------------------T
176 tr|W0THH8|W0THH8_KLUMD         TLD-----TV-NPLIEASATAA-AKSCSGGT--------------------
177 tr|B2AEF2|B2AEF2_PODAN         TRD-----KI-LPVLRTSTEAA-VKQCTGGD--------------------T
178 tr|Q96TX1|Q96TX1_NEUCS         TAP-----VI-LPVLKTSAEAA-VQTCTGEQ--------------------N
179 tr|E9DYA6|E9DYA6_METAQ         TAE-----RI-LPVLQKSAQAA-VTQCTGGA--------------------N
180 tr|B9WAI2|B9WAI2_CANDC         TEP-----VI-TKWLVDSANAA-AGSCTGGS--------------------
181 tr|Q6C171|Q6C171_YARLI         LAS-----DI-MWYLEASAYAA-ARTCSGGR--------------------
182 tr|A6ZMV1|A6ZMV1_YEAS7         TRD-----TI-DDLIKTSAEAA-AKSCNGGT--------------------
183 tr|B6K7X7|B6K7X7_SCHJY         TYD-----TI-MPLMKTSATAA-ALACSGGS--------------------

184 tr|I1RSA2|I1RSA2_GIBZE         TSA-----RI-IPLLKTSAQAA-AKVCTGS-----------GDGYKGP
```

```
185 tr|F7VWZ9|F7VWZ9_SORMK          TYE-----RI-NPWIRSTAIAA-AAVCTGPVGAAAPQVDSGGIQPGFRGI
186 tr|A0A1B2J789|A0A1B2J789_PICPA TAE-----TI-MNWMNTSAVAV-AQSCSGGT-------------------
187 tr|G0V8N4|G0V8N4_NAUCC         TQP-----TI-MKWLNASAYAA-AQSCSGGT-------------------
188 tr|H8WXN8|H8WXN8_CANO9         TKG-----TI-YKWLVDSANAA-AGACSGGF-------------------
189 tr|A6ZZS0|A6ZZS0_YEAS7         TRN-----QI-MSWLNTSAIAA-AKSCSGGT-------------------
190 tr|G9MQD1|G9MQD1_HYPVG         TKD-----KI-LPILRKSTEAA-VKQCTGGD-------------------S
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH ARE-----KI-MDWIRTSAQGA-AQSCSGGY-------------------
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM TAA-----RI-FPVLQGSALAL-SKQCSKAP-------------------
193 tr|A0A254UET1|A0A254UET1_ASPNG TYD-----TI-MARLQPTAKAA-AAQCVGGK-------------------
194 tr|A1D587|A1D587_NEOFI          TSS-----EI-VPRIQQSALAA-AKQCSGGS-------------------
195 tr|Q1K7I4|Q1K7I4_NEUCR         TAP-----VI-LPVLKTSAEAA-VQTCTGEQ-------------------N
196 tr|G8JS88|G8JS88_ERECY         SRE-----IA-MKWIRDSAKGA-ALSCSGGR-------------------
197 tr|A0A136J7D3|A0A136J7D3_9PEZI TYD-----TL-MPILRTNAAAA-ARQCTGSP----------PASKYRGP
198 tr|E9E4X8|E9E4X8_METAQ         TTE-----TI-RPVLQTSAEAA-VKQCTGAA-------------------L
199 tr|Q6CP42|Q6CP42_KLULA         TRD-----TV-LALLQASAQGA-ADSCSGGT-------------------
200 tr|G0WHL4|G0WHL4_NAUDC         TSD-----TV-DPLIKTSAAAA-AKSCTGGT-------------------
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU TRN-----QI-MSWLNTSAIAA-AKSCSGGT-------------------
202 tr|F9WYX6|F9WYX6_ZYMTI         SDDSVYRQNI-TAVLRASATAA-AAACVYGE-------------------
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 IHD-----KT-LAVLRKSTAAA-VETCTGGD-------------------N
204 tr|J3P147|J3P147_GAGT3         MRG-----TI-MPVMRKSIEAG-LRQCTGGP-------------------R
205 tr|I2H842|I2H842_TETBL         TRD-----TI-MQWIDTSAYAA-AQACSGGT-------------------
206 tr|S6E3R3|S6E3R3_ZYGB2         VSD-----QI-DTLLTASAAAA-AKSCSGGT-------------------
207 tr|E5AD94|E5AD94_LEPMJ         THD-----LI-MPKLQVSAKAA-AQTCEGPSQHN----------------G
208 tr|Q6FJM7|Q6FJM7_CANGA         TSD-----TI-DPLLKASAMAA-AGSCDGGT-------------------
209 tr|A5DUW0|A5DUW0_LODEL         TEP-----TI-NKWLVDSANGA-AGSCSGGS-------------------
210 tr|G8BF46|G8BF46_CANPC         TSG-----II-YKWLVDSANAA-AGACSGGI-------------------
211 tr|W3XAF6|W3XAF6_PESFW         VYD-----TI-MPLLQTTAVAA-AEHCVS-H----------TSGFKGH
212 tr|B6H7I2|B6H7I2_PENRW         TAA-----QI-KPKLASSAEAA-AKSCTGNN-------------------
213 tr|Q5ATF9|Q5ATF9_EMENI         TYD-----RI-FAKLQTSAQAA-ALSCSGAG-------------------
214 tr|A0A167CDD3|A0A167CDD3_9ASCO LQD-----EI-MPLIQSSAKAA-AASCSGGT-------------------
215 tr|G0RXS3|G0RXS3_CHATD         TRD-----TI-MRTLRSSTEGC-VSSCLADG-------------------T
216 tr|B8NX26|B8NX26_ASPFN         TKD-----QI-VPKLQGSVESI-SKMCNGQS-------------------
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM TRD-----RL-LPKLQASATAA-ALQCSGP--------------------
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM TRD-----TI-ATRLRASATAA-AAQCVGGK-------------------
219 tr|A1CMM3|A1CMM3_ASPCL         TYD-----LL-MPKMRASAEAA-ALQCSGP--------------------
220 tr|B6HU84|B6HU84_PENRW         TYD-----LI-YPQLKTSAQAA-AKQCSGGD-------------------
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS TAP-----VI-LPVLKTSAEAA-VQTCTGEQ-------------------N
222 tr|G2WHY6|G2WHY6_YEASK         TRN-----QI-MSWLNTSAIAA-AKSCSGGT-------------------
223 tr|G0W7G8|G0W7G8_NAUDC         TRS-----KI-LPLINASAIAA-GKSCSGGT-------------------
224 tr|G4NGV8|G4NGV8_MAGO7         TFN-----TI-QPILRNNAEAA-AVACSGSP----------AA-GFSGK
225 tr|W3X855|W3X855_PESFW         TYD-----TI-YPLLEASAQGA-AKQCDGTG----------EGTAGRTL
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA IAD-----KV-LPVLRKSTAAC-VKQCTGGE-------------------G
227 tr|J8Q2K7|J8Q2K7_SACAR         TRK-----QI-MSWLNTSAIAA-AKSCSGGT-------------------
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO TYD-----II-HSWLVDSANAA-ASSCSGGL-------------------
229 tr|J6EGN8|J6EGN8_SACK1         TSD-----TI-DDLIKTSAEAA-AKSCNGGT-------------------
230 tr|A0A100ISD9|A0A100ISD9_ASPNG TYD-----LV-MPKLRASAKAA-AETCTGGE-------------------
231 tr|G8YQC0|G8YQC0_PICSO         TSD-----TI-MNLLDKSASAA-GESCSGGT-------------------
232 tr|F7VVP8|F7VVP8_SORMK         MKD-----TI-LPVLKTSAEAA-AKQCTGGA-------------------T
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH TRD-----TI-DDLIKTSAEAA-AKSCNGGT-------------------
234 tr|F8MBP4|F8MBP4_NEUT8         TYE-----RI-NPWIRATAAAA-VATCTGPVGAAAPQVDSGGIQPGFKGI
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG TSS-----DI-LPRLQGSAVGA-AKQCSGGS-------------------
236 tr|A1CD27|A1CD27_ASPCL         TAD-----RI-YPVLQGSALAI-SKQCS-AP--------------------
237 tr|A5DHF3|A5DHF3_PICGU         MYD-----EM-RPYVEASAEGA-AKSCSGGS-------------------
238 tr|D4ATT7|D4ATT7_ARTBC         TIE-----KI-NPLLRESAKAA-ALQCNGP--------------------
239 tr|E2PT42|E2PT42_ASPNC         TYD-----LV-MPKLRASAKAA-AETCTGGE-------------------
240 tr|C7ZPE5|C7ZPE5_NECH7         VSD-----KI-LPVLQTSAEAA-IKQCTGGT-------------------T
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA TAE-----QL-MPKLQASALAA-AKVCTGST-------------------
242 tr|Q4WKP7|Q4WKP7_ASPFU         TSS-----EV-VPRIQQSALAA-AKQCSGGQ-------------------
243 tr|F8MXT4|F8MXT4_NEUT8         TAP-----VI-LPVLKTSAEAA-VQTCTGEQ-------------------N
244 tr|A0A177AJ74|A0A177AJ74_9PEZI TAD-----TI-SKFLLTSAKAA-AAQCTGGV-------------------
245 tr|J7RQR5|J7RQR5_KAZNA         TAS-----TI-DPLMQSSAQAA-AQSCSGGT-------------------
246 tr|Q75CW7|Q75CW7_ASHGO         TKA-----RI-EPNIRKSAEAA-AKSCSGGR-------------------
247 tr|G3AKK4|G3AKK4_SPAPN         TES-----VI-APYLEASSAAA-AQSCSGGS-------------------
248 tr|B2W762|B2W762_PYRTR         TRP-----QI-DPLLQTSAKAA-AKQCNAGV-------------------
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH TRN-----QI-MSWLNTSAIAA-AKSCSGGT-------------------
250 tr|C4XWK1|C4XWK1_CLAL4         HFD-----QIYNNWIVPSAKAA-AASCSGGS-------------------

001 tr|A0A254U2J9|A0A254U2J9_ASPNG N---NTCG-VRWWP--K-----KWDG-----W-NGMEEEIAVTNVLSSAL
```

```
002 sp|Q6FLP9|DCW1_CANGA                    D--GHTCG-LNWFR--D-----DWDG-----MY-GLGEQMAALEVMVNTQ
003 tr|A1CCM5|A1CCM5_ASPCL                  N---SSCG-VRWHT--S-----SWDG-----Q-TGMEEQISVTDVLSVNL
004 tr|C1H190|C1H190_PARBA                  ---DNSCG-LRWTK--Q---AEFDG-----ET-GVGEQMSALEVIQANL
005 tr|Q2UR85|Q2UR85_ASPOR                  S--KTDCG-RSWYK-QD-----KWDG-----S-KSLESDMSALSVLSSTM
006 tr|A5DV30|A5DV30_LODEL                  D--GMTCG-EDWST--S-----GWDG-----VY-GLPEQISALEVIMSLV
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO          D--GHTCG-MSWSD--G-----DWDG-----FY-GLGEQIAALEIIQNSV
008 tr|J3PGN0|J3PGN0_GAGT3                  ---GRACG-FYWTG--G-----EYIDPAID-KTS-GAGERMNVLAAVSSLL
009 tr|A0A124BXU6|A0A124BXU6_ASPNG          N---NTCG-VRWWN--S-----TWDG-----W-SGLEEQMSVTSVFSSNM
010 tr|C5P4A1|C5P4A1_COCP7                  D--GKSCG-LKWST--G-----GFDG-----SV-GIGEQMAALEVIQSNL
011 tr|Q0CG55|Q0CG55_ASPTN                  E--HGMCA-LKWTE--R-----KWDG-----ME-DVGLQIGALEVIQSTL
012 tr|C4Y2X5|C4Y2X5_CLAL4                  D--GHTCG-LSWTN--G-----TWDG-----NY-GLGEQMSALEVMNNLR
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG          L--QATCG-LKWTD--R-----KWDG-----MD-DVGIQMAALEVMQSTL
014 tr|C7Z068|C7Z068_NECH7                  ---GRACG-FYWSD--G-----EFVDPAVD-ETS-GAGEQMNVLAAVSSLL
015 tr|G8BXX2|G8BXX2_TETPH                  D--GHTCG-INWFY--G-----GWDG-----KY-GLGEQMAALEILVNTR
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR          D--GHTCG-LSWTN--G-----TWDG-----VY-GLGEQMNALEVIQNLL
017 tr|A1CSC2|A1CSC2_ASPCL                  S--HTQCG-RRWHQ--D-----TWDG-----T-TSLEEDMSALSVFSSTM
018 sp|Q05031|DFG5_YEAST                    D--GHTCG-LNWQK--Q-----TNDG-----YY-GLGEQMSALEVIQNLL
019 tr|H2AQJ6|H2AQJ6_KAZAF                  D--GHTCG-LNWQD--K-----TCDG-----NY-GLGEQMSALEVIQNLL
020 tr|G2XFH9|G2XFH9_VERDV                  ---GRVCG-FYWDR--G-----EFIDPAVD-KTT-GAGEQMNVLAAVSSLL
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH          D--GHTCG-LNWQA--K-----GWDG-----KY-GLGEQMSALEALVVTQ
022 sp|Q5AD78|DCW1_CANAL                    D--GVTCG-LSWTDWSQ-----GWDG-----KW-GLGEQMSALEVMQNLM
023 tr|B6H7E8|B6H7E8_PENRW                  D--GKHCG-IRWTM-KS-----EWDG-----T-MGLEQQMSVLGVLNAVM
024 tr|B6GZU2|B6GZU2_PENRW                  N--KNVCG-LKWTE--Q-----KWDK-----TK-DFGQQMAALEVIQANL
025 tr|E9E4W7|E9E4W7_METAQ                  ---GRQCG-FYWTE--G-----KFIDPRVD-KTS-GAGEAMDVLAAVSSML
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO          L--GTTCG-LKWTD--Q-----KWDH-----TK-DFGQQMAALEVVQANL
027 tr|G2R7G8|G2R7G8_THITE                  -------CG-FRWNT--G-----SYDGDNV---VG-PAGQEMSALAALSTLL
028 tr|G3JBY1|G3JBY1_CORMM                  D--GTACG-FSWLE--P-----KYDG-----LA-GVGPQMNALAAVFYNL
029 tr|A5DNV5|A5DNV5_PICGU                  D--GHTCG-INWFL--G-----HWDG-----VY-GLGEQMCALEVITALR
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI          N--GRTCG-LKWSN--W---GNWDG-----SN-GIGQQMAALEVLQSNL
031 tr|Q2TWC1|Q2TWC1_ASPOR                  ---ENLCG-ETWGK--D-----TWDG-----M-KGLEVQMAALGGITSNL
032 tr|J5RXY9|J5RXY9_SACK1                  D--GHTCG-LNWFN--G-----TWDG-----MY-GLGEQMAALEVMVNTR
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA          D--GTGCG-FSWLN--G-----TNDG-----LM-GVGPQMDALAAVFYNL
034 tr|A0A254U5V1|A0A254U5V1_ASPNG          S--GTQCG-RRWYQ--A-----QWDG-----E-TSLESDMSALSVFSSNM
035 tr|A0A061B6H7|A0A061B6H7_CYBFA          D--GHTCG-MNWWH--S-----GWDG-----FY-GLGEQMSALEIMQNLR
036 tr|G8ZQ93|G8ZQ93_TORDC                  D--GHTCG-INWFY--D-----GWDG-----MY-GLGEQMSALEALVNTQ
037 tr|Q9C2J1|Q9C2J1_NEUCS                  D--GTACG-FKWTQ--------TFDG-----WA-GVGAQMNALSAVMYTL
038 tr|C5NZK5|C5NZK5_COCP7                  ---DNACG-LRWTK--G---EAYDG-----ST-GVGEQMAALEIIQSNL
039 tr|Q6C0T7|Q6C0T7_YARLI                  D--GHTCS-LDWLT--G-----KYDD-----QYIGLGEQLSVLGAILNTQ
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB          D--GHTCG-LNWFR--D-----DWDG-----MY-GLGEQMAALEVMVNTQ
041 tr|F7VZ72|F7VZ72_SORMK                  ---GRMCG-FNWRK--K-----KYDG-----SH-GVGQQMNALGAVSALL
042 tr|A3LN37|A3LN37_PICST                  D--GHTCG-LNWFN-ST-----GWDG-----YW-GLGEQISALEVIQNLR
043 tr|Q6CAI2|Q6CAI2_YARLI                  D--GHTCS-IDWLT--G-----SYSN-----DWIGLGEQMSALEVLQNNL
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO          D--WHTCG-LNWQN--G-----TWDG-----YY-GLGEQMSALEVVTNLR
045 tr|G9MJS5|G9MJS5_HYPVG                  ---GRQCG-IKWNT--G-----KYDG-----RT-GAGQEMNVVGAVSSLL
046 tr|G9MHI4|G9MHI4_HYPVG                  D--GTGCG-FAWLP--K---GTYDG-----KN-GVGSQMNALDAVMYTM
047 tr|D4AP27|D4AP27_ARTBC                  E--GTTCG-LKWTT--G-----KFDG-----ST-GVGEQLAAMEIFQSNL
048 tr|B6GZT8|B6GZT8_PENRW                  N---GTCG-VQWYR--S-----TWDG-----W-MGMEEQISATKVFIANL
049 sp|O74556|YCZ2_SCHPO                    D--GVTCG-VQWWW--N--NDTWDG-----LY-GLGEQMSALEAIQAPL
050 tr|A0A0C4E991|A0A0C4E991_MAGP6          ---GRACG-FYWTS--G-----EYIDPAID-KTS-GAGERMNVLAAVSSLL
051 tr|B8MHG0|B8MHG0_TALSN                  ----SACG-MQWFS--Q-----KYDG-----T-SAIEQEMSAMSIFSNAL
052 tr|A3LMV8|A3LMV8_PICST                  D--GITCG-MNWAS--G-----NWDG-----IY-GLGEQTSALEVMNALI
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU          D--GKTCG-INWGI--N-----GWDG-----LY-GLGEQISALEIIQNSL
054 sp|Q9P6I3|YHG7_SCHPO                    D--GVTCG-YMWYWNNG-----TWDD-----H-YGLGEQISAVETFQALL
055 tr|Q2US57|Q2US57_ASPOR                  N---NSCG-VVWYN--S-----TWDG-----W-SGLEEQMSVTSILAANM
056 tr|W7N622|W7N622_GIBM7                  A--GTACG-FTWLT--N-----GFDG-----IV-GVGPQMSSLQSIMYTL
057 sp|Q9P6I4|YHG6_SCHPO                    S--GTSCG-IYWFWNNG-----TWDD-----N-YGVQEQFSALQAVQMLM
058 tr|C7GRB3|C7GRB3_YEAS2                  D--GHTCG-LNWQK--Q-----TNDG-----YY-GLGEQMSALEVIQNLL
059 tr|C7GP28|C7GP28_YEAS2                  D--GHTCG-LNWFN--G-----TWDG-----MY-GLGEQMSALEVMVNTR
060 tr|G3JGZ5|G3JGZ5_CORMM                  ---GRTCG-FYWSG--G-----KFVDPSVD-KTT-GAGEYMSALAAVSSLL
061 tr|C4QXV4|C4QXV4_KOMPG                  D--GHTCG-MNWLA--D-----GWDG-----FY-GLGEQMSALETLQNTR
062 tr|I1S0Z5|I1S0Z5_GIBZE                  ---GRVCG-FKWAS--G-----KYDG-----KT-GAGQQMNVLAAVSSLL
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM          D--GAACG-LKWTT--G-----VWDG-----SE-DVGLQMSALEVIQNLL
064 tr|G0W5X4|G0W5X4_NAUDC                  D--GHTCG-LNWFD--N-----GWDG-----YY-GLGEQMSALDVLVNTR
065 tr|G8ZPC1|G8ZPC1_TORDC                  D--GHTCG-LNWNK--Q-----ANDG-----YY-GLGEQMSALEVMQQLL
066 tr|A0A167CDC5|A0A167CDC5_9ASCO          D--GHTCG-LNWNV--G-----KWDG-----VW-GLGEQMSALEVMQNTR
067 tr|R9XAT7|R9XAT7_ASHAC                  D--GHTCG-INWLT--N-----SWDG-----TW-GLGEQMAALEVMQNLR
068 tr|B2B747|B2B747_PODAN                  G--QRACG-FKWST--G-----TFDG-----SQ-GAGQTMNVLGAVSSLL
069 tr|G0SE99|G0SE99_CHATD                  ---GRTCG-FQWVK--R-----QYDG-----SK-GAGQQMNVLGAVSALL
070 tr|J8Q6F0|J8Q6F0_SACAR                  D--GHTCG-LNWEK--Q-----TNDG-----YY-GLGEQMNALEVIQNLL
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH          D--GHTCG-LDWST-TS-----GWDG-----LY-GLGEQASALEIMNQLL
```

```
072 tr|J3K6E2|J3K6E2_COCIM          D--GKSCG-LKWST--G----GFDG-----SV-GIGEQMAALEVIQSNL
073 tr|G4N3E1|G4N3E1_MAGO7          ---GRACG-FYWTR--G----TYIDPSVD-KTT-GAGERMDALAAVSSLL
074 tr|C5M5J2|C5M5J2_CANTT          D--GVTCG-EDWST--G----SWDG-----VY-GLGEQMSSLEVIMSLI
075 tr|A0A254U0X0|A0A254U0X0_ASPNG  DGLENLCG-VRWHQ--A----TWDG-----S-MGLEVQMSALGGVTGAL
076 sp|Q75DG6|DCW1_ASHGO            D--GHTCG-LNWLI--N----GWDG-----KW-GLGEQMAALEIIQNLR
077 tr|Q6BZF0|Q6BZF0_DEBHA          D--WHTCG-LNWQN--G----TWDG-----YY-GLGEQMCALEVVTNLR
078 tr|E9E3Q1|E9E3Q1_METAQ          ---QRQCG-FYWSL--G----RFVDPAAD-KTT-GAGEQMNVLAAVSSLL
079 sp|P36091|DCW1_YEAST            D--GHTCG-LNWFN--G----TWDG-----MY-GLGEQMSALEVMVNTR
080 tr|W0TAH6|W0TAH6_KLUMD          D--GHTCG-LNWFY--D----GWDG-----KY-GLGEQMAALEVMQNLR
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA  D--GHTCG-LNWQL--G----SNDG-----NY-GLGEQISALEVFDNLL
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI  A--NTACG-QKWYT--G----SFDG-----IN-GVGEQMASLAAVMYNL
083 tr|Q2TYU3|Q2TYU3_ASPOR          N---NTCG-VRWHE--S----KWDG-----W-VGMEEQISATDVLSSVL
084 tr|G2XGF6|G2XGF6_VERDV          G--GTACG-LAWQP--R---GQFDG-----LV-GVGEQMSALSAVMYTL
085 tr|G3AMT4|G3AMT4_SPAPN          D--GHTCG-LDWTNWTR----GWDG-----MY-GLGEQMSALEVMQNLM
086 tr|A0A1C1D213|A0A1C1D213_9EURO  N--GTSCG-SKWWK--G---AIFDG-----ET-GVGQQMCALEVIQSNL
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS  D--GRACG-FRWTT--G----GYDG-----LT-GAGQEMSVLAALSSLL
088 tr|A0A100I5A6|A0A100I5A6_ASPNG  DGLENLCG-VRWHQ--D----TWDG-----S-MGLEVQMSALGGVTGAL
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH  D--GHTCG-INWEK--G----SWDG-----MY-GLGEQASALEVIQSLL
090 tr|G2Q8A7|G2Q8A7_MYCTT          ---GRTCG-FQWSK--R----QYDG-----SH-GAGQQMNVLAAVSSLL
091 tr|B8NPZ1|B8NPZ1_ASPFN          ---QNLCG-ERWYS--T----EPVG-----P-TGLEVQMAALGGITSNL
092 tr|Q7SAB2|Q7SAB2_NEUCR          D--GRACG-FRWTT--G----GYDG-----LT-GAGQEMSVLAALSSLL
093 tr|A2R8R5|A2R8R5_ASPNC          N---NTCG-VRWWN--S----TWDG-----W-SGLEEQMSVTSVFSSNM
094 tr|C5DHG0|C5DHG0_LACTC          D--GHTCG-LNWQL--G----YHDG-----YY-GLGEQACALEGIQSLM
095 tr|G0SFA3|G0SFA3_CHATD          P--GTACG-FRWTV--N----GFDG-----QT-GVGEQMNALAAVMYSM
096 tr|K1WJG5|K1WJG5_MARBU          D--GVTCG-GHWYL--------DYDG-----NY-GLGQQMSALSVIQAML
097 tr|H2B005|H2B005_KAZAF          D--GHTCG-INWFY--G----GWDG-----YY-GLGEQMSALEIMVNTR
098 tr|C5DYD7|C5DYD7_ZYGRC          D--GHTCG-LNWFK--G----SWDN-----QY-GLGEQMSALEVIQNTL
099 tr|G8YM23|G8YM23_PICSO          D--GHTCG-TNWFK--H----SWDG-----NY-GLGEQMCALEVIQNLR
100 tr|W3X8E3|W3X8E3_PESFW          ---GRTCG-FGWAS--G----TFDG-----SV-GAGQTMNVLGAISSLL
101 tr|A0A254U3K7|A0A254U3K7_ASPNG  N---NTCG-VRWWN--S----TWDG-----W-SGLEEQMSVTSVFSSNM
102 tr|Q6CER8|Q6CER8_YARLI          D--GHTCS-INWLT--G----EYNV-----SYIGLGEQLSALEVILNTQ
103 tr|G2QT05|G2QT05_THITE          ---QRTCG-FKWST--K----AFDG-----SV-GAGQQMNVLAAVSSLL
104 sp|Q5ACZ2|DFG5_CANAL            D--GVTCG-ENWAI--D----KWDG-----VY-GLGEQTSALEVMMALI
105 tr|G2WKU6|G2WKU6_YEASK          D--GHTCG-LNWQK--Q----TNDG-----YY-GLGEQMSALEVIQNLL
106 tr|A0A0C7N752|A0A0C7N752_9SACH  D--GHTCG-LNWQN--G----SWDG-----WY-GLGEQASALEVIQSSL
107 tr|W3WN68|W3WN68_PESFW          T--GTACG-FSWLD--N---STWDG-----NS-GVGEQLNAMSIVMVNL
108 tr|G2QFS1|G2QFS1_MYCTT          ---GRRCG-FYWSE--G----VFVDPAVD-KTS-GAGEAMSVLAAVSSLL
109 tr|G3J9G4|G3J9G4_CORMM          D--GTGCG-FSWLN--G----TNDG-----LM-GVGPQMDALAAVFYNL
110 tr|Q2UJ03|Q2UJ03_ASPOR          D--HSTCG-MKWTS--G----QWDQ-----SD-DVGVQMSVLEVMQATL
111 tr|G8YRT2|G8YRT2_PICSO          D--GVTCG-TSWTE--G----KWDN-----SW-GLGQQMSALETMLGTI
112 tr|W3X554|W3X554_PESFW          A--GTACG-FQWTK--S---PEFDN-----MV-GVGEQQSALAAVMYNL
113 tr|Q1K7A8|Q1K7A8_NEUCR          D--GTACG-FKWTQ--------TFDG-----WA-GVGAQMNALSAVMYTL
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM  N--RTLCG-RRWHQ--D----TFDG-----T-SSLEEQMSALSVFSSSM
115 tr|A7F3P0|A7F3P0_SCLS1          S--GRACG-EHWTA--G---SAYDG-----KY-GVGEQMSALSVIQATL
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA  D--GTACG-FTWLE--G---KKYDG-----LA-GVGPQMNALAAVFYNL
117 tr|Q6CIP9|Q6CIP9_KLULA          D--GHTCG-LNWFT--G----DWDG-----QY-GLGEQACALEIINQLL
118 tr|H2AMQ5|H2AMQ5_KAZAF          D--GHTCG-LNWFL--G----SWDG-----YY-GLGEQMSALEVLVNTR
119 tr|G8JMI3|G8JMI3_ERECY          D--GETCG-LDWTT--G----SWDG-----YY-GLGEQASALEVIQNLL
120 tr|C5DGT1|C5DGT1_LACTC          D--GHTCG-LNWFY--D----GWDG-----KW-GLGEQMAALEIMQNVL
121 tr|K1XJR4|K1XJR4_MARBU          ---GRTCG-LSWSS--G----VYDG-----TV-GVGQQMAAMSALFVNI
122 tr|A0A124BYC5|A0A124BYC5_ASPNG  S--KTTCG-QHWYK--S----SWDG-----T-EGIEEEMSATSIFTSNL
123 tr|A7TLL2|A7TLL2_VANPO          D--GHTCG-INWFY--G----GWDG-----YY-GLGEQMCALEVMINTR
124 tr|Q0CB86|Q0CB86_ASPTN          N---NSCG-VRWYG--N----KWDG-----W-HGMEEQISVADVISSAM
125 tr|K1XLH0|K1XLH0_MARBU          S--GTQCG-LQWTL--K----QNDG-----SL-GVGEQMAVLEVIQSNL
126 tr|J7S438|J7S438_KAZNA          D--GHTCG-LNWFK--D----GWDG-----MY-GLGEQMSALECILNTR
127 tr|Q4WG09|Q4WG09_ASPFU          ---DNTCG-VRWWQ--P----TWDG-----FTPGLETQMAALAGITANL
128 tr|H8WZ09|H8WZ09_CANO9          D--GVTCG-ENWSA--S----GWDG-----VY-GLGEQMSSLEVILSLI
129 tr|I1RL09|I1RL09_GIBZE          ---GRACG-FYWSG--G----EFVDVAVD-GTS-GAGEQMNVLAAVSSLL
130 tr|Q4W985|Q4W985_ASPFU          D--GAACG-LKWTT--G----VWDG-----SE-DVGLQMSALEVIQNLL
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU  D--GHTCG-LDWEK--Q----TNDG-----YY-GLGEQMNALEVIQNLL
132 tr|W6QCW3|W6QCW3_PENRF          N--KNVCG-MKWTA--Q----KWDK-----TK-DFGQQMAALEVIQANL
133 tr|W3WMD3|W3WMD3_PESFW          ---GRFCG-FHWTT--G----TFDG-----KT-GAGQQMNVLGGLTSLL
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH  D--GHTCG-LNWQA--K----GWDG-----KY-GLGEQMSALEALVVTQ
135 tr|I2GYH1|I2GYH1_TETBL          D--GHTCG-LNWQA--S----THDG-----YY-GLGEQMSALEVIQNLL
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH  D--GHTCG-LNWFK--G----AWDG-----QY-GLGEQMSALEVIQNTL
137 tr|Q752P3|Q752P3_ASHGO          H--GYTCG-LNWHK--G----TYDG-----VY-GLGEQASALEIIQNVL
138 tr|B8MYP3|B8MYP3_ASPFN          S--KTDCG-RSWYK-QD----KWDG-----S-KSLESDMSALSVLSSTM
139 tr|Q2TXL6|Q2TXL6_ASPOR          P--DKICG-QRWFL--D----KYDG-----V-TGLREHMCALSVFTANM
140 tr|Q0CN17|Q0CN17_ASPTN          Q--QSECG-RRWYQ--A----TYDG-----T-NSLETDMSALSVFSSNM
141 tr|J3NHD5|J3NHD5_GAGT3          ---QRACG-FKWST--R----AFDG-----SV-GAGQQMNVLGAVTSLL
```

```
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI  K--GRACG-LRWVH--DGQIGVWDG-----TT-GVGEEMSALEVIQSTL
143 tr|A1CVB0|A1CVB0_NEOFI         D--GAACG-LKWTT--G-----VWDG-----SE-DVGLQMSALEVMQNLL
144 tr|B8NVI3|B8NVI3_ASPFN         P--DKICG-QRWFL--D-----KYDG-----V-TGLREHMCALSVFTANM
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH D--GHTCG-QNWFY--N-----GWDG-----YY-GLGEQMAALEIMQNTF
146 tr|A1DJ54|A1DJ54_NEOFI         ---DNTCG-IRWWQ--P-----TWDG-----FTPGLESQMAALAGITANL
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS ---GRVCG-FYWSG--G-----VFVDPAVD-KTT-GAGEAMDVLAAVSSLL
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA D--GHTCG-LNWWY--S-----GWDG-----YY-GLGEQMSALEVMQNLR
149 tr|A0A177A618|A0A177A618_9PEZI V--GAECG-LRWYE--G-----ATDG-----RE-GVGQQMGAMDVMGALL
150 tr|Q75CW6|Q75CW6_ASHGO         D--GHTCG-MVWRN--Y-----TWDG-----KY-GLGEQLSALEIIQNTL
151 tr|Q7S4K4|Q7S4K4_NEUCR         ---GRVCG-FYWSG--G-----VFVDPAVD-KTT-GAGEAMDVLAAVSSLL
152 tr|G2QB99|G2QB99_MYCTT         P--GTACG-FRWTT--G-----AFDG-----SA-GVGEQMNALSAVMYPL
153 tr|G0S3F2|G0S3F2_CHATD         ---GRRCG-FYWKS--G-----KFVDPSVD-HTS-GAGEAMSVLAAVSSLL
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO D--GVTCG-LNWFT--G-----GWDG-----KF-GLGEQMCALEAMTNLQ
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB D--GHTCG-LDWQL--K-----TNDG-----YY-GLGEQMSALEVIQQLL
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM N---NSCG-VRWYT--S-----KWDG-----W-TGMEEQISVTDVLSVSL
157 tr|G2QKJ0|G2QKJ0_MYCTT         -------CG-FRWNT--G-----EYDGDTA---AG-PAGQEMSALAALSTML
158 tr|Q4WFX5|Q4WFX5_ASPFU         T--GTYCG-MRWTT--G-----EFDG-----TT-GVGQQLSALEVVQANL
159 tr|Q5BGD7|Q5BGD7_EMENI         EGKSNLCG-SRWYQ--E-----TWDG-----I-QGLEVQQAALGGITANL
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH D--GHTCG-LNWFY--G-----GWDG-----KY-GLGEQMAALEVMQNVR
161 tr|G8BYN9|G8BYN9_TETPH         D--GVTCG-LSWFD--S-----TNDG-----YY-GLGEQMSALEIINQLL
162 tr|F7VVT4|F7VVT4_SORMK         -------CG-YRWNV--G-----KYDGDVD---NG-PAGQGMSALAAFSTYL
163 tr|G8YB80|G8YB80_PICSO         D--GHTCG-TNWFK--N-----YWDG-----NY-GLGEQMSALEVMQNLR
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA D--GHTCG-LNWEY--D-----GWDG-----YY-GLGEQVSALETIQSHL
165 tr|G2WRS7|G2WRS7_VERDV         ---GRTCG-FQWYS--G-----QYDG-----KT-GAGQTMNVLSAVSSLL
166 tr|A0A124BVB5|A0A124BVB5_ASPNG T--GTYCG-MQWTT--G-----AYDG-----TV-GVGQQMAALEVVQANL
167 tr|A7EIG7|A7EIG7_SCLS1         N--GRTCG-LNWRT--G-----GFDG-----FY-GVGEQMSALEVIQSNL
168 tr|C5M5U7|C5M5U7_CANTT         D--GVTCG-LSWTNWSQ-----GWDG-----KW-GLGEQMSALEVMQNLM
169 tr|B9WAA8|B9WAA8_CANDC         D--GVTCG-ENWAL--G-----SWDG-----VY-GLGEQTSALEVMMALI
170 tr|F8MRU1|F8MRU1_NEUT8         ---GRVCG-FYWSG--G-----VFVDPAVD-KTT-GAGEAMDVLAAVSSLL
171 tr|A1DJS0|A1DJS0_NEOFI         N---NSCG-VRWYT--S-----KWDG-----W-TGMEEQISVTDVLSVNL
172 tr|B6HLM8|B6HLM8_PENRW         K---DLCG-NRWYG--------GYDG-----Q-NSMENAISGSQMMSAVM
173 tr|A0A117DWS0|A0A117DWS0_ASPNG S--GTQCG-RRWYQ--A-----QWDG-----E-TSLETDMSALSVFSSNM
174 tr|A0A124BXM1|A0A124BXM1_ASPNG N---NTCG-VRWYP--K-----KWDG-----W-NGMEEEIAVTNVLASAL
175 tr|J3NZQ6|J3NZQ6_GAGT3         ---GRACG-FFWTK--G-----EFVDPALTSDTL-GVGTRLDALSAVMSLL
176 tr|W0THH8|W0THH8_KLUMD         D--GHTCG-LNWFK--G-----SWDG-----QY-GLGEQASALEIMNQLL
177 tr|B2AEF2|B2AEF2_PODAN         ---GRRCG-FYWRE--G-----VYVDPAVD-KTS-GAGEQMNVLAAVSSLL
178 tr|Q96TX1|Q96TX1_NEUCS         ---GRMCG-FSWKK--K-----KYDG-----SH-GVGQQMNVLGAVSSLL
179 tr|E9DYA6|E9DYA6_METAQ         ---GRQCG-LKWAD--G-----KYDG-----KT-GAGQEMSVLAAVQSLL
180 tr|B9WAI2|B9WAI2_CANDC         D--GVTCG-LSWTNWSE-----GWDG-----KW-GLGEQMSALEVMQNLM
181 tr|Q6C171|Q6C171_YARLI         D--GHTCS-LNWLT--G-----QYSN-----DYIGLGEQLSAMEVIQNTQ
182 tr|A6ZMV1|A6ZMV1_YEAS7         D--GHTCG-LNWQK--Q-----TNDG-----YY-GLGEQMSALEVIQNLL
183 tr|B6K7X7|B6K7X7_SCHJY         D--AVTCG-TRWYW--N--NGTWDS-----NY-GVGQQLSALEVIQSLL
184 tr|I1RSA2|I1RSA2_GIBZE         A--GTACG-FSWTT--N-----TFDG-----SL-GVGPQMNALDIFMYTL
185 tr|F7VWZ9|F7VWZ9_SORMK         D--GTACG-FKWAE--------PFDG-----WS-GVGAQMNALSAVMYTL
186 tr|A0A1B2J789|A0A1B2J789_PICPA D--GHTCG-LNWLY--D-----GWDG-----FY-GLGEQMSALETLQNTR
187 tr|G0V8N4|G0V8N4_NAUCC         D--GHTCG-LNWFS--G-----GWDG-----MY-GLGEQMSALEALVNTR
188 tr|H8WXN8|H8WXN8_CANO9         D--GHTCG-LSWTNWTV-----GYDG-----YY-GLGEQMSALEVIQSVM
189 tr|A6ZZS0|A6ZZS0_YEAS7         D--GHTCG-LNWFN--G-----TWDG-----MY-GLGEQMSALEVMVNTR
190 tr|G9MQD1|G9MQD1_HYPVG         ---GRRCG-FYWSS--G-----KYIDTAVD-HTS-GAGEAMNVLAAVSSLL
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH D--GHTCG-QNWFY--G-----GWDG-----FY-GLGEQMAALEIMQNVL
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM ---DNTCG-VRWWQ--P-----TWDG-----FTPGLETQMAALAGITANL
193 tr|A0A254UET1|A0A254UET1_ASPNG T--GTYCG-MQWTT--G-----TYDG-----TV-GVGQQMAALEVVQANL
194 tr|A1D587|A1D587_NEOFI         N--HTLCG-RRWHQ--D-----TFDG-----T-SSLEGQMSALSVFSSSM
195 tr|Q1K7I4|Q1K7I4_NEUCR         ---GRMCG-FSWKK--K-----KYDG-----SH-GVGQQMNVLGAVSSLL
196 tr|G8JS88|G8JS88_ERECY         D--GHTCG-MNYSR--G-----SYDE-----SY-GLGEQLAALETIQNLR
197 tr|A0A136J7D3|A0A136J7D3_9PEZI A--GTACG-FTWTG--------TFDN-----LV-GVGPQMSALSALQYTL
198 tr|E9E4X8|E9E4X8_METAQ         ---GRQCG-FKWAS--G-----VYDG-----KT-GAGQEMSVLSAVMSLL
199 tr|Q6CP42|Q6CP42_KLULA         D--GHTCG-LNWFY--G-----GWDG-----KY-GLGEQMAALEVMQNLR
200 tr|G0WHL4|G0WHL4_NAUDC         D--GHTCG-LNWAK--G-----SHDG-----YY-GLGEQMCALEVMQNLL
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU D--GHTCG-LNWFN--G-----TWDG-----MY-GLGEQMAALEVMVNTR
202 tr|F9WYX6|F9WYX6_ZYMTI         N--GTSCA-NSWYDPVD-----DTED-----YYGQINQNLNALEVFLANL
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 ---KRTCG-FKWST--K-----AFDG-----SV-GAGQQMNVLGAVTSLL
204 tr|J3P147|J3P147_GAGT3         ---GRTCG-FYWSK--G-----VFRDPNAD-GTT-GVGEQMNVLSGVMSLL
205 tr|I2H842|I2H842_TETBL         D--GHTCG-MDWNW--N-----GWDG-----YY-GLGEQMAALEFMVNTR
206 tr|S6E3R3|S6E3R3_ZYGB2         D--GHTCG-LNWFK--G-----AWDG-----QY-GLGEQMSALEVIQNTL
207 tr|E5AD94|E5AD94_LEPMJ         G--DHQCG-MKWWS--P-----GFDG-----IG-GVGPQMTALNVISVLN
208 tr|Q6FJM7|Q6FJM7_CANGA         D--GHTCG-LDWQL--K-----TNDG-----YY-GLGEQMSALEVIQQLL
209 tr|A5DUW0|A5DUW0_LODEL         D--GHTCG-LSWTNWTV-----GWDG-----LY-GLGEQMSALEVMQNLM
210 tr|G8BF46|G8BF46_CANPC         D--GHTCG-LSWTNWTV-----GYDG-----YY-GLGEQMSALEVIQSVM

211 tr|W3XAF6|W3XAF6_PESFW         A--GTACT-FSWYD--D---STYED-----ST-GVGEQMNAMSVIYAML
```

```
212 tr|B6H7I2|B6H7I2_PENRW          N---NTCG-ITWYQ--N-----KWDG-------T-TGMEQEISATNVFLANL
213 tr|Q5ATF9|Q5ATF9_EMENI          N---NTCG-IKWYK--S-----SWDG-------S-IGMEQQIIATDVLSSVL
214 tr|A0A167CDD3|A0A167CDD3_9ASCO  D--GHTCG-LNWFK--N-----GWDG-------VW-GLGEQICALDTFNTLL
215 tr|G0RXS3|G0RXS3_CHATD          -------CG-FRWNI--G-----KYDEDVA---NG-PAGQQMSAVAALSTLL
216 tr|B8NX26|B8NX26_ASPFN          DGQSNLCG-LTWGA--N-----KFDG-------V-KGLEAQMSALGGVTGNL
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM  ---DNACG-LRWTK--G---EAYDG-------ST-GVGEQMAALEIIQSNL
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM  T--GTYCGGMRWTT--G-----EYDG-------TT-GVGQQLSALEVVQANL
219 tr|A1CMM3|A1CMM3_ASPCL          ---DNACG-LRWNK--N---ESYDG-------ST-GVGEQMSALDIFQAHL
220 tr|B6HU84|B6HU84_PENRW          D--GKHCG-IRWYK-QA-----SWDG-------T-KGLEQQMAVLGVLAANL
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS  ---GRMCG-FSWKK--K-----KYDG-------SH-GVGQQMNVLGAVSSLL
222 tr|G2WHY6|G2WHY6_YEASK          D--GHTCG-LNWFN--G-----TWDG-------MY-GLGEQMSALEVMVNTR
223 tr|G0W7G8|G0W7G8_NAUDC          D--GHTCG-LNWFI--G-----DWDG-------MY-GLGEQMCALEVMVNTR
224 tr|G4NGV8|G4NGV8_MAGO7          P--GTACG-FRYKA--G---KDFDG-------MV-GVGEQMNALNALMYTL
225 tr|W3X855|W3X855_PESFW          S--GYACG-QHWTW--G---AQNDG-------TS-GVGQQMSALSAVFYTL
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA  ---GRQCG-FYWQS--G-----AWQAPDP--PNS-GASQQMNAVAAVSSLL
227 tr|J8Q2K7|J8Q2K7_SACAR          D--GHTCG-LNWFN--G-----TWDG-------MY-GLGEQMSALEVMVNTR
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO  D--GHTCG-LSWTN--G-----SWDG-------YY-GLGEQMCALEVMTNLR
229 tr|J6EGN8|J6EGN8_SACK1          D--GHTCG-LDWGK--Q-----THDG-------YY-GLGEQMSALEVIQNLL
230 tr|A0A100ISD9|A0A100ISD9_ASPNG  L--QATCG-LKWTD--R-----KWDG-------MD-DVGIQMAALEVMQSTL
231 tr|G8YQC0|G8YQC0_PICSO          D--GVTCG-TTWTE--G-----KWDD-------TW-GLGQQMSALETMLGTI
232 tr|F7VVP8|F7VVP8_SORMK          ---GRVCG-FYWSE--G-----AFVDPAVD-KTT-GAGEAMDVLAAVSSLL
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH  D--GHTCG-LNWQK--Q-----TNDG-------YY-GLGEQMSALEVIQNLL
234 tr|F8MBP4|F8MBP4_NEUT8          D--GTACG-FKWTE-------TFDG-------WA-GVGAQMNALSAVMYTL
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG  S--KTTCG-QRWYQ--S-----TWDG-------T-EGIEEEISATSIFTSNL
236 tr|A1CD27|A1CD27_ASPCL          ---DNTCG-MRWYQ--P-----TWDG-------YTPGLETQMAALAGVTSNL
237 tr|A5DHF3|A5DHF3_PICGU          D--GVTCG-QNWST--G-----KWDE-------VY-GLGEQQCALEVMMALL
238 tr|D4ATT7|D4ATT7_ARTBC          ---DNACG-LYWTK--G---ADWGT-------SKT-GVGEQMAALEVIQSNL
239 tr|E2PT42|E2PT42_ASPNC          L--QATCG-LKWTD--R-----KWDG-------MD-DVGIQMAALEVMQSTL
240 tr|C7ZPE5|C7ZPE5_NECH7          ---GRACG-FKWSS--G-----KYDG-------KT-GAGQQMNVLAAVSSLL
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA  V--GQQCP-LKWTT--E-----TFEG-------ET-GLGEQMAALEVIQSNL
242 tr|Q4WKP7|Q4WKP7_ASPFU          N--RTLCG-RRWHQ--D-----TFDG-------T-SSLEEQMSALSVFSSSM
243 tr|F8MXT4|F8MXT4_NEUT8          ---GRMCG-FSWKK--K-----KYDG-------SH-GVGQQMNVLGAVSSLL
244 tr|A0A177AJ74|A0A177AJ74_9PEZI  K--GRACG-LRWVH--NGQTGVWDG-------TT-GVGEEMSALEVIQSTL
245 tr|J7RQR5|J7RQR5_KAZNA          D--GHTCG-LNWQH--S-----GWDG-------MY-GLGEQMSALEVMNYLL
246 tr|Q75CW7|Q75CW7_ASHGO          D--GHTCG-MVWRN--Y-----TWDG-------KY-GLGEQLSALEIIQNTL
247 tr|G3AKK4|G3AKK4_SPAPN          D--GVTCG-ENWAA--N-----GWDG-------VY-GLGEQTSALEVLMSLI
248 tr|B2W762|B2W762_PYRTR          N--QTSCG-LRWVD--N---GVNDG-------SF-GVGEQMAAMEVIQGLL
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH  D--GHTCG-LNWFN--G-----TWDG-------MY-GLGEQMSALEVMVNTR
250 tr|C4XWK1|C4XWK1_CLAL4          D--HITCG-ENWSV--D-----GWDN-------KY-GLGEQISAGETIMALI

001 tr|A0A254U2J9|A0A254U2J9_ASPNG  VNTK----------------KTAPVTS--STGGNS-SSDP----NAGTG
002 sp|Q6FLP9|DCW1_CANGA            A-LK----------------RAPPYNA--TNGGNS-TGDG----AAGTK
003 tr|A1CCM5|A1CCM5_ASPCL          ITTK----------------HGGPVTS--TTGGNS-TSDP----TAGQG
004 tr|C1H190|C1H190_PARBA          V--DS----------V----PGPA-N--ERTGTS-KGDP----NAGLS
005 tr|Q2UR85|Q2UR85_ASPOR          IAHKKE----------------HQAPLTA--ETGGTS-KSNP----SAGSG
006 tr|A5DV30|A5DV30_LODEL          V---------------------KEPLTV--ATGGSN-RSDF----SAGTD
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO  I-GS----------------IGVPCTE--ETCGDKLDSS-----SFVTV
008 tr|J3PGN0|J3PGN0_GAGT3          IDQA----------------AAPVTN--STGGTS-AGDP----LAGTI
009 tr|A0A124BXU6|A0A124BXU6_ASPNG  IAFNS---------------SSAPLTS--STGGNS-TSNP----SAGTD
010 tr|C5P4A1|C5P4A1_COCP7          I--DS----------V----DPPVTE--KNGGTS-KGNP----GAGV-
011 tr|Q0CG55|Q0CG55_ASPTN          I--TR----------V----DPPVTN--DNGGTS-KGDP----GGGAE
012 tr|C4Y2X5|C4Y2X5_CLAL4          VADF----------------PHFPLTA--DTGGTS-IGNP----AAGYP
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG  I--SR----------V----DPPVTQ--DTGGTS-QGNP----AGGEP
014 tr|C7Z068|C7Z068_NECH7          IGDA----------------EPPVTN--RTGGIS-KGDP----DAGKE
015 tr|G8BXX2|G8BXX2_TETPH          A-LD----------------RPAPYTS--SNGGSS-TGDG----AAGTE
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR  VAER----------------P-PPLTA--DTGGSS-IGNP----AAGYA
017 tr|A1CSC2|A1CSC2_ASPCL          ITHKQMQ---------------QAQGPLTS--DTGGTS-KSNS----SAGTG
018 sp|Q05031|DFG5_YEAST            I-HDR----------------PAPYKE--DNGGTS-KGDA----NAGMN
019 tr|H2AQJ6|H2AQJ6_KAZAF          I-HDR----------------PAPYTA--STGGSS-VGDA----SAGLN
020 tr|G2XFH9|G2XFH9_VERDV          IAGA----------------NSPTTN--DTGGTS-RGDS----NAGRD
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH  A-LK----------------RPAPYNS--TDGGNS-TGNG----AAGLE
022 sp|Q5AD78|DCW1_CANAL            V-HK----------------RPAPYTA--DTGGSS-IGNP----AAGYG
023 tr|B6H7E8|B6H7E8_PENRW          VPFT----------------ADGPYNA--DNGGKS-KSDP----HGGSN
024 tr|B6GZU2|B6GZU2_PENRW          I--TR----------V----AAPVTK--DDGGTS-KGDP----NAGGE
025 tr|E9E4W7|E9E4W7_METAQ          IGDV----------------APPVTN--ATGGTS-KGNP----NAGGK
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO  I--TQ----------V----APPVTS--DNGGTS-KGNP----NAGGK
027 tr|G2R7G8|G2R7G8_THITE          IDQQK---------------VLNGPLTN--TTGGTS-QGDP----SAGGT
028 tr|G3JBY1|G3JBY1_CORMM          LEDA----------------KPPATG--EDRGTS-KGNP----NAGST
```

```
029 tr|A5DNV5|A5DNV5_PICGU          V-HD------------------RPPPYTA--NNGGSS-KGNP----GGGYG
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI I--HA----------T----KGPVRE--EDGGTS-QGNP----SAGGD
031 tr|Q2TWC1|Q2TWC1_ASPOR         MLLE----------------SKSPQTI--DTNPDA-AEHH----I----
032 tr|J5RXY9|J5RXY9_SACK1         A-LD----------------KPAPYTA--DEGGSS-AGDG----AAGTQ
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA LPNV----------------EAPATE--QDRGTS-KGDA----GAGST
034 tr|A0A254U5V1|A0A254U5V1_ASPNG ITHRQSQG--------G---QSQGPLTS--DTGGTS-QSNP----NAGTG
035 tr|A0A061B6H7|A0A061B6H7_CYBFA I-RD----------------RPAPCSN--FTCGTS-IGDP----AAGTY
036 tr|G8ZQ93|G8ZQ93_TORDC         A-LN----------------RPPPYNS--TNGGSS-TGDG----AAGTE
037 tr|Q9C2J1|Q9C2J1_NEUCS         THKGV----------GKAAKGPVTT--AQGGTS-KGDP----GAGVT
038 tr|C5NZK5|C5NZK5_COCP7         I--DL----------V----AGPA-D--NSTGIS-RGNP----SAGTG
039 tr|Q6C0T7|Q6C0T7_YARLI         I-MS----------------SEGPLAD--ITGGDS-KGNG----SAGTV
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB A-LK----------------RAPPYNA--TNGGNS-TGDG----AAGTK
041 tr|F7VZ72|F7VZ72_SORMK         INQQG----------------TKAPVTN--TTGGTS-KGNP----NAGSQ
042 tr|A3LN37|A3LN37_PICST         V-RD----------------FPPPLTA--NTGGSS-KGNP----AAGYS
043 tr|Q6CAI2|Q6CAI2_YARLI         I-FNPKMAPKPTTNENGTTSGGVGPLTH--DTGGTS-EGSP----DAGSS
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO V-HD----------------LPPPYTA--NNGGSS-KGNP----SGGYA
045 tr|G9MJS5|G9MJS5_HYPVG         IGQA----------------HVPVTN--STGGTS-QGNP----NAGSK
046 tr|G9MHI4|G9MHI4_HYPVG         VLQA----------------PSPATA--DKGGTS-KGNP----GAGMG
047 tr|D4AP27|D4AP27_ARTBC         I--RK----------V----VPPVTQ--SSGGIS-LGGP----GGSGK
048 tr|B6GZT8|B6GZT8_PENRW         INFNK----------------TAPVTS--TTGGNS-TSNP----TAGED
049 sp|O74556|YCZ2_SCHPO           LLKS----------------LQVFKA--SNGGSS-TGDP----NAGLY
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 IDQT----------------PAPVTN--STGGTS-AGDP----LAGTI
051 tr|B8MHG0|B8MHG0_TALSN         VGFSAPTANSGSSYTAP---EAPAPVTA--DTGGNS-TSNP----SGGQS
052 tr|A3LMV8|A3LMV8_PICST         V---------------------EEPLKRSADPEVAD-KIDY----NAGLN
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU I-PQ----------------VPPPCEE--DTCGEKLASDV----SIVVV
054 sp|Q9P6I3|YHG7_SCHPO           AQQS----------------ATILTL--DTGASS-ESNPD----AGTD
055 tr|Q2US57|Q2US57_ASPOR         IGLNT----------------SGAPVTS--TTGGNS-TSDP----TAGES
056 tr|W7N622|W7N622_GIBM7         APAA----------------KAPVTT--KTGGTS-KGNP----GGGQT
057 sp|Q9P6I4|YHG6_SCHPO           IEYA----------------PEIATL--ASSTDN-RSNSTYASNVVIN
058 tr|C7GRB3|C7GRB3_YEAS2         I-HDR----------------PAPYKE--DNGGTS-KGDA----NAGMN
059 tr|C7GP28|C7GP28_YEAS2         A-LD----------------KPAPYTA--ENGGSS-VGDG----AAGTQ
060 tr|G3JGZ5|G3JGZ5_CORMM         IKDA----------------NPPITN--STGGTS-KGSP----NGGFA
061 tr|C4QXV4|C4QXV4_KOMPG         A-LV----------------RPAPYTA--QTGGSS-QGDP----AAGLG
062 tr|I1S0Z5|I1S0Z5_GIBZE         MENT----------------PPPVTA--KKGGTS-KGNP----NAGNV
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM V--DR----------V----DPPVTD--ATGGTS-VGDP----SGGME
064 tr|G0W5X4|G0W5X4_NAUDC         V-LD----------------RPAPYNS--TNGGSS-IGDG----AAGTE
065 tr|G8ZPC1|G8ZPC1_TORDC         I-HQK----------------AAPFTE--QTGGSS-AGDA----NAGLN
066 tr|A0A167CDC5|A0A167CDC5_9ASCO A-LQ----------------IAKPLTA--HTGGSS-KGNP----AAGSD
067 tr|R9XAT7|R9XAT7_ASHAC         C-LE----------------RPGPYTA--MNGGTS-PGDP----AAGTK
068 tr|B2B747|B2B747_PODAN         IGQS----------------RPPVTN--STGGTS-KGDP----NAGSQ
069 tr|G0SE99|G0SE99_CHATD         VEYS----------------KPPLTN--GTGGTS-KGDP----MAGSR
070 tr|J8Q6F0|J8Q6F0_SACAR         I-HDR----------------PAPYTE--NTGGTS-QGDA----NAGMN
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH I-HQR----------------PGPLTG--STGGSS-EGDA----NAGLN
072 tr|J3K6E2|J3K6E2_COCIM         I--GS----------V----DPPVTE--KDGGTS-KGNP----GAGV-
073 tr|G4N3E1|G4N3E1_MAGO7         VSEV----------------AAPLTG--KDGGTS-QGNP----NAGIP
074 tr|C5M5J2|C5M5J2_CANTT         V---------------------EDPISV--KTGGTN-RTNY----AAGTD
075 tr|A0A254U0X0|A0A254U0X0_ASPNG MMLGDG----------------SSGPQTI--DENPDA-AEHH----I----
076 sp|Q75DG6|DCW1_ASHGO           C-LE----------------RPAPYTA--MNGGTS-PGDP----AAGTK
077 tr|Q6BZF0|Q6BZF0_DEBHA         V-HD----------------LPPPYTA--NNGGSS-KGNP----AGGYA
078 tr|E9E3Q1|E9E3Q1_METAQ         IESA----------------EQPVTN--KTGGTS-VGNP----NAGGK
079 sp|P36091|DCW1_YEAST           A-LD----------------KPAPYTA--ENGGSS-VGDG----AAGTQ
080 tr|W0TAH6|W0TAH6_KLUMD         C-LD----------------RPAPLTA--NTGGTS-VGNP----AAGTE
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA I-KD----------------RPAPYTN--ETGGTS-EGDA----SAGTS
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI YPSS----------------HLPLSA--ASGGTS-TGNV----DAGGS
083 tr|Q2TYU3|Q2TYU3_ASPOR         VTEKK----------------GSGPLTS--TTGGNS-TSNP----NAGSG
084 tr|G2XGF6|G2XGF6_VERDV         VQKER----------T----APVTA--QTGGTS-KGNP----QAGM-
085 tr|G3AMT4|G3AMT4_SPAPN         A-HK----------------RPAPYTA--SNGGSS-LSNP----AAGYG
086 tr|A0A1C1D213|A0A1C1D213_9EURO I--TE----------V----SGPVTE--GSGGTS-QGDP----NAGTS
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS VFNEN--ASG----GDGKAGNKGAPLTG--ETGGTS-KGNP----HAGTG
088 tr|A0A100I5A6|A0A100I5A6_ASPNG MMLGDG----------------SSGPQTI--DENPDA-AEHH----I----
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH F-DQR----------------PAPYTD--SDGGTS-VGDS----SAGLN
090 tr|G2Q8A7|G2Q8A7_MYCTT         IDQS----------------PPPYTN--MSGGTS-KGDP----LAGTN
091 tr|B8NPZ1|B8NPZ1_ASPFN         MLFE----------------AQSPKTI--ESNPNA-TETE----I----
092 tr|Q7SAB2|Q7SAB2_NEUCR         VFNEN--ASG----GDGKAGNKGAPLTG--ETGGTS-KGNP----HAGTG
093 tr|A2R8R5|A2R8R5_ASPNC         IAFSNS----------------SAAPLTS--STGGNS-TSNP----SAGTD
094 tr|C5DHG0|C5DHG0_LACTC         Y-SKR----------------PAPYTS--DNGGSS-VGDT----EAALN
095 tr|G0SFA3|G0SFA3_CHATD         VVDPN----------LSQFAPVPLTA--VSGGTS-KGDP----SAGTS
096 tr|K1WJG5|K1WJG5_MARBU         I--DE----------A----PQLLTN--TTGGTS-QGDV----NAGTG
097 tr|H2B005|H2B005_KAZAF         A-LH----------------KPGPYTS--ENGGSS-MGDG----AAGTQ

098 tr|C5DYD7|C5DYD7_ZYGRC         I-HNS----------------DPPYRE--STGGSS-KGDS----AAGLN
```

```
099 tr|G8YM23|G8YM23_PICSO          V-HD------------------LPPPYTA--HEGGSS-IGNP----AAGYS
100 tr|W3X8E3|W3X8E3_PESFW          IGES------------------KVPVTN--TTGGTS-GGNY----NAGQD
101 tr|A0A254U3K7|A0A254U3K7_ASPNG  IAFSNS----------------SAAPLTS--STGGNS-TSNP----SAGTD
102 tr|Q6CER8|Q6CER8_YARLI          I-LN------------------STGPLTD--ITGGTS-KGNG----SAGTQ
103 tr|G2QT05|G2QT05_THITE          IDQV------------------PPPVTN--STGGTS-KGDP----TAGMH
104 sp|Q5ACZ2|DFG5_CANAL           V---------------------EPPLSV--KTGGTN-RTDY----SAGTN
105 tr|G2WKU6|G2WKU6_YEASK          I-HDR-----------------PAPYKE--DNGGTS-KGDA----NAGMN
106 tr|A0A0C7N752|A0A0C7N752_9SACH Y-NQR-----------------PAPFTS--DNGGSS-VGDA----NAGLN
107 tr|W3WN68|W3WN68_PESFW          LADS------------------PSPYTS--TTGGSS-VGNA----NAGAS
108 tr|G2QFS1|G2QFS1_MYCTT          IDEA------------------PPPATN--AT-GIS-RGDP----NAGSR
109 tr|G3J9G4|G3J9G4_CORMM          LPNV------------------KVPATE--QDRGTS-KGDA----GAGSS
110 tr|Q2UJ03|Q2UJ03_ASPOR          V--GG-------------V----APPVTE--DNGGTS-KGDP----SAGTE
111 tr|G8YRT2|G8YRT2_PICSO          V-EH------------------FHAPYTR--HTGGTS-KSEP----NAGLN
112 tr|W3X554|W3X554_PESFW          VKRTT-------------Q----APVTA--DTGGTS-KGDV----NAGAS
113 tr|Q1K7A8|Q1K7A8_NEUCR          THKGV---------------GKAAKGPVTT--AQGGTS-KGDP----GAGVT
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM IAHRM----------------QAQAPLTV--DTGGTS-KSNA----SAGTG
115 tr|A7F3P0|A7F3P0_SCLS1          I--GT-------------A----PNLVTN--DTGGTS-VGNA----AGGSG
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA LEDA------------------KPPATS--EDRGTS-KGNP----NAGSN
117 tr|Q6CIP9|Q6CIP9_KLULA          V-HDL-----------------PAPLTE--STGGSS-SGDA----SAGLN
118 tr|H2AMQ5|H2AMQ5_KAZAF          A-LD------------------KPGPYNS--TNGGSS-QGDG----AAGTE
119 tr|G8JMI3|G8JMI3_ERECY          I-HTK-----------------APPLTK--NAGGTS-NGDP----GAGLG
120 tr|C5DGT1|C5DGT1_LACTC          C-LD------------------RPAPYTA--ATGGSS-VGNG----AAGTE
121 tr|K1XJR4|K1XJR4_MARBU          LPLQE-----------------VPPPLTN--NTGGTS-VGNP----DAGAR
122 tr|A0A124BYC5|A0A124BYC5_ASPNG VAYKHL----------------SPATK--TTATNV-T-SSAGT-DTGNG
123 tr|A7TLL2|A7TLL2_VANPO          A-LH------------------VPGPYNS--TNGGSS-TGDG----AAGTE
124 tr|Q0CB86|Q0CB86_ASPTN          LKFQ------------------AKAPLTS--STGGTS-SSDP----NAGGT
125 tr|K1XLH0|K1XLH0_MARBU          V--DQ-------------A----QPWKSL-VAGTGTS-EGNV----NAGQD
126 tr|J7S438|J7S438_KAZNA          A-LM------------------KPAPYTA--TNGGSS-KGES----SRWYG
127 tr|Q4WG09|Q4WG09_ASPFU          MYYK------------------SSAPKTI--QSNPDG-KEHQ----I----
128 tr|H8WZ09|H8WZ09_CANO9          A---------------------EEPISV--QTGGTN-RSNY----AAGTD
129 tr|I1RL09|I1RL09_GIBZE          IEDA------------------EPPATN--KTGGIS-KGDP----NAGKD
130 tr|Q4W985|Q4W985_ASPFU          V--DR-------------V----DPPVTD--ATGGTS-VGDP----SGGME
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU I-HDR-----------------PAPYKE--DNGGTS-KGDA----NAGMN
132 tr|W6QCW3|W6QCW3_PENRF          I--TH-------------V----AAPVTN--DDGGTS-KGNP----NAGDK
133 tr|W3WMD3|W3WMD3_PESFW          AA--------------------TPPLTN--TTGGTS-VGDP----NAGSE
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH A-LK------------------RPAPYNS--TDGGNS-TGNG----AAGLE
135 tr|I2GYH1|I2GYH1_TETBL          I-HDR-----------------PAPYTS--IDGGSS-EGDA----NAGLN
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH V-HTR-----------------PAPFTA--SDGGSS-VGDA----SAGLN
137 tr|Q752P3|Q752P3_ASHGO          I-HTK-----------------APPLTA--DTGGQS-EGNP----DAGLD
138 tr|B8MYP3|B8MYP3_ASPFN          IAHKKE----------------HQAPLTA--ETGGTS-KSNP----SAGSG
139 tr|Q2TXL6|Q2TXL6_ASPOR          VPFKTG----------N---RDQGPLTA--DTGGTS-KGDP----SAGTG
140 tr|Q0CN17|Q0CN17_ASPTN          IKFKKG----------------DVSPLTA--DTGGKS-KSNP----DAGTG
141 tr|J3NHD5|J3NHD5_GAGT3          VDDV------------------RPPLTN--DTGGTS-QGDP----NAGTH
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI I--GS-------------T----AGPLTN--TTGGTS-QGNY----NAGSN
143 tr|A1CVB0|A1CVB0_NEOFI          V--DR-------------V----DPPVTD--ATGGTS-VGDP----SGGME
144 tr|B8NVI3|B8NVI3_ASPFN          VPFKTG----------N---RDQGPLTA--DTGGTS-KGDP----SAGTG
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH C-LD------------------RPAPYTA--QDGGTS-QGNG----AAGTE
146 tr|A1DJ54|A1DJ54_NEOFI          MYYK------------------STAPNTI--QTNPEG-KEHQ----I----
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS IDEA------------------DPPVTN--TTGGTS-KGDP----NAGTG
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA I-RD------------------RPAPCTN--STCGTS-QGDG----AAGTQ
149 tr|A0A177A618|A0A177A618_9PEZI V--TG-------------A----RDLVTN--KTGGTS-LGDY----GDSVD
150 tr|Q75CW6|Q75CW6_ASHGO          I-HTK-----------------PPPTTA--NEGAKS-KGDP----SAGLY
151 tr|Q7S4K4|Q7S4K4_NEUCR          IDEA------------------DPPVTN--TTGGTS-KGDP----NAGTG
152 tr|G2QB99|G2QB99_MYCTT          VVRGA-----------------AAPPLTA--DTGGTS-KGNP----GGGAA
153 tr|G0S3F2|G0S3F2_CHATD          IEYA------------------EPPATN--ET-GIS-RGDP----NAGMR
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO A--N------------------KFPIYTA--TTGGSS-QGNP----AGGYA
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB I-HER-----------------PAPYRA--DNGGTS-VGDA----AAGLN
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM ITTK------------------HKGPVTS--TTGGNS-TSNP----TAGSG
157 tr|G2QKJ0|G2QKJ0_MYCTT          LEQEK----------------VLKGPLTN--TTGGTS-QGDP----NAGQK
158 tr|Q4WFX5|Q4WFX5_ASPFU          Y--NT-------------V----AGPLTN--STGGTS-KRNS----AVERG
159 tr|Q5BGD7|Q5BGD7_EMENI          MLLT------------------DVVAKTI--DTNPGA-KEQF----L----
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH C-LD------------------RPAPYTA--NTGGSS-IGNP----AAGTE
161 tr|G8BYN9|G8BYN9_TETPH          V-SSS-----GAAAEAASSGSASSPYTS--SSGGNS-TGDA----SAGLN
162 tr|F7VVT4|F7VVT4_SORMK          ITEEQ----------------AVFKPLVTN--STGGES-RGNP----NAGEK
163 tr|G8YB80|G8YB80_PICSO          V-HD------------------LPPPYTA--DDGGSS-IGNA----AAGYS
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA T-PI------------------SPAPLTQ--ENEGVA-VGDV----NAGLN
165 tr|G2WRS7|G2WRS7_VERDV          INSA------------------KPPLTN--STGGTS-KGDF----NAGAD
166 tr|A0A124BVB5|A0A124BVB5_ASPNG L--QA-------------N----PGPLTH--TTGGSS-QGNS----AAGTQ
167 tr|A7EIG7|A7EIG7_SCLS1          I--EQ-------------V----AGPVTN--TTGGTS-KGDP----SAGSG
168 tr|C5M5U7|C5M5U7_CANTT          S-NS------------------RPAPYTA--DTGGSS-IGNP----AAGYG
```

```
169 tr|B9WAA8|B9WAA8_CANDC          V---------------------EPPLSV--KTGGTN-RTNY----AAGTN
170 tr|F8MRU1|F8MRU1_NEUT8          IDEA---------------------DPPVTN--TTGGTS-KGDP----NAGTG
171 tr|A1DJS0|A1DJS0_NEOFI          ITTK---------------------HKGPVTS--TTGGNS-TSDP----TAGSG
172 tr|B6HLM8|B6HLM8_PENRW          LKFLGS---------------SSKPVST--ATGGNG-TSDP----NAGTR
173 tr|A0A117DWS0|A0A117DWS0_ASPNG ITHRQSQG---------G---QSQGPLTS--DTGGTS-QSNP----NAGTG
174 tr|A0A124BXM1|A0A124BXM1_ASPNG YSTK---------------------QTAPVTS--STGGNS-TSDP----NAGTG
175 tr|J3NZQ6|J3NZQ6_GAGT3          VEFDD---------------------AQSPTTG--STGGNS-AGDT----GAGLR
176 tr|W0THH8|W0THH8_KLUMD          I-HER---------------------PPPLTE--KTGGSA-RGDP----NAGLN
177 tr|B2AEF2|B2AEF2_PODAN          IDDA---------------------PPPANN--IT-GLS-KPNY----DAGSR
178 tr|Q96TX1|Q96TX1_NEUCS          IEQPG---------------------KAPVTN--STGGTS-KGNP----NAGSQ
179 tr|E9DYA6|E9DYA6_METAQ          IGKA---------------------RPPVTH--DSGGTS-TGNT----DGGQG
180 tr|B9WAI2|B9WAI2_CANDC          V-HK---------------------RPAPYTA--DTGGSS-IGNP----AAGYG
181 tr|Q6C171|Q6C171_YARLI          V-LK---------------------SSGPLTD--ITGGTS-KGNG----SAGAL
182 tr|A6ZMV1|A6ZMV1_YEAS7          I-HDR--------------------PAPYKE--DNGGTS-KGDA----NAGMN
183 tr|B6K7X7|B6K7X7_SCHJY          VDKM---------------------QVPLTQ--STGGTS-ASNP----NAGSA
184 tr|I1RSA2|I1RSA2_GIBZE          VGNA---------------------KAPYTS--KTGGTS-KGDP----GGGNT
185 tr|F7VWZ9|F7VWZ9_SORMK          THKGL---------------GKAAKGPATE--KLGGTS-KGDP----GAGVT
186 tr|A0A1B2J789|A0A1B2J789_PICPA A-LV---------------------RPAPYTA--QSGGSS-QGDP----AAGLG
187 tr|G0V8N4|G0V8N4_NAUCC          V-LH---------------------LPGPYNS--TDGGSS-IGNS----AAGTE
188 tr|H8WXN8|H8WXN8_CANO9          IYHD---------------------APAPYTA--ETGGSS-KSQP----ASGYG
189 tr|A6ZZS0|A6ZZS0_YEAS7          A-LD---------------------KPAPYTA--ENGGSS-VGDG----AAGTQ
190 tr|G9MQD1|G9MQD1_HYPVG          INEA---------------------PPPVTN--SSGGIS-KGNP----NAGHG
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH C-LQ---------------------KPAPYTA--KDGGTS-IGNG----AAGTE
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM MYYK---------------------SSAPKTI--QSNPDG-KEHQ----I----
193 tr|A0A254UET1|A0A254UET1_ASPNG L--QA------------N-----PGPLTH--TTGGSS-QGNS----AAGTQ
194 tr|A1D587|A1D587_NEOFI          IAHKM----------------QAQAPLTA--DTGGTS-KGNA----SAGTG
195 tr|Q1K7I4|Q1K7I4_NEUCR          IEQPG---------------------KAPVTN--STGGTS-KGNP----NAGSQ
196 tr|G8JS88|G8JS88_ERECY          C-LE---------------------KGSPLTA--SEGGTS-IGNP----GAGGN
197 tr|A0A136J7D3|A0A136J7D3_9PEZI IKKNN------------R---AAVPATA--NNGGTS-IGNV----NAGAR
198 tr|E9E4X8|E9E4X8_METAQ          IPQA---------------------KAPVTE--KDGGTS-KGNP----NAGGS
199 tr|Q6CP42|Q6CP42_KLULA          C-LD---------------------RPAPYTA--DDGGTS-AGNP----AAGTE
200 tr|G0WHL4|G0WHL4_NAUDC          I-HDR--------------------PAPLTG--NTGGSS-KGDA----LAGLN
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU A-LD---------------------RPAPYTA--EDGGSS-VGDG----AAGTQ
202 tr|F9WYX6|F9WYX6_ZYMTI          PSGDIKM-------------------GV-G--STNGTG-VGGSLGGGANGTE
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 IDDV---------------------SPPLTN--ATGGTS-IGDP----NAGTH
204 tr|J3P147|J3P147_GAGT3          HEFPN---------------NPVEGPVTL--ETGGTS-RGDP----NAGLR
205 tr|I2H842|I2H842_TETBL          A-LD---------------------LPPPYTS--SNGGSS-TGSG----AAGTE
206 tr|S6E3R3|S6E3R3_ZYGB2          V-HTR--------------------PAPFTA--SDGGSS-VGDA----SAGLN
207 tr|E5AD94|E5AD94_LEPMJ          V--DR------------V-----PPPYSS--VTGGSS-QGNP----NLGTE
208 tr|Q6FJM7|Q6FJM7_CANGA          I-HER--------------------PAPYRA--DNGGTS-VGDA----AAGLN
209 tr|A5DUW0|A5DUW0_LODEL          S-TK---------------------RPAPYTA--NDGGSS-KSHP----ASGYG
210 tr|G8BF46|G8BF46_CANPC          IYHD---------------------APAPYTA--ETGGSS-ESQP----DSGYG
211 tr|W3XAF6|W3XAF6_PESFW          VDDA---------------------ITPYTT--TTGGSS-SSNV----DGGTS
212 tr|B6H7I2|B6H7I2_PENRW          INYNTG---------------SFGPVTS--KTGGNS-ASDP----NAGSE
213 tr|Q5ATF9|Q5ATF9_EMENI          VSEK---------------------SNPPLTT--KTGGNS-TSDP----NAGTS
214 tr|A0A167CDD3|A0A167CDD3_9ASCO I-NT---------------------RPGPLTE--KSGGPS-GGFG----AAGLN
215 tr|G0RXS3|G0RXS3_CHATD          LDQPH----------------VLRGPLTN--DTGGTS-RGDP----NAGLV
216 tr|B8NX26|B8NX26_ASPFN          MLMA---------------------AESPNTI--DTNPDA-KEHN----V----
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM I--DL------------V-----AGPA-D--NSTGIS-RGNP----SAGTG
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM Y--NT------------V-----AGPLTT--STGGTS-KRNS----AVERG
219 tr|A1CMM3|A1CMM3_ASPCL          VDYV------------E-----KRVT-A--SSGGIS-QGDP----NAGTG
220 tr|B6HU84|B6HU84_PENRW          IPFS---------------------GKAPLTA--SSGGTS-KSNP----NAGTN
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS IEQPG---------------------KAPVTN--STGGTS-KGNP----NAGSQ
222 tr|G2WHY6|G2WHY6_YEASK          A-LD---------------------KPAPYTA--ENGGSS-VGDG----AAGTQ
223 tr|G0W7G8|G0W7G8_NAUDC          C-LH---------------------KPAPYTS--TNGGSS-VGDG----AAGTE
224 tr|G4NGV8|G4NGV8_MAGO7          IRNNT------------K---EVVPVTS--DTGGTS-KGDP----TAGII
225 tr|W3X855|W3X855_PESFW          LDGA---------------------PTPYTS--VTGGTS-TGDS----SGGTA
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA YPNA---------------------KPAVTA--DTGGTS-KSDP----NAGVG
227 tr|J8Q2K7|J8Q2K7_SACAR          A-LD---------------------KPAPYTA--EDGGSS-TGDG----AAGTQ
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO V-RD---------------------VDPPYTA--NTGGSS-IGDP----ASGYS
229 tr|J6EGN8|J6EGN8_SACK1          I-HDR--------------------PAPYKE--SNGGTS-KGDA----NAGMN
230 tr|A0A100ISD9|A0A100ISD9_ASPNG I--SR------------V-----DPPVTQ--DTGGTS-KGNP----AGGEP
231 tr|G8YQC0|G8YQC0_PICSO          I-EH---------------------FHAPYTD--RTGGTS-KSEP----NAGID
232 tr|F7VVP8|F7VVP8_SORMK          IEQA---------------------DPPATN--QTGGIS-KGDP----NAGTG
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH I-HDR--------------------PAPYKE--DNGGTS-KGDA----NAGLN
234 tr|F8MBP4|F8MBP4_NEUT8          THTGV---------------GNVTKGPVTA--AQGGTS-KGDP----GAGVT
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG VVYKHQ-------------------SLATK--ATATNG-TSGSTGT-ESGNG
236 tr|A1CD27|A1CD27_ASPCL          MFFK---------------------SAAPKTI--ESNPDG-KEHQ----I----

237 tr|A5DHF3|A5DHF3_PICGU          A-RD---------------------VDPPYTS--KNGGTS-KSQP----NAGID
238 tr|D4ATT7|D4ATT7_ARTBC          WYKKQ------------A-----SGPA-N--SKTGTS-KGDP----NAGND
```

```
239 tr|E2PT42|E2PT42_ASPNC          I--SR----------V----DPPVTQ--DTGGTS-QGNP----AGGEP
240 tr|C7ZPE5|C7ZPE5_NECH7         MEHT----------------EPPVTA--KKGGIS-KGNP----NAGS-
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA I--SK----------T----PPPLTN--LTGGTS-PGDP----NAGLQ
242 tr|Q4WKP7|Q4WKP7_ASPFU         IAHRM----------QAQAPLTV--DTGGTS-KSNA----SAGTG
243 tr|F8MXT4|F8MXT4_NEUT8         IEQPG----------KAPVTN--STGGTS-KGNP----NAGSQ
244 tr|A0A177AJ74|A0A177AJ74_9PEZI I--GQ----------T----PGPVTN--TTGGTS-QGNY----KTGSN
245 tr|J7RQR5|J7RQR5_KAZNA         V-HDR----------PAPYTA--AAGGSS-VGDA----SAGLN
246 tr|Q75CW7|Q75CW7_ASHGO         I-HTK----------PPPTTA--NEGAKS-QGDP----GAGLQ
247 tr|G3AKK4|G3AKK4_SPAPN         V---------------QEPLSV--KTGGSN-KTDF----DAGTK
248 tr|B2W762|B2W762_PYRTR         Y--PT----------V----PGPATQ--QSGGIS-KSNP----DAGSD
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH A-LD----------KPAPYTA--ENGGSS-VGDG----AAGTQ
250 tr|C4XWK1|C4XWK1_CLAL4         T-SK----------N-TPLTP--KTGGSDAGASV----DAGNS

001 tr|A0A254U2J9|A0A254U2J9_ASPNG DNTGSS-GST---ESKITTGDKAG--------AS-ILTIAF-VGMWGGMIA
002 sp|Q6FLP9|DCW1_CANGA          PHPT------NLAPLHITGGSRAG--------AG-IITAII-GISIIACAL
003 tr|A1CCM5|A1CCM5_ASPCL        DKSGNT-PT---THTVTTGDKAG--------AS-ILTVGF-AVGWVSLMA
004 tr|C1H190|C1H190_PARBA        DGRNDL---PKL--SEITMGDKVG--------AG-FMTSII-LLGVLGGAW
005 tr|Q2UR85|Q2UR85_ASPOR        HKDQQT-GT---PKPITTGDRAG--------AS-IVTFFF-ACGWMASVS
006 tr|A5DV30|A5DV30_LODEL        SQDS-----VNKNEITVTGKDKAG--------AG-VLTAIV-LAILLGGSI
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO PAPT------VTRHFTKTSTDVHT--------ET-HTTSVVTELFYTAK--
008 tr|J3PGN0|J3PGN0_GAGT3        ----P---SKDRTFKEITTGDRAG--------AA-ILTMLL-LGSATSTFG
009 tr|A0A124BXU6|A0A124BXU6_ASPNG DSSDDK-TI---LSTITTGDRAG--------AG-IVTVAF-AGGWIGLMA
010 tr|C5P4A1|C5P4A1_COCP7        STGDA----IGM-QRGIKTADKAG--------AW-ITTMIL-IGVILGGFY
011 tr|Q0CG55|Q0CG55_ASPTN        PAPP---T-PGILTMNITTADRAG--------AG-ILTVLM-SAFVIGSTG
012 tr|C4Y2X5|C4Y2X5_CLAL4        KAAT------NASPLVLDSGDKAG--------AG-IITAVI-GAAIVGASV
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG GPPPAP-V-PEGLRLEITKADRAG--------AG-MMTVML-SMIVIGSTG
014 tr|C7Z068|C7Z068_NECH7        ---SH---D-MPEPDPITQADKAG--------AG-VLTFLI-LSSALGTFV
015 tr|G8BXX2|G8BXX2_TETPH        TSAT------NLSPLNITAGSRAG--------AG-IITAVI-GISIVACAC
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR IAKT------DAEPLDLHAKDKAG--------AS-IITAII-GVAIIGTAT
017 tr|A1CSC2|A1CSC2_ASPCL        TEH-SA-PQ---ALPVTTGDRAG--------AG-VLTVVF-LGAWIGAVT
018 sp|Q05031|DFG5_YEAST          SST----TNVLQNNLNIKKGDRAG--------AA-IITAVI-LSVLTGGAV
019 tr|H2AQJ6|H2AQJ6_KAZAF        TTTS--SADVLENNLTITSKDRAG--------AA-IITAVI-LGCMAGGAI
020 tr|G2XFH9|G2XFH9_VERDV        ----S---NPFQDPEPITTADRAG--------AG-IVTFII-LASGLSSWY
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH QAAT------NLSPLHITAGSRAG--------AG-IITCVV-GISIIASLL
022 sp|Q5AD78|DCW1_CANAL          KLTS------DATPLSIDGGDKAG--------AG-IITAII-GASLVGSCV
023 tr|B6H7E8|B6H7E8_PENRW        PEGSTG-PA-----PITSGDRVG--------AG-ILTAVF-VLVWVGAIS
024 tr|B6GZU2|B6GZU2_PENRW        PQKA---I-PPALDYDITTGDKAG--------AG-VLTVLF-LIGIGGSTG
025 tr|E9E4W7|E9E4W7_METAQ        ----D---NGERPVKPVTAAGKAG--------APPVLTILL-LVGAVSLFV
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO DPEP---K-PKALSYKITTADRAG--------AG-ILTIMM-LATLGGSCG
027 tr|G2R7G8|G2R7G8_THITE        ----F---NALAPPKPITTADRAG--------AG-ILTAVV-LATFLGGLV
028 tr|G3JBY1|G3JBY1_CORMM        AA--------QPLPQPRAIVTSDRAG--------AA-ILTMLC-TVGIVGGSF
029 tr|A5DNV5|A5DNV5_PICGU        ATNI------NATPLNLGSGDKAG--------AG-IITAII-GMSIVGATV
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI GRGNKD-VKAGPWSQPAGTGDRVG--------AG-ALTAVM-AAGVVCGVW
031 tr|Q2TWC1|Q2TWC1_ASPOR        DNNENS-SKDPTKAKPIETADRAG--------AW-ILTVMI-AAGAIGAVG
032 tr|J5RXY9|J5RXY9_SACK1        AQPT------NLAPLNITKGSKAG--------AG-IITAVI-GISIVACAL
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA QA--------QPLPEPRAITGGDRAG--------AA-ILTILI-AAGILAGSF
034 tr|A0A254U5V1|A0A254U5V1_ASPNG PKDAPN-KL---P-DITTGDRAG--------AS-ILTVLF-VGSWAGAIT
035 tr|A0A061B6H7|A0A061B6H7_CYBFA EDST------NLSPLSIDGGDKAG--------AA-IITCII-GASLVGSAA
036 tr|G8ZQ93|G8ZQ93_TORDC        ESPT------NLSPLDITAGSRAG--------AG-IITCVI-GISILSCAL
037 tr|Q9C2J1|Q9C2J1_NEUCS        DPAS---RGGLAALKPITMADRVG--------AG-IVTAIL-AISIVGGSV
038 tr|C5NZK5|C5NZK5_COCP7        --SDPT---FEL--SDITTGDRVG--------AG-FLTSVV-LIGILGGAW
039 tr|Q6C0T7|Q6C0T7_YARLI        FESPYAN--TIAAPLTIGPGDRAG--------AG-IITAVV-CIVLIGSGW
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB PHPT------NLAPLHITGGSRAG--------AG-IITAII-GISIIACAL
041 tr|F7VZ72|F7VZ72_SORMK        S---D---SFGAHNRPISTADRAG--------AG-IVTFLL-SAGALGTFA
042 tr|A3LN37|A3LN37_PICST        TLHT------ITSPLELETKDLAG--------AG-IITAVV-GVSLVAAGV
043 tr|Q6CAI2|Q6CAI2_YARLI        SANP------LLMNLDIKGSDKAG--------AG-VITAVL-LIGVLGSGW
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO VTNT------NTSPLNLDKGDTAG--------AS-IITVII-AASLIGSGV
045 tr|G9MJS5|G9MJS5_HYPVG        P---N---SFQRPLTPVTAGDRAG--------AG-IVTIVI-IGSLCTGLA
046 tr|G9MHI4|G9MHI4_HYPVG        SIDP---AGDLDKITKPTTGGKVG--------AA-ILTLLL-LLSVLGGSA
047 tr|D4AP27|D4AP27_ARTBC        GKDNG----PKI-LRPITGADTAG--------AA-ILTLLM-FCSLSGVSY
048 tr|B6GZT8|B6GZT8_PENRW        DTDGSN-QM---RTTTTTADKAG--------AG-IVTAVF-VAGWIGLMS
049 sp|O74556|YCZ2_SCHPO          TAPV------SFANKNFENLRKHW--------ML-LG------FFLLVPTL
050 tr|A0A0C4E991|A0A0C4E991_MAGP6 ----P---SRDRTFKEITTADRAG--------AG-ILTMLL-LGSATGTFG
051 tr|B8MHG0|B8MHG0_TALSN        STGYQP-PK---AMKITGGDRAG--------AA-ILTLVF-ATAWIGMMV
052 tr|A3LMV8|A3LMV8_PICST        SHDT-----VNQNEITVTGKDRAG--------AG-VLTAVV-LGILLGGGA
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU PTPT------IFRTFTTTNTVHKT--------AT-LIHSSVENVFYTGTTS
054 sp|Q9P6I3|YHG7_SCHPO          DGDTVT-------ITPATKSDKGW--------AG-FLTFAF-SFVFLLFSI
055 tr|Q2US57|Q2US57_ASPOR        DNEGSS-ST---TTKVTTGDVVG--------AS-IVTIVL-VAVPIAMVV
```

```
056 tr|W7N622|W7N622_GIBM7          NMDV---DVSKPKYAKITTMDKVG-------AG-AITCMI-LACVIGGSV
057 sp|Q9P6I4|YHG6_SCHPO           DTNTTT----------TIVVKEKDRGG-------AG-FLTFLS-AIFILGASI
058 tr|C7GRB3|C7GRB3_YEAS2         SST-----TNVLQNNLNIKKGDRAG-------AA-IITAVI-LSVLTGGAV
059 tr|C7GP28|C7GP28_YEAS2         AQPT------NLAPLNITKGSKAG-------AG-IITAVI-GISIVACAL
060 tr|G3JGZ5|G3JGZ5_CORMM         G--GD---NGLKIPKPITVADKAG-------AG-ILTFML-VFGASGTFF
061 tr|C4QXV4|C4QXV4_KOMPG         VKTE------AVPPLKLTNADVAG-------AA-IITAII-GLSVIAGAI
062 tr|I1S0Z5|I1S0Z5_GIBZE         GD--G---KIDKTYGPLTTADKAG-------AG-ILTFLV-LSFACGMFG
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM QPDP---R-PPVLTMTITGADRAG-------AG-LLTAML-GVLMIGTTG
064 tr|G0W5X4|G0W5X4_NAUDC         ATPT------NLAPLHITKGSKVG-------AA-FITAAI-GLSIVFTSI
065 tr|G8ZPC1|G8ZPC1_TORDC         TSS-----ANVLQNQLDIQTKDRAG-------AA-IITAVI-LMVLVGGSI
066 tr|A0A167CDC5|A0A167CDC5_9ASCO SGTS-----NFLTNSLDISSKDRAG-------AG-ILTALL-AVGIVGTGW
067 tr|R9XAT7|R9XAT7_ASHAC         TKAE------NLPPLNIKAGDRAG-------AG-IITALI-GSSLLACTL
068 tr|B2B747|B2B747_PODAN         S---D---NFKDKYLPPTTGDKAG-------AG-ILTVLI-LVSVVGTFG
069 tr|G0SE99|G0SE99_CHATD         S---D---TLTQEVKPPTGGDKAG-------AA-ILTFVV-LATTVGAFT
070 tr|J8Q6F0|J8Q6F0_SACAR         SST-----TNVLQNNLNITKGGRAG-------AA-IITAVV-LSVLVGGAV
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH VST-----TNVLQKKINITTGSRAG-------AS-IITAIV-LGVIILGGV
072 tr|J3K6E2|J3K6E2_COCIM         STGNA-----IGM-QRGIKTADKAG-------AW-ITTMIL-IAVILGGFY
073 tr|G4N3E1|G4N3E1_MAGO7         ----K---SNDRTLSEITSGDKAG-------AS-FLTILT-MVSAASLFG
074 tr|C5M5J2|C5M5J2_CANTT         SEDS------VNQNELNITGKDKAG-------AG-ALTAIV-LAIILGGTI
075 tr|A0A254U0X0|A0A254U0X0_ASPNG DTSEDE--KDPTAKEKIDTGDRAG-------AW-ILTVVL-AGLMLGAVG
076 sp|Q75DG6|DCW1_ASHGO           TKAE------NLPPLDIKAGDRAG-------AG-IITALI-GSSFLACTL
077 tr|Q6BZF0|Q6BZF0_DEBHA         VTNT------NSSPLNLDKGDTAG-------AS-IITVII-AASLVGSGV
078 tr|E9E3Q1|E9E3Q1_METAQ         ----D---NGERPVKPITTADKAG-------AG-ILTFLL-LGGAAGMFV
079 sp|P36091|DCW1_YEAST           AQPT------NLAPLNITKGSKAG-------AG-IITAVI-GISIVACAL
080 tr|W0TAH6|W0TAH6_KLUMD         QSPS------NLSPLTITHGDKAG-------AG-IITAVI-GITIIGSVV
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA STTT-----QLDTAALNITGGDKAG-------AG-IITAIV-LAVIIAGCV
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI STN-----ADGKTLSKITTGDKVG-------AA-FLTMLV-ISGLLGGMF
083 tr|Q2TYU3|Q2TYU3_ASPOR         DDSSSD-KSQ---LKSITTGDKAG-------AS-IVTIAF-VGIWGGLIA
084 tr|G2XGF6|G2XGF6_VERDV         -E--------PHSWEFKPIQQAERVA-------AG-FLTVGI-LTSLMAGVF
085 tr|G3AMT4|G3AMT4_SPAPN         TSIT------NTSPLHLGPGDKAG-------AG-IITAIV-GVSLIGTCV
086 tr|A0A1C1D213|A0A1C1D213_9EURO SPIGP----GDLHRHEVTAADKAG-------AG-ILTVLL-IFFIVGGAW
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS KG--V---VDPTELRELTKGDKVG-------AW-ILTAFT-LGTLGAAFV
088 tr|A0A100I5A6|A0A100I5A6_ASPNG DTSEDE--KDPTAKEKIETGDRAG-------AW-ILTVLL-AGLMLGAVG
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH KTT-----TNVLQSDLKITHKDRVG-------AA-VLTAVV-LAIILLGGSI
090 tr|G2Q8A7|G2Q8A7_MYCTT         S---R---STLNEPRQITGGDRAG-------AG-ILTILL-IAAGVGTFG
091 tr|B8NPZ1|B8NPZ1_ASPFN         DHHSDE---EPNKPKPITTGDRAG-------AG-IITVVV-AIAVAGTVV
092 tr|Q7SAB2|Q7SAB2_NEUCR         KG--V---VDPTELRELTKGDKVG-------AW-ILTAFT-LGTLGAAFV
093 tr|A2R8R5|A2R8R5_ASPNC         DSSDDK-TI----LSTITTGDRAG-------AG-IVTVVF-AGGWVGLMA
094 tr|C5DHG0|C5DHG0_LACTC         STS-----TNVLQSDIKITHKDKVG-------AS-IITAVI-LAVLVGGSI
095 tr|G0SFA3|G0SFA3_CHATD         PRGAEE-ERKRLLDGEITIKDQVA-------AG-FLTSAI-VLGVVMGCI
096 tr|K1WJG5|K1WJG5_MARBU         DDASE-----IVVITPATTADKAG-------AA-ILTAAV-LAGLVGGVG
097 tr|H2B005|H2B005_KAZAF         ATAT------NLSPLNVTPGSKAA-------AG-IITAVI-GISIFACVL
098 tr|C5DYD7|C5DYD7_ZYGRC         TTT-----TNVLQKKLKIQSKDRAG-------AA-IITGVV-LMVVVGGSV
099 tr|G8YM23|G8YM23_PICSO         VTTA------NASPLKLDGGDKAG-------AA-IITIII-SASLIGSGI
100 tr|W3X8E3|W3X8E3_PESFW         S---D---SFTNSLTPITQGDRAG-------AG-ILTVLI-LASAAGTFG
101 tr|A0A254U3K7|A0A254U3K7_ASPNG DSSDDK-TI----LSTITTGDRAG-------AG-IVTVVF-AGGWVGLMA
102 tr|Q6CER8|Q6CER8_YARLI         GDNPFGTSQDDASPLDIHGGDRAG-------AG-IITAAV-GIMLILVFW
103 tr|G2QT05|G2QT05_THITE         S---D---TFQNTLAPPTAADKAG-------AS-ILTILI-LCSAVGTFG
104 sp|Q5ACZ2|DFG5_CANAL           SEDN------ANKNELTITGKDKAG-------AG-VLTAIV-LAVILGGAI
105 tr|G2WKU6|G2WKU6_YEASK         SST-----TNVLQNNLNIKKGDRAG-------AA-IITAVI-LSVLTGGAV
106 tr|A0A0C7N752|A0A0C7N752_9SACH KST-----TNVLHSNLKITHKDKVG-------AA-ILTAVV-IAILVGGSI
107 tr|W3WN68|W3WN68_PESFW         DS--------SKIATARAITTGDRAG-------AG-ILTALV-LASLLSSLV
108 tr|G2QFS1|G2QFS1_MYCTT         SRGP------SEPLAPITTADRAG-------AG-ILTMLI-LGGFIGTWS
109 tr|G3J9G4|G3J9G4_CORMM         QA--------QPLPEPRAITAGDRAG-------AA-ILTILI-AAGILAGSF
110 tr|Q2UJ03|Q2UJ03_ASPOR         SGGP---R-PHSRRTDTSTANRAG-------AG-ILTILM-AMLIFYTVW
111 tr|G8YRT2|G8YRT2_PICSO         THET------ANRNLLKISGKDKAG-------AG-VLTAVI-LAIVLGCAV
112 tr|W3X554|W3X554_PESFW         TE--------DKMPVLEPITLKEKVA-------AG-FITSAL-ALSVLGGSI
113 tr|Q1K7A8|Q1K7A8_NEUCR         DPAS---RGGLAALKPITMADRVG-------AG-IVTAIL-AISIVGGSV
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM SEQ-AA-PQ----PQPVTTGDRAG-------AG-IVTVVF-LSGWIAAVT
115 tr|A7F3P0|A7F3P0_SCLS1         TSTNS----DGTVEKPVTSGDRAG-------AG-FVTALI-ISGVLGGVG
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA VA--------QPLPEPRKLVTADRAG-------AA-ILTMLV-AAGIVGGSL
117 tr|Q6CIP9|Q6CIP9_KLULA         VST-----TNVLQKNINITAGDKAG-------AG-IITTIV-LGVLVLGTV
118 tr|H2AMQ5|H2AMQ5_KAZAF         TTAT------NLAPLNITGGSKAA-------AA-IITAVI-GISLIACTL
119 tr|G8JMI3|G8JMI3_ERECY         TPM---D-NVLHSRLNIGSKDRIK-------AA-LITAVC-LGLLVGTAL
120 tr|C5DGT1|C5DGT1_LACTC         TSAT------NLSPLDITTGSRAA-------AG-IITAVI-GLSIVACTV
121 tr|K1XJR4|K1XJR4_MARBU         S---V---TDLRMLLPPSEQDKAG-------AA-VITAVW-VAASVVLFG
122 tr|A0A124BYC5|A0A124BYC5_ASPNG TTTAST-AAGSNSVVPTNGANSLSCGLFGVALA-IVAGAF-AIA------
123 tr|A7TLL2|A7TLL2_VANPO         QHPT------NLSPLNITSGSRAG-------AG-IITAII-GISIIACGL
124 tr|Q0CB86|Q0CB86_ASPTN         DSNSNS-GN----LKPITTGDKAG-------AG-ILTVVF-VGLWGGMVA
```

```
125 tr|K1XLH0|K1XLH0_MARBU          SPEDV----AALTEMKITSGDRVG--------AG-ILTALV-ICGVIAGSG
126 tr|J7S438|J7S438_KAZNA          SIPN-------QLVPIAYHKRVSGR--------CR-YHHCHHRYFNRIVCIV
127 tr|Q4WG09|Q4WG09_ASPFU          DTHEDE---APDALAPINTGDRAG--------AW-ILTVII-VVAVGGSVG
128 tr|H8WZ09|H8WZ09_CANO9          TEDT------TNKNKIDITGKDRAG--------AG-VLTAIV-LAIILGGGI
129 tr|I1RL09|I1RL09_GIBZE          ---SH---D-NPEPTPITTADRAG--------AG-ILTFLV-LAGGLGTFV
130 tr|Q4W985|Q4W985_ASPFU          QPDP---R-PPVLTMTITGADRAG--------AG-LLTAML-GVLMIGTTG
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU  SST----TSVLQNNLNITKGDRAG--------AA-IITAVI-LSVLIGGAV
132 tr|W6QCW3|W6QCW3_PENRF          PPKA---I-PPGLDYDITSGDKAG--------AG-ILTVLF-MLAIGGSTG
133 tr|W3WMD3|W3WMD3_PESFW          E---------STLTPLAEITTADRAG--------AG-ILTCII-LASSLGAFA
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH  QAAT------NLNPLHITAGSRAG--------AG-IITCVV-GISIIASLL
135 tr|I2GYH1|I2GYH1_TETBL          KTS----KNVLSNDLDIKTGDKVG--------AA-ILTAGV-LAILVGGAI
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH  SST----TNVLQNKLSINTGDRAG--------AA-VVTAVV-LMIVVGGSV
137 tr|Q752P3|Q752P3_ASHGO          HDS--QPAILRGMHNITGKDRVG--------AS-LITAIS-LGLMLGCAL
138 tr|B8MYP3|B8MYP3_ASPFN          HKDQQT-GT----PKPITTGDRAG--------AS-IVTFFF-ACGWMASVS
139 tr|Q2TXL6|Q2TXL6_ASPOR          SRKPEK-DK----PREITTGDRVG--------AS-IATVVV-VGIWLGIAA
140 tr|Q0CN17|Q0CN17_ASPTN          TGEPPA-SE----PAPITTGDRAG--------AS-ILTILF-ISGWVAAVT
141 tr|J3NHD5|J3NHD5_GAGT3          S---T---PH-QTFAPITGGDRAG--------AA-FLTIAV-LGGAAFMFG
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI  AKTAA----SLLDDWTPTTGDKAG--------AG-IVTAVI-LAGVIGGVA
143 tr|A1CVB0|A1CVB0_NEOFI          QHDP---R-PQVLTMTITGADRAG--------AG-LLTAML-GVLMIGTTG
144 tr|B8NVI3|B8NVI3_ASPFN          SRKPEK-DK----PREITTGDRVG--------AS-IATVVV-VGIWLGIAA
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH  TQAT------NLSPLTITTGSRAA--------AG-IITAVI-GLSIVACTV
146 tr|A1DJ54|A1DJ54_NEOFI          DTHEDA---APDALDPINTGDRAG--------AW-ILTVII-VVAVGGSVG
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS  SRHA------TEPAKPITTADKAG--------AA-MCTILL-IAGGIAIWI
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA  QSQT------NLSPLTIDKGDEAG--------AA-IITVVV-GATIVGAFA
149 tr|A0A177A618|A0A177A618_9PEZI  TAAKVR-----GSIEVTTAGRAA--------AG-MLTAGM-CALVGVAVW
150 tr|Q75CW6|Q75CW6_ASHGO          DNSP--N-IITLSTRQIGLKDRVG--------AG-FMTVVL-LIPFITGSI
151 tr|Q7S4K4|Q7S4K4_NEUCR          SRHA------TEPAKPITTADKAG--------AA-MCTILL-IAGGIAIWI
152 tr|G2QB99|G2QB99_MYCTT          PAHEE----VGIGYAEITLQDRVA--------AG-FVTSAL-ALGVVAGSA
153 tr|G0S3F2|G0S3F2_CHATD          SRGA------AQHFREINAGDRVG--------AA-ILTMLI-LGGAIATWS
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO  VTNT------NASPLKLGAGDTAG--------AG-IITAII-GVSIVASGV
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB  STT----TNVLKNNLKITGGDRAG--------AA-IVTTIV-LGIIIGGAA
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM  DKSGQT-TP----TRKITMGDKAG--------AS-ILTIGL-VVGWVSLIA
157 tr|G2QKJ0|G2QKJ0_MYCTT          ----F---EGVSPPREITAGDRAG--------AG-ILTAVV-LASFLGSLV
158 tr|Q4WFX5|Q4WFX5_ASPFU          SSERIA---VSDDSSIVTMADRVV--------AW-VATGGL-GVLLGGYLY
159 tr|Q5BGD7|Q5BGD7_EMENI          DTYNDD---TPDALPHITTGDRVG--------SW-ILTVLW-GSGIVAAGW
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH  QSPT------NLAPLTITKGDQAG--------AG-IITAVI-GITIIGSVV
161 tr|G8BYN9|G8BYN9_TETPH          TTS----TNVLQNNLDIESKDKAG--------AA-VATAVI-LAVLVGGSV
162 tr|F7VVT4|F7VVT4_SORMK          ----P---ATLLSMSELTGKDKAG--------AG-VLTAFV-VCSILGTYF
163 tr|G8YB80|G8YB80_PICSO          VTTE------NASPLKLDGGDKAG--------AA-IITIVI-SASLIGSGI
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA  TTRN----ELTYSDMNIDSGDVAG--------AA-IVTAVV-LLLAVGGAI
165 tr|G2WRS7|G2WRS7_VERDV          S---Q---DIMDHHKPIKMGDRAG--------AG-ILTFMV-LATALSTWA
166 tr|A0A124BVB5|A0A124BVB5_ASPNG  TSDSSR---DPSELRPITTSDRIG--------AG-VVTGVT-ACLVLGVTA
167 tr|A7EIG7|A7EIG7_SCLS1          KSSSS-----FSSTKPVTTADKVG--------AG-ILTTVL-LLAFIGSIW
168 tr|C5M5U7|C5M5U7_CANTT          TPIT------DARPLSLDSGDKAG--------AG-ILTAVI-GATLVGTCI
169 tr|B9WAA8|B9WAA8_CANDC          SEDN-----ANKNELTITGKDKAG--------AG-VLTAIV-LAVILGGAI
170 tr|F8MRU1|F8MRU1_NEUT8          SRHA------TEPAKPITTADKAG--------AA-MCTILL-IAGGIAIWI
171 tr|A1DJS0|A1DJS0_NEOFI          DKSGQT-TP----TRKITTGDKAG--------AS-ILTIGF-AVGWVGLMA
172 tr|B6HLM8|B6HLM8_PENRW          GRSGSS-----ALPPITTADKAG--------AG-ILAVVF-VGAMVGGVV
173 tr|A0A117DWS0|A0A117DWS0_ASPNG  PTDAPN-KL----P-DITTGDRAG--------AS-ILTVLF-VGGWAGAIT
174 tr|A0A124BXM1|A0A124BXM1_ASPNG  DNTD-S-GST---ESKITTGDKAG--------AS-ILTIAF-VGMWGGMIA
175 tr|J3NZQ6|J3NZQ6_GAGT3          ----H---TPATGYSDITTGDRAG--------AA-ILTTVI-IGLALGTFG
176 tr|W0THH8|W0THH8_KLUMD          VST----TNVLQKKLTITGSDKAG--------AG-VITAVV-IGVLVLGSV
177 tr|B2AEF2|B2AEF2_PODAN          SRGP------SEPLAPITNGDRAG--------AA-ILTILI-LGSAVGSWA
178 tr|Q96TX1|Q96TX1_NEUCS          S---D---SFT-HDKPVTTADKAG--------AG-IVTFLL-AASTLGTFV
179 tr|E9DYA6|E9DYA6_METAQ          D---G---SVMPNQKPVTAGDKVG--------AS-IVTILL-LGGACGMFG
180 tr|B9WAI2|B9WAI2_CANDC          KLTS------DATPLSIDGGDRAG--------AG-IITAII-GASLVGSCV
181 tr|Q6C171|Q6C171_YARLI          GFHPFGE--DEAHPLNIGGGDRAG--------AA-IITVIV-GVMMIAAGW
182 tr|A6ZMV1|A6ZMV1_YEAS7          SST----TNVLQNNLNIKKGDRAG--------AA-IITAVI-LSVLTGGAV
183 tr|B6K7X7|B6K7X7_SCHJY          SASS------TVVISPATTKDKHW--------SR-FFTALL-VFLLLSCSI
184 tr|I1RSA2|I1RSA2_GIBZE          DAND---NDGTPKRKAITGADKAG--------AG-ILTFLF-VAGIIGGVS
185 tr|F7VWZ9|F7VWZ9_SORMK          DPAS---RGGLAALKPITTADRVG--------AG-IVTAIL-AISIVGGSV
186 tr|A0A1B2J789|A0A1B2J789_PICPA  TKTE------AVPPLKLTNADVAG--------AA-IITAII-GLSVIAGAI
187 tr|G0V8N4|G0V8N4_NAUCC          QEAT------NLAPLHITDGSKAA--------AA-IITSVV-GISIVAAGV
188 tr|H8WXN8|H8WXN8_CANO9          TTAT------NAKPLSLDSGDKAG--------AG-IITAII-GASLVGCGV
189 tr|A6ZZS0|A6ZZS0_YEAS7          AQPT------NLAPLNITKGSKAG--------AG-IITAVI-GISIVACAL
190 tr|G9MQD1|G9MQD1_HYPVG          ---GD---NGEVELKPIGAADKAG--------AV-IVTLIL-LGSATAAFI
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH  TQAS------NLSPLTITAGSRAA--------AG-IITAVI-GLSIVACTV
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM  DTHEDE---APDALAPINTGDRAG--------AW-ILTVII-VVAVGGSVG
193 tr|A0A254UET1|A0A254UET1_ASPNG  ASDSSR---DPSELRPITTSDRIG--------AG-VVTGVT-ACLVLGVTA
194 tr|A1D587|A1D587_NEOFI          TEQ-AA-PQ----PQPVTAGDRAG--------AG-ILTVVF-LSGWIAAVT
```

```
195 tr|Q1K7I4|Q1K7I4_NEUCR          S---D---SFT-HDKPVTTADKAG--------AG-IVTFLL-AASTLGTFV
196 tr|G8JS88|G8JS88_ERECY          RNLG-------TLSPLDISNADRAG--------AG-IITALV-GFSLISCIV
197 tr|A0A136J7D3|A0A136J7D3_9PEZI DE-------TKMVMFEPITVGERVA--------AG-FLTTAA-TLSLIAAVV
198 tr|E9E4X8|E9E4X8_METAQ          G---D---NAQKKSKPITTADKAG--------AG-ILTILV-LGSACGVFG
199 tr|Q6CP42|Q6CP42_KLULA          QSPT-------NLAPLTIGKGDQAG--------AG-IITAVI-GITIIGSVV
200 tr|G0WHL4|G0WHL4_NAUDC          STT----TNVLENKMTITNGDRAG--------AA-IVTAII-LGVLTGGAV
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU AQPT-------NLSPLNITKGSKAG--------AG-IITAVI-GISIVACAL
202 tr|F9WYX6|F9WYX6_ZYMTI          RGDG-N-GSGTGSNSPQSSGDGTGA-------AS-TLMVSS-MLAVVGMFV
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6 S---T---PH-QVFAPITDGDRAG--------AA-FLTIAV-LGGAAFMFG
204 tr|J3P147|J3P147_GAGT3          ----H---TPTRSFKEITGGDRAG--------AG-ILTTLL-IGSGLGMFG
205 tr|I2H842|I2H842_TETBL          QKPT-------NLGPLNITAGSKAG--------AG-IITCVV-GISILACSM
206 tr|S6E3R3|S6E3R3_ZYGB2          SST----TNVLQNKLSINTGDRAG--------AA-VVTAVV-LMIVVGGSV
207 tr|E5AD94|E5AD94_LEPMJ          NDDDR----SPLRQSEITTADKAG--------AA-IMTILL-AVATAGGAY
208 tr|Q6FJM7|Q6FJM7_CANGA          STT----TNVLKNNLKITGGDRAG--------AA-IVTTIV-LGIIIGGAA
209 tr|A5DUW0|A5DUW0_LODEL          TTKA-------NAQPLSLDGGDKAG--------AG-IITAVI-GISIIGCGV
210 tr|G8BF46|G8BF46_CANPC          TVAT-------NAKPLSLDSGDKAG--------AG-IITAII-GASLIGCCV
211 tr|W3XAF6|W3XAF6_PESFW          AS-----TKYQYLRDITTADKVG--------AG-FVTTII-MACLVAALV
212 tr|B6H7I2|B6H7I2_PENRW          NDDEKK-------MKPITTGDKAG--------AS-ILTLIF-VFGWAGAVA
213 tr|Q5ATF9|Q5ATF9_EMENI          DSKHNT-GE----AKPITTGDRAG--------AG-ILTVVF-VGLWGAMIA
214 tr|A0A167CDD3|A0A167CDD3_9ASCO SSDN-----VLTVDELTIGTKDKAG--------AG-VITAIV-IIGLLATSW
215 tr|G0RXS3|G0RXS3_CHATD          ----N---NDPEPMRPITVGDRAG--------AG-VLTAVI-LGSFLGGLL
216 tr|B8NX26|B8NX26_ASPFN          PSSGE--GKDPSKPKPITTADRAG--------AW-ILMVMM-IAAVTGSVG
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM --SDPT---FEL--SDITTGDRVG--------AG-FLTSVV-LIGILGGAW
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM SSERIA---VSDDSSIVTMADRVV--------AW-VATGGL-GVLLGGYLY
219 tr|A1CMM3|A1CMM3_ASPCL          VSDEDS---VGIHEKLISTADRAG--------AG-ILTGLV-VAAAISAVY
220 tr|B6HU84|B6HU84_PENRW          SSDSNV-PI----LNSIGTGDRAG--------AG-ILTVIF-VSGWAVAVT
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS S---D---SFT-HDKPVTTADKAG--------AG-IVTFLL-AASTLGTFV
222 tr|G2WHY6|G2WHY6_YEASK          AQPT-------NLAPLNITKGSKAG--------AG-IITAVI-GISIVACAL
223 tr|G0W7G8|G0W7G8_NAUDC          HHPT-------NLAPLHITKGSRAG--------AG-IITAVI-GISILACAL
224 tr|G4NGV8|G4NGV8_MAGO7          KA-------PEIGVFQPITTAGRVA--------AG-FLTTGT-ILSLVGACI
225 tr|W3X855|W3X855_PESFW          KSE------SDVRGTKPITTGDKVG--------AG-ILTALV-LGGLVGGVC
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA QA--N---AYHKDPKPITTADRAG--------AG-IITFVL-LSASLGMFG
227 tr|J8Q2K7|J8Q2K7_SACAR          AQPT-------NLAPLNITKGSKAG--------AG-IITAVI-GISIVACAL
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO TTAV-------DATPLSLDAGDRAG--------AG-IITAIL-GSALIGTGV
229 tr|J6EGN8|J6EGN8_SACK1          SST----TNVLQNNLNITKGGRAG--------AA-IITAII-LSVLIGGAV
230 tr|A0A100ISD9|A0A100ISD9_ASPNG GPAPAP-V-PEGLRLEITKADRAG--------AG-MMTVML-SMIVIGSTG
231 tr|G8YQC0|G8YQC0_PICSO          THET------TNRNLLTISGKDKAG--------AG-VLTAGV-LAIVLGCAV
232 tr|F7VVP8|F7VVP8_SORMK          SRGSS---SEEEEMKPITTADKAG--------AA-ICTVLL-LAGGIAVWG
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH SST----TNVLQNNLNIKKGDRAG--------AA-IITAVI-LSVLTGGAV
234 tr|F8MBP4|F8MBP4_NEUT8          DPAS---KGGLASLKAITMADRVG--------AG-IVTAIL-AISIVGGSV
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG TTTAST-TAGSNSVVPTNGANSLAYGPLGVAAA-ILAGVV-ALA------
236 tr|A1CD27|A1CD27_ASPCL          ATPGDE---RPDQLAPINTGDRAG--------AW-ILTVVI-VVLVGGSVG
237 tr|A5DHF3|A5DHF3_PICGU          SQDN------INENELKITGKDKAG--------AG-VLTAIV-LAIILGGAV
238 tr|D4ATT7|D4ATT7_ARTBC          SGNNQT---KQP--SEINSGDRAG--------AG-ILTTLI-LLGIAGGVW
239 tr|E2PT42|E2PT42_ASPNC          GPPPAP-V-PEGLRLEITKADRAG--------AG-MMTVML-SMIVIGSTG
240 tr|C7ZPE5|C7ZPE5_NECH7          GS--T---TPGKHYKPITTADKAG--------AG-ILTFLV-LSIACGMFG
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA RGHQT----PDIATRKITAGDRAG--------AV-ILTVLL-ILVLLVSPY
242 tr|Q4WKP7|Q4WKP7_ASPFU          SEQ-AA-PQ----PQPVTTGDRAG--------AG-IVTVVF-LSGWIAAVT
243 tr|F8MXT4|F8MXT4_NEUT8          S---D---SFT-HDKPVTTADKAG--------AG-IVTFLL-AASTLGTFV
244 tr|A0A177AJ74|A0A177AJ74_9PEZI AKTAA----SLINDWTPTTGDKAG--------AG-VVTAVI-LAGLVGGVG
245 tr|J7RQR5|J7RQR5_KAZNA          STA----TNVFANSLNIRGKDKAG--------AA-IVTTVI-LGTLAGGSI
246 tr|Q75CW7|Q75CW7_ASHGO          LSSP--N-SITLPTRQIRLKDRVG--------AG-FATVVF-LLPFITGSI
247 tr|G3AKK4|G3AKK4_SPAPN          AKDA------INKNEIVVTGKDRAG--------AA-VLTAVV-LAVILGGAI
248 tr|B2W762|B2W762_PYRTR          ATDKPI----TFNVVTTGDKAG--------AS-ILTFLV-LVSILVGAW
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH AQPT-------NLAPLNITKGSKAG--------AG-IITAVI-GISIVACAL
250 tr|C4XWK1|C4XWK1_CLAL4          TLSS------VNRNKITITGKDRAG--------AG-VLTAIV-LGVILAGAV

001 tr|A0A254U2J9|A0A254U2J9_ASPNG WMVIGG----------------------------------------
002 sp|Q6FLP9|DCW1_CANGA           WLVY------------------------------------------
003 tr|A1CCM5|A1CCM5_ASPCL         FMVIGG----------------------------------------
004 tr|C1H190|C1H190_PARBA         WMMA------------------------------------------
005 tr|Q2UR85|Q2UR85_ASPOR         WMVYGG----------------------------------------
006 tr|A5DV30|A5DV30_LODEL         WMIF------------------------------------------
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO ------------------------------------------
008 tr|J3PGN0|J3PGN0_GAGT3         WMSWDR----------------------------------------
009 tr|A0A124BXU6|A0A124BXU6_ASPNG WLMLG-----------------------------------------
010 tr|C5P4A1|C5P4A1_COCP7         FSVS------------------------------------------
011 tr|Q0CG55|Q0CG55_ASPTN         WVLYE-----------------------------------------
```

```
012 tr|C4Y2X5|C4Y2X5_CLAL4          WLVL----------------------------------------
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG  WMVYE---------------------------------------
014 tr|C7Z068|C7Z068_NECH7          WMCAFD--------------------------------------
015 tr|G8BXX2|G8BXX2_TETPH          WLVI----------------------------------------
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR  WLVI----------------------------------------
017 tr|A1CSC2|A1CSC2_ASPCL          WMVYGR--------------------------------------
018 sp|Q05031|DFG5_YEAST            WMLF----------------------------------------
019 tr|H2AQJ6|H2AQJ6_KAZAF          WMLF----------------------------------------
020 tr|G2XFH9|G2XFH9_VERDV          WMSFWD--------------------------------------
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH  WLIL----------------------------------------
022 sp|Q5AD78|DCW1_CANAL            WLIL----------------------------------------
023 tr|B6H7E8|B6H7E8_PENRW          WMLVGGK-------------------------------------
024 tr|B6GZU2|B6GZU2_PENRW          WLIWD---------------------------------------
025 tr|E9E4W7|E9E4W7_METAQ          WMSFFDPVVP----------------------------------
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO  WLIWD---------------------------------------
027 tr|G2R7G8|G2R7G8_THITE          WMGMGWSEGS----------------------------------
028 tr|G3JBY1|G3JBY1_CORMM          FLMSDLYD------------------------------------
029 tr|A5DNV5|A5DNV5_PICGU          WLLL----------------------------------------
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI  WIVC----------------------------------------
031 tr|Q2TWC1|Q2TWC1_ASPOR          WLIKTQ--------------------------------------
032 tr|J5RXY9|J5RXY9_SACK1          WLVF----------------------------------------
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA  FVTSNVFSYA------------------------------------
034 tr|A0A254U5V1|A0A254U5V1_ASPNG  WLVYGG--------------------------------------
035 tr|A0A061B6H7|A0A061B6H7_CYBFA  WLLL----------------------------------------
036 tr|G8ZQ93|G8ZQ93_TORDC          WLIL----------------------------------------
037 tr|Q9C2J1|Q9C2J1_NEUCS          FLTI----------------------------------------
038 tr|C5NZK5|C5NZK5_COCP7          WMVS----------------------------------------
039 tr|Q6C0T7|Q6C0T7_YARLI          WLLTDGMYKGFKI----------KRRDPSE--------------
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB  WLVY----------------------------------------
041 tr|F7VZ72|F7VZ72_SORMK          WMGLGK--------------------------------------
042 tr|A3LN37|A3LN37_PICST          WLVI----------------------------------------
043 tr|Q6CAI2|Q6CAI2_YARLI          WMLA----------------------------------------
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO  WLIL----------------------------------------
045 tr|G9MJS5|G9MJS5_HYPVG          WMSLGP--------------------------------------
046 tr|G9MHI4|G9MHI4_HYPVG          VLLL----------------------------------------
047 tr|D4AP27|D4AP27_ARTBC          WMSV----------------------------------------
048 tr|B6GZT8|B6GZT8_PENRW          FMMLGG--------------------------------------
049 sp|O74556|YCZ2_SCHPO            VLY-----------------------------------------
050 tr|A0A0C4E991|A0A0C4E991_MAGP6  WMSWDR--------------------------------------
051 tr|B8MHG0|B8MHG0_TALSN          WLVIGGAN------------------------------------
052 tr|A3LMV8|A3LMV8_PICST          WMIF----------------------------------------
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU  SA------------------------------------------
054 sp|Q9P6I3|YHG7_SCHPO            WLYF----------------------------------------
055 tr|Q2US57|Q2US57_ASPOR          LLIFT---------------------------------------
056 tr|W7N622|W7N622_GIBM7          FATI----------------------------------------
057 sp|Q9P6I4|YHG6_SCHPO            WALVEDEEGKIPSRGKKGIAIS----------S------------
058 tr|C7GRB3|C7GRB3_YEAS2          WMLF----------------------------------------
059 tr|C7GP28|C7GP28_YEAS2          WLVF----------------------------------------
060 tr|G3JGZ5|G3JGZ5_CORMM          WMGFMV--------------------------------------
061 tr|C4QXV4|C4QXV4_KOMPG          WLLL----------------------------------------
062 tr|I1S0Z5|I1S0Z5_GIBZE          WMSLGK--------------------------------------
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM  WLLYE---------------------------------------
064 tr|G0W5X4|G0W5X4_NAUDC          WIIL----------------------------------------
065 tr|G8ZPC1|G8ZPC1_TORDC          WMLF----------------------------------------
066 tr|A0A167CDC5|A0A167CDC5_9ASCO  WMMKQ---------------------------------------
067 tr|R9XAT7|R9XAT7_ASHAC          WLII----------------------------------------
068 tr|B2B747|B2B747_PODAN          WMSTGV--------------------------------------
069 tr|G0SE99|G0SE99_CHATD          WMSMGE--------------------------------------
070 tr|J8Q6F0|J8Q6F0_SACAR          WMLF----------------------------------------
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH  WMIY----------------------------------------
072 tr|J3K6E2|J3K6E2_COCIM          FSVS----------------------------------------
073 tr|G4N3E1|G4N3E1_MAGO7          WMSWDSR-------------------------------------
074 tr|C5M5J2|C5M5J2_CANTT          WMMF----------------------------------------
075 tr|A0A254U0X0|A0A254U0X0_ASPNG  WLVKAQ--------------------------------------
076 sp|Q75DG6|DCW1_ASHGO            WLII----------------------------------------
077 tr|Q6BZF0|Q6BZF0_DEBHA          WLIL----------------------------------------
078 tr|E9E3Q1|E9E3Q1_METAQ          WMSAFD--------------------------------------
079 sp|P36091|DCW1_YEAST            WLVF----------------------------------------
080 tr|W0TAH6|W0TAH6_KLUMD          WLVL----------------------------------------
```

```
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA  WMVI----------------------------------------
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI  FVFYESM---------------------------------------
083 tr|Q2TYU3|Q2TYU3_ASPOR          FMVLGT----------------------------------------
084 tr|G2XGF6|G2XGF6_VERDV          FVMKEETK--------------------------------------
085 tr|G3AMT4|G3AMT4_SPAPN          WLIL------------------------------------------
086 tr|A0A1C1D213|A0A1C1D213_9EURO  WIIL------------------------------------------
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS  LMASGSFEKLEGGATPATAAAGEKKKKEKKEKLSIWYKIARESNFFRHQG
088 tr|A0A100I5A6|A0A100I5A6_ASPNG  WLVKAQ--------------------------------------
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH  WMVF------------------------------------------
090 tr|G2Q8A7|G2Q8A7_MYCTT          WMSTGV----------------------------------------
091 tr|B8NPZ1|B8NPZ1_ASPFN          WMIVP-----------------------------------------
092 tr|Q7SAB2|Q7SAB2_NEUCR          LMASGSFEKLEGGATPATAAAGEKKKKEKKEKLSIWYKIARESNFFRHQG
093 tr|A2R8R5|A2R8R5_ASPNC          WLMLG-----------------------------------------
094 tr|C5DHG0|C5DHG0_LACTC          WMIF------------------------------------------
095 tr|G0SFA3|G0SFA3_CHATD          FVII------------------------------------------
096 tr|K1WJG5|K1WJG5_MARBU          CLVLG-----------------------------------------
097 tr|H2B005|H2B005_KAZAF          WLIY------------------------------------------
098 tr|C5DYD7|C5DYD7_ZYGRC          WMLF------------------------------------------
099 tr|G8YM23|G8YM23_PICSO          WLIL------------------------------------------
100 tr|W3X8E3|W3X8E3_PESFW          WMSLGM----------------------------------------
101 tr|A0A254U3K7|A0A254U3K7_ASPNG  WLMLG-----------------------------------------
102 tr|Q6CER8|Q6CER8_YARLI          FLIV------------------------------------------
103 tr|G2QT05|G2QT05_THITE          WMSVGE----------------------------------------
104 sp|Q5ACZ2|DFG5_CANAL           WMIF------------------------------------------
105 tr|G2WKU6|G2WKU6_YEASK          WMLF------------------------------------------
106 tr|A0A0C7N752|A0A0C7N752_9SACH  WMMF------------------------------------------
107 tr|W3WN68|W3WN68_PESFW          LMIKD-----------------------------------------
108 tr|G2QFS1|G2QFS1_MYCTT          WMSIGD----------------------------------------
109 tr|G3J9G4|G3J9G4_CORMM          FVTSDVLSFA------------------------------------
110 tr|Q2UJ03|Q2UJ03_ASPOR          WGINE-----------------------------------------
111 tr|G8YRT2|G8YRT2_PICSO          WMAI------------------------------------------
112 tr|W3X554|W3X554_PESFW          FVMK------------------------------------------
113 tr|Q1K7A8|Q1K7A8_NEUCR          FLTI------------------------------------------
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM  WMVYGR----------------------------------------
115 tr|A7F3P0|A7F3P0_SCLS1          FMVIG-----------------------------------------
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA  FLMSNLFD----------------------------------------
117 tr|Q6CIP9|Q6CIP9_KLULA          WMVL------------------------------------------
118 tr|H2AMQ5|H2AMQ5_KAZAF          WIIF------------------------------------------
119 tr|G8JMI3|G8JMI3_ERECY          WMVI------------------------------------------
120 tr|C5DGT1|C5DGT1_LACTC          WLVL------------------------------------------
121 tr|K1XJR4|K1XJR4_MARBU          WMSVGES---------------------------------------
122 tr|A0A124BYC5|A0A124BYC5_ASPNG  ------------------------------------------------
123 tr|A7TLL2|A7TLL2_VANPO          WLIV------------------------------------------
124 tr|Q0CB86|Q0CB86_ASPTN          WLFIGG----------------------------------------
125 tr|K1XLH0|K1XLH0_MARBU          FMIIGS----------------------------------------
126 tr|J7S438|J7S438_KAZNA          ASVLRKQE--------------------------------------
127 tr|Q4WG09|Q4WG09_ASPFU          WLIKT-----------------------------------------
128 tr|H8WZ09|H8WZ09_CANO9          WMIF------------------------------------------
129 tr|I1RL09|I1RL09_GIBZE          WMCAFD----------------------------------------
130 tr|Q4W985|Q4W985_ASPFU          WLLYE-----------------------------------------
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU  WMLF------------------------------------------
132 tr|W6QCW3|W6QCW3_PENRF          WLIWD-----------------------------------------
133 tr|W3WMD3|W3WMD3_PESFW          WMNLETM---------------------------------------
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH  WLIL------------------------------------------
135 tr|I2GYH1|I2GYH1_TETBL          WMLF------------------------------------------
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH  WMLF------------------------------------------
137 tr|Q752P3|Q752P3_ASHGO          WMIF------------------------------------------
138 tr|B8MYP3|B8MYP3_ASPFN          WMVYGG----------------------------------------
139 tr|Q2TXL6|Q2TXL6_ASPOR          FMVTGG----------------------------------------
140 tr|Q0CN17|Q0CN17_ASPTN          WLVYGG----------------------------------------
141 tr|J3NHD5|J3NHD5_GAGT3          WMNKGD----------------------------------------
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI  WMSMAD----------------------------------------
143 tr|A1CVB0|A1CVB0_NEOFI          WLLYE-----------------------------------------
144 tr|B8NVI3|B8NVI3_ASPFN          FMVTGG----------------------------------------
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH  WLVI------------------------------------------
146 tr|A1DJ54|A1DJ54_NEOFI          WLIKT-----------------------------------------
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS  FMNLGD----------------------------------------
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA  WTVL------------------------------------------
149 tr|A0A177A618|A0A177A618_9PEZI  FMVGSDGWVA--------------------------------------
150 tr|Q75CW6|Q75CW6_ASHGO          WMSI------------------------------------------
```

```
151 tr|Q7S4K4|Q7S4K4_NEUCR          FMNLGD----------------------------------------
152 tr|G2QB99|G2QB99_MYCTT          FVIL------------------------------------------
153 tr|G0S3F2|G0S3F2_CHATD          WMSIGD----------------------------------------
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO  WLVL------------------------------------------
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB  WMMF------------------------------------------
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM  FMIIGGK---------------------------------------
157 tr|G2QKJ0|G2QKJ0_MYCTT          WMGMGWSEK-------------------------------------
158 tr|Q4WFX5|Q4WFX5_ASPFU          LVVT------------------------------------------
159 tr|Q5BGD7|Q5BGD7_EMENI          WLVKQQ----------------------------------------
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH  WLII------------------------------------------
161 tr|G8BYN9|G8BYN9_TETPH          WMLF------------------------------------------
162 tr|F7VVT4|F7VVT4_SORMK          FLVTGVGERK------------------------------------
163 tr|G8YB80|G8YB80_PICSO          WLIL------------------------------------------
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA  WMVL------------------------------------------
165 tr|G2WRS7|G2WRS7_VERDV          WMSVGDA---------------------------------------
166 tr|A0A124BVB5|A0A124BVB5_ASPNG  FMVS------------------------------------------
167 tr|A7EIG7|A7EIG7_SCLS1          WMIVD-----------------------------------------
168 tr|C5M5U7|C5M5U7_CANTT          WLIL------------------------------------------
169 tr|B9WAA8|B9WAA8_CANDC          WMIF------------------------------------------
170 tr|F8MRU1|F8MRU1_NEUT8          FMNLGD----------------------------------------
171 tr|A1DJS0|A1DJS0_NEOFI          FMVIGG----------------------------------------
172 tr|B6HLM8|B6HLM8_PENRW          FLFMDP----------------------------------------
173 tr|A0A117DWS0|A0A117DWS0_ASPNG  WLVYGG----------------------------------------
174 tr|A0A124BXM1|A0A124BXM1_ASPNG  WMVIGG----------------------------------------
175 tr|J3NZQ6|J3NZQ6_GAGT3          WILWERSE--------------------------------------
176 tr|W0THH8|W0THH8_KLUMD          WMIY------------------------------------------
177 tr|B2AEF2|B2AEF2_PODAN          WMSFGD----------------------------------------
178 tr|Q96TX1|Q96TX1_NEUCS          WMGIGK----------------------------------------
179 tr|E9DYA6|E9DYA6_METAQ          WMSYEASGR-------------------------------------
180 tr|B9WAI2|B9WAI2_CANDC          WLIL------------------------------------------
181 tr|Q6C171|Q6C171_YARLI          WLLV------------------------------------------
182 tr|A6ZMV1|A6ZMV1_YEAS7          WMLF------------------------------------------
183 tr|B6K7X7|B6K7X7_SCHJY          WLFL------------------------------------------
184 tr|I1RSA2|I1RSA2_GIBZE          FLVLDL----------------------------------------
185 tr|F7VWZ9|F7VWZ9_SORMK          FLTI------------------------------------------
186 tr|A0A1B2J789|A0A1B2J789_PICPA  WLLL------------------------------------------
187 tr|G0V8N4|G0V8N4_NAUCC          WLII------------------------------------------
188 tr|H8WXN8|H8WXN8_CANO9          WLIL------------------------------------------
189 tr|A6ZZS0|A6ZZS0_YEAS7          WLVF------------------------------------------
190 tr|G9MQD1|G9MQD1_HYPVG          WICSFD----------------------------------------
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH  WLVV------------------------------------------
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM  WLIKT-----------------------------------------
193 tr|A0A254UET1|A0A254UET1_ASPNG  FMVS------------------------------------------
194 tr|A1D587|A1D587_NEOFI          WMVYGR----------------------------------------
195 tr|Q1K7I4|Q1K7I4_NEUCR          WMGIGK----------------------------------------
196 tr|G8JS88|G8JS88_ERECY          WLVS------------------------------------------
197 tr|A0A136J7D3|A0A136J7D3_9PEZI  FLLRDETDIEARM----------DKLENEKTVKWGV----------
198 tr|E9E4X8|E9E4X8_METAQ          WMSVGV----------------------------------------
199 tr|Q6CP42|Q6CP42_KLULA          WLII------------------------------------------
200 tr|G0WHL4|G0WHL4_NAUDC          WMFF------------------------------------------
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU  WLVF------------------------------------------
202 tr|F9WYX6|F9WYX6_ZYMTI          ELFVLA----------------------------------------
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6  WMNKGH----------------------------------------
204 tr|J3P147|J3P147_GAGT3          WMSYD-----------------------------------------
205 tr|I2H842|I2H842_TETBL          WLVV------------------------------------------
206 tr|S6E3R3|S6E3R3_ZYGB2          WMLF------------------------------------------
207 tr|E5AD94|E5AD94_LEPMJ          WILA------------------------------------------
208 tr|Q6FJM7|Q6FJM7_CANGA          WMMF------------------------------------------
209 tr|A5DUW0|A5DUW0_LODEL          WLIL------------------------------------------
210 tr|G8BF46|G8BF46_CANPC          WLIL------------------------------------------
211 tr|W3XAF6|W3XAF6_PESFW          FMNLD-----------------------------------------
212 tr|B6H7I2|B6H7I2_PENRW          FMMLGG----------------------------------------
213 tr|Q5ATF9|Q5ATF9_EMENI          WMVLGEGL--------------------------------------
214 tr|A0A167CDD3|A0A167CDD3_9ASCO  WMLI------------------------------------------
215 tr|G0RXS3|G0RXS3_CHATD          WMGMGWSEG-------------------------------------
216 tr|B8NX26|B8NX26_ASPFN          WLVKP-----------------------------------------
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM  WMVS------------------------------------------
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM  LVVT------------------------------------------
219 tr|A1CMM3|A1CMM3_ASPCL          FMVI------------------------------------------

220 tr|B6HU84|B6HU84_PENRW          WMIRGG----------------------------------------
```

```
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS  WMGIGK----------
222 tr|G2WHY6|G2WHY6_YEASK         WLVF------------
223 tr|G0W7G8|G0W7G8_NAUDC         WLVF------------
224 tr|G4NGV8|G4NGV8_MAGO7         FVVT------------
225 tr|W3X855|W3X855_PESFW         FLILT-----------
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA WMSVGI----------
227 tr|J8Q2K7|J8Q2K7_SACAR         WLVF------------
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO WLVI------------
229 tr|J6EGN8|J6EGN8_SACK1         WMLF------------
230 tr|A0A100ISD9|A0A100ISD9_ASPNG WMVYE-----------
231 tr|G8YQC0|G8YQC0_PICSO         WMAI------------
232 tr|F7VVP8|F7VVP8_SORMK         FMSLGD----------
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH WMLF------------
234 tr|F8MBP4|F8MBP4_NEUT8         FLTI------------
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG ----------------
236 tr|A1CD27|A1CD27_ASPCL         WLIKT-----------
237 tr|A5DHF3|A5DHF3_PICGU         WMLF------------
238 tr|D4ATT7|D4ATT7_ARTBC         WMVS------------
239 tr|E2PT42|E2PT42_ASPNC         WMVYE-----------
240 tr|C7ZPE5|C7ZPE5_NECH7         WMSMGK----------
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA WLLCV-----------
242 tr|Q4WKP7|Q4WKP7_ASPFU         WMVYGR----------
243 tr|F8MXT4|F8MXT4_NEUT8         WMGIGK----------
244 tr|A0A177AJ74|A0A177AJ74_9PEZI WMIMAD----------
245 tr|J7RQR5|J7RQR5_KAZNA         WMLF------------
246 tr|Q75CW7|Q75CW7_ASHGO         WMSI------------
247 tr|G3AKK4|G3AKK4_SPAPN         WMIF------------
248 tr|B2W762|B2W762_PYRTR         WMVS------------
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH WLVF------------
250 tr|C4XWK1|C4XWK1_CLAL4         WMLF------------
```

```
001 tr|A0A254U2J9|A0A254U2J9_ASPNG
002 sp|Q6FLP9|DCW1_CANGA
003 tr|A1CCM5|A1CCM5_ASPCL
004 tr|C1H190|C1H190_PARBA
005 tr|Q2UR85|Q2UR85_ASPOR
006 tr|A5DV30|A5DV30_LODEL
007 tr|A0A1E3NP59|A0A1E3NP59_9ASCO
008 tr|J3PGN0|J3PGN0_GAGT3
009 tr|A0A124BXU6|A0A124BXU6_ASPNG
010 tr|C5P4A1|C5P4A1_COCP7
011 tr|Q0CG55|Q0CG55_ASPTN
012 tr|C4Y2X5|C4Y2X5_CLAL4
013 tr|A0A254TKQ9|A0A254TKQ9_ASPNG
014 tr|C7Z068|C7Z068_NECH7
015 tr|G8BXX2|G8BXX2_TETPH
016 tr|A0A0L0P6S8|A0A0L0P6S8_CANAR
017 tr|A1CSC2|A1CSC2_ASPCL
018 sp|Q05031|DFG5_YEAST
019 tr|H2AQJ6|H2AQJ6_KAZAF
020 tr|G2XFH9|G2XFH9_VERDV
021 tr|A0A1S7HQM3|A0A1S7HQM3_9SACH
022 sp|Q5AD78|DCW1_CANAL
023 tr|B6H7E8|B6H7E8_PENRW
024 tr|B6GZU2|B6GZU2_PENRW
025 tr|E9E4W7|E9E4W7_METAQ
026 tr|A0A1S9RH77|A0A1S9RH77_9EURO
027 tr|G2R7G8|G2R7G8_THITE
028 tr|G3JBY1|G3JBY1_CORMM
029 tr|A5DNV5|A5DNV5_PICGU
030 tr|A0A1B8GPA6|A0A1B8GPA6_9PEZI
031 tr|Q2TWC1|Q2TWC1_ASPOR
032 tr|J5RXY9|J5RXY9_SACK1
033 tr|A0A2N6NN14|A0A2N6NN14_BEABA
034 tr|A0A254U5V1|A0A254U5V1_ASPNG
035 tr|A0A061B6H7|A0A061B6H7_CYBFA
036 tr|G8ZQ93|G8ZQ93_TORDC
037 tr|Q9C2J1|Q9C2J1_NEUCS
```

038 tr|C5NZK5|C5NZK5_COCP7
039 tr|Q6C0T7|Q6C0T7_YARLI
040 tr|A0A0W0ENZ4|A0A0W0ENZ4_CANGB
041 tr|F7VZ72|F7VZ72_SORMK
042 tr|A3LN37|A3LN37_PICST
043 tr|Q6CAI2|Q6CAI2_YARLI
044 tr|A0A0V1PWA0|A0A0V1PWA0_9ASCO
045 tr|G9MJS5|G9MJS5_HYPVG
046 tr|G9MHI4|G9MHI4_HYPVG
047 tr|D4AP27|D4AP27_ARTBC
048 tr|B6GZT8|B6GZT8_PENRW
049 sp|O74556|YCZ2_SCHPO
050 tr|A0A0C4E991|A0A0C4E991_MAGP6
051 tr|B8MHG0|B8MHG0_TALSN
052 tr|A3LMV8|A3LMV8_PICST
053 tr|A0A099P0Y5|A0A099P0Y5_PICKU
054 sp|Q9P6I3|YHG7_SCHPO
055 tr|Q2US57|Q2US57_ASPOR
056 tr|W7N622|W7N622_GIBM7
057 sp|Q9P6I4|YHG6_SCHPO
058 tr|C7GRB3|C7GRB3_YEAS2
059 tr|C7GP28|C7GP28_YEAS2
060 tr|G3JGZ5|G3JGZ5_CORMM
061 tr|C4QXV4|C4QXV4_KOMPG
062 tr|I1S0Z5|I1S0Z5_GIBZE
063 tr|A0A0J5PSQ0|A0A0J5PSQ0_ASPFM
064 tr|G0W5X4|G0W5X4_NAUDC
065 tr|G8ZPC1|G8ZPC1_TORDC
066 tr|A0A167CDC5|A0A167CDC5_9ASCO
067 tr|R9XAT7|R9XAT7_ASHAC
068 tr|B2B747|B2B747_PODAN
069 tr|G0SE99|G0SE99_CHATD
070 tr|J8Q6F0|J8Q6F0_SACAR
071 tr|A0A0A8LDJ2|A0A0A8LDJ2_9SACH
072 tr|J3K6E2|J3K6E2_COCIM
073 tr|G4N3E1|G4N3E1_MAGO7
074 tr|C5M5J2|C5M5J2_CANTT
075 tr|A0A254U0X0|A0A254U0X0_ASPNG
076 sp|Q75DG6|DCW1_ASHGO
077 tr|Q6BZF0|Q6BZF0_DEBHA
078 tr|E9E3Q1|E9E3Q1_METAQ
079 sp|P36091|DCW1_YEAST
080 tr|W0TAH6|W0TAH6_KLUMD
081 tr|A0A0H5CF34|A0A0H5CF34_CYBJA
082 tr|A0A2C5WNS2|A0A2C5WNS2_9PEZI
083 tr|Q2TYU3|Q2TYU3_ASPOR
084 tr|G2XGF6|G2XGF6_VERDV
085 tr|G3AMT4|G3AMT4_SPAPN
086 tr|A0A1C1D213|A0A1C1D213_9EURO
087 tr|A0A0B0DST1|A0A0B0DST1_NEUCS
088 tr|A0A100I5A6|A0A100I5A6_ASPNG
089 tr|A0A1G4JM92|A0A1G4JM92_9SACH
090 tr|G2Q8A7|G2Q8A7_MYCTT
091 tr|B8NPZ1|B8NPZ1_ASPFN
092 tr|Q7SAB2|Q7SAB2_NEUCR
093 tr|A2R8R5|A2R8R5_ASPNC
094 tr|C5DHG0|C5DHG0_LACTC
095 tr|G0SFA3|G0SFA3_CHATD
096 tr|K1WJG5|K1WJG5_MARBU
097 tr|H2B005|H2B005_KAZAF
098 tr|C5DYD7|C5DYD7_ZYGRC
099 tr|G8YM23|G8YM23_PICSO
100 tr|W3X8E3|W3X8E3_PESFW
101 tr|A0A254U3K7|A0A254U3K7_ASPNG
102 tr|Q6CER8|Q6CER8_YARLI
103 tr|G2QT05|G2QT05_THITE
104 sp|Q5ACZ2|DFG5_CANAL
105 tr|G2WKU6|G2WKU6_YEASK
106 tr|A0A0C7N752|A0A0C7N752_9SACH
107 tr|W3WN68|W3WN68_PESFW

108 tr|G2QFS1|G2QFS1_MYCTT
109 tr|G3J9G4|G3J9G4_CORMM
110 tr|Q2UJ03|Q2UJ03_ASPOR
111 tr|G8YRT2|G8YRT2_PICSO
112 tr|W3X554|W3X554_PESFW
113 tr|Q1K7A8|Q1K7A8_NEUCR
114 tr|A0A0J5PM92|A0A0J5PM92_ASPFM
115 tr|A7F3P0|A7F3P0_SCLS1
116 tr|A0A2N6NK60|A0A2N6NK60_BEABA
117 tr|Q6CIP9|Q6CIP9_KLULA
118 tr|H2AMQ5|H2AMQ5_KAZAF
119 tr|G8JMI3|G8JMI3_ERECY
120 tr|C5DGT1|C5DGT1_LACTC
121 tr|K1XJR4|K1XJR4_MARBU
122 tr|A0A124BYC5|A0A124BYC5_ASPNG
123 tr|A7TLL2|A7TLL2_VANPO
124 tr|Q0CB86|Q0CB86_ASPTN
125 tr|K1XLH0|K1XLH0_MARBU
126 tr|J7S438|J7S438_KAZNA
127 tr|Q4WG09|Q4WG09_ASPFU
128 tr|H8WZ09|H8WZ09_CANO9
129 tr|I1RL09|I1RL09_GIBZE
130 tr|Q4W985|Q4W985_ASPFU
131 tr|A0A0L8RDQ3|A0A0L8RDQ3_SACEU
132 tr|W6QCW3|W6QCW3_PENRF
133 tr|W3WMD3|W3WMD3_PESFW
134 tr|A0A1S7HZQ4|A0A1S7HZQ4_9SACH
135 tr|I2GYH1|I2GYH1_TETBL
136 tr|A0A1S7HIA4|A0A1S7HIA4_9SACH
137 tr|Q752P3|Q752P3_ASHGO
138 tr|B8MYP3|B8MYP3_ASPFN
139 tr|Q2TXL6|Q2TXL6_ASPOR
140 tr|Q0CN17|Q0CN17_ASPTN
141 tr|J3NHD5|J3NHD5_GAGT3
142 tr|A0A1B8GPI3|A0A1B8GPI3_9PEZI
143 tr|A1CVB0|A1CVB0_NEOFI
144 tr|B8NVI3|B8NVI3_ASPFN
145 tr|A0A0C7N2N9|A0A0C7N2N9_9SACH
146 tr|A1DJ54|A1DJ54_NEOFI
147 tr|A0A0B0DFT3|A0A0B0DFT3_NEUCS
148 tr|A0A0H5C0E8|A0A0H5C0E8_CYBJA
149 tr|A0A177A618|A0A177A618_9PEZI
150 tr|Q75CW6|Q75CW6_ASHGO
151 tr|Q7S4K4|Q7S4K4_NEUCR
152 tr|G2QB99|G2QB99_MYCTT
153 tr|G0S3F2|G0S3F2_CHATD
154 tr|A0A1E4SC85|A0A1E4SC85_9ASCO
155 tr|A0A0W0CYJ3|A0A0W0CYJ3_CANGB
156 tr|A0A0J5PQ80|A0A0J5PQ80_ASPFM
157 tr|G2QKJ0|G2QKJ0_MYCTT
158 tr|Q4WFX5|Q4WFX5_ASPFU
159 tr|Q5BGD7|Q5BGD7_EMENI
160 tr|A0A0A8L8U3|A0A0A8L8U3_9SACH
161 tr|G8BYN9|G8BYN9_TETPH
162 tr|F7VVT4|F7VVT4_SORMK
163 tr|G8YB80|G8YB80_PICSO
164 tr|A0A0H5CB47|A0A0H5CB47_CYBJA
165 tr|G2WRS7|G2WRS7_VERDV
166 tr|A0A124BVB5|A0A124BVB5_ASPNG
167 tr|A7EIG7|A7EIG7_SCLS1
168 tr|C5M5U7|C5M5U7_CANTT
169 tr|B9WAA8|B9WAA8_CANDC
170 tr|F8MRU1|F8MRU1_NEUT8
171 tr|A1DJS0|A1DJS0_NEOFI
172 tr|B6HLM8|B6HLM8_PENRW
173 tr|A0A117DWS0|A0A117DWS0_ASPNG
174 tr|A0A124BXM1|A0A124BXM1_ASPNG
175 tr|J3NZQ6|J3NZQ6_GAGT3
176 tr|W0THH8|W0THH8_KLUMD
177 tr|B2AEF2|B2AEF2_PODAN

178 tr|Q96TX1|Q96TX1_NEUCS
179 tr|E9DYA6|E9DYA6_METAQ
180 tr|B9WAI2|B9WAI2_CANDC
181 tr|Q6C171|Q6C171_YARLI
182 tr|A6ZMV1|A6ZMV1_YEAS7
183 tr|B6K7X7|B6K7X7_SCHJY
184 tr|I1RSA2|I1RSA2_GIBZE
185 tr|F7VWZ9|F7VWZ9_SORMK
186 tr|A0A1B2J789|A0A1B2J789_PICPA
187 tr|G0V8N4|G0V8N4_NAUCC
188 tr|H8WXN8|H8WXN8_CANO9
189 tr|A6ZZS0|A6ZZS0_YEAS7
190 tr|G9MQD1|G9MQD1_HYPVG
191 tr|A0A1G4JVK5|A0A1G4JVK5_9SACH
192 tr|A0A0J5PNX6|A0A0J5PNX6_ASPFM
193 tr|A0A254UET1|A0A254UET1_ASPNG
194 tr|A1D587|A1D587_NEOFI
195 tr|Q1K7I4|Q1K7I4_NEUCR
196 tr|G8JS88|G8JS88_ERECY
197 tr|A0A136J7D3|A0A136J7D3_9PEZI
198 tr|E9E4X8|E9E4X8_METAQ
199 tr|Q6CP42|Q6CP42_KLULA
200 tr|G0WHL4|G0WHL4_NAUDC
201 tr|A0A0L8RFQ5|A0A0L8RFQ5_SACEU
202 tr|F9WYX6|F9WYX6_ZYMTI
203 tr|A0A0C4DZP6|A0A0C4DZP6_MAGP6
204 tr|J3P147|J3P147_GAGT3
205 tr|I2H842|I2H842_TETBL
206 tr|S6E3R3|S6E3R3_ZYGB2
207 tr|E5AD94|E5AD94_LEPMJ
208 tr|Q6FJM7|Q6FJM7_CANGA
209 tr|A5DUW0|A5DUW0_LODEL
210 tr|G8BF46|G8BF46_CANPC
211 tr|W3XAF6|W3XAF6_PESFW
212 tr|B6H7I2|B6H7I2_PENRW
213 tr|Q5ATF9|Q5ATF9_EMENI
214 tr|A0A167CDD3|A0A167CDD3_9ASCO
215 tr|G0RXS3|G0RXS3_CHATD
216 tr|B8NX26|B8NX26_ASPFN
217 tr|A0A0E1RZZ1|A0A0E1RZZ1_COCIM
218 tr|A0A0J5SN29|A0A0J5SN29_ASPFM
219 tr|A1CMM3|A1CMM3_ASPCL
220 tr|B6HU84|B6HU84_PENRW
221 tr|A0A0B0E7F0|A0A0B0E7F0_NEUCS
222 tr|G2WHY6|G2WHY6_YEASK
223 tr|G0W7G8|G0W7G8_NAUDC
224 tr|G4NGV8|G4NGV8_MAGO7
225 tr|W3X855|W3X855_PESFW
226 tr|A0A2N6NY19|A0A2N6NY19_BEABA
227 tr|J8Q2K7|J8Q2K7_SACAR
228 tr|A0A1E4RBW2|A0A1E4RBW2_9ASCO
229 tr|J6EGN8|J6EGN8_SACK1
230 tr|A0A100ISD9|A0A100ISD9_ASPNG
231 tr|G8YQC0|G8YQC0_PICSO
232 tr|F7VVP8|F7VVP8_SORMK
233 tr|A0A0L8VJB2|A0A0L8VJB2_9SACH
234 tr|F8MBP4|F8MBP4_NEUT8
235 tr|A0A254TXZ7|A0A254TXZ7_ASPNG
236 tr|A1CD27|A1CD27_ASPCL
237 tr|A5DHF3|A5DHF3_PICGU
238 tr|D4ATT7|D4ATT7_ARTBC
239 tr|E2PT42|E2PT42_ASPNC
240 tr|C7ZPE5|C7ZPE5_NECH7
241 tr|A0A0A2VM24|A0A0A2VM24_PARBA
242 tr|Q4WKP7|Q4WKP7_ASPFU
243 tr|F8MXT4|F8MXT4_NEUT8
244 tr|A0A177AJ74|A0A177AJ74_9PEZI
245 tr|J7RQR5|J7RQR5_KAZNA
246 tr|Q75CW7|Q75CW7_ASHGO
247 tr|G3AKK4|G3AKK4_SPAPN

248 tr|B2W762|B2W762_PYRTR
249 tr|A0A0L8VLX1|A0A0L8VLX1_9SACH
250 tr|C4XWK1|C4XWK1_CLAL4

X - Insufficient data - the calculation for this site was performed on less than 10% of the sequences.