JOURNAL OF TROPICAL LIFE SCIENCE

2023, Vol. 13, No. 3, 503 – 516 http://dx.doi.org/10.11594/jtls.13.03.09

Research Article

Characterization of β-glucosidases from *Meridianimaribacter* sp. CL38

Clarine Wan Ling Hong ¹, Sye Jinn Chen ¹, Kok Jun Liew ¹, Ming Quan Lam ², Muhammad Ramziuddin Zakaria ¹, Kheng Loong Chong ¹, Chun Shiong Chong ¹

¹ Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, Skudai 81310, Malaysia

² Department of Biological Science, Faculty of Science, Universiti Tunku Abdul Rahman, Kampar 31900, Malaysia

Article history: Submission November 2022 Revised August 2023 Accepted August 2023

**Corresponding author:* E-mail: cschong@utm.my

ABSTRACT

The production of second-generation biofuel requires a huge amount of freshwater. It is estimated that at least three gal of freshwater is used to produce one gal of biofuel. The replacement of freshwater with seawater serves as a potential alternative in biofuel generation. Therefore, salt-tolerant enzymes play an important role in saccharification and fermentation process. Halophilic β -glucosidase is one of the key enzymes for the process. In this study, the β-glucosidase of halophile Meridianimaribacter sp. CL38 isolated from mangrove soil was characterized. Strain CL38 achieved maximum production of $\beta\mbox{-glucosidase}$ at 12^{th} hour of growth. The $\beta\mbox{-glu-}$ cosidase showed highest activity at 2% (*w/v*) NaCl while highly stable at salt concentration ranging from 1-2% (w/v) (more than 96% of relative activity). Its β -glucosidase activity remained active in the presence of 5mM Mn²⁺, Mg²⁺, Ca²⁺ ions, and 1% (v/v) Tween-20 and Tween-80. The draft genome sequence of strain CL38 was retrieved from GenBank database and submitted to dbCAN meta server for CA-Zymes annotation. Strain CL38 harbors 44 GHs and GH3 are annotated as β-glucosidases. The β-glucosidases of Meridianimaribacter flavus (99.61%) and Mesoflavibacter sabulilitoris (97.44%) showed the closest identity with Bgl3a and Bgl3b protein sequences from strain CL38, respectively. Glycoside hydrolase family 3 domain was identified in both enzymes via InterPro scan server. The presence of signal peptides indicated that both enzymes were secreted extracellularly. Five motifs were identified in Bgl3a and Bgl3b, with the active site (nucleophile) found at Asp296 and Asp297, respectively. Collectively, these β -glucosidases could be potentially used in the biofuel production, in particular the lignocellulosic biomass pretreatment process. This is the first attempt to characterize the β-glucosidase in genus Meridianimaribacter as so far none of the lignocellulolytic enzymes from this genus were characterized.

Keywords: β-glucosidase, GH3, Halophiles, Lignocellulosic biomass, Meridianimaribacter

Introduction

Fossil fuel is the main source of energy that has been widely used worldwide [1, 2]. Due to the extensive use of fossil fuels, it is currently facing a depletion crisis [3]. The combustion of fossil fuels is also known to cause air pollution and greenhouse effect [1]. Collectively, this encouraged scientists to seek new energy sources. Second-generation biofuel seems to be a good replacement for fossil fuels [4]. The production of the second-generation biofuel from lignocellulosic biomass is a sustainable process. It is not only able to ensure a continuous supply of energy source but also help to reduce the environmental pollution arising from agricultural waste processing[5].

The production of biofuel requires a vast amount of freshwater. Around three to four gal of freshwater is needed to produce one gal of biofuel [6]. Hence, seawater has been proposed to replace freshwater in the process [7]. Seawater has a general salinity of 3.2% (*w/v*) [8]. As the saccharifica-

How to cite:

Hong CWL, Chen SJ, Liew KJ, et al. (2023) Characterization of β -glucosidases from *Meridianimaribacter* sp. CL38. Journal of Tropical Life Science 13 (3): 503 – 516. doi: 10.11594/jtls.13.03.09.

tion and fermentation of the lignocellulosic biomass involve microbial degradation and enzymatic reaction, seawater uses biofuel production requires halophiles or enzymes with halotolerant characteristics [9–11]. A wide range of places promotes halophiles growth, for instance, oceans, salt lakes, coastal lagoons, and estuaries [12]. Besides, the mangrove environment is also a place with a high possibility of isolating salt-tolerant microorganisms that secrete halotolerant enzymes [13, 14]. As the mangrove environment consists of a high amount of organic matter (i.e., weathered log and leaf litter), it is also a good place for discovering novel lignocellulolytic enzymes [15].

Lignocellulose biomass consists mainly of cellulose, hemicellulose, and lignin [16]. In among, cellulose occupies nearly 40-60% of the lignocellulose biomass [17]. Due to this reason, cellulase is the key enzyme for the conversion of lignocellulosic waste into biofuel [18]. Cellulase is a group of enzymes, including endoglucanase, exoglucanase, and β -glucosidase. These three enzymes work synergistically to achieve complete saccharification of the cellulose into glucose [19]. β-glucosidase plays the most significant role as it governs the rate-limiting enzyme reaction during cellulose complete degradation [19]. Without β -glucosidase, the intermediate products (i.e., cellobiose and cellotriose) will accumulate and this will cause feedback inhibition on endoglucanase and exoglucanase activities [20].

Meridianimaribacter sp. CL38 was isolated from mangrove soil collected from Tanjung Piai, Johor, Malaysia (1°16′06.0″N 103°30′31.2″E) [11]. It is a bacterium identified as Gram-stained negative, yellow-pigmented, rod-shaped, and considered moderate halophile that can thrive in salt concentrations up to 9% (w/v) [11]. According to a previous genomic study, strain CL38 is closely related to Meridianimaribacter flavus NH57N, suggesting that both strains belong to the same genus and species [21]. Nonetheless, the genome of strain CL38 was found to encode different lignocellulolytic genes [11]. Based on the preliminary screening, strain CL38 showed positive results in some lignocellulose degradation tests, including β-glucosidase, endoglucanase, and exoglucanase [11]. When writing, none of the lignocellulolytic enzymes in Meridianimaribacter genus had been biochemically characterized and reported. Therefore, in this project, the β -glucosidase of strain CL38 was characterized and bioinformatics analysis of these β -glucosidases was also carried out.

Material and Methods

Bacterial culture of Meridianimaribacter sp. CL38

Meridianimaribacter sp. CL38 was isolated from mangrove soil by Lam et al. [11] and preserved in 20% (ν/ν) glycerol stock. In this project, the bacterium was resuscitated from the stock and was inoculated onto marine agar (pH 7.6). The agar plate culture was incubated at 35°C for 3–7 days before the bacterial colonies were formed.

Bacterial growth and β -glucosidase activity

A loopful of Meridianimaribacter sp. CL38 culture was inoculated into production medium (pH7.5), which consists of 5.0g/L peptone, 1.0g/L yeast extract, 5.0g/L MgCl₂, 2.0g/L MgSO₄·7H₂O, 0.5g/L CaCl₂, 1.0g/L KCl, and 20.0g/L NaCl [11]. After overnight incubation at 35°C with 180 rpm shaking, 5% (v/v) of the overnight culture was transferred into a fresh production medium supplemented with 1% (w/v) carboxymethyl cellulose (CMC). The culture was incubated at 35°C, 180 rpm. Samples were collected from the culture at 3hour intervals until the 24th hour and bacterial growth was measured by OD_{600nm}. Nevertheless, crude enzyme of each sampling point was harvested by centrifugation (4°C, 4500rpm) was subjected to β -glucosidase assay protocol.

β-glucosidase assay

The β -glucosidase activity was determined by *p*-nitrophenyl-β-D-glucopyranoside using (pNPG) as substrate. 5 mM pNPG was prepared by dissolving the substrate in 50 mM Tris-HCl buffer (pH 7.0). Then, 0.5 mL crude enzyme was mixed with 0.5 mL substrate. The enzyme reaction was started by incubating the mixture at 35°C. After 30 mins of incubation, the enzyme reaction was stopped by adding 1 mL of 1M NaOH. The β glucosidase activity was determined by releasing the *p*-nitrophenol (*p*NP) compound after the reaction and the absorbance changes were measured at OD_{405nm} using a spectrophotometer. One unit of enzyme activity was defined as the amount of β -glucosidase that can release one µmol of *p*NP per minute per mL under the assay condition. All enzyme reactions were conducted in at least triplicate.

JTLS | Journal of Tropical Life Science

Effects of salt concentration on β -glucosidase activity and stability

The β -glucosidase activity of CL38 was conducted by incubating the mixture of enzyme-substrate at different concentrations of salt (1-10% NaCl, *w*/*v*). The stability of β -glucosidase was carried out by preincubating the crude enzyme of CL38 with different concentrations of NaCl as mentioned, for 1 hour at 35°C. Then, the activity and stability of the enzyme were measured by standard enzyme assay. Crude enzyme without any salt added (1% NaCl, *w*/*v*) was treated as the assay control with its relative activity expressed as 100%. The experiment was conducted in at least triplicate.

Effects of metal ions, detergents and organic solvents on β -glucosidase stability

The crude enzyme of CL38 was preincubated for 1 hour at 35°C with 5mM of metal ion (Ca²⁺, Mg²⁺, Co²⁺, Ni²⁺, Cu²⁺, Mn²⁺), 1% detergent (sodium dodecyl sulfate in w/v, Triton X-100, Tween-20, Tween-40, Tween-60, Tween-80 in v/v) and 25% organic solvent (methanol, acetone, isopropanol, chloroform, ethanol in v/v). After preincubation, an enzyme assay was carried out. Crude enzyme without any additives added was treated as the assay control with its relative activity expressed as 100%. The experiment was conducted in at least triplicate.

Protein sequence analysis

Draft genome sequence of Meridianimaribacter sp. CL38 was retrieved from NCBI GenBank database with the accession number of QKWS00000000.1. Then, the genome sequence of strain CL38 was submitted to the dbCAN meta server to identify the Carbohydrate-active enzymes (CAZymes) [22, 23]. Sequences listed as glycoside hydrolase (GH) family that encode for β-glucosidase were selected for further analysis. The protein sequence was submitted to the InterPro v86.0 server to determine the functional domain [24]. β -glucosidase sequences from other species were selected from the NCBI database and used in multiple sequence alignment via Cluster Omega [25]. A phylogenetic tree was constructed using MEGA-X software with a neighbor-joining method, and was bootstrapped 1000 times [26, 27].

Results and Discussions

Bacterial growth and β -glucosidase activity

The bacterial growth of *Meridianimaribacter* sp. CL38 and β -glucosidase activity in production medium supplemented with carboxymethyl cellulose (CMC) were monitored (Figure 1). According to the growth curve, the lag phase was observed from 0 to the 3rd hour, and then the bacterium growth achieved an exponential phase from the 3rd to 12th hour and entered the stationary phase thereafter. As for enzyme activity, it was low during the lag and early of the exponential phase. The β -glucosidase activity reached its peak at 12 hours of bacterial growth (Figure 1).

Effects of salt concentration on β -glucosidase activity and stability.

The β -glucosidase activity and stability in the presence of salt were determined (Figure 2). The β-glucosidase demonstrated optimal activity at 2% NaCl (w/v). When the salt concentration was increased (from 2% to 10%), the enzyme activity decreased. The enzyme activity remains more than 50% at 10% NaCl (w/v) (64.19 ± 1.18%). Whereas the enzyme is most stable at 1-2% NaCl (w/v) and the stability decreased with the increase of salinity. The halotolerant ability of β-glucosidase produced by strain CL38 could be due to the adaptation of strain CL38 to its habitat. The mangrove which connected the sea and the land create a salty environment that require the bacteria to secrete halotolerant hydrolytic enzyme to degrade the carbon source into simple sugar [28].

Effects of metal ions, detergents, and solvents on β -glucosidase stability

Table 1 summarizes the results of the β -glucosidase stability when the crude enzymes were preincubated with different metal ions, detergents, and solvents. The β -glucosidase activity of strain CL38 was relatively stable in the presence of Ca²⁺, Mg²⁺, and Mn²⁺ ions as compared to other metal ions, which retains more than 80% of relative activities. Ca²⁺ and Mg²⁺ were commonly found in inorganic salts that use in the chemical pretreatment of lignocellulosic biomass [29, 30]. These metal salts have lower corrosivity as compared to acid solution and this makes it a more environmental friendly pretreatment method [31]. However, trace amount of metal ions could be found in

JTLS | Journal of Tropical Life Science



Figure 1. β-glucosidase activity and growth curve of *Meridianimaribacter* sp. CL38 cultured in production medium supplemented with 1% (*w/v*) CMC.



 $\Box \beta$ -glucosidase activity $\Box \beta$ -glucosidase stability

Figure 2. The effect of different salt concentration to the β -glucosidase activity and stability

the biomass after chemical pretreatment [32]. After the chemical pretreatment, the lignocellulosic biomass will go for enzymatic pretreatment. The β -glucosidase of strain CL38 that can withstand these trace metal ions will outperform other intolerance enzymes during the saccharification process of the biomass. Based on Table 1, the Cu²⁺ showed strong inhibition on the enzyme activity (24.47 ± 0.49%), this may be due to the autooxi dation catalyzed by the Cu²⁺ on the cysteine residues of the enzyme, and in turn causing disulfide interchange and dysfunctional protein structure [33]. The inhibition caused by Co²⁺ (54.01 \pm 0.85%) might be due to the formation of cobaltous-amino acid complexes, which are irreversible oxidation compounds that formed between Co²⁺ ions and amino acid residues [34].

The β -glucosidase of strain CL38 showed the

Chemicals	Concentration of chemicals	Relative activity (%)
Control	-	100.00 ± 0.49
Ca ²⁺	5mM	81.43 ± 0.36
Mg^{2+}	5mM	92.97 ± 0.28
Co ²⁺	5mM	54.01 ± 0.85
Ni ²⁺	5mM	73.28 ± 0.52
Cu^{2+}	5mM	24.47 ± 0.49
Mn^{2+}	5mM	99.86 ± 0.81
Sodium dodecyl sulfate (SDS)	1% (w/v)	13.20 ± 0.44
Triton X-100	1% (v/v)	27.45 ± 0.31
Tween-20	1% (v/v)	94.21 ± 0.88
Tween-40	1% (v/v)	36.20 ± 1.15
Tween-60	1% (v/v)	56.23 ± 0.67
Tween-80	1% (v/v)	97.48 ± 0.28
Methanol	25% (v/v)	3.19 ± 0.49
Acetone	25% (v/v)	0
Isopropanol	25% (v/v)	0
Chloroform	25% (v/v)	35.71 ± 0.16
Ethanol	25% (v/v)	0

Table 1. Effect of metal ions, detergents, and organic solvents on β -glucosidase activity.

Table 2. Selected β -glucosidase sequence from closely related species compared to Bgl3a

Species	Origin	Protein Length (aa)	Per- cent iden- tity (%)	Accession	Ref.
<i>Meridianimaribacter</i> sp. CL38 (Bgl3a)	Mangrove soil	761	100	WP_131505993.1	[11]
Meridianimaribacter flavus	Marine sediment	761	99.61	WP_134198609.1	[21]
<i>Flavobacteriaceae</i> bacterium 144Ye	Near-bottom water and surface sediments of Benham Bank	761	98.55	RYH75728.1	[40]
Winogradskyella vidalii	Seawater surface	759	74.54	WP_179353224.1	[41, 42]
Kordia zhangzhouensis	Freshwater surface	753	74.34	WP_046745035.1	[43]
Flavobacterium sediminis	Tidal flat sediment	760	69.80	WP_109570126.1	[44]
Olleya aquimaris	Seawater	766	68.69	WP_111659610.1	[45]
Mangrovimonas xylaniphaga	Estuarine mangrove sediment	776	48.85	WP_053978191.1	[46]
Winogradskyella wichelsiae	Seawater surface	780	47.07	WP_179374618.1	[41, 42]
Winogradskyella ludwigii	Seawater surface	780	47.01	WP_179337019.1	[41, 42]
Meridianimaribacter sp. CL38 (Bgl3b)	Mangrove soil	780	46.76	WP_131507911.1	[11]
Mesoflavibacter sabulilitoris	Seashore sand	780	46.25	WP_106680313.1	[47]
Winogradskyella litoriviva	Coastal seawater	780	46.06	WP_173299877.1	[48]
Winogradskyella echinorum	Sea urchin	780	45.66	WP_186844354.1	[49]

JTLS | Journal of Tropical Life Science

highest stability against Tween-80 followed by Tween-20, which retained more than 94% enzyme activity after an hour of preincubation with the detergents. Detergents like Tween-40, Tween-60, Triton X-100, and sodium dodecyl sulfate (SDS) inhibited the enzyme activity after preincubation, which the relative activities are less than 60% as compared to the control. SDS is an ionic surfactant that is more likely to disrupt and denature enzyme structure due to the interaction of SDS with positively charged amino acids [35]. In contrast, the non-ionic detergents (i.e. Triton and Tween family) are less likely to affect the enzyme activity because they will only bind to the enzyme with weak hydrophobic interactions and hydrogen bonds [36].

According to Table 1, all organic solvents tested were inhibitors to the β -glucosidase of strain CL38. The organic solvents act as chaotropic agents which denatured the enzyme by disrupting the hydrogen bonds between protein subunits [37]. Nonetheless, the β -glucosidase of *Olleya aquimaris* was also found to be inhibited by organic solvents [38]. More research on physical or chemical modification on the enzyme is needed if the non-tolerance β -glucosidase needs to be involved

in a process with high concentration of organic solvents.

Analysis of protein sequence related to β -glucosidase activity

Draft genome of strain CL38 was retrieved from GenBank database with accession number of QKWS000000.1 and the genome sequence were uploaded to dbCAN meta server for carbohydrateactive enzymes (CAZymes) identification. Based on the result, there were 157 CAZymes identified in CL38. Among the glycoside hydrolases (GHs), GH2, 3, 5, 9, 16, and 144 were the possible GH families that were detected in the dbCAN meta server which encoded for β -glucosidases. Since GH3 is the most common GH family encoded for β -glucosidases [39], two of the protein sequences identified as GH3 in the strain CL38 genome were selected for further study. The two sequences with the code of Bgl3a (TBV27841.1) and Bgl3b (TBV26231.1) were used thereafter. Closely related sequences with Bgl3a and Bgl3b were selected by NCBI BlastP and tabulated in Table 2.

Based on Table 2, most of the closely related species were halophile isolated from saline environments such as seawater, marine sediment, sea-





shore sand, etc. Hence the β -glucosidase isolated from these halophiles have high possibilities to be halotolerant [50]. Although Bgl3a and Bgl3b belong to the same species, there is only 46.76% identity between them (Table 2). β -glucosidase of *Meridianimaribacter flavus* (99.61%) showed the highest percent identity with Bgl3a followed by Flavobacteriaceae bacterium 144Ye (98.55%). Whereas for Bgl3b, the closest sequences are from *Mesoflavibacter sabulilitoris* and *Winogradskyella echinorum* with 97.44% and 90.26% identity respectively. To further analyze the relationship among Bgl3a, Bgl3b, and β -glucosidases of other species, a phylogenetic tree was constructed (Figure 3).

The phylogenetic tree (Figure 3a) showed that Bgl3a amino acid sequence has the closest relationship with the β-glucosidase of *Meridianimari*bacter flavus [21]. Moreover, the Bgl3a protein sequence also formed a big cluster with the β -glucosidases from Flavobacteriaceae bacterium 144Ye, Olleya aquimaris, Kordia zhanqzhouensis, Winogradskyella vidalii, and Flavobacteriaceae sediminis [21, 45, 51, 52]. Whereas for Bgl3b amino acid sequence, it has the closest relationship with β-glucosidase produced by *Mesoflavibacter sabu*lilitoris and was clustered with another group of GH3 β-glucosidase from bacteria such as Winogradskyella SDD. and *Mangrovimonas* xvlaniphaga [46, 47]. Interestingly, all the aforementioned bacteria belong to the same family taxon Flavobacteriaceae, indicating that the protein sequence of either Bgl3a or Bgl3b is most probably encoded by similar genes that found in the members of the family; these two GH3 proteins might be part of the core genome shared among the family members.

Nonetheless, regardless of Bgl3a or Bgl3b groups of β -glucosidases, these β -glucosidases have the same domain arrangements as shown in Figure 3b. All sequences consist of a glycoside hydrolase family 3 domain which involves in the enzyme hydrolysis function of the glycosidic bond formed between the two D-glucose monomers [53]. Besides, most of these β -glucosidases consist of a signal peptide sequence at the N-terminal, indicating that these enzymes are most probably expressed extracellularly [54]. A fibronectin type-III like domain was found in all sequences at the Cterminal. This domain is involved in ligand-binding function, protein-protein interaction, and could be a novel carbohydrate binding module-like domain [55–57].

As shown in the multiple sequence alignment (Supplementary 1), there are 5 motifs present among the sequences. These motifs could connect to form glycoside hydrolase family 3 domain and are used as the signature for GH3 family identification [58]. By referring to the well-studied β -glucosidase of *Chaetomium thermophilum* (GenBank Accession: ABR57325.2) [39], the active sites of Bgl3a and Bgl3b are found at Asp296 and Asp297, respectively. This active site is functioning as the nucleophile during the hydrolysis reaction and is highly conserved throughout different species. The segregating site indicated that mutation might occur among species to adapt to different habitats or conditions [59].

Conclusion

To the extent of our knowledge, this is the first β-glucosidase characterization of genus Meridian*imaribacter*. The β-glucosidases of *Meridiani*maribacter sp. CL38 remained stable in the presence of 1-2% NaCl (w/v) (96.28-100%), 5mM (99.86%), Mg²⁺ (92.97%), Ca²⁺ ions Mn^{2+} (81.43%), and 1% (v/v) Tween-20 (94.21%) and Tween-80 (97.48%). Two protein sequences (Bgl3a and Bgl3b) related to the β-glucosidase activity from Meridianimaribacter sp. CL38 were identified with closest identity to Meridianimaribacter flavus (99.61%) and Mesoflavibacter sabulilitoris (97.44%), respectively. A total of five motifs were found in Bgl3a and Bgl3b with the active site (nucleophile) found at Asp296 and Asp297, respectively. Taken together, the β -glucosidases of strain CL38 could serve as good candidates in the production of biofuel that involves Mg²⁺ and Ca²⁺ ions as inorganic metal salt in the lignocellulosic biomass seawater-based pretreatment process.

Acknowledgement

This work was supported by the Ministry of Higher Education Malaysia under Fundamental Research Grant Scheme (FRGS/1/2019/WAB13/UTM/02/1) as the main funding body. The authors would like to thank Universiti Teknologi Malaysia for providing facilities to carry out this research work. Clarine Wan Ling Hong acknowledges UTM Zamalah Scholarship from Universiti Teknologi Malaysia. Kheng

JTLS | Journal of Tropical Life Science

Loong Chong acknowledges UTMNexus Scholarship. The authors also would like to acknowledge Johor National Parks Corporation for sampling permit (CJB G No. 887005) at Tanjung Piai National Park, Johor.

References

- 1. Martins F, Felgueiras C, Smitkova M, Caetano N (2019) Analysis of fossil fuel energy consumption and environmental impacts in European countries. Energies. doi: 10.3390/en12060964.
- Bimanatya TE, Widodo T (2018) Fossil fuels consumption, carbon emissions, and economic growth in Indonesia. International Journal of Energy Economics and Policy 8 (4 SE-Articles): 90–97.
- 3. Höök M, Tang X (2013) Depletion of fossil fuels and anthropogenic climate change—A review. Energy Policy 52 797–809. doi: 10.1016/j.enpol.2012.10.046.
- Claassen PAM, van Lier JB, Lopez Contreras AM et al. (1999) Utilisation of biomass for the supply of energy carriers. Applied Microbiology and Biotechnology 52 (6): 741–755. doi: 10.1007/s002530051586.
- 5. Ahorsu R, Medina F, Constantí M (2018) Significance and challenges of biomass as a suitable feedstock for bioenergy and biochemical production: A review. Energies. doi: 10.3390/en11123366.
- 6. Keeney D, Muller M (2006) Water use by ethanol plants: potential challenges. Minneapolis, MN.
- Indira D, Jayabalan R (2020) Saccharification of lignocellulosic biomass using seawater and halotolerant cellulase with potential application in second-generation bioethanol production. Biomass Conversion and Biorefinery 10 (3): 639–650. doi: 10.1007/s13399-019-00468-4.
- Mainka T, Weirathmüller D, Herwig C, Pflügl S (2021) Potential applications of halophilic microorganisms for biological treatment of industrial process brines contaminated with aromatics. Journal of Industrial Microbiology and Biotechnology 48 (1–2): kuab015. doi: 10.1093/jimb/kuab015.
- Liew KJ, Ngooi CY, Shamsir MS et al. (2019) Heterologous expression, purification and biochemical characterization of a new endo-1,4-β-xylanase from *Rhodothermaceae bacterium* RA. Protein Expression and Purification 164 105464. doi: 10.1016/j.pep.2019.105464.
- Liew KJ, Teo SC, Shamsir MS et al. (2018) Complete genome sequence of *Rhodothermaceae bacterium* RA with cellulolytic and xylanolytic activities. 3 Biotech 8 (8): 376. doi: 10.1007/s13205-018-1391-z.
- Lam MQ, Oates NC, Thevarajoo S et al. (2020) Genomic analysis of a lignocellulose degrading strain from the underexplored genus *Meridianimaribacter*. Genomics 112 (1): 952–960. doi: 10.1016/j.ygeno.2019.06.011.
- Corral P, Amoozegar MA, Ventosa A (2020) Halophiles and their biomolecules: Recent advances and future applications in biomedicine. Marine Drugs. doi: 10.3390/md18010033.
- Thatoi H, Behera BC, Mishra RR, Dutta SK (2013) Biodiversity and biotechnological potential of microorganisms from mangrove ecosystems: A review. Annals of Microbiology 63 (1): 1–19. doi: 10.1007/s13213-012-0442-7.

- 14. Amoozegar MA, Safarpour A, Noghabi KA et al. (2019) Halophiles and their vast potential in biofuel production. Frontiers in Microbiology. doi: 10.3389/fmicb.2019.01895.
- 15. Arfi Y, Chevret D, Henrissat B et al. (2013) Characterization of salt-adapted secreted lignocellulolytic enzymes from the mangrove fungus *Pestalotiopsis sp.* Nature Communications 4 (1): 1810. doi: 10.1038/ncomms2850.
- Zoghlami A, Paës G (2019) Lignocellulosic biomass: Understanding recalcitrance and predicting hydrolysis. Frontiers in Chemistry 7 874. doi: 10.3389/fchem.2019.00874.
- 17. Putro JN, Soetaredjo FE, Lin S-Y et al. (2016) Pretreatment and conversion of lignocellulose biomass into valuable chemicals. RSC Adv 6 (52): 46834–46852. doi: 10.1039/C6RA09851G.
- Zhao X, Zhang L, Liu D (2012) Biomass recalcitrance. Part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. Biofuels, Bioproducts and Biorefining 6 (4): 465–482. doi: 10.1002/bbb.1331.
- Singh G, Verma AK, Kumar V (2015) Catalytic properties, functional attributes and industrial applications of bglucosidases. 3 Biotech 6: 1–14. doi: 10.1007/s13205-015-0328-z.
- 20. Wang M, Lu X (2016) Exploring the synergy between cellobiose dehydrogenase from *Phanerochaete chrysosporium* and cellulase from *Trichoderma reesei*. Frontiers in Microbiology. doi: 10.3389/fmicb.2016.00620.
- 21. Wang B, Sun F, Du Y et al. (2010) *Meridianimaribacter flavus* gen. nov., sp. nov., a member of the family Flavobacteriaceae isolated from marine sediment of the South China Sea. International Journal of Systematic and Evolutionary Microbiology 60 (1): 121–127. doi: 10.1099/ijs.0.009845-0.
- 22. Zhang H, Yohe T, Huang L et al. (2018) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Research 46 (W1): W95–W101. doi: 10.1093/nar/gky418.
- 23. Yin Y, Mao X, Yang J et al. (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Research 40 (W1): W445–W451. doi: 10.1093/nar/gks479.
- 24. Blum M, Chang H-Y, Chuguransky S et al. (2021) The InterPro protein families and domains database: 20 years on. Nucleic Acids Research 49 (D1): D344–D354. doi: 10.1093/nar/gkaa977.
- 25. Madeira F, Park Y mi, Lee J et al. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research 47 (W1): W636–W641. doi: 10.1093/nar/gkz268.
- 26. Kumar S, Stecher G, Li M et al. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35 (6): 1547–1549. doi: 10.1093/molbev/msy096.
- 27. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39 (4): 783– 791. doi: 10.2307/2408678.
- 28. Mah MH, Lam MQ, Tokiman L et al. (2022) Revealing the Potential of Xylanase from a New Halophilic *Microbulbifer* sp. CL37 with Paper De-Inking Ability. Arabian Journal for Science and Engineering 47 (6): 6795– 6805. doi: 10.1007/s13369-021-06400-1.

JTLS | Journal of Tropical Life Science

- 29. Li H, Xu J (2013) Optimization of microwave-assisted calcium chloride pretreatment of corn stover. Bioresource Technology 127: 112–118. doi: 10.1016/j.biortech.2012.09.114.
- 30. Loow Y-L, Wu TY, Tan KA et al. (2015) Recent advances in the application of inorganic salt pretreatment for transforming lignocellulosic biomass into reducing sugars. Journal of Agricultural and Food Chemistry 63 (38): 8349–8363. doi: 10.1021/acs.jafc.5b01813.
- Xu J, Xu J, Zhang S et al. (2018) Synergistic effects of metal salt and ionic liquid on the pretreatment of sugarcane bagasse for enhanced enzymatic hydrolysis. Bioresource Technology 249: 1058–1061. doi: 10.1016/j.biortech.2017.10.018.
- 32. Williams PT, Horne PA (1994) The role of metal salts in the pyrolysis of biomass. Renewable Energy 4 (1): 1–13. doi: 10.1016/0960-1481(94)90058-2.
- Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. Microbiology and Molecular Biology Reviews 65 (1): 1–43. doi: 10.1128/MMBR.65.1.1-43.2001.
- 34. Bodansky Oscar (1949) The influence of magnesium and cobalt on the inhibition of phosphatases of bone, intestine, and osteogenic sarcoma by amino acids. Journal of Biological Chemistry 179 (1): 81–102. doi: 10.1016/S0021-9258(18)56815-0.
- Otzen DE (2002) Protein unfolding in detergents: Effect of micelle structure, ionic strength, pH, and temperature. Biophysical Journal 83 (4): 2219–2230. doi: 10.1016/S0006-3495(02)73982-9.
- 36. Jelińska A, Zagożdżon A, Górecki M et al. (2017) Denaturation of proteins by surfactants studied by the Taylor dispersion analysis. PLOS ONE 12 (4): e0175838. doi: 10.1371/journal.pone.0175838.
- 37. Sinha RP, Khare SK (2014) Protective role of salt in catalysis and maintaining structure of halophilic proteins against denaturation. Frontiers in Microbiology. doi: 10.3389/fmicb.2014.00165
- 38. Hwang E-J, Lee Y-S, Choi Y-L (2018) Cloning, purification, and characterization of the organic solvent tolerant β-glucosidase, OaBGL84, from *Olleya aquimaris* DAU311. Applied Biological Chemistry 61 (3): 325–336. doi: 10.1007/s13765-018-0361-9.
- 39. Mohsin I, Poudel N, Li D-C, Papageorgiou AC (2019) Crystal structure of a GH3 β-glucosidase from the thermophilic fungus *Chaetomium thermophilum*. International Journal of Molecular Sciences. doi: 10.3390/ijms20235962
- 40. M. RS, R. RAR, D. MA et al. (2019) Draft genome sequences of six bacteria isolated from the Benham Bank, Philippine Rise, Philippines. Microbiology Resource Announcements 8 (29): e00777-19. doi: 10.1128/MRA.00777-19.
- 41. Alejandre-Colomo C, Harder J, Fuchs BM et al. (2020) High-throughput cultivation of heterotrophic bacteria during a spring phytoplankton bloom in the North Sea. Systematic and Applied Microbiology 43 (2): 126066. doi: 10.1016/j.syapm.2020.126066.
- 42. Alejandre-Colomo C, Viver T, Urdiain M et al. (2020) Taxonomic study of nine new Winogradskyella species occurring in the shallow waters of Helgoland Roads, North Sea. Proposal of *Winogradskyella schleiferi* sp. nov., *Winogradskyella costae* sp. nov., *Winogradskyella helgolandensis* sp. nov., *Winogradskyella* v. Systematic

and Applied Microbiology 43 (6): 126128. doi: 10.1016/j.syapm.2020.126128.

- 43. Du J, Liu Y, Lai Q et al. (2015) *Kordia zhangzhouensis* sp. nov., isolated from surface freshwater. International Journal of Systematic and Evolutionary Microbiology 65 (Pt_10): 3379–3383. doi: 10.1099/ijsem.0.000424.
- 44. Bae SS, Kim MR, Jung Y et al. (2018) *Flavobacterium sediminis* sp. nov., a starch-degrading bacterium isolated from tidal flat sediment. International Journal of Systematic and Evolutionary Microbiology 68 (12): 3886–3891. doi: 10.1099/ijsem.0.003081.
- 45. Lee S-Y, Park S, Oh T-K, Yoon J-H (2010) Description of *Olleya aquimaris* sp. nov., isolated from seawater, and emended description of the genus *Olleya* Mancuso Nichols et al. 2005. International Journal of Systematic and Evolutionary Microbiology 60 (4): 887–891. doi: 10.1099/ijs.0.014563-0.
- 46. Dinesh B, Furusawa G, Amirul AA (2017) *Mangrovimo-nas xylaniphaga* sp. nov. isolated from estuarine mangrove sediment of Matang Mangrove Forest, Malaysia. Archives of Microbiology 199 (1): 63–67. doi: 10.1007/s00203-016-1275-8.
- 47. Park S, Park J-M, Jung Y-T et al. (2014) *Mesoflavibacter* sabulilitoris sp. nov., isolated from seashore sand. International Journal of Systematic and Evolutionary Microbiology 64 (Pt_11): 3743–3748. doi: 10.1099/ijs.0.067223-0.
- 48. Nedashkovskaya OI, Kukhlevskiy AD, Zhukova N V et al. (2015) Winogradskyella litoriviva sp. nov., isolated from coastal seawater. International Journal of Systematic and Evolutionary Microbiology 65 (Pt_10): 3652– 3657. doi: 10.1099/ijsem.0.000470.
- 49. Nedashkovskaya OI, Vancanneyt M, Kim SB, Zhukova N V (2009) *Winogradskyella echinorum* sp. nov., a marine bacterium of the family Flavobacteriaceae isolated from the sea urchin *Strongylocentrotus intermedius*. International Journal of Systematic and Evolutionary Microbiology 59 (6): 1465–1468. doi: 10.1099/ijs.0.005421-0.
- 50. Kumar S, Karan R, Kapoor S et al. (2012) Screening and isolation of halophilic bacteria producing industrially important enzymes. Brazilian journal of microbiology: [publication of the Brazilian Society for Microbiology] 43 (4): 1595–1603. doi: 10.1590/S1517-838220120004000044.
- 51. Qi F, Huang Z, Lai Q et al. (2016) *Kordiaulvae* sp. nov., a bacterium isolated from the surface of green marine algae *Ulva* sp. International Journal of Systematic and Evolutionary Microbiology 66 (7): 2623–2628. doi: 10.1099/ijsem.0.001098.
- 52. Lee JH, Kang JW, Shin SB, Seong CN (2017) *Winograd-skyella flava* sp. nov., isolated from the brown alga, *Sar-gassum fulvellum*. International Journal of Systematic and Evolutionary Microbiology 67 (9): 3540–3546. doi: 10.1099/ijsem.0.002161.
- 53. Geronimo I, Ntarima P, Piens K et al. (2019) Kinetic and molecular dynamics study of inhibition and transglycosylation in *Hypocrea jecorina* family 3 β-glucosidases. Journal of Biological Chemistry 294 (9): 3169–3180. doi: 10.1074/jbc.RA118.007027.
- 54. E. DR, Peng W, Maarten van DJ (2012) Membrane proteases in the bacterial protein secretion and quality control pathway. Microbiology and Molecular Biology Reviews 76 (2): 311–330. doi: 10.1128/MMBR.05019-11.

JTLS | Journal of Tropical Life Science

- 55. Sidar A, Albuquerque ED, Voshol GP et al. (2020) Carbohydrate binding modules: Diversity of domain architecture in amylases and cellulases from filamentous microorganisms. Frontiers in Bioengineering and Biotechnology. doi: 10.3389/fbioe.2020.00871
- 56. Koide A, Bailey CW, Huang X, Koide S (1998) The fibronectin type III domain as a scaffold for novel binding proteins. Journal of Molecular Biology 284 (4): 1141–1151. doi: 10.1006/jmbi.1998.2238.
- 57. Hansen CK (1992) Fibronectin type III-like sequences

and a new domain type in prokaryotic depolymerases with insoluble substrates. FEBS Letters 305 (2): 91–96. doi: 10.1016/0014-5793(92)80871-D.

- 58. Mackenzie CO, Grigoryan G (2017) Protein structural motifs in prediction and design. Current Opinion in Structural Biology 44 161–167. doi: 10.1016/j.sbi.2017.03.012.
- 59. Fu YX (1995) Statistical properties of segregating sites. Theoretical Population Biology 48 (2): 172–197. doi: 10.1006/tpbi.1995.1025.

Supplementary 1

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter_sp._CL38_GH3b Flavobacterium_sediminis Olleva autimate Olleya_aquimaris Kordia_zhangzhouensis Winogradskyella_vidalii WinogradsKyella_vidaili Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GH3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus Mangrovimonas_xylaniphaga

Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter_sp._CL38_GH3b Flavobacterium_sediminis Olleya aquimaris Kordia_zhangzhouensis Winogradskyella_vidalii Windgradskyella_vidalii Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GH3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter_sp._CL38_GH3b Flavobacterium_sediminis Olleya aquimaris Kordia_zhangzhouensis Winogradskyella_vidalii Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GH3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter_sp._CL38_GH3b Flavobacterium_sediminis Olleya_aquimaris Kordia_zhangzhouensis Winogradskyella_vidalii Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GH3a Meridianimaribacter_flavus Chaetonium_thermophilum Cochliobolus_heterostrophus

Mangrovimonas xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter sp._CL38_GH3b Flavobacterium_sediminis Olleya_aquimaris Kordia_tangrhomensis Kordia_zhangzhouensis Winogradskyella_vidalii Flavobacteriaceae bacterium 144Ye Meridianimaribacter sp._CL38_GH3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Mesoflavibacter_sp._CL38_GH3b Flavobacterium_sediminis Olleya_aquimaris Kordia_zhangzhouensis Winogradskyella_vidalii Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GH3a Meridianimaribacter_flavus Chaetomium thermophilum Cochliobolus_heterostrophus

MKT-LSTOLKTIAVGVFVVSLLGCNOOESKTEESFSFS	37
MKKLQFAKIKYFTLIIGVLTVLSCNTNDKKTTASK	35
MTTLKLTNIKSLAIGFGLLTAISCNTNSKKTTSTK	35
MTTLKLTNIKSLAIGFGLLTAVSCNTNSNKIKSKNSKN	35
KKRTIILTLLIMGSYAHAQKKKKMKKMK	25
KTSKTSKTSKTSKTSKTSKTS	35
EVMEVMEVM	21
MFLKNKL-FALLIISVLGHINQNVYAQSKPKP	29
HKTIKITR-VCAVAVILLCLVNTKTHAQMEE	30
MKTIKITR-VCAVAVILLCLVNTKTHAQMEEEE	30
MKTIKITR-VCAVAVILLCLVNTKTHAQMEE	30
MKLAVGLACLVGSHLAATGSLKSKAKLEKKAWATSEPSTPSPWMNPNAA	49
MSILPLSLSGLAAGAAIQQRDPVPAGIVAPPIIPAPH-GGWVS	45
QENPFQSKVDSVLGLMTLEEKLGQLNLPSSGDITTGQAQSSNIAQKIEQGQVGGL	92
SDNPFEEKVDSIISLMTLDEKLGQLNLPATGPITTGSSKSTDVVSKIEEGKIGGL	90
NDNPYQTKVDSIIGLMTLDEKIGQLNLPATGPITTGASKSTDVVKKIEDGKIGGL	90
TONPIQARVDSTIGLMTLDERIGQLNEPATGPITTGSSRSTDVVRRIEDGRIGGE	90
WTLS TEDKYDALLAKMTLEEKT GOMNOYNGEWDYTGDYDKDGDOA - KKYDHLEKGYVGSM	94
KETSVDORVDSLLAKMTLEEKVGOMNLVNGEWDVTGPTDOGGSAK-LKVEHLKKGVGAM	80
SELNTEOK IDALLALMTVEEKVGOLNLYNGFWNVTGPAPEHGGAA-TKYKHLKNGWVGAM	88
TKVD INOK IES LLSOMT I EEKVGOMNLYNG FWDVTGPAPEEGNAA-LKYDHLKKGWVGAM	89
TKVD INQKVESLLSQMTI EEKVGQMNLYNGFWDVTGPAPEEGNAA-LKYDHLKKGWVGAM	89
TKVDINQKVESLLSQMTIEEKVGQMNLYNGFWDVTGPAPEEGNAA-LKYDHLKKGWVGAM	89
GWEQAYQWAKDFVSQMTLLEKVNLTTGVGWEGEQ	83
DWAESYRKASLLVSNMTLAEKTNLTAGTGIFMGR	77
MOTIT 1 Motif 2	140
ENTROTHNIVEVORTAVENSRUATEDITORDVIHGIETTFFIPUGLSASWDMDMI	14/
FNIKDPKNIYDVOKIAVEKSRLGIPLIFGEDVI HOFKTFFDIDIGIGLSASWDMEMI	145
FNIKDPKNIYEVOKIAVEKSRLGIPLIFGMDVIHGFKTTFPIPLGLSASWDMEMT	145
LTVRGVKEVKAVOKIAVEETRLGIPLIIGEDVIHGYKTLSPIPLAEAASWDMEAI	139
LNVRGVKDVKAVOKIAVEETRLGIPLIIGFDVIHGYETLTPIPLAESASWDLEAI	149
LNIRGVENVKEVQRIAVEETRLGIPLLIGFDVIHGYKTLSPIPLAESASWDLEAI	135
LNVRGTEEVNKVQRIAVEESRLGIPLLIGFDVIHGYKTLSPIPLAESASWDLKAI	143
LTVRGAEEVRNVQRIAVEESRLGIPLLIGFDVIHGYKTLSPIPLAESASWDLQAI	144
LTVRGAEEVRNVQRIAVEESRLGIPLLIGFDVIHGYKTLSPIPLAESASWDLQAI	144
LTVRGAEEVRKVQRIAVEESRLGIPLLIGFDVIHGYKTLSPIPLAESASWDLQAI	144
CVGQTGAIPRLGLRSLCLHDSPLGIRGTDYNSAFPSGQTTAATWDRTLF	132
CVGNTGSALRVGIPQLCLQDGPLGVRNTDHNTAFPAGITAGATWDKDLI	126
· · · · · · · · · · · · · · · · · · ·	
Motif 3	
THE REPORT OF A DESCRIPTION OF A DESCRIP	000
KKTAQMAAQEATADGINWTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG	206
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPILASEIAKAMVAGYQ EKTAQIAASEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPIGSEIAKAMVNGYQQ EXTAQIA NEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPIGSEIAKAMVNGYQQ	206
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSEIAKAMVNGYQQ EKTAQIAANEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINWTFSPMVD-ISRDPRWGRVSEGNGENDEFLGSAIARAMVRGYQQ	206 204 204
KKTAQMAAQEATADGINNTFSPMVD-ISKDPKWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISKDPKWGRVSEGNGEDPFLGSALARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISKDPRWGRVSEGNGEDPFLGSALARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISKDPKWGRVSEGNGEDPFLGSALARAMVRGYQQ KKSAQVAAGEASASGINNFFSPMVD-VSNDAWGRVWEGAGENDFLGSAVAFA	206 204 204 204
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSEIAKAMVNGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAAGEASAGINNTFGPMVD-VSRDARWGRVMEGAGEDAFLGSKVAEARVKGFQG KNSAEMAAKESSAAGINNTFGPMVD-ISRDARWGRVMEGAGEDAFLGSKIAVARVNGFQG	206 204 204 204 196 208
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFGPMVD-VSRDARWGRVMEGAGEDAFLGSKVAEARVKGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDTYLGSKIAVARVNGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDTYLGSKIAVARVNGFQG	206 204 204 204 196 208
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAANEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSEIAKAMVNGYQQ EKTAQIAANEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAAGEASASGINWTFAPMVD-VSRDARWGRVMEGAGEDAFLGSKVAAEARVKGFQG RNSARVAATEASAAGINWTFAPMVD-ISRDPRWGRVMEGAGEDPYLGSKIAKARVLGFQG RNSARVAATEASAAGINWTFAPMVD-ISRDPRWGRVMEGAGEDPYLGSKIAKARVLGFQG EASARIAAIEASSAGINWTFAPMVD-VSRDARWGRVMEGAGEDPYLGSLIAIRAVNGFQG	206 204 204 204 196 208 194 202
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPFLGSXIAKARVLGFQG RNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSXIAKARVLGFQG EASARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAIARVNGFQQ ESSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAIARVNGFQG ESSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAIARVNGFQG	206 204 204 204 204 208 208 208 208 202 203
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVNGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVNGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVNGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDAFLGSKVAEARVKGFQG KNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSKIAKARVLGFQG EASARIAATEASAAGINNTFAPMVD-VSRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-VSRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG	206 204 204 198 208 194 202 203 203
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAACEASASGINNTFGPMVD-VSRDARWGRVMEGAGEDPYLGSKIAVARVNGFQG RNSARMAARESSAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPYLGSKIAVARVNGFQG EASARIAATEASAAGINNTFAPMVD-VSRDARWGRVMEGAGEDPYLGSKIAVARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG	206 204 204 198 208 194 202 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPTLGSRVAERAKVGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLAIARAVUNGFQG EASARIAAIESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG	206 204 204 198 208 194 202 203 203 203 203
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVMEGAGEDAFLGSKVAEARVKGFQG KNSAEMAAEESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDAFLGSKIAVARVNGFQG EASARIAAEESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EASARIAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG YRRGYAMGEARGKAINVLGPVAGLGRMPAGGRNMEGFSPDPVLGSLIAKARVNGFQG	206 204 204 196 208 194 202 203 203 203 203 192 186
KKTAQUAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIARAMVNGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVMEGAGEDPTIGSKIAKARVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDTIGSKIAVARVNGFQG RNSARWAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSKIAVARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVIGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIALARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIALARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVIGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG YRGYAMGKEARGKGINVILGEPVGPLGRKPAGGRNMEGFS DPVIJGIAMATEVKGIQD YRRAVAIGEFRGKCVNIHGPSVELGRKPRCGRNMEGFS GSDPVIJASISVEGIQS	206 204 204 196 208 194 202 203 203 203 203 192 186
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVNGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVNGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVNGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVMEGAGEDFILGSAIARAMVNGYQQ KNSAQNAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSKIAVARVNGFQG RNSARNAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSLIAKARVLGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSLIAKARVNGFQG SKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDFYLGSLIAKARVNGFQG XKRQYAMGKEARGGINVLLGFVGPLGRKPAGGRNWEGFSPDFVLGJLAKARVNGFQG 	206 204 204 204 204 208 208 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ KKSAQVAACEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPTLGSKIAKARVLGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSKIAKARVLGFQG EASARIAAIESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNWEGFSDPVLGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNWEGFSDPVLGIASLIAKARVNGFQG . * *:* * * * * * * * * * * * * * * * *	206 204 204 204 208 194 202 203 203 203 203 203 203 203 203 203
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDRAWGRVMEGAGEDPTIGSIAARAMVRGYQQ EKTAQIAANEASANGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAIARANVRGYQQ EKSARVAATEASAAGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAIARANVGFQG ENSARVAATEASAAGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDRAWGRVMEGAGEDPTIGSIAKARVNGFQG MKGYAMGREARGKGINVLIGPVAGJLGRNAGGRNMEGFGSDPVLJAFGSISVEGIQS **********************************	206 204 204 204 196 208 194 202 203 203 203 203 203 203 203 203 203
KKTAQHAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPYLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPFLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDAFLGSKVAEARVKGFQG KNSAEMAAESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EASARIAAIEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVLGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPFLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GARMAGRVMEGGSDPVLGSLIAKARVNGFQG DYRGYAMGEARGKGINUTLGPVGPLGRKMPAGGRNMEGFSDPVLGAFGGSLSVEGIQS *.*:***** Motif 4 DDLTADNTMACVKHFALYGAPEAGRDYNTVDMSRIRMYNEYFPYKAA DDMAANNTLMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYLPYKAA	206 204 204 196 208 194 202 203 203 203 192 186 255 253 253 253
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPTIGSAIARAMVRGYQG KNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVIGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVIGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG YRRGYAMGKEARGKGINVLIGPVAGPLGRMPAGGRNWEGFS DDPULTGIAMAETVKGIQD YRRAVAIGEFRGKCVNIHLGPSVGPLGRKPRGGNNEGFSSDPVLAGSLSVEGIQS 	206 204 204 204 208 194 203 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPTLGSRIARAMVRGYQQ KNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSRIAKARVLGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EASARIAAIEASSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPYLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GALGRNARGRNMEGFSDPVLGSLIAKARVNGFQG SKGARVAATEASASGINNTFAPMVD-GALGRNARGRNMEGFSDPVLGSLIAKARVNGFQG DYRGYAAGEERGKGINVLLGPVGGLGRKPRGGRNMEGFSDPVLGSLIAKARVNGFQG DYRGYANGEERGKGVNTHLGPSVGPLGRKPRGGRNMEGFSDPVLGSLIAKARVNGFQG DDMAANNTIMGCVKHFALYGAPEAGRDYNTVDMSHIRMYNEYFPYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYMEYLAPYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYMEYLAPYKAA DDMAANNTIMSCVKHFALYGPEAGRDYNTVDMSRIRMYMEYLAPYKAA ETKEDLGLPNTIACAKHFAAYGPEAGRDYNTVDMSRIRMYNEYLAPYKAA	206 204 204 204 208 194 202 203 203 203 203 203 203 203 203 255 253 253 253 253 255 253
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARAMVRGYQQ KKSAQVAAGEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVRGYQQ EKTAQIAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVLGFQG EASARIAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GADARWGRVMEGAGEDPTIGSIAKARVNGFQG DYRGYAMGREARGKGINVLIGPVAGJIGRNPGGGSNPGFGSDPVLQAFGGSLSVEGIQS * .** * * * * * * * * * * * * *	206 204 204 208 194 202 203 203 203 203 203 203 203 203 203
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVMEGAGEDPTIGSAIARAMVRGYQQ EKSAQVAAGEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSSIAKARVLGFQG ENSARWAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSSIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSSIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GARMPAGGNMEGFSDPVILGSIAKARVNGFQG DYRGYAMGEARGKGINULGGVAGELGRMPAGGNNMEGFSDPVILGIAMAETYKGIOD YRRAVAIGEFRGKGVNIHLGPSVGPLGRKPRGGRNMEGFSDPVILQFFGGSLSVEGIQS *.*:**.**** Motif 4 DDLTADNTMMCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYIPYKAA DDDMANNTLMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYIPYKAA DDDMANNTLMSCVKHFALYGFPAGRDYNTVDMSRIRMYNEYIPYKAA DDLSAHNSTAACAKHFAAYGFPAGRDYNTVDSRIRMYNEYLPYKAA DDLSAHNTLACAKHFAAYGFPAGRDYNTVDGSRIRMYNEYLPYKAA DDLSAHNTLACAKHFAAYGFPAGRDYNTVDGSRIRMYNEYLPYKAA DDLSAHHSIAACAKHFAAYGFPAGRDYNTVDGSRIRMYNEYLPYKAA	206 204 204 204 203 203 203 203 203 203 203 203 203 203
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVMEGAGEDPTLGSAIARAMVRGYQQ KNSAEMAAKESSAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTLGSKIAKARVLGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTLGSLIAKARVNGFQG KNSARVAATEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTLGSLIAKARVNGFQG KSARVAATEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTLGSLIAKARVNGFQG YRRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNWEGFS DDPVLTGIAMAETVKGIQD YRRAVAIGEFRGKGVNIHLGPSVGPLGRKPRGGRNWEGFS DDPVLTGIAMAETVKGIQD YRRAVNIGEFRGKGVNIHLGPVG-GARPRGGRNWEGFS DDPVLTGIAMAETVKGIQD DUSANNTMMGCVKHFALYGAPEAGRDYNTVDMSHIRMYNEYF PYKAA DDLTANNTMMGCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYF PYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDMAANNTIMSCVKHFALYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHNSIAACAKHFAAYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHNTIAACAKHFAAYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHNTIAACAKHFAAYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHNTIAACAKHFAAYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHNTIAACAKHFAAYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHNTIAACAKHFAAYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHNTIAACAKHFAAYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHPFIAFAYGFPEAGRDYNTVDMSRIRMYNEYLAPYKAA DDLSAHFFIAGYGFPEAGRDYNTADGTSTLNNIILPFFKAA DDLSAHFFAAY	206 204 204 204 203 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPTIGSAIARAMVRGYQQ KNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVLGFQG EASARIAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG CKSARVAATEASASGINNTFAPMVD-GAPDARWGRVMEGAGEDPTIGSLIAKARVNGFQG DYRGYAMGREARGKGINVILGPVAGLIGRNAGGRNMEGFSDPVIJAFGGSLSVEGIDD YRRAVAIGEEFRGKCVNIHLGPVGPLIGRKPRGGRNMEGFSDPVIJAFGGSLSVEGIDD YRRAVAIGEEFRGKCVNIHLGPVGFLGRKPRGGRNMEGFSDPVIJAFGGSLSVEGIDD ODDMANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYIPPYKAA DDMANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYILAPYKAA DDMANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYILAPYKAA DDMANNTIMSCVKHFALYGFAEAGRDYNTDMSRIRMYNEYILAPYKAA DDLTANNTIAACAKHFAAYGFAEAGRDYNTDMSRIRMYNEYILAPYKAA DDLTANNTIAACAKHFAAYGFAEAGRDYNTDMSRIRMYNEYILAPYKAA DDLTANNTIAACAKHFAAYGFAEAGRDYNTDMSRIRMYNEYILAPYKAA DDLTANNTIAACAKHFAAYGFAEAGRDYNTDMSRIRMYNEYILAPYKAA DDLTANNTIACCHFFAGYGFAEAGRDYNTDMSRIRMYNEYILAPYKAA DDLTANNTIACCHFFAAYGFAEAGRDYNTDMGSRIRMYNEYILAPYKAA DDLTANNTIACCHFFAAYGFAEAGRDYNTDMGTSTLYNTUPPFKAA DDLTANNTIACCHFFAAYGFAEAGRDYNTDMGTYTNUFFLYNTUPFKAA DDLSANNTIACCHFFAAYGFAEAGRDYNVADIGNTLYNTUPPFKAA	2066 2044 2044 2044 2022 2033 2033 2033 2033
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARAMVRGYQQ EKTAQIAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVLGFQG EASARIAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASASGINNTFAPMVD-ISRDPRWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GADARWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GADARWGRVMEGAGEDPTIGSIAKARVNGFQG DYRGYAMGREARGKGINVLIGPVAGEIGRNAGGRNNEGFSDPVUJGAFGGSLSVEGIQS * .** * * * * * * * * * * * * <u>Motif 4</u> DDMANNTIMSCVKHFALYGAPEAGRDYNTVDMSKIRMFNEYIPYFXAA DDMANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYIPYFXAA DDMANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYIPYFXAA DDMANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRMYNEYIAPYKAA ETKEDGJENTIAACAKHFAAYGFAEAGRDYNTVDMSRIRMYNEYIAPYKAA DDLSAHHSIAACAKHFAAYGFAEAGREYNTDVGIGTSTINNIILPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGREYNTDVGIGTSTINNIILPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGREYNTDVGIGTSTINNIILPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGREYNVDAUGRGTINNVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGREYNVDAUGRGTINVULPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGREYNVDAUGRGTINVULPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGREYNVDAUGNGGTINVULPPFKAA	206 204 204 204 202 203 203 203 203 203 203 203 203 203
<pre>KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDARWGRVMEGAGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVIGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVIGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG YRRGYAMGKEARGKGINVLLGPVAGPIGRMPAGGRNWEGFSPDPULTGIAMAETVKGIQD YRRAVAIGEFRGKGVNIHLGPVGCHGRKPRGGRNMEGFSDPVLAGLIAKARVNGFQG </pre>	2066 2044 2044 2044 2044 2044 2044 2044
<pre>KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVWEGAGEDPTIGSAIARAMVRGYQQ KNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVWEGAGEDPTIGSXIAKARVLGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNMEGFSDPVLTGIAMAETVKGIQD YRRAVNIGEFFGKKVNIHLGPSVGPLGRKPRGGRNMEGFSDPVLTGIAMAETVKGIQD YRRAVNIGEFFGKKVNIHLGPSVGPLGRKPRGGRNMEGFSDPVLTGIAMAETVKGIQD YRRANNTIGSVKHFALYGAPEAGRDYNTVDMSHIRMYNEYFPYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSRVRMFFEYFPYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSRVRMFFEYLPYKAA DDMAANNTIMSCVKHFALYGFAEGRDYNTVDMSRVRMFFEYLPYKAA DDLSAINNTMACKHFFAYGFAEGRDYNTVDMSRVRMFFEYLPYKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNTVDMSRVRMFFEYLPYKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNTVDMSRVRMFFEYLPFFKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNTVDIGTSTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNVDIGTSTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNVDIGTSTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNVDIGTSTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNVDIGTSTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNVDIGTYLYNVDIFFKAA DDLSAINNTIAACAKHFAAYGFAEGRDYNVDIGTYLYNVDFFKAA A</pre>	2000 2004 2004 2004 2005 2005 2005 2005
<pre>KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPTLGSXIARAKVGFQG KNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSXIAKARVLGFQG EASARIAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASSAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GADARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GADARWGRVMEGAGEDPTLGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GALGRARAGGNNEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGLGRKPRGGRNMEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGLGRKPRGGRNMEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGLGRKPRGGRNMEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGLGRKPRGGRNMEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGLGRKPRGGRNMEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGLGRKPRGGRNMEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGLGRKPRGGRNMEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGLGRKPRGGRNMEGFSDPVLQAFGGSLSVEGIDD YRRAVAIGEERGKCVNIHLGPVGGLGRKPGGRNMEGFSTLYNTVLSPYKAA DDLTANNTIMSCVKHFALYGAPEAGRDYNTVDMSRIRNYNEYLAPYKAA DDLMANNTIMSCVKHFALYGREAGRDYNTDMSRIRNYNEYLAPYKAA DDLMANNTIMSCVKHFALYGREAGRDYNTDMSRIRNYNEYLAPYKAA DDLSANNTIAACAKHFAAYGREAGRDYNVADIGNGTLYNTVLPFKAA DDLSANNTIAACAKHFAAYGREAGRDYNVADIGNGTLYNTVLPFKAA DDLSANNTIAACAKHFAAYGREAGRDYNVADIGNGTLYNTVLPFKAA DDLSANNTIAACAKHFAAYGREAGRDYNVADIGNGTLYNTVLPFKAA DDLSANNTIAACAKHFAAYGREAGRDYNVADIGNGTLYNTVLPFKAA DDLSANNTIAACAKHFAAYGREAGRDYNVADIGNGTLYNTVLPFKAA DDLSANNTIAACAKHFAAYGREAGRDYNVADIGNGTLY</pre>	2066 2044 2044 2044 2022 2033 2033 2033 2033
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQWAAGEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSRVARARVKGFQG KNSAEWAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSRVARARVKGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-GADARWGRVMEGAGEDPTIGSLIAKARVNGFQG MKGYAMGKEARGKGINVLIGPVAGLIGRVAGGRNMEGFGSDPVLQAFGGSLSVEGIQS 	2066 2044 2044 2044 2082 2033 2033 2033 2033 2033 2033 2033
KKTAQMAAQEATADGINNTFSPMVD-ISKDPKNGKVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISKDPRNGKVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISKDPRNGKVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISKDPRNGKVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISKDARNGKVMEGAGEDPTIGSAIARAMVRGYQQ KKSAQVAACGASASGINNTFAPMVD-ISKDARNGRVMEGAGEDPTIGSAIARAMVRGYQG ENSARNAATEASAAGINNTFAPMVD-ISKDARNGRVMEGAGEDPTIGSKIAKARVIGFQG RNSARVAATEASAAGINNTFAPMVD-ISKDARNGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISKDARNGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISKDARNGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISKDARNGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISKDARNGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISKDARNGRVMEGAGEDPTIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRCGRNMEGFSDPULTGIAMAETVKGIQD YRRAVAIGEFFAGKCVNIHLGPSVGLGRKPRCGRNMEGFSDPULTGIAMAETVKGIQD YRRAVNIGEFFASSAGINNTLAGPVG-GARPEAGGRNYEGFSDPULTGIAMAETVKGIQD YRRAVNIGEFFAGKVNIHLGPSVGLGRKPRCGRNMEGFSDPULTGIAMAETVKGIQD CDDAANNTMACVKIFFALYGAPEAGRDYNTVDMSRIKMYNEYLPYKAA DDDKANNTMACVKIFFALYGAPEAGRDYNTVDMSRIKMYNEYLPYKAA DDLSANNTMACVKIFFALYGAPEAGRDYNTVDMSRIKMYNEYLPYKAA DDLSANNTMACVKIFFALYGAPEAGRDYNTVDMSRIKMYNEYLPYKAA DDLSANNTMACVKIFFALYGAPEAGRDYNTVDMSRIKMYNEYLPYKAA DDLSANNTMACXKIFFAAYGAPEAGRDYNTVDMSRIKMYNEYLPYKAA DDLSANNTIACAKHFAAYGFAEGRDYNVDDIGNGTLYNTVLPFFKAA DDLSANNTIACAKHFAAYGFAEGRDYNVDDIGNGTLYNTVLPFFKAA DDLSANNTIACAKHFAAYGFAEGRDYNVDDIGNGTLYNTVLPFFKAA DDLSANNTIACAKHFAAYGFAEGRDYNVDDIGNGTLYNTVLPFFKAA DDLSANNTIACAKHFAAYGFAEGRDYNVDDIGNGTLYNTVLPFFKAA DDLSANNTIACAKHFAAYGFAEGRDYNVDDIGNGTLYNTVLPFFKAA DDLSANNTIACAKHFAAYGFAEGRDYNVDDIGNGTLYNTVLPFFKAA A	2066 2044 2044 2046 2088 1944 2023 2033 2033 2033 2033 2033 2033 20
KKTAQMAAQEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPIGSAIARAMVRGYQQ KNSACMAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPIGSAIARAMVRGYQQ EKSAQVAAGEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPIGSXIAKARVLGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPIGSIIARAKNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPIGSIIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPIGSIIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPIGSIIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPIGSIIAKARVNGFQG YRGYAMGKEARGKGINVLLGPVAGPIGRMPAGGRNWEGFSPDPVLTGIAMAETVKGIQD YRRAVAIGEFRGKGVNIHLGPSVGPIGRKPRGGRNMEGFSSDPVLAGSIAKARVNGFQG 	206 204 204 204 205 203 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASSIAKAMVAGYQG EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASSIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASSIARAMVRGYQQ KKSAQVAAGEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASSIAKARVLGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASSIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASSIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASSIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASSIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASSIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASSIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIASIAKARVNGFQG D'KRGYAMGREARGKGINVLLGPVAGLARMAGGNMEGFSDPVIJAFGGSLSVEGIOS 	206 204 204 204 203 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTLASEIAKAMVAGYQG EKTAQIAASEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINWTFSPMVD-ISRDPRWGRVSEGNGEDPTLGSAIARAMVRGYQQ EKTAQIAANEASANGINWTFSPMVD-ISRDARWGRVMEGAGEDTYLGSKIAKARVLGFQG KNSACMAAKESSAAGINWTFAPMVD-ISRDARWGRVMEGAGEDTYLGSKIAKARVLGFQG ENSARVAATEASAAGINWTFAPMVD-ISRDARWGRVMEGAGEDTYLGSLIAKARVNGFQG EKSARVAATEASAAGINWTFAPMVD-ISRDARWGRVMEGAGEDTYLGSLIAKARVNGFQG EKSARVAATEASAAGINWTFAPMVD-ISRDARWGRVMEGAGEDTYLGSLIAKARVNGFQG EKSARVAATEASAAGINWTFAPMVD-ISRDARWGRVMEGAGEDTYLGSLIAKARVNGFQG EKSARVAATEASAAGINWTFAPMVD-ISRDARWGRVMEGAGEDTYLGSLIAKARVNGFQG EKSARVAATEASASGINWTFAPMVD-ISRDARWGRVMEGAGEDTYLGSLIAKARVNGFQG EKSARVAATEASASGINWTFAPMVD-ISRDARWGRVMEGAGEDTYLGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRMPAGGRNWEGAGEDTYLGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRMPAGGRNWEGFGSDPULTGIAMAETVKGIOD YRRAVAIGEEFRGKCVIHLGPSVGPLGRNPGGRNWEGFGSDPULTGIAMAETVKGIOD YRRAVAIGEFRGKCVIHLGPSVGPLGRNPGGRNWEGFGSDPULTGIAMAETVKGIOD YRRAVAIGEFRGKVNHFALYGAPEAGRDYNTVDMSHIRMYNEYIPYKAA DDMAANNTMMCVKHFALYGAPEAGRDYNTVDMSHIRMYNEYIPYKAA DDMAANNTMMCVKHFALYGAPEAGRDYNTVDMSHIRMYNEYIPYKAA DDLSANNTMMCVKHFALYGFAEAGRDYNTVDMSHIRMYNEYILAPYKAA EDLSANNTMACXKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA NDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYGFAEAGRDYNVDDIGNGTIYNVLPPFKAA DDLSANNTIACAKHFAAYG	206 204 204 204 208 194 202 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVLGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNMEGFSDDFULGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNMEGFSDDFULGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNMEGFSDDFULGSLIAKARVNGFQG DDLTADNTMMACVKHFALYGAPEAGRDYNTVDMSHIRMYNEYFPYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSRVRMFFEYFPYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSRVRMFFEYFPYKAA DDLSANNTMMACVKHFALYGFAEGRDYNTVDMSRVRMFFEYLPYKAA DDLSANNTIMSCVKHFALYGFAEGRDYNTVDMSRVRMFFEYLPYKAA DDLSANNTIMACXKHFAAYGFAEGRDYNTVDMSRVRMFFEYLPYKAA DDLSANNTIACACHFFAAYGFAEGRDYNTVDIGTSTLYNTUPFFEAA EDLSANNTIACACHFFAAYGFAEGRDYNVADIGNGTLYNTVLPFFKAA DDLSANNTIACACHFFAAYGFAEGRDYNVADIGNGTLYNTVLPFFKAA DDLSANNTIACACHFFAAYGFAEGRDYNVADIGNGTLYNTVLPFFKAA AGFAEGRDYNVADIGNGTLYNTVLPFFKAA A	206 204 204 204 202 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPTIGSXIARAKVGFQG KNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVLGFQG EASARIAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSILAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSILAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSILAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSILAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSILAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSILAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSILAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GALGRARAGGNMEGFSDPVIJGILAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GALGRARAGGNMEGFSDPVIJGILAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GALGRARAGGNMEGFSDPVIJGILAKARVNGFQG 	2066 2044 2044 2044 208 203 203 203 203 203 203 203 203 203 203
<pre>kkthQMAQEArDADGINNTFSPMVD-ISRDPRMGRVSEGNGEDPTIASEIAKAMVAGYQG ekthQIAANEASANGINWTFSPMVD-ISRDPRMGRVSEGNGEDPTIGSAIARAMVRGYQQ ekthQIAANEASANGINWTFSPMVD-ISRDPRMGRVSEGNGEDPTIGSAIARAMVRGYQQ ekthQIAANEASANGINWTFSPMVD-ISRDPRMGRVMEGAGEDPTIGSAIARAMVRGYQQ kksAQVAAGEASANGINWTFAPMVD-ISRDPRMGRVMEGAGEDPTIGSAIARAMVRGYQQ kksAQVAAGEASANGINWTFAPMVD-ISRDPRMGRVMEGAGEDPTIGSXIAKARVLGFQG exsARVAATEASAAGINWTFAPMVD-ISRDPRMGRVMEGAGEDPTIGSLIAKARVLGFQG eksARVAATEASAAGINWTFAPMVD-ISRDPRMGRVMEGAGEDPTIGSLIAKARVLGFQG eksARVAATEASAAGINWTFAPMVD-ISRDPRMGRVMEGAGEDPTIGSLIAKARVNGFQG eksARVAATEASAAGINWTFAPMVD-ISRDARMGRVMEGAGEDPTIGSLIAKARVNGFQG eksARVAATEASAAGINWTFAPMVD-ISRDARMGRVMEGAGEDPTIGSLIAKARVNGFQG eksARVAATEASAAGINWTFAPMVD-ISRDARMGRVMEGAGEDPTIGSLIAKARVNGFQG eksARVAATEASAAGINWTFAPMVD-ISRDARMGRVMEGAGEDPTIGSLIAKARVNGFQG eksARVAATEASASGINWTFAPMVD-ISRDARMGRVMEGAGEDPTIGSLIAKARVNGFQG eksARVAATEASASGINWTFAPMVD-GAPAGRNMEGFSDPVLJGIAKARVNGFQG exsARVAATEASASGINWTFAPMVD-GAPAGAGNNEGFSDPVLJAFGGSLSVEGIQS </pre>	206 204 204 204 208 196 202 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVLGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNWEGFSDPVLIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRNPAGGRNYTVDNSKIRMYNEFFEYPYKAA DDLTADNTMMACVKHFALYGAPEAGRDYNTVDNSKIRMYNEFFEYPYKAA DDLTADNTMMACVKHFALYGAPEAGRDYNTVDNSKIRMYNEFYLPYKAA DDLSANNTIMSCVKHFALYGAPEAGRDYNTVDNSKIRMYNEFYLPYKAA DDLSANNTIMACXKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPYFKAA DDLSANNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPFKAA DDLSANNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPFKAA DDLSANNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPFKAA AGVIACAKHLIGDEQHFRQVGDGFDIDESLSSNIDDRTMHELYLWFFADA T	206 204 204 204 208 196 202 203 203 203 203 203 203 203 203 203
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFGPMVD-ISRDPRWGRVBEGAGEDPTIGSXIARARVAGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVLGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNMEGFSDDPULGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNMEGFSDDPULGSLIAKARVNGFQG YKRGYAMGKEARGKGINVLLGPVAGPLGRKPRGGRNMEGFSDDPULGSLIAKARVNGFQG DDLTANNTMACVKHFALYGAPEAGRDYNTVDMSRVRMFEYLPYKAA DDMAANNTLMSCVKHFALYGAPEAGRDYNTVDMSRVRMFEYLPYKAA DDMAANNTLMSCVKHFALYGAPEAGRDYNTVDMSRVRMFEYLPYKAA DDMAANNTLMSCVKHFALYGFAEAGRDYNTVDMSRVRMFEYLPYKAA DDLTANNTMACVKHFALYGFAEAGRDYNTVDMSRVRMFEYLPYKAA DDLTANNTLACAKHFAAYGFAEAGRDYNVDDIGNGTLYNVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVDDIGNGTLYNVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVDDIGNGTLYNVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVDDIGNGTLYNVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVDDIGNGTLYNVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVDDIGNGTLYNVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVDDIGNGTLYNVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVDDIGNGTLYNVLPPFKAA C	2066 2044 2044 2044 2022 2033 2032 2033 2033
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSIAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDARWGRVMEGAGEDPTIGSIAIARAMVRGYQQ KKSAQVAAGEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAIARAMVRGYQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAIARANVGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-GAPAGAGNNEGFSDPVIJAFIAKARVNGFQG UKRGYAMGREARGKGINVLIGPVAGLIGRNAGGRNMEGFSDPVIJAFIAKARVNGFQG D'KRGYAMGREARGKGINVLIGPVAGLIGRNAGGRNMEGFSDPVIJAFGGSLSVEGIQS **********************************	2066 2044 2044 2044 2022 2033 2033 2033 2033
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFSPMVD-ISRDPRWGRVBEGAGEDPTIGSXIAKARVIGFQG RNSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVIGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSXIAKARVIGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDPTIGSLIAKARVNGFQG YKRGYAMGKEARGKGINVILGPVGPIGRWPAGGRNWEGFSDPULTGIAMAETVKGIOD YRRAVAIGESPKGKVNIHLGPSVGLGRKPRGGRNWEGFSDPULTGIAMAETVKGIOD YRRAVNIGEPFKKVNIHGPSVGLGRKPRGGRNWEGFSDPULTGIAMAETVKGIOD YRRAVNIGEPFKKVNIHGPSVGLGRKPRGGRNWEGFSDPULTGIAMAETVKGIOD YRRAVNIGEPFKKVNIHGPSVGLGRKPRGGRNWEGFSDPULTGIAMAETVKGIOD YRRAVNIGEPFKKVNIHGPSVGLGRKPRGGRNWEGFSDPULTGIAMAETVKGIOD YRRAVNIGEPFKKVNIHGPSVGLGRKPRKGRNNEGFSDPULTGIAMAETVKGIOD YRRAVNIGEPFKKVNIHGPSVGLGRKPRKGRNWEGFSDPULTGIAMAETVKGIOD YRRAVNIGEPFKKVNIFFALYGAPEAGRDYNTVDMSHIRMYNEYIAPYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSHIRMYNEYIAPYKAA DDMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSHIRMYNEYIAPYKAA DDLSAHNSIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPPFKAA ADLSAHNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPPFKAA DDLSAHNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPPFKAA CGVIATIKHLIADLLRDDWGFDGFVYDVGSYLMASINELIAGMGD-E VDEGVGSVMASFNELDGVPATGNKKLLTDLLEDDWGFDGFVYDVGSTLYMAHGMGD-E VDEGVGSVMASFNELDGVPATGNFLQRDIKGW	2066 2044 2044 2044 2022 2033 2033 2033 2033
KKTAQMAAGEATADGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIASEIAKAMVAGYQG EKTAQIAASEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ EKTAQIAANEASANGINNTFSPMVD-ISRDPRWGRVSEGNGEDPTIGSAIARAMVRGYQQ KKSAQVAAGEASASGINNTFAPMVD-ISRDARWGRVWEGAGEDTIGSVAEAKVKGFQG KNSAEMAAKESSAAGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSVAEAKVKGFQG ENSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSJIATARVNGFQG EKSARVAATEASAAGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSJIATARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSJIATAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSJIATAKARVNGFQG EKSARVAATEASASGINNTFAPMVD-ISRDARWGRVMEGAGEDTIGSJIATAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRKPRGGRNMEGFSDDFULGSIATAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRKPRGGRNMEGFSDDFULGSIATAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRKPRGGRNMEGFSDDFULGSIATAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRKPRGGRNMEGFSDDFULGSIATAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRKPRGGRNMEGFSDDFULGSIATAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRKPRGGRNMEGFSDDFULGSIATAKARVNGFQG YKRGYAMGKEARGKGINVLIGPVAGPLGRKPRGGRNMEGFSDDFULAGSIAFKAGUD YRRANTIGSVKHFALYGAPEAGRDYNTVDMSRIRMYNEYI JAPYKAA DDLTADNTMMACVKHFALYGAPEAGRDYNTVDMSRVRNFFEILPYKAA DDLMAANNTIMSCVKHFALYGAPEAGRDYNTVDMSRVRNFFEILPYKAA DDLMAANNTIMSCVKHFALYGFAEAGRDYNTADIGNGTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPFFKAA DDLSAINNTIAACAKHFAAYGFAEAGRDYNVADIGNGTLYNTVLPFKAA CGVIACAKHLIGDEQEHFRQVGGGDGIDESLSSNIDDRTHELYLWFFADA T	206 204 204 204 202 203 203 203 203 203 203 203 203 203

continued

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter_sp._CL38_GH3b Flavobacterium_sediminis Olleya_aquimaris Kordia_zhangzhouensis Winogradskyella_vidalii Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GH3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter_sp._CL38_GH3b Flavobacterium_sediminis Olleya_aquimaris Kordia_zhangzhouensis Winogradskyella_vidalii Flavobacteriaceae bacterium 144Ve Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GH3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter_sp._CL38_GH3b Flavobacterium_sediminis Olleya_aquimaris Kordia_bharahouranis Kordia_zhangzhouensis Korota_znangznouensis Winogradskyella_vidalii Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GB3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochlicbolus_heterostrophus

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibacter_sabulilitoris Meridianimaribacter_sp._CL38_GB3b Flavobacterium_sediminis Olleya_aquimaris Kordia_zhangzhouensis Winogradskyella_vidalli Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GB3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Minogradskydria_riobilvid Mesoflavibacter_sabuliitoris Meridianimaribacter_sp._CL38_GB3b Flavobacterium_sediminis Olleya_aquimaris Kordia_zhangzhouensis Notura_nanyzhouenst Winogradskyella_vidalii Flavobacteriaceae_bacterium_144Ye Meridianimaribacter_sp._CL38_GH3a Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus

Mangrovimonas_xylaniphaga Winogradskyella_litoriviva Mesoflavibactor_sabulilitoris Meridlanimaribacter_sp__CL38_GH3b Flavobacterium_sediminis Flavobacteriaceaebacteria Kordia_zhangzhouensis Winogradskyella_vidalii Flavobacteriaceaebacterium_144Ye Meridianimaribacter_flavus Chaetomium_thermophilum Cochliobolus_heterostrophus

Characterization of β -glucosidases from Meridianimaribacter s	sp. C
QQVSALALKAGIDMDMVGEGLSGTLKKSLDEGKVSMEDIDTAVKRILTA	363
VAVTSLALKAGLDMDMAGDSPNISAAFIKNLKTALDNGDISINDVDKAVARVLTA	367
YDVTSLALKAGLDMDMAGDSPN1KASFTHALKGALDQGR1SMED IDTAVARVLTA	367
IDVTSLALKAGLDMDMAGDSPNIKASTTHALKGALDQGKISVEDIDTAVAKVLTA	307
TEASLAND EAGADMDMESHLIVIELAALWWEGKVDERLIDDAVRAILRV	366
ROAD TIDDERGEDUCES IL	350
KURATIARINGSDIDMEGHVTIUELVURVURVUESTIDDAVIRTIKV	352
	300
RQAAQIAANAGSDMDMEGIVIIQELAQUVADGVVAESILDDAVRAILAV	261
	301
- POUN CAUS CIDEMUNES IV	247
COMPONE ACT DECEMBER DATE OF A SCHOOL STRUCTURE OF DECIMAL ACTION AND AND AND AND AND AND AND AND AND AN	241
- 3073112632632632632632703122713318222713728372773728777277877172181	341
	401
KERLOLFEDPIRIGDLERPRRDVFINENRAFARRVSAE	405
RIQVGLPEDPIRICNEEKARNDIIIQENKUIRKUVGRE	405
KYOLGI KUDD	405
KEELGLEDDD	307
KEELCIEDDD	404
KEELOLEDDF	300
KYELCI KOND	308
KIEDOLPAR	300
	300
KEELCI EEDE	300
FERVICE OF THE STREET OF COTHER FOR THE STREET OF THE STRE	406
FERVINIQUALDER SEMILUTIOFINGARES - RQTINEAVDVADDSALTALAAA YFQMGQDQNY PRPNFDTNTQNAEGPLYPGALISPSGVVNEFVDVQGNHAEVAREVARD	399
an u	
SMVLLKNDNNLLPLKKSG-TIALIGPLANTAVNMAGTWSVATKQDKSNPVLEGLKTVVGD	460
STVLLKNENQTLPLKKSG-TIAVIGPLAKASNNMAGTWSVSTDHENCISVWDGLQQTVGD	464
SSVLLKNDNQLLPLKKSG-TIAVIGPLAKANSNMAGTWSVSTDHQSSISVWDGLQQTVGN	464
SSVLLKNDNQLLPLKKSG-TIAVIGPLAKANSNMAGTWSVSTDHKSSISVWDGLQQTVGN	464
SIVLLKNDKQLLPLPKTGKRIALIGALAADKTSPLGSWRIASDDETAVSVLEGMQQYTGN	457
SIVLLKNENNLLPLKKEGQKIALIGALAADKTSPLGSWRIAAKDNTAVSVLEGLNQYTGN	464
SIVLLKNQGNLLPLQKEGQKIALIGPLAADKNSPLGSWRIASDDHTAVSVLEGMQQYTGN	450
SIVLLKNEGDLLPLKKKGQRIALIGDLANDKSSPLGSWRLASDDETAVSVLEGMQKYKGN	458
SIVLLKNDNHLLPLKKNGQTIALIGDLADDKSSPLGSWRIASDDDTAVSVLEGMSFYKGN	459
SIVLLKNDNNLLPLKKNGQTIALIGDLADDKSSPLGSWRIASDDDTAVSVLEGMSSYKGN	459
SIVLLKNDNNLLPLKKNGQTIALIGDLADDKSSPLGSWRIASDDDTAVSVLEGMSSYKGN	459
GTVLLKN-EGALPLDKPK-FLAVIGEDAGPNHNGPNSCDDRGCVGGTLAMGWGSGT	460
AITLLKNQDNILPLAASA-PLKIFGTDAEKNPDGINSCADQGCNKGTLGMGWGSGS	454
OANT LYAKGSNUDUDUEKET PMEGKET DEDNKSDKELLNEALAVAAKADVVVAA TGES	520
DUNILLYAKGSNUDYDFELEKRATMEGKTIDEDNETDOOMIDEAVAVAKKSDULLATIGES	524
DVNLLVAKGSNVDYDLDLEKRATMEGKDIPROGRTDOOMLNEALATAKKSDVIVATIGES	524
DANLIYAKGSNUDYDIDIEKRATMEGKDIERDGRTDOOMIDEALATAKKSDUIVATIGES	524
RLVYEKGADUTIGSVSFIOEV-OINTTOKSSFEAAKKTAOESDIVWWIGEH	508
TUT	515
TLTHHKGVELVVCNTAFTEEV-OINTTNRTGIDEAVAAAOTVDVVVMVLGEH	501
TINFENGDEVEKCTSTFLEEL_DINTTOYSGEEENVETAKNADUUWUUGEU	500
TINFSRGDEVFEGTSTFLOEV-KINTTOTTGFDEATATAKTADVVVVULGEH	510
TINFSRGFEVFEGTSTFLOEV-KINTIDIIGFDEATATAKIADVVVNUGED	510
TINFSRGFLYEEGISTELQEV-RININDISGEDEAIRIARADYYWWUSEN	510
ANEDVI UNDER I OR OF TWO CERY EVENTION SUPERIOR DE ALE TRANSPORT	510
ANT FILVTPDAALQAQAIKDGSKIESVLSNHAMETIKKVVSQDNVTAVVFVNANS	513
-ARIPIMDSPIDGERARGANIQFFNIDNFPGNSNPSPNDIAIVEVIADS	302
AFISGFSSSUTNIOTDOAODDIIBAIIOTGEDUUMUIETGEDIAT-UFF-NEKU	572
SELSGESSSRTD IGIPOVOKDLLNALLKTGKDVVLVLFTGRDLTL-VFP-SENV	576
AEFSGESSSRTDLGIPOVOKDLLOALLKTGKDVVLVLETGRDLTL-VEP-CET	576
	576
CFOTCEXPERTENCE DEVOCET FEVELUENDENT UNIVER TORENT ALL DER ANUT	560
GPQIGERISTIELSBEGVQQELLEEVIKVNEWIVLVIVINOKELEB-FWA-RRII	567
GFOSGEGRARIDIGHEGYQQEHERYIKYNANIYIYINAKAFHII-IWA-DANI	553
GPOSCECROSPENTELDGLOOPLIOEVINEVINEVINEVINERIAL ENA TENT	561
GEQSGEGRSETNIDI.DCI.OCFLIEFVYEVNDNUVI.VI.NNCRDIAT-FWA-SEUT	562
GF-=OS-=G-EGESETNIDLPGLOOFLLEEVYKVNTNVVIVINCEDIAT_EWA-SENT	562
GEOSGEGESETNIDLEGLOOFLLEEVYKUNDNUUTULINNCBDLAT_EMA-SERI	5602
CECVITODONCODONNITUNCODZI IZNUN CHOONTUUT DEVODUT I TODVODDNI	572
GENY TEVED NDCDERS ANT NI WINCOPLITEN VASION OF TIVE TO VERY ULTUWINE POUR	560
	302
PATLNVWFPGSEAGLATSDVLFGDVNDSGKLTATEDMNVCOVDTEVNUKNTCODTEN	630
PSILNVWFPGSEAGLSISDVLFGAVNPSGKLTATFPRNVGOVPLFTNNKATGAPLONG	634
PAILNVWFPGSEAGLSISDVLFGDVNPSGKLTATFPRNVGOVPLFYNHKNTGRDLGND	634
PAILNVWFPGSEAGLSISDVLFGDVNPSGKLTATFPRNVGQVPLFVNHKNTCPDLCND	634
PAIVEAWOLGTETGNAVAOVLYGDYNPSGKLPMSFDRNVGOCDIVYNYYATGRDIDK-	617
PAIVEAWOLGTOSGHALAOVLYGDYNPSGKLPMTFDRNVCOVDIVVNVKNPCDDVLDC	625
DAIVEGWOLGTOSGNAIAOVI.YGDYNDSGKI.DMTEDDNUGOTDIVVNVKemcBBUDM_	610
PAIVEAWOLGTETGNAVAOVLYGDYNPSGKLPMTFPRNVGOMPTYVNVKNTCPDTDK_	61.8
PAIVEAWHLGTESGNAIANVLYGDYNPSGKLDMSFDRNVGOADIYYNNYDTGRDTDK-	61 9
PAIVEAWHLGTESGNAIANVLYGDYNPSGKLPMSFPRNVGOADTVVNNVDTGBDTDK-	61 9
PAIVEAWHIGTESGNAIANVLYGDYNPSGKLPMSFPRNVGGADIYYNNYDPGPDPDY_	
	61 9
TAILWAGLPGOESGNAITDVLYGKVNPAGRSDFTWGATREGYGADVLVDDDDARVDOONE	619
TAILWAGLPGQESGNAITDVLYGKVNPAGRSPFTWGATREGYGADVLYDPDDARVPQQNF KAIVFAHLPGOEAGNSLMOVLYGDVSPSGHLPYTLPNAEDDFGNSVKLVGYOLGOPODTF	619 633 622
TAILMAGLPGQESGNAITDVIJGKVNPAGESPETWGATEEGYGADVLJDPDDARVPQQNF KAIVFAHLPGQEAGNSLMQVLYGDVSPSGHLPYTLPNAEDDFGNSVKLVGYQLGQPQDTF :*: *::*::::::::::::::::::::::::::::::	619 633 622

continued

Mangrovimonas_xylaniphaga	EGKFEKFRSNYLDVRNEPLYPFGYGLSYTKFEYGNFNISTTSLSM	675
Winogradskyella_litoriviva	EGHFEKFKSNYIDARNEPLYPFGYGLSYTTFKYGDVKLDKTSMSQ	679
Mesoflavibacter_sabulilitoris	EGHFEKFKTNYLDVRNEPLYPFGYGLSYTTFEYGELKLDKTSMSQ	679
Meridianimaribacter_spCL38_GH3b	EGHFEKFKTNYLDVRNEPLYPFGYGLSYTTFEYSELKLDKTSMSQ	679
Flavobacterium_sediminis	DNNVFWTHYSDVEKTPLWSFGHGLSYTTFEYSEITLDKESYKK	660
Olleya_aquimaris	EDVVFWSHFQDEKNDPLYVFGHGLSYTSFDYSNLKVDNT-FSN	667
Kordia_zhangzhouensis	NDNVFWSHYTDVDNTPLYPFGHGLSYTTFQYSNFKVHGTEFKI	653
Winogradskyella_vidalii	DNNVFYSHYSDVDKTPLFVFGHGLSYTTFSYSNLSVSNT-YKT	660
Flavobacteriaceae_bacterium_144Ye	DGNVFWSHYTDVEKTPLYPFGHGLSYTTFEYSNLKVTNT-YKE	661
Meridianimaribacter_spCL38_GH3a	DGNVFWSHYTDVEKTPLYPFGHGLSYTTFEYSNLKVTNT-YKE	661
Meridianimaribacter_flavus	DGNVFWSHYTDVEKTPLYPFGHGLSYTTFEYSNLKVTNT-YKE	661
Chaetomium thermophilum	TEGVFIDYRYFDKENTRVIYEFGHGLSYTTFEYRNLQIQKENVSAYI-PTTGLTEPAPTF	692
Cochliobolus_heterostrophus	TEGLYIDYRHFHRANITPRIAFGHGLSYTTFSFSDATITPVTPLTRTPPARPARGTTPSY	682
		101103
Mangrovimonas_xylaniphaga		675
Winogradskyella_litoriviva	***************************************	679
Mesoflavibacter_sabulilitoris	***************************************	679
Meridianimaribacter_spCL38_GH3b	***************************************	679
Flavobacterium_sediminis		660
Olleya_aquimaris		667
Kordia_zhangzhouensis		653
winogradskyella_vidalil		060
Manidianimanihaata arta arta		061
Meridianimaribacter_spCL38_GH3a		661
Meridianimaribacter_flavus		661
Chaetomium_thermophilum	GEHSTNYSDYLYPEGFHRANRYIYPYVNSTDLELASGDPYYGQTADQFLPPNATSSDPQP	752
Cochliobolus_heterostrophus	STALPPASEATWPANFNRIWKILISWLDKNDADAAAQVGTSPSTIPIPAGISNEQKP	/39
Mangrovimonas_xylanipnaga	DGTLDVSVDVSNGGDIDGREVVQLIIRDLVG-SVTRPVRELKG	717
Winogradskyeila_litoriviva	KDMVTVSVD1TNTGD1DGKEVAQLY1RDV1G-S1TRPVKELKG	721
Mesofiavibacter_sabuilitoris		721
Meridianimaribacter_spcL38_GH3D	CEDUCA CUMUNICAL DESCRIPTION AUTOMAL	721
Ollows amimonia		702
Ulleya_aquimaris	NNTVTVSVDDTNSGKVRGREVVQLFIRDLIA-SVTRPVRELRG	709
Kordia_znangznouensis	GDTIKVTVDVTNTGDIDGREVVQLIIRDLFG-SITKPVRELKG	700
Winogradskyella_vidalli	NORWAY CONTRACT AND A CONTRACTACT AND A CONTRACT AND A CONTRACT AND A CONTRACT AND A CONTRACT AN	702
Maridianianniheatan an CT20 CH2	NOTWUS PER LANGEL AND A LANGEL AND A LANGER AND A LANG	703
Meridianimaribacter_spcb36_GH3a	NOPUKUS PRI KNIGALKOKE VAQLI IKDI VG-SVIRPLKELKG	703
Chastomium thermorbilum	I I DE EGENEDGENDOT VDVI VTUTAD TTNTGAL FEDEVUOT VUEL GEDDD DEVOT DD	809
Cochligholus heterostrophus	GPAAGGAOGONDAT.FDVAVDTAVTVTNTGNESGKAVAOLVLOFDSFSTVDTDTTOLED	797
cochrisbords_heterostrophus	- GERROINQUOMERILEDVAIDIRVIVIMIONKISKAVAQUILUTEISSIVDIEINQUKS ; ;,* * *, *,***; * *;.	131
Mangnowimonas wulaninhaga		773
Winogradskyalla literiviya	FORTE-LARGETATIVEFIT SUPPLY FUNCTION TO F-VALID TO TO TO TO THE SUPPLY FOR THE SUPPLY FUNCTION TO THE SUPPLY FUNCTION FOR THE SUPPLY FOR THE	777
Moroflavibator sabulilitoris	FURTE - LENGETRIVSFILSVEDLEFINARID-FVAEFGDFUVATGINSNVELNEIFE	777
Meridianimaribacter en CL38 CU2b	FORVE-TERCETUTUTERTSVEDIRETNSNID-FVARCEDOURTCOMODURT ENGER	777
Flavebacterium sodiminis	EXATE - INCOMPTONE TO FRANCE PARTICIPAL CONTRACTOR FOR THE STREET OF THE	759
Ollows aggimanic	FEIVE_IEBCETTERIOITITERELOFINGESE FIVEDORETVICCCCCETTION	765
Vordia zbangzhouonrie	PERVE-LEPOEIRRIGHIDIRDEDOFFRAQUE-FIVEDODFEVIICOSSRIILQURFI	751
Winogradskyella widalii	FELVE-IFACESKEITETET TEKEICEVNIKOF-FVUESCOFFIFUCCSSFFUT FACTO	759
Flavebactoriaceae bacterium 144Ve	PERVERSESSALITIKE PERSESSENT PROPERTY TO THE PERSON AND THE PERSON	750
Moridianimaribactor en CT38 GW3a	FELLE-LEASASKIVEFVLIKKELSFIINGSE-F-VVESSDERVFVGGRSNAELEQDEI	750
Meridianimaribacter_spcbs0_003a	FELTE-LEAGASKIVEFULTEKELGEVENOGE-E-VVESGDEKVEVGGDSNAELEQGE	750
Chaotamium thormophilum	FADIL-TERGRANT VEFY STREED FINGE FVVESODERVEVOUD AND AND AND AND AND AND AND AND AND AN	866
Cochlipholus beterostrophus	FARMA-IRFORT TRESCHIERROLS INDVIDOUVERNING AVERDISERIUM DORID	857
cochiliobolius_necelos clophus	* : * : . :: :* : : : : : : : *	057
Mangrovinonas vulaninhaga	776	
Winogradskyells literiyiya	LTD 780	
Moreflavibactor eshulilitaria	TTD 780	
Meridianimaribactor en (130 cush	110 780	
Flavobactarium andiminia	18 760	
Olleve aggineric	760	
Olieya_aquimaris	1 /00 TP	
Kordia_zhangzhouensis	LR 753	
WinogradsKyella_vidalli	159 18 761	
Manidianimanihaata an 0139 003-	18 761 I.P 761	
Meridianimaribacter_spCL38_GH3a	161 IF761	
Chaetomium thermorbilum	New 967	
Cochlichelus heteroctrochus	A THE COCTO A STY 870	
cocurrent undereros crophus	HIGHOVODVARTA 010	

Figure 4. Multiple sequence alignment of Bgl3a and Bgl3b with β-glucosidase of other species. β-glucosidases of *Cochliobolus heterostrophus* (Genbank Accession: AAB82946.1) and *Chaetomium thermophilum* (ABR57325.2) were used for motifs and active site identification, respectively. Region of motifs were drawn with black lines, and the active site was shown in red box. Asterisks indicates fully conserved amino acids; colon indicates strongly similar properties between group and full stop indicate weakly similar properties between groups.

This page is intentionally left blank.