Homology modeling and molecular dynamics study on *Schwanniomyces occidentalis* alpha-amylase

Yahya Sefidbakht¹,²*⃝, Omid Ranaei Siadat¹,² and Fatemeh Taheri²

¹Protein Research Center, Shahid Beheshti University, G.C., Velenjak, Tehran, Iran; ²Biological Engineering Laboratories (BEL), Faculty of New Technologies and Energy Engineering, Biotechnology Department, Shahid Beheshti University, Velenjak, Tehran, Iran

Communicated by Ramaswamy H. Sarma

(Received 28 September 2015; accept 12 February 2016)

With consumers growing increasingly aware of environmental issues, industries find enzymes as a reasonable alternative over physical conditions and chemical catalysts. Amylases are important hydrolase enzymes, which have been widely used in variety of industrial process such as pharmaceutical, food, and fermentation industries. Among amylases α-Amylase is in maximum demand due to its wide range of applications. The homology modeling study on *Schwanniomyces occidentalis* amylase (AMY1, UniProt identifier number: P19269) was performed by Modeller using *Aspergillus oryzae* (6TAA) as the template. The resulting structure was analyzed for validity and subjected to 14 ns of molecular dynamics (MD) simulation trough GROMACS. The validity of obtained model may represent that utilized OPLS force field is suitable for calcium-containing enzymes. DSSP secondary structure and contact map analysis represent the conservation of domain A TIM barrel feature together with calcium ion coordination sphere. Investigating the covariance matrix followed by principle component analyses for the first five eigenvectors of both trajectories indicate a little more flexibility for AMY1 structure. The electrostatic calculation for the final structures shows similar isoelectric point and superimposed buffering zone in the 5–8 pH range.

**Keywords:** homology modeling; alpha-amylase; *Schwanniomyces occidentalis*; molecular dynamics

Enzymes are proteins capable of transforming their specific substrate into the product in shorter time and milder thermal conditions compared with chemical methods (Tawfik, 2014). Due to the fact that micro-organisms possess rapid growth, it is possible to clone and scale up the expression of enzymes for bioprocess. Alpha-amylases (EC 3.2.1.1), found in bacterial, fungal, plant, and animal sources, are responsible for catalyzing the hydrolysis of α-1,4 glycosidic linkage of glycogen, starch, linear amylase, and branched amylopectin to generate smaller sugars. Alternatively known as 1, 4-α-D-glucan glucanohydrolase or glycogenase, they are generally calcium metalloenzymes (Kandra, 2003). This endoamylase plays an important role in starch processing industry and has been proved as a green substrate for hydrochloric or oxalic acid and high temperature conditions (pH 2 and 150°C) needed in traditional processing units (Vieille & Zeikus, 2001). With the advent of various novel strategies enzymes are now involved in the pharmaceutical, clinical diagnostics, food, and detergent industries (Feitkenhauer, 2003; Zeman & McCrea, 1985). Alpha-amylases play an outstanding role in all the above-mentioned applications, meanwhile considered as a valuable model enzyme for studying biophysical properties and thermal adaptation in proteins (Fitter, 2005; Prakash & Jaiswal, 2010).

Alpha-amylase is made up of three domains; domain A forms the (β/α) 8 or TIM barrel motif with the catalytic acidic triads of Asp 217, Glu 241, and Asp 308. The protruding calcium-binding domain B (about 70 residues including Asn 132, Glu 173, Asp 186, Asp 217, His 221, and Glu 241) is located between the third beta strand and helix and the C-terminal Greek key domain (Da Lage, Feller, & Janecek, 2004; MacGregor, 1988). Although, protein structures related to several alpha-amylases are currently available in PDB database which shows the importance of this enzyme, there is no three-dimensional (3D) structure of *Schwanniomyces occidentalis* alpha-amylase (Strasser et al., 1989). It is essential to have enough knowledge about 3D structure for understanding the biophysical properties, however obtaining X-ray diffraction quality crystals of flexible multi-domain monomeric protein targets is often difficult (Harris, 2003). The homology modeling is a reliable and an efficient method for the prediction of 3D structure of proteins (Krieger, Nabuurs, & Vriend, 2003; Šali & Blundell, 1993).

In the present study, the 3D feature of AMY1 was obtained by a homology modeling procedure based on the crystal structure of alpha-amylase from *Aspergillus Oryzae* (TAKA), deposited in protein data bank (PDB...
code 6TAA with 55% identity with the target sequence) and 2.10 Å resolution (Swift et al., 1991). The models were generated and a stable model was selected with regard to discrete optimized protein energy (DOPE) (Šali & Blundell, 1993; Shen & Sali, 2006). The generated 3D structure was subjected to molecular dynamics (MD) simulations (GROMACS 4.6.7) (Berendsen, van der Spoel, & van Drunen, 1995; Hess, Kutzner, Van Der Spoel, & Lindahl, 2008).

The modeled structure validity was tested through Procheck, verified 3D and packing quality methods and GROMACS trajectory tools (RMSD, Rg and RMSF). Secondary structure, contact map and Principle component analysis were performed for both structures. Subsequently, homology-modeled AMY1 was used for comparing the conservation of structural features and calcium coordination sphere.

Materials and methods

Homology modeling and structural validation

The homology model was used to build the initial model of AMY1. The first step was searching a number of related sequences to find a related protein as a template.

The sequences of alpha-amylase isolated from the S. occidentalis (AMY1) were retrieved from UniProtKB (http://www.uniprot.org/) that found in 1991, under identifier number of P19269 which consist 512 amino acids (1–25 are signal peptide and 487 residues, 26-512 build the main chain) (Wu, Wang, & Hsu, 1991). The modeled sequence in comparison with the presented UniProt one is a natural variant and possesses A and S instead of D and L in positions 350 and 479, respectively (in strains: CCRC 21164 and ATCC 26077/CBS 2863, i.e. it is due to the few variations in the different strains). An experimentally resolved 3D structure for AMY1 is not available. In order to find the most similar sequence with known 3D structure (template) to each of the target sequence, sequence similarity search was performed on the BLASTp online server (http://blast.ncbi.nlm.nih.gov) (Altschul et al., 1997). The 3D structures of the template, 6TAA were extracted from the Protein Data Bank (PDB) (http://www.rcsb.org/) (Bernstein et al., 1977). The automated sequence alignment and analysis of the template and target were carried out by EMBOSS Needleman–Wunsch method (Rice, Longden, & Bleasby, 2000; Needleman & Wunsch, 1970). This tool reports residues that are identical (percentage of identity) and conserved (percentage of similarity/positive). Modeller 9.14 program was performed to build the 3D structure of AMY1. Modeller is an implementation of an automated approach to comparative modeling by satisfaction of spatial restraints (Eswar et al., 2007; Fiser & Šali, 2003). The predicted structures (100 models) were saved in the PDB format and sorted according to scores calculated from DOPE. The best model was selected with regard to DOPE. The quality of predicted structure was confirmed by PROCHECK (Colovos & Yeates, 1993; Laskowski, MacArthur, Moss, & Thornton, 1993) and ProSA (calculates an overall quality score for 3D structures and shows whether the Z-score of the query input structure lies within the range of scores found for experimentally determined structures of similar size (Wiederstein & Sippl, 2007)). Normality of the local environment of amino acids was checked by WHAT IF coarse packing quality control, which should stay above −5 (Vriend & Sander, 1993). VERIFY-3D method produces average data points for each residue to evaluate the structural quality based on the energetic and empirical methods. Protein structure with more than 80% residues with score of >2 is considered to be of high quality (Eisenberg, Lüthy, & Bowie, 1997).

Molecular dynamic simulation

Energy minimization and MD were performed through MD simulations by GROMACS (4.6.7) and Optimized Potential for Liquid Simulations (OPLS) force field (Jorgensen, Maxwell, & Tirado-Rives, 1996). The obtained structure from modeller was used as an input. Water box was created with at least 10 Å distances from protein using SPC water model and applying boundary conditions. System neutralization was done by adding sodium ions. The MD simulation was carried out to examine the quality of the model structures by checking their stability via performing 50 ps simulations at a constant temperature 300 K (NVT). Energy minimization was performed in 2000 steps to avoid any kind of bad contacts generated while solvating the system. Then NPT optimization was done for 500 ps. To increase the likelihood of achieving the appropriate structure, the MD simulation was performed for 14 ns using OPLS force field (Jorgensen et al., 1996). GROMACS implemented tools carried out for trajectory analysis. The RMSD and RMSF were calculated for the protein backbone using least-squares fitting. The protein radius of gyration for template and model structures was also obtained. DSSP V 2.1.0 used to follow changes in secondary structure during MD (Joosten et al., 2011). The principle component analyses were performed and the first five normal modes obtained.

Calculation of protein electrostatics

The resulting PDB structure was obtained by last 1 ns average of the simulation trajectory. pKa values of titratable side chains were predicted by the Virginia Tech H++server (http://biophysics.cs.vt.edu/H++) at pH 7.0
Hardware and software

The homology modeling (MODELLER 9v14) and visualization of AMY1 and 6TAA structures (VMD 1.9.2) were carried out with MAC 10.10.1 environment. The MD simulations were performed using GROMACS 4.6.7 in Centos environment. For high duty jobs, one node of a cluster consisting 12 cores (1.4 GHz) AMD CPU with 32 GB RAM were used. The structure validation analysis using WHATIF (http://swift.cmbi.ru.nl/servers/html/index.html), VERIFY-3D (http://nihserver.mbi.ucla.edu/SAVS/), and ProSA (https://prosa.services.came.sbg.ac.at/prosa.php) were carried out online. Contact map analysis was performed on the average structures obtained by GROMACS tools. The contact maps were visualized for carbon alpha (cut off = 1.5 nm) and compared utilizing CMView software freely available by Vehlow et al. (Vehlow et al., 2011). The obtained structure could also be further refined using the fully automated i3Drefine software as the final step in computational structure prediction pipeline (Bhattacharya & Cheng, 2013).

Results

Homology modeling and validation

An optimal sequence alignment is essential to the success of homology modeling. In this study, homology modeling technique was used to construct models of AMY1 based on the appropriate identity among the structures of AMY1 and 6TAA. The homology modeling approach was used to construct the structure of AMY1. The homology model of AMY1 showed root mean square deviation value of 2.1 Å with the template. The sequence alignment used for homology modeling annotated by structural elements is shown in Figure 1. It seems that the two models (AMY1 and 6TAA) have a very similar tertiary structure. The TIM barrel fold β-strands as the key features of alpha-amylase can be observed to be well conserved and superimposed in both structures. Schellman motif amino acids capping the alpha helix c-terminal of the A domain, (Leu 118, His 119, Met 123, Figure 1) are also conserved in both structures. It should be noted that Gly 122 is replaced with Ser in AMY1, which could be a candidate for point mutation studies. However there are some little differences which well demonstrate the criteria declaring homologous sequences could most probably share same structural features in order to provide necessary functional conformation which is more conserved than sequence alone. Figure 1 shows that the alignment results indicate 55% sequence identity between 6TAA and AMY1. MODELLER software applies an inbuilt evaluation score called DOPE to assess the quality of generated structures. The modeled structure with a lowest DOPE score of −55783.08 was selected for validations and simulation studies. PROCHECK was used to evaluate the stereochemical properties of the model (Ramachandran plot, Supplementary Figure 1). In this study, PROCHECK showed 94.2% of amino acids in the core and 4.9% are in allowed regions as presented in Supplementary Figure 2(A). ProSA analysis showed small difference in Z-score of template and model structure as depicted in Supplementary Figure 2(B). VERIFY-3D score per residue were in the average score of >2 for 90% of residues. Figure 2(A) illustrates the VERIFY-3D score for the modeled and template structures. VERIFY-3D determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar, etc.) and comparing the results to good structures (Lüthy, Bowie, & Eisenberg, 1992). As it can be inferred, there are three regions where the score is lower than .2 which are residues 163–173, 312–314, and 384–395 (overlay 26 from 471 which are nearly 5.45% of residues) and therefore the obtained structure meets the criteria for a good model, which states that at least 80% of residues should have VERIFIED-3D score larger than .2. In addition, packing quality obtained using WHAT-IF evaluation criteria is demonstrated in Figure 5, the packing score for nearly all residues were higher than the least score of −5 for predicted structure. Thus, validation methods indicate that the model was acceptable and was chosen for subsequent MD simulation.

MD simulations

The RMSD of protein backbone was followed for 14 ns (Figure 3(A)) After the initial rise, the system reached to a stable state and variations are reduced less than .04 nm in the last ns, indicating a good overall structure alignment with 6TAA. Another structural future of protein MD trajectory is the ability to follow the protein radius ($R_g$) in curse of simulation time (Figure 3(B)). The $R_g$ for predicted structure starts with a little higher value while for the template the protein radius shows a gradual increase in the first ns of simulation. However, as the corresponding data show both structures equilibrate around similar radii. The DSSP secondary structure analysis was performed on the trajectory file in course of simulation time as shown in Figure 4. The average numbers of amino acids present in each secondary structure.
are depicted with corresponding standard deviations in Figure 5. The number of coil forming residues only seems to be meaningfully different which can represent higher flexibility in AMY1 structure.

The 3D structure of AMY1 was superimposed with 6TAA as shown in Figure 6(a). The amino acids in the vicinity (\( \pm 0.5 \) nm) of central calcium (in front of the TIM barrel); Arg 215, Asp 217, Ser 218 (AMY1), Glu 241, His 307, and Asp 308 are identical except for position 218. However, it is replaced with the similar residue, Ter, in 6TAA structure.

The amino acids in the proximity of second calcium ion (distal form the TIM barrel, domain B), Asn132, His 158, Glu173, Asp 174, Asp 186, His 221 are exactly similar in both structures. The adjacent hydrophobic residues (Phe 157, Val 172, Cys 175, Trp176, Leu 187, and Val 222) are also identical in both sequences. The advantage of studying MD is to identify regions and amino acids that have the greatest movement (fluctuation) during the simulation. These fluctuations have an important role in protein flexibility, enzyme activity, and stability.

The amino acid fluctuations along the trajectory are shown in RMSF Figure 6(B). Most of them are less than 1 angstrom and less than 5% of fluctuations are between 1 and 3 angstrom. Interestingly, both structures show nearly similar pattern, especially in the first two domains.

It is well demonstrated that the presence of charge–charge interactions could stabilize folded protein structure, however the role of these interactions during folding and in unfolded state are neglected in most computational methods which focus only on the folded or nearly folded states (Azia & Levy, 2009). Both structures possess same number of His (7), Met (9), Tyr (34), and aspartic acid (Asp, 42) while AMY1 possesses 17 glutamic acid (Glu) compared to 12 for 6TAA as it can be inferred from Figure 1.

As mentioned, both structures possess seven His residues where six pairs are well matched in the alignment, five pairs are well conserved, and are in proximity of CA ions. The unmatched His residues are present in the surface and are in contact with water molecules in both structures (Figures 7–9).

Contact map based on global alignment is represented in figure. The common contacts are in majority and there are minor variations in both structures, however the contact maps associated with alpha helix and beta sheet secondary structures are well superimposed. Contact map analysis is helpful in investigating structural personalities of proteins while providing information about the separation of contacting residues, which cannot be inferred from the 3D images alone. The difference distance map visualizes conformational changes between two structures. For a pair of residues \((i, j)\) the distance map visualizes how much the distance between \(i\) and \(j\) has changed from 6TAA structure to AMY1. This highlights regions of conformational change as red hot-spots, while regions that remain unchanged are colored in blue. An arrow and the circle show the Gly 35 and Asp 157, respectively.

The amount of available hydrophilic and hydrophobic surfaces of AMY1 is shown in Figure 9(A). The proportion of the hydrophilic surface \((128 \pm 2 \text{ nm}^2)\) to the hydrophobic surface \((104 \pm 1.7 \text{ nm}^2)\) in AMY1 is about 55 and 45%, respectively. Small variations in this curve also show that the time of simulation was enough, because the variations are low after the first few ns.

The modeled structure isolectric point and enzyme net charge (electron charge) in neutral conditions were calculated to be 4 and \(-23\), and for the template was 3.8 and \(-19\), respectively. This difference in net charge can well be inferred from Figure 9(B). The graph also illustrates how the charge of a protein is buffered over the range of pH 5–8, which also represents the pH range of template and AMY1 amylase activity. The final structure and the amount of protein surface remained nearly constant. The predicted 3D structure of the \(S. occidentalis\) alpha-amylase was submitted to the Protein Model Database (PMDB, bioinformatics.cineca.it/PMDB/) and was assigned the accession number PM0080148.

### Discussion

In this study, we predicted 3D structure of a fungal alpha-amylase that possesses 55% sequence identity (70% similarity) with the template protein sequence TAKA amylase (6TAA). The advantage of using fungi for the production of amylases is ease of manipulation in addition with economical bulk production and application for bioconversion of solid substrates due to hyphal mode of growth (solid-state fermentation, SSF), tolerance to high osmotic pressure, and low water activity \((a_w)\) (Hernández, Rodríguez, Guerra, & Rosés, 2006). The fungal amylases are also generally regarded as safe in the food industry. Fungal a-amylases have been permitted as bread additives since 1955 in the US, which shows its preference over microbial sources (Gupta, Gigras, Mohapatra, Goswami, & Chauhan, 2003).

The protein sequence alignment with template is needed because Modeller can discover structural information based on sequence. This information is applied to simulation of target structure (Shenoy, 2010). Homology modeling, referred to as comparative or knowledge-based modeling, develops a 3D model from a protein sequence based on the structures of homologous proteins (Kaczanowski & Zielenkiewicz, 2010). Here, several
Figure 1. Structure and sequence alignment of Schwanniomyces occidentalis alpha-amylase AMY1 with the template crystal structure 6TAA.

Figure 2. (A) VERIFY-3D score for the homology model (AMY1, red) and template (6TAA, blue) and refined AMY1(green). (B) WHAT-IF packaging quality score of the homology model (red) and template (blue).
Structure validation tools were applied in order to qualify the predicted model. Ramachandran plot related to protein structure validation is based on the angles of backbone’s covalent bonds (Chen et al., 2010). In this case, Ramachandran plot indicated that nearly all amino acids are in allowed regions. The ProSA structure quality method showed small difference in Z-score between AMY1 and 6TAA, which indicates significant structural closeness between model and template. Other validation methods such as WHATIF packing quality (most residues were $\geq -5$) and VERIFY-3D confirmed the validity of the predicted structure in comparison with template structure. The structures with verified 3D score of larger than .2 for 80% of residues are considered acceptable. The VERIFY-3D score for refined model were improved 5% as it can be inferred from Figure 2(A).

From the MD viewpoint, variations were comparable and converged along simulation time especially in case of $R_g$. Although the amino acid compositions of both structures are 45% different (Figure 1), RMSD and RMSF scores show similar patterns. The secondary structure analysis over the trajectory (Figure 4) is also very similar considering the fact that 6TAA structure lacks 21 N-terminal residues. The average numbers of amino acids present in each secondary structure are represented in Figure 5. The comparison of two RMSF data shows, in average, higher fluctuations in AMY1, especially in the c-terminal domain. PCA or essential

![Figure 3](image1.png)

**Figure 3.** (A) MD trajectory RMSD during simulation time of the AMY1 and template. (B) Radius of gyration over the simulation time for both predicted model and template (6TAA).

![Figure 4](image2.png)

**Figure 4.** Trajectory DSSP secondary structure analysis in course of simulation time.
dynamics aims at identifying large-scale collective motions of atoms and thus reveal the structures underlying the atomic fluctuations.

CMView software was used to compare the contact map of template and target structures regarding the global Needleman–Wunsch alignment algorithm with standard parameters (BLOSSUM50, Gap-open: 10, Gap extend: 0.5).

As it can be inferred from alignment file, RMSF and contact map, the domain A is much more conserved than the other two domains which seems to be a pattern in alpha-amylase family. The sequence and structural analysis of calcium ion coordination sphere also show the conserved microenvironments, especially for the distal calcium ion, which is surrounded by residues from the B domain.

The presence of surface charge–charge interaction may increase the folding rate and hence protein stability as discussed recently (Tripathi, Garcia, & Makhatadze, 2015; Tzul, Schweiker, & Makhatadze, 2015). The surface properties of AMY1 is an important MD approach and shows that the change in the content of hydrophilic and hydrophobic surfaces were in equilibrium state from the first ns of trajectory and remained stable for the predicted structure. From different point of view, the decreases in the number of ionic interactions cause greater flexibility in alpha-amylase family and adopt them for activity in lower temperatures (Hiteshi & Gupta, 2014). AMY1 remains active in acidic conditions (pH 4) while it only keeps 20% of the original activity at slightly alkaline conditions of pH 8. The optimum pH for TAKA amylase is reported to be around 5.3 which is close to 6 for AMY1 (Takahasgi, Toda, Nishibe, & Yamamoto, 1982).

Comparison of thermal stability shows that AMY1 (expressed in Saccharomyces cerevisiae) optimum temperature and pH had been reported to be 40°C (Wang, Lin, & Hsu, 1989). The AMY1 lost its activity completely at 70°C. Similarly, the catalytic activity has been showed to be maximal at 40°C, for Taka-amylase (Wang, Gu, Cui, & Liu, 2011).

The residues involved in interacting with calcium are well conserved in both structures. Therefore, the effect of calcium ion on the stability of taka amylase can also be expected for AMY. Taka-amylase A, and the liquefying amylase of Bacillus subtilis are sensitive to EDTA treatment.

The 6TAA melting temperature ($T_m$) has been determined to be 71°C in calcium saturation (CS) and 57°C in calcium depleted (CD) conditions. It is interesting to note that the stability of B. licheniformis (BLA, 1BLI, 483 residues and 58.27 Da) with 3 Ca ions is 102°C in

Figure 5. Comparison of average numbers (with standard deviation) for residues taking part in forming secondary structure elements in course of simulation.

![Figure 5](image)

Figure 6. (A) Superimposition of the homology model and template structures. For template structure (6TAA), alpha and beta structures are showed in different colors. (B) The average RMSF value obtained from trajectory MD for the model (AMY1) and template (6TAA).
CS condition and 51°C for CD situation. Therefore the melting temperature of TAKA amylase is nearly 6°C higher in CD condition. B. amyloliquefaciens (BAA) $T_m$ is also much lower, 38°C in CD conditions (Duy & Fitter, 2005). Therefore studying the TAKA and AMY1 amylase structures could be helpful in designing thermostable amylase which could be active under calcium depletion conditions.

It is interesting to note that, Declerck et al. have also build a homology model of BLA based on the X-ray structure of taka-amylase A from A. oryzae (Matsura, Kusunoki, Harada, & Kakudo, 1984). Although they conclude that no extensive interpretation could be done because of the low sequence similarity between BLA and taka-amylase (about 25%) (Declerck, Joyet, Trosset, Garnier, & Gaillardin, 1995), later comparison with the determined X-ray structure of BLA reveals that the overall conformation of the protein had been properly modeled (Declerck et al., 1997). Site-directed mutagenesis of taka-amylase A catalytic active-site residues have been also studied before, consisting on the importance of homology modeling in elucidating structural features of amylase enzymes (Nagashima et al., 1992).

The validity of obtained model may represent that the utilized OPLS force field is suitable for calcium-containing enzymes. Despite the powerful background and high performance MD calculations, the starting structure and the time course of simulation might affect the resulting structure. Therefore, the major structural characteristics with regard to performed analyses were discussed here.

Figure 7. (A) Representation of the five first eigenvectors of trajectories, the flash and circle indicate Gly 35 and ASP 157, respectively. (B) First two eigenvectors for 6TAA (solid lines) and AMY1 (doted lines) against simulation time.
Conclusion
Fungal amylase has been shown to have advantages in terms of safety and solid-state applications over other microbial sources. Taka amylase is also indicated to have high thermal stability under calcium depletion conditions. The 3D structure of AMY1 has not been reported yet. Production and optimization of activity are required to understand the relationship between enzyme structure and sequence. In this investigation, the 3D structure of the Schwanniomyces alpha-amylase was built by homology modeling. Then energy minimization and MD were used to provide an insight regarding the structural and flexibility of both model and template. The obtained 3D structure can be considered reliable not only due to the computational methods but because of the conserved structural features associated with amylase enzymes. Usefulness of homology modeling method could be inferred from comparisons with the template. The 3D structure of AMY1, as many other industrial enzymes, may facilitate protein engineering studies.

Supplementary material
The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2016.1154892.

Acknowledgments
We would like to thank Dr Justin A. Lemkul, from Department of Pharmaceutical Sciences University of Maryland, for his support and help toward applying GROMACS.

Disclosure statement
No potential conflict of interest was reported by the authors.

ORCID
Yahya Sefidbakht http://orcid.org/0000-0003-0538-0252

References
Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, 25, 3389–3402. PubMed PMID: 9254694; PubMed Central PMCID: PMC146917.

Anandakrishnan, R., Aguilar, B., & Onufriev, A. V. (2012). H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research*, 40, W537–W541.

Azia, A., & Levy, Y. (2009). Nonnative electrostatic interactions can modulate protein folding: Molecular dynamics with a grain of salt. *Journal of Molecular Biology*, 393, 527–542.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences*, 98, 10037–10041.
Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., ... Tasumi, M. (1977). The protein data bank: A computer-based archival file for macromolecular structures. *Journal of Molecular Biology*, 112, 535–542.

Berendsen, H. J., van der Spoel, D., & van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91, 43–56.

Bhattacharya, D., & Cheng, J. (2013). i3Drefine software for protein 3d structure refinement and its assessment in CASP10. *PLoS ONE*, 8, e69648. doi:10.1371/journal.pone.0069648. PubMed PMID: 23894517; PubMed Central PMCID: PMC3716612.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., ... Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography*, 66(Pt 1):12–21. doi:10.1101/s0979044090040273. PubMed PMID: 20057044; PubMed Central PMCID: PMC2803126.

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*, 2, 1511–1519. doi:10.1002/pro.5560020916. PubMed PMID: 8401235; PubMed Central PMCID: PMC2142462.

Da Lage, J. L., Feller, G., & Janecek, S. (2004). Horizontal gene transfer from Eukarya to bacteria and domain shuffling: The alpha-amylase family. *Cellular and Molecular Life Sciences (CMLS)*, 61, 97–109. doi:10.1007/s00018-003-3334-y

Declerck, N., Joyet, P., Troset, J. Y., Garnier, J., & Gaillardin, C. (1995). Hyperthermostable mutants of *Bacillus licheniformis* α-amylase: Multiple amino acid replacements and molecular modelling. *Protein Engineering, Design and Selection*, 8, 1029–1037.

Declerck, N., Machius, M., Chambert, R., Wiegand, G., Huber, R., & Gaillardin, C. (1997). Hyperthermostable mutants of *Bacillus licheniformis* alpha-amylase: Thermodynamic studies and structural interpretation. *Protein Engineering Design and Selection*, 10, 541–549.

Declerck, N., & Fitter, J. (2005). Thermostability of irreversible unfolding α-amylases analyzed by unfolding kinetics. *Journal of Biological Chemistry*, 280, 37360–37365.

Eisenberg, D., Lüthy, R., & Bowie, J. U. (1997). VERIFY3D: Assessment of protein models with three-dimensional profiles. *Methods in enzymology*, 277, 396–404.

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., et al. (2007). Comparative protein structure modeling using MODELLER. Current protocols in protein science/editorial board, John E Coligan [et al]; Chapter 2:Unit 2 9. doi:10.1002/0471408648. ps0209s50. PubMed PMID: 18429317.

Feitknecht, H. (2003). Anaerobic digestion of desiring wastewater: Influence of pretreatment and anionic surfactant on degradation and intermediate accumulation. *Enzyme and Microbial Technology*, 33, 250–258.

Fiser, A., & Sali, A. (2003). Modeller: Generation and refinement of homology-based protein structure models. *Methods in enzymology*, 374, 461–491. doi:10.1016/S0076-6879(03)74020-8

Fitter, J. (2005). Structural and dynamical features contributing to thermostability in α-amylases. *Cellular and Molecular Life Sciences*, 62, 1925–1937. doi:10.1007/s00018-005-5079-2.

Gupta, R., Gigras, P., Mohapatra, H., Goswami, V. K., & Chauhan, B. (2003). Microbial α-amylases: A biotechnological perspective. *Process Biochemistry*, 38, 1599–1616.

Harris, T. (2003). PDK1 and PKB/Akt: Ideal targets for development of new strategies to structure-based drug design. *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)*, 55, 117–126.

Hernández, M. S., Rodriguez, M. R., Guerra, N. P., & Rosés, R. P. (2006). Amylase production by *Aspergillus niger* in submerged cultivation on two wastes from food industries. *Journal of Food Engineering*, 73, 93–100.

Hess, B., Kutzner, C., Van Der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4, 435–447.

Hiteshi, K., & Gupta, R. (2014). Thermal adaptation of α-amylases: A review. *Extremophiles*, 18, 937–944. doi:10.1007/s00792-014-0674-5

Joosten R. P., te Beek T. A., Krieger E., Hekkelman M. L., Hooft R. W., Schneider R., et al. (2011). A series of PDB related databases for everyday needs. *Nucleic Acids Research*, 39(Database issue):D411–D419. doi:10.1093/nar/gkq1105. PubMed PMID: 21071423; PubMed Central PMCID: PMC3013697.

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118, 11225–11236.

Kaczanowski, S., & Zielenkiewicz, P. (2010). Why similar protein sequences encode similar three-dimensional structures? *Theoretical Chemistry Accounts*, 125, 643–650.

Kandra, L. (2003). α-Amylases of medical and industrial importance. *Journal of Molecular Structure: THEOCHEM*, 666–667, 487–498.

Krieger, E., Nabuurs, S. B., & Vriend, G. (2003). Homology modeling. Methods of biochemical analysis, 44, 509–524.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26, 283–291.

Lüthy, R., Bowie, J. U., & Eisenberg, D. 1992. Assessment of protein models with three-dimensional profiles. *Nature*, 356, 83–85. doi:10.1038%2F356083a0.

MacGregor, E. A. (1988). Alpha-amylase structure and activity. *Journal of Protein Chemistry*, 7, 399–415.

Matsusuru, Y., Kusumoki, M., Harada, W., & Kakudo, M. (1984). Structure and possible catalytic residues of Taka-amylase A. *Journal of Biochemistry*, 95, 697–702.

Myers, J., Grothus, G., Narayanan, S., & Önmüfiev, A. (2006). A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins: Structure, Function, and Bioinformatics*, 63,928–938.

Nagashima, T., Tada, S., Kitamoto, K., Gomi, K., Kumagai, C., & Toda, H. (1992). Site-directed mutagenesis of catalytic active-site residues of Taka-amylase A. *Biosciences, Biotechnology, and Biochemistry*, 56, 207–210. doi:10.1271/bbb.56.207

Needleman S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48, 443–453.
