PIBB 生物化学与生物物理进展
Progress in Biochemistry and Biophysics
2013, 40(6): 565~577
www.pibb.ac.cn

Insights Into The Working Mechanism and Unfolding Property of *Arthrobacter chlorophenolicus* Amylosucrase

LIU Ming[1)], LI Ping[2)], LIU Bin[3)], SU Ji-Guo[4)*], WANG Cun-Xin[3)*]

([1)] *Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing* 100050, China;
[2)] *Institute of Process Engineering, Chinese Academy of Sciences, Beijing* 100190, China;
[3)] *College of Life Science and Bioengineering, Beijing University of Technology, Beijing* 100124, China;
[4)] *College of Science, Yanshan University, Qinhuangdao* 066004, China)

Abstract Amylosucrase (AS) is a kind of glucosyltransferase (E.C. 2.4.1.4) belonging to the glycoside hydrolase (GH) family 13. In the presence of an activator polymer, *in vitro*, AS is able to catalyze the synthesis of an amylose-like polysaccharide composed of only α-1,4-linkages using sucrose as the only energy source. Unlike AS, other enzymes responsible for the synthesis of such amylose-like polymers require the addition of expensive nucleotide-activated sugars. These properties make AS an interesting enzyme for industrial applications. In spite of the great application potential, industrial applications of AS were largely hampered by its low stability. In this work, the structure of *Arthrobacter chlorophenolicus* amylosucrase (AcAS) was modeled according to crystal structures of *Neisseria polysaccharea* amylosucrase (NpAS) and *Deinococcus geothermalis* amylosucrase (DgAS). The structural difference of the three AS were investigated for the sake of exploring the working mechanism of AcAS. The dynamics and functional motions of AcAS were also investigated by the Gaussian network model (GNM). Based on GNM, it was found that AcAS can be divided into two regions with different moving directions. Finally, the unfolding property of AcAS was studied by the iterative GNM, and several kinetically weak regions were identified on the basis of the predicted unfolding pathway. These discoveries may be helpful to the industrial development of AS.

Key words amylosucrase, homology modeling, unfolding, GNM, iterative GNM
DOI: 10.3724/SP.J.1206.2013.00044

Amylosucrase(AS) is a kind of glucosyltransferase (E.C.2.4.1.4) belonging to the glycoside hydrolase (GH) family 13 according to the carbohydrate-active enzymes classification[1–3]. The first AS was identified in *Neisseria perflava* as early as 1946 [4]. Several decades later, MacKenzie *et al*.[5] identified intracellular AS in six other *Neisseriae* species, and later an extracellular *Neisseria polysaccharea* AS (NpAS) was discovered[6]. Until recently AS has only been found in bacteria from the genus of *Neisseria*[7], *Deinococcus*[8–9], *Alteromonas*[10] and *Arthrobacter*[11]. NpAS is the only AS for which several structures, alone or in complex with sucrose substrate or products are available to date[12–15]. According to crystal structures, AS possesses the characteristic $(\beta/\alpha)_8$-barrel catalytic A domain, a B domain between β3 and α3 of the catalytic A domain, and a C-terminal domain consisting of a sandwich of two Greek key motifs. In addition to these common structural features of GH13, AS also possesses two unique domains: an α-helical N-terminal domain and a B' domain between β7 and α7 of the catalytic core, which is suggested to be involved in the polymerase activity of this enzyme. The B and B' domains contribute largely to the

*Corresponding author.
SU Ji-Guo. Tel: 86-335-8057027, E-mail: jiguosu@ysu.edu.cn
WANG Cun-Xin. Tel: 86-10-62757456, E-mail: cxwang@bjut.edu.cn
Received: February 1, 2013 Accepted: April 1, 2013

formation of the active site pocket[7].

Unlike many other enzymes of the GH13, which mainly degrade polyglucan, AS is a very exceptional one since its main role is to catalyze the produce of the insoluble polysaccharides. In the presence of an activator polymer *i.e*. glycogen, *in vitro*, AS is able to catalyze the synthesis of an amylose-like polysaccharide composed of only α-1, 4-linkages using sucrose as the only energy source[16]. In the absence of glycogen, the reaction pathways however become much more complicated, including polymer synthesis, hydrolysis of sucrose, synthesis of smaller maltosaccharides, and synthesis of sucrose isoforms[7]. Hydrolysis, however, is always a minor side reaction under each circumstance. The function of AS *in vivo* is undoubtedly the extension of glycogen-like oligosaccharides, which is clearly demonstrated by the formidable increase in k_{cat} observed when glycogen is present[17]. In contrast to AS, other enzymes responsible for the synthesis of such amylose-like polymers require the addition of expensive nucleotide-activated sugars[18]. Amylosucrase can also be used to modify the structure of polysaccharides such as glycogen by the addition of α-1, 4-linked glucosyl units [19]. These properties make AS a kind of interesting and promising enzyme for industrial applications.

Among the AS discovered so far, NpAS is the one which is studied most. Although of great potential for industrial applications, NpAS suffers from a low catalytic efficiency on sucrose alone (k_{cat}=1·s^{-1}) and a weak thermostability ($t_{1/2}$[50℃]=3 min), limiting its industrial development[9]. Directed evolution has been attempted to improve catalytic efficiency and thermostability of NpAS [20]. Searching for more thermostable and efficient enzymes in the natural diversity is another alternative that has motivated the biochemical characterization of the AS from *Deinococcus geothermalis* (DgAS)[9], and *Deinococcus radiodurans* (DrAS)[8]. DrAS possesses similar stability and activity properties with NpAS. With a specific activity of 44 U·mg^{-1} at the optimal temperature of 50℃, the recombinant DgAS is the most thermostable AS characterized to date[9]. Previously, we thoroughly investigated the structure, dynamics and unfolding properties of NpAS and DgAS to explain the reason why DgAS is more stable than NpAS and reveal novel insights into rational design on thermostability [21]. Recently, Seo *et al*[11]. found a new AS (AcAS) from *Arthrobacter chlorophenolicus* A6 which is a Gram-positive actinobacterium capable of surviving under psychrophilic conditions. Their studies indicates that AcAS is considerably heat-labile and exerting its function as a monomer. Besides, AcAS produce more soluble maltooligosaccharides and less insoluble glucan than NpAS. Until now, the crystal structure of AcAS has not been experimentally determined. In order to figure out the detailed working mechanism of AcAS, the 3D structure of AcAS was built by homology modeling and the structures of AcAS, NpAS and DgAS were carefully and thoroughly investigated. In addition, the Gaussian network model (GNM) and iterative GNM methods are also applied to gain insight into the dynamics of three AS and unfolding property of AcAS, and hopefully the results revealed in the current work can be used as guides for practical developments on the AS in the future.

1 Materials and methods

1.1 AcAS structure modeling

The sequence of the *Arthrobacter chlorophenolicus* amylosucrase (AcAS) was obtained from the Universal Protein Resource Knowledgebase (UniProtKB) with accession number B8H6N5. The crystal structures of NpAS and DgAS with the highest resolution were selected as templates. The sequence of AcAS was aligned with NpAS and DgAS with CLUSTALX [22]. The 3D model of AcAS was generated according to the standard protocol of MODELER [23]. The model with the best score was then minimized by NAMD [24]. Altogether, 1 000 steps of steepest descent and 2 000 steps of conjugate gradient optimization were performed. Finally, a Ramachandran plot was obtained using the PROCHECK[25] program to verify the quality of the model.

1.2 GNM

The GNM was first proposed by Bahar *et al*. [26] within such a minimalist mindset to explore the role and contribution of purely topological constraints, defined by the 3D structure of protein. The GNM describes a three-dimensional structure of protein as an elastic network of Cα atoms connected by harmonic springs within a certain cutoff distance. The cutoff

distance is usually taken as 7.0 Å, based on the radius of the first coordination shell around residues observed in PDB structures [27–28]. GNM is more accurate when evaluating the deformation magnitudes, or the distribution of motions of individual residues [29]. The detailed descriptions for GNM could be found in our previous study[21].

1.3 Iterative GNM

The traditional GNM can only explore the dynamics of the target protein with a specific conformation. By breaking each "weak" contact between nodes step by step, we have shown the unfolding processes of proteins can be predicted by the iterative GNM, with a reasonable accuracy[30-31].

In this work, the iterative GNM was employed to predict the unfolding behavior of the two AS. When applying the iterative GNM, the interactions between residues are divided into the covalent and the non-covalent ones. In contrast to the traditional GNM, the interactions between the covalent and the non-covalent bonded ones are treated differently. The spring constants between all pairs of non-bonded residues within the cutoff distance are treated equally, *i.e*. a single force constant γ is employed. The strengths of the interactions between all covalently bonded pairs are defined by $c\gamma$, where c is coefficient and can be determined by fitting predicted fluctuations against the crystallographic B-factors [30]. Considering all contacting residues, the internal Hamiltonian of the system can be defined by Equation(1),

$$V=\frac{1}{2}\gamma[\Delta R^T(\boldsymbol{\Gamma}\otimes\boldsymbol{E})\Delta R] \quad (1)$$

where γ is the force constant of the harmonic spring connecting each node; ΔR represents the 3N-dimensional column vector for fluctuations; $\boldsymbol{E}$ is the third-order identity matrix; $\otimes$ represents the operation of direct product; and $\boldsymbol{\Gamma}$ is the $N \times N$ symmetric Kirchhoff matrix in which the elements are defined as:

$$\boldsymbol{\Gamma}=\begin{cases} -c & \text{if } |i-j|=1 \\ -1 & \text{if } |i-j|>1 \text{ and } R_{ij}\leqslant r_c \\ 0 & \text{if } |i-j|>1 \text{ and } R_{ij}>r_c \\ -\sum_{j,j\neq i}\boldsymbol{\Gamma}_{ij} & \text{if } i=j \end{cases} \quad (2)$$

As the temperature of the system is gradually increased, the native contacts among residues are expected to break in a fluctuation-dependent manner [30]. The philosophy held by this method, therefore, is that the weakest interaction will be broken first. In order to take the stochastic nature of thermal denaturation into consideration, one can incorporate some noise into the calculations during the real simulation processes[30]. In this work, the native contact to be removed in each step was randomly selected from contacts with the first three largest fluctuations, and the calculations were performed 100 times for each AS to get more reasonable results. The detailed information for this method could also be found in our previous study[21].

2 Results and discussion

2.1 Homology modeling of AcAS

To improve the accuracy of modeling structure, the crystal structure of NpAS (PDB:1G5A) and DgAS (PDB:3UCQ) with highest resolutions were selected as templates to build the structural model of AcAS. The sequence alignment of AcAS with NpAS and DgAS is shown in Figure 1. The asterisk represents identical amino acids present in three protein sequences. The sequence identities between AcAS and NpAS, DgAS are up to 57% and 44%, respectively.

In total, one hundred 3D models of AcAS were generated using the known structure of NpAS and DgAS with MODELER. The model having the best score was considered as initial model of AcAS. The initial model was refined using energy minimization in NAMD, to allow the structure to relax gradually into a final conformation. The refined model was then further validated using PROCHECK to perform full geometric analysis as well as stereo chemical quality of a protein structure. According to the results from PROCHECK 91.2% residues of the refined model were located in the favored region, 8.3% in the additional allowed region, 0.5% in the generously allowed region and no non-Gly residues were found in the disallowed region (Figure 2). Therefore, the refined model was considered as reasonable and could be employed for further analyses.

```
AcAS MQESVGVEQPVAGQDGPRDVVLAALSRER--AAGLAADPAFERRLDRYLPLLTGLFRPLYGARADWQDQLAALVVQAARSWQERTPELKA  88
NpAS ------SP-N--SQYLKTRILDIYTPEQRAGIEKSEDWRQFSRRMDTHFPKLMNELDSVYGNNEALLPMLEMLLAQAWQSYSQRNSSLKD  81
DgAS ------MLKDVLTSELAAQVRDAF---------DDDRDAETFLLRLERYGEDLWESLRAVYGDQVRALPG--RLLEVMLHAYHARPAELRR  74
                   .     :                      * *:: :   *   : .:** .         *:    :::  * ..*:

AcAS LDAEREADPHWFQSNGMLGGVCYVDRYAESLEGVRRSIPYFKELGLTYLHLMPLFLAPEPHSDGGYAVSSYRQVNPRLGTMEQLRSLAAE 178
NpAS IDIARENNPDWILSNKQVGGVCYVDLFAGDLKGLKDKIPYFQELGLTYLHLMPLFKCPEGKSDGGYAVSSYRDVNPALGTIGDLREVIAA 171
DgAS LDEARLLRPDWLQRPEMVGYVAYTDRFAGTLKGVEERLDYLEGLGVKYLHLMPLLRPREGENDGGYAVQDYRAVRPDLGTMDDLSALARA 164
      :*  *   *.*:     :* *.*.* :*  *:*:.  : *:: **:.*******:   * ..******..** *.* ***: :*  :

AcAS LRSHGISLVVDFIFNHTSDEHEWARKAASGDPEFSDYYWIFPDRTMPDAFEQNAREIFPENHPGSFIRMED-----GRWVWATFHTYQWD 263
NpAS LHEAGISAVVDFIFNHTSNEHEWAQRCAAGDPLFDNFYYIFPDRRMPDQYDRTLREIFPDQHPGGFSQLED-----GRWVWTTFNSFQWD 256
DgAS LRGRGISLVLDLVLNHVAREHAWAQKARAGDPKYRAYFHLFPDRRGPDAFEATLPEIFPDFAPGNFSWDEEIGEGEGGWVWTTFNSYQWD 254
      *:  *** *:*:::**.: ** **::.  :*** :  :: :****  ** :: .  ****:  **.*   *:     * ***:**:::***

AcAS LNYANPDVFRAMAGEMLFLANQGVDILRMDAVAFIWKQLGTTCENLPEAHSLLRAFNAVCRLAAPSLLFKSEAIVHPDEVALYIDP---- 349
NpAS LNYSNPWVFRAMAGEMLFLANLGVDILRMDAVAFIWKQMGTSCENLPQAHALIRAFNAVMRIAAPAVFFKSEAIVHPDQVVQYIGQ---- 342
DgAS LNWANPDVFLEFVDIILYLANRGVEVFRLDAIAFIWKRLGTDCQNQPEVHHLTRALRAAARIVAPAVAFKAEAIVAPADLIHYLGTRAHH 344
      **::** **  :.. :*:*** **:::*:**:*****::** *:* *:.* * **:.*.  *:.**:: **:**** * ::  *:.

AcAS -AECQLSYNPLQMALIWESLATRDVSLLAQALERRHNIPAGTAWVNYVRSHDDIGWTFADEDAAELGINGFDHRRFLNSFYVNRFPGSFA 438
NpAS -DECQIGYNPLQMALLWNTLATREVNLLHQALTYRHNLPEHTAWVNYVRSHDDIGWTFADEDAAYLGISGYDHRQFLNRFFVNRFDGSFA 431
DgAS GKVSDMAYHNSLMVQLWSSLASRNTRLFEEALRAFPPKPTSTTWGLYVRCHDDIGWAISDEDAARAGLNGAAHRHFLSDFYSGQFPGSFA 434
       .::.*:   *. :*.:**:*:. *: :**      *  *:*  ***.******:::*****  *:.*  **:**. *: .:* ****

AcAS RGVPFQDNPRTGDCRISGTTASLCGLEDS-----PA-----EAVERILLAHSVPFSTGGIPLLYLGDEVGQVNDYAYAGEPGHGEDSRWVHR 520
NpAS RGVPFQYNPSTGDCRVSGTAAALVGLAQD-----DP-----HAVDRIKLLYSIALSTGGLPLIYLGDEVGTLNDDDWSQDSNKSDDSRWAHR 513
DgAS RGLVFQYNPVNGDRRISGSAASLAGLEAALETGDPGRIEDAVRRLLLLHTVILGFGGVPLLYMGDELALLNDYAFEDVPEHAPDNRWVHR 524
      **: ** ** .** *:**::*:* **        *    .** *: * ::: :. **:**:*:***:. :**  :   . :. *.**.**

AcAS PHYPAEQYAR-RTDPSTPAGAVYAGLRRMAEVRAGTPELAGTQLIDF-ATHN-RSVLAYQRPGSPPESTRVLALANFSDSAQNLPAETFS 607
NpAS PRYNEALYAQ-RNDPSTAAGQIYQDLRHMIAVRQSNPRFDGGRLVTF-NTNN-KHIIGYIRNN------ALLAFGNFSEYPQTVTAHTLQ 594
DgAS PQMDWALAERVRQEPSSPAGRVNTGLRHLLRVRRDTPQLHASIESQVLPSPDSRALL-LRRDHP---LGGMVQVYNFSEETVMLPSHVLR 610
      *:       : * :**:.** :  .**::  ** ..*.: .     .  : : : ::   *         :: . ***: .  :.:..:

AcAS GFS-PAAVDLLSEAAVQLDSGLT-LLPRQYLWLRVTPL-- 643
NpAS AMP-FKAHDLIGGKTVSLNQDLT-LQPYQVMWLEIA---- 628
DgAS DVLGDHVQDRLSGSAFRLDRPTVRLEGYRALWLTAGEAPA 650
       .    . * :.  :. *:   . *   : :**
```

Fig. 1　Sequence alignment result of AcAS with the template NpAS and DgAS

"*" represents identical amino acids, ":" represents strong similarity, "." represents weaker similarity, and blank represents different amino acids or gap.

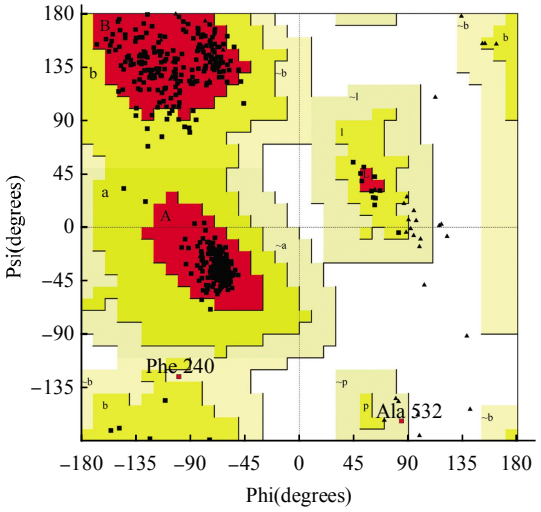

Fig. 2　Ramachandran plot of the structure of AcAS

2.2　Detailed comparison between the structures of AcAS and NpAS/DgAS

AcAS, NpAS and DgAS are constituted by 643, 628 and 650 residues, respectively. The full length AcAS can also be divided into five domains: A (97～191, 268～402, 468～560), B (192～267), B' (403～467), N (1～96) and C domains (561～643). According to sequence alignment, AcAS is more similar with NpAS. Among the five regions, A, B and B' domains are much more conserved than N and C domains. For example, the B' domain of AcAS shares 75% and 62 % identical residues with those of NpAS and DgAS. As to the N and C domains of AcAS, however, they share less than 30% identical residues with those of NpAS and DgAS. Based on this

fact, it is suggested that the structural model for the A, B and B' domains of AcAS should be more accurate than those for the N and C domains. In order to improve the accuracy of the structural model for the N and C domains, we also tried to employ some other fragments-assemble methods, such as Rosetta and I-TASSER. On the whole, these fragment-based methods performed no better than the traditional comparative modeling method according to the results of PROCHECK (data not shown). Since the N and C domains are generally believed to be not critical to the function of AS, it is suggested that the structure model for the N and C domains of AcAS are acceptable. The detailed identity between each domain of AcAS and NpAS/DgAS are listed in Table 1.

Table 1 The sequence identity of five domains between AcAS and NpAS/DgAS

Identity		NpAS					DgAS				
		N	A	B	B'	C	N	A	B	B'	C
AcAS	N	24%					25%				
	A		66%					47%			
	B			63%					52%		
	B'				75%					62%	
	C					28%					23%

The catalytic domain A of AcAS adopts the typical $(\beta/\alpha)_8$-barrel of GH family 13. The active site is highly conserved with respect to that of NpAS and DgAS (Figure 3a). D286 and E328 of NpAS are critical to its catalytic functions. They are responsible to the formation of a covalent glucosyl-enzyme intermediate. D286 acts as nucleophile and E328 as a general acid/base catalyst (proton donor)[32–33]. In AcAS and DgAS, the corresponding positions are also occupied by the same type of residues (E335 and D293 in AcAS, and E326 and D284 in DgAS). Beside the two catalytic residues, in NpAS, there are another

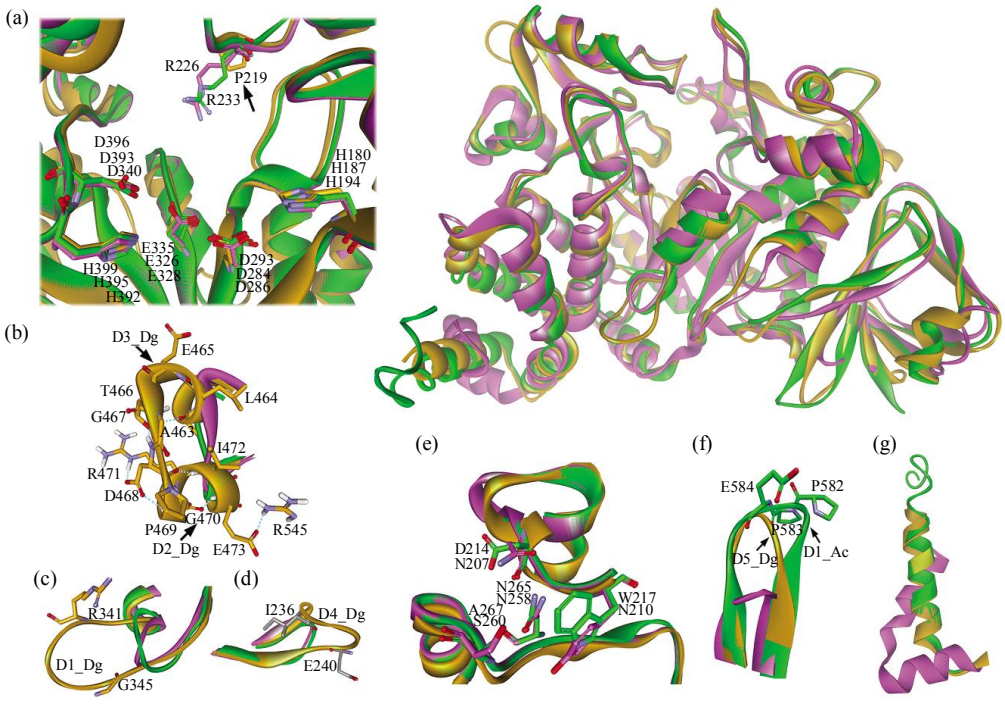

Fig. 3 The structure superimpose of AcAS (green), NpAS (purple) and DgAS (orange)

Hydrogen bonds were represented by dotted line (cyan).

three important residues: H187, H392 and D393, which are known to play assistant roles in the catalysis process. Particularly, D393 has been proved to act as a stabilizer of the glucosyl-enzyme intermediate. In AcAS and DrAS, all the corresponding positions are also conserved, indicating they share the similar catalytic mechanism with NpAS. Above the catalytic site, there is an oligosaccharide binding site (OB1) which is composed of surface residues of B, B' and A domains. Most residues that constitute OB1 are conserved among the three AS except R226 (NpAS numbering). R233 of AcAS is corresponding to R226 of NpAS, whereas the corresponding position in DgAS is substituted by P219 (Figure 3a). It is pointed out that the proline substitution in DgAS is responsible for yielding more sucrose isomers (with almost equivalent amounts of trehalulose and turanose) than NpAS [34]. Because of the conserved R233, the isomerization activity of AcAS is similar to NpAS, rather than DgAS[11].

According to the NpAS in complex with maltoheptaose (PDB code: 1MW0), there is another oligosaccharide binding site (OB2) which is nearby the catalytic site and OB1. In the OB2 site, A362, E365, D413, H414, F417, D427, G428, S429, R432, and E583 participate in the oligosaccharide binding, either through hydrogen bonds or VDW interactions. Most of these residues in AcAS are conserved with respect to NpAS except E365, D427 and E583. In NpAS, the negatively charged D427 forms a hydrogen bond with the maltoheptaose with its sidechain carboxyl. This corresponding position of AcAS is however occupied by a proline residue (P434) (Figure 4). The substitution destroys the hydrogen bond interaction and may introduce unfavorable effect to binding owing to its different physiochemical property. E365 and E583 of NpAS also interact with the oligosaccharide through hydrogen bonds. In AcAS, both of these two positions are substituted by aspartate. Although glutamate and aspartate are similar in physiochemical property, the latter's sidechain is a carbon atom shorter than that of the former. These shorter side-chains cannot form the hydrogen bonding interactions with maltoheptaose (Figure 4). All of these changes may reduce the affinity for oligosaccharide and lead to a weaker oligosaccharide binding at OB2 site in AcAS. It is reported that the OB2 site provides an anchoring platform to capture the polymer and direct the branches towards the OB1 acceptor site for elongation[35]. Therefore, compared with NpAS, the elongation activity of AcAS is weakened owing to the weaker affinity between OB2 site and oligosaccharide. This property may make AcAS to synthesize more soluble oligosaccharides and less insoluble glucan by AcAS than NpAS[11]. Interestingly, DgAS and DrAS also have lower affinity towards maltooligosaccharides due to the replacement of D413 and D427 in NpAS by an alanine and a proline[8, 36]. That probably explains why the polymerization activity of AcAS is more analogous to DgAS and DrAS rather than to NpAS.

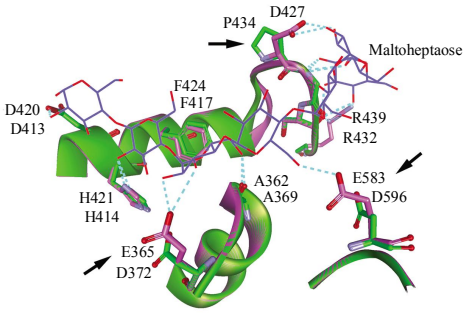

Fig. 4 The oligosaccharide binding site OB2 on AcAS (green) and NpAS (purple)
Hydrogen bonds were represented by dotted line (cyan).

As far as the length of A domain is concerned, AcAS is the same as NpAS but are 10 residues shorter than that of DgAS. There are mainly five insertions in DgAS, namely ^{341}RAHHG^{345} (D1_Dg, abbreviation for difference region 1 in DgAS), ^{470}GRIE^{473} (D2_Dg), ^{464}LETG^{467}(D3_Dg), ^{236}IGEGE^{240} (D4_Dg) and ^{587}PLG^{589} (D5_Dg). D1_Dg is a part of the extended loop region (^{338}LGTRAHHGKV^{347}, see Figure 3c) which is important to the formation of DgAS dimer. Particularly, the salt-bride formed between R341 of one monomer and D84 of another, together with that formed between R74 and E25, are critical to the stabilization of the DgAS homodimer[34]. However, the vital salt bridge formers E25, R74 and D84 in DgAS dimer are not conserved in AcAS and are substituted by P37, A88 and H98, respectively. Therefore, AcAS is not able to form a dimer owing to lack of the extended loop and the two pivotal salt-bridges. D2_Dg and D3_Dg prolong two helixes of DgAS (Figure 3b) a turn, respectively. This makes the ^{464}LETGDPGRIE^{473} region form an additional hydrogen bonds network (Figure 3b). The hydrogen bonds network reinforces

the interaction between A and B' domain, which may therefore contribute to the stability of DgAS. The length of B and B' in AcAS is the same as NpAS, but they are found shorter than that in DgAS. Domain B is composed of 76 amino acid residues in AcAS (*vs* 81 residues in DgAS). The five longer residues ^{236}IGEGE^{240} (D4_Dg) in DgAS is an additional loop organized in two antiparallel β-strands (Figure 3d). This insertion make the turn between the two antiparallel β-strands flatter. In addition, according to the complex structure of NpAS and sucrose molecules (PDB code: 1MW3), there is a surface sucrose binding site in B domain and residues N207, Y210, N258 and S260 are important for the sucrose binding. In AcAS, N207, Y210 and S260 are however substituted by D214, W217 and A267, respectively (Figure 3e). It is suggested that the surface sucrose binding site may be absent from AcAS.

The C domain of AcAS is composed of 83 amino acid residues, whose length is similar to DgAS (85 amino acid residues). However, the C domain of NpAS is only composed of 75 amino acid residues. In sequence, AcAS has a conspicuous six residues insertion region ^{580}GSPPES^{585} (D1_Ac, abbreviation for difference region 1 in AcAS) to NpAS, and DgAS has a three residues insertion region ^{587}PLG^{589} (D5_Dg). The D5_Dg belong to an eight residues loop (^{584}RDHPLGGM^{591}) of DgAS which is found to interact through polar interactions with D1_Dg (^{341}RAHHG^{345}) from A domain of the second monomer and play a key role in the dimer formation [34]. In AcAS, the corresponding loop (^{582}PPE^{584}, see Figure 3f) is three residues longer than that of DgAS. The longer loop may introduce the steric hindrance for the dimerization, and it may partially explain why AcAS exert its function as a monomer.

Additionally, another oligosaccharide binding site (OB3) was found to be situated at the surface of domain C in NpAS[14]. In OB3 site, N560, N562 and F599 are identified as key residues for oligosaccharide binding. However, in AcAS, N560 and N562 were respectively replaced by an alanine and a histidine, which break the hydrogen bonding interactions with oligosaccharide. It is therefore presumed that AcAS probably doesn't have the third oligosaccharide binding site. The defectiveness of OB3 site could also lead to the weaker affinity for oligosaccharide in AcAS and might partially contribute to the weaker elongation activity of AcAS.

2.3 Comparison among the dynamics of the three AS

Motions captured by the elastic models of a protein, specifically the first few slowest modes are global in nature and have been reported to capture the biologically and functionally relevant motions of a protein[37–39]. They are able to represent conformational changes around the native state and allow capturing more energetically accessible structural rearrangements[40]. The slowest modes have also been shown to correlate with the experimental observations of conformational changes of proteins[39, 41–42].

In order to capture the global motions of the three AS, the cross correlation for the slow motions of the three AS were calculated. Since the modes with low frequency correspond to functional motions and those with high frequency correspond to localized motions, only the slow modes are used to improve the signal/noise ratio. Here the first 40 modes are used for AcAS, NpAS and DgAS. The results for AcAS, NpAS and DgAS are depicted in Figure 5a～c, respectively. From Figure 5a, we found the cross correlation map for the first slowest mode of the AcAS can be roughly divided into two regions based on correlations. The first region is constituted by residues from 1 to 343. This region includes the N, B and the major half of the A-domain (α1 ～α5 and β1 ～β5). These parts, therefore, may move with a concerted manner since they are topologically connected. The second region is composed of the residues 344 to 643. These residues constitute the B', C and the minor half of A-domain (α6～α8 and β6～β8). Probably, the motions of these parts are also weakly coupled together according to the cross correlation map. Besides, it is noticed that the motions of the C domain are highly concerted, with an average correlation over 0.6. This just reflects the topology of the C domain, which is purely constituted by tightly packed β-sheets. According to Figure 5a, the motions of the first region and the second region of AcAS are negatively correlated, which means the two regions may move along with two different directions. Except these features, we also notice that the residues from 100 to 260, representing the B and a part of the A-domain (α1 ～α3 and β1 ～β3) are weakly correlated with the second region, reflecting the topology of the local structure. Although the overall feature of the cross correlation map of three AS are closely similar, it is found that the cross correlation for the first region in AcAS are the weakest one and DgAS

has a slightly stronger cross correlation whatever in the first region or the second region. It is suggested that the packing of the first region in AcAS is the loosest and DgAS has the tightest packing of local structures. These may partially explain why AcAS is heat-labile[11] and DgAS is the most thermostable AS among the three AS[9].

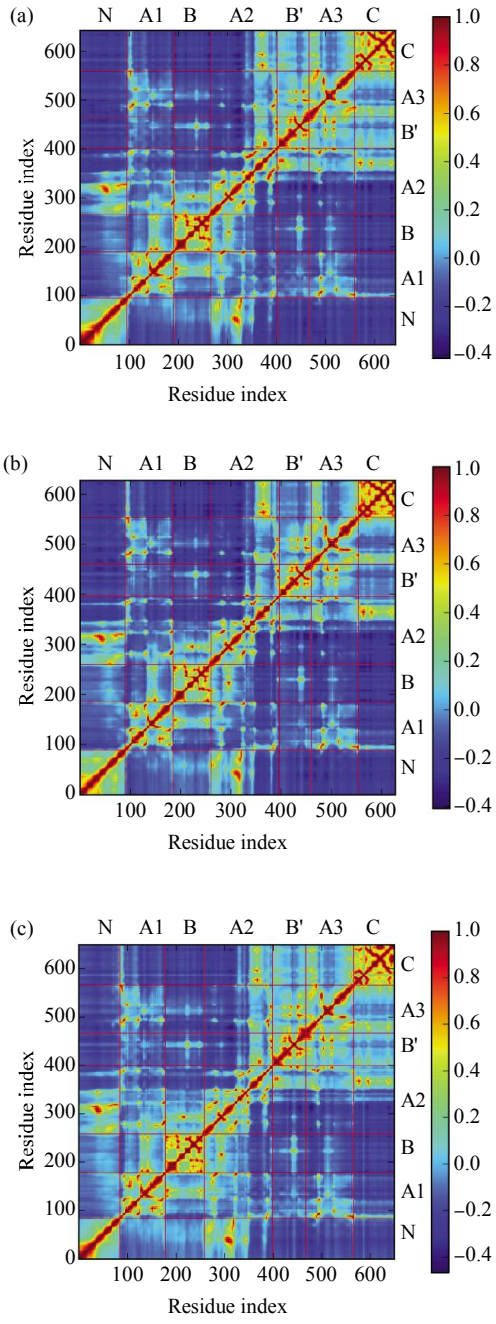

Fig. 5 The cross-correlation maps for the first 40 slow modes of AcAS (a), NpAS (b) and DgAS (c)

The A domain is decomposed into three parts because of the two insertion domains, *i.e*. B and B'. A1 part is composed of α1～α2 and β1～β3. A2 part is composed of α3～α6 and β4～β7. A3 part is composed of α7～α8 and β8.

2.4 The unfolding property of AcAS

Understanding the unfolding process of a protein is of great importance for the rational design of kinetic stability. Kinetic stability is a measure of how rapidly a protein unfolds. It is a particularly important consideration for proteins that unfold very slowly or denature irreversibly, such as aggregation or proteolytic degradation. In cases such as these, it is not the free energy difference between the folded and unfolded state that is important. That will only influence the equilibrium between the folded and unfolded state, however, it is not an equilibrium process anymore. The important thing is the free energy difference between the folded and the transition states (activation energy), for it is the magnitude of this difference that determines the rate of unfolding and hence inactivation. Clearly, it would be very helpful for the rational design if we could gain some information of the unfolding sequence of AcAS ahead of the time consuming design and validation cycles. With the prior knowledge of the unfolding properties of AcAS, one can reinforce the weak parts specifically, thus slowing down the unfolding paces and hence the inactivation tendency. In order to figure out the unbinding processes of AcAS, the iterative GNM was performed as described in the method section. For the sake of clarity, the connections between nodes are decomposed into two types: the intra-domain connections and the inter-domain connections. Obviously, the broken of the intra-domain connections in one domain mainly describe the unfolding of the domain itself, namely the unfolding of secondary structures. In contrast, the broken of the inter-domain connections stand for the unfolding around the interface between domains. Here, the unfolding curves for each domain and the interfaces between domains of AcAS are displayed in Figure 6a and b, respectively. The disconnection order for the intra and inter domain connections together describe the whole unfolding process of each domain.

According to Figure 6, before loss-number-of-native contact (LNNC) = 265, the inner connections of B and N domain disconnect quickly. The inter connections of B domain with A domain were also broken very fast, however, the connections between the N and A domain remained untouched until around the LNNC =500. The inter connections of N domain still remained 58% until the 1350th step and began to decrease quickly within 130 steps to zero. These

suggest that N domain is coupled well with A domain in AcAS. The C-domain collapsed rapidly at LNNC = 265 and LNNC =750 and became fully unfolded around the 850th step, at which the other domains still remained more than 50% intra contacts. The inter-domain connections between C and A domains were also disconnected simultaneously. It is indicated that the topology structure of C domain in AcAS are the least compact among the five domains. On the contrary, the secondary structures of B and B' domains broke slowly until the 1090th and 1540th step, respectively. This suggests the topology structures of B and B' domains are more compact than the others.

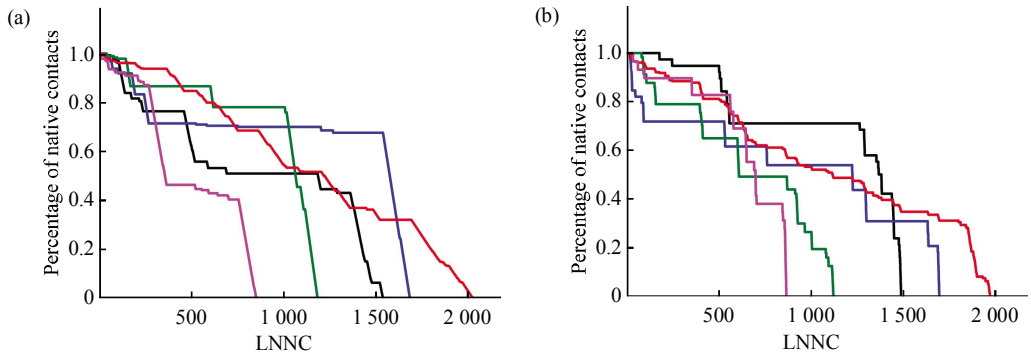

Fig. 6 The unfolding curves for each domain of AcAS

The unfolding curves for the intra and the inter domain contacts of AcAS are displayed in (a) and (b), respectively. —: N; —: B; —: B'; —: A; —: C.

Although the unfolding curves can tell the overall unfolding sequence for each domain, it cannot display the details of the unfolding process. In order to monitor the details in the unfolding process of AcAS, the contact map was plot and saved every 10 steps. Since there are about two hundred plots for the whole unfolding process of each AS, these plots are mainly submitted to the supplementary materials for the sake of saving space. Only the contact maps for the native states and the LNNC = 100, 200 and 500 states of AcAS are given by Figure 7a～d, respectively. Here we arbitrarily define a contact as the weak if it was broken within 500 steps, i.e. breaking approximately 20% of the total contacts. According to these figures, the kinetically weak regions of the AcAS were identified, and the results are displayed in Figure 8. Since the contacts within each secondary structure are relatively strong than those between two close secondary structures, most weak regions of each domain are found to be located around the interface between two domains

Understanding the information about these weak regions ahead of time will be quite helpful for our following design and validation cycles. The kinetically weak regions of AcAS are depicted by Figure 8. Being the starting point of the unfolding (Figure 6a), the contacts associated with the N and C domain of AcAS are of course apt to be broken. In domain N, the contacts among three α-helixes (region Ⅰ), i.e. ^{25}LSRER^{29}, ^{35}ADPAF^{39} and ^{70}ALVVQAA^{76}, were disconnected at the very beginning of the unfolding process. For domain C, there are three weak regions, namely P641～P642(region Ⅱ), P582～E584(region Ⅴ) and A598～L630 (region Ⅶ). Weak regions of N and C domain are mainly located in the intra-domain, which is consistent with the unfolding curve. Based on the unfolding curves for these two domains, it is found that they do not show large scale unfolding before the 500th step (Figure 6b). To exemplify the applications of the predicted weak regions on stability engineering, some suggestions on design were given as follow. A75 and A76 locate at an α-helix, forming a hydrophobic core with F39 at another α-helix. In DgAS, the corresponding positions are M61 and L62. According to the structural model of AcAS, it is found the packing between A75/A76 and F39 is not good enough. Based on calculation, substituting A75 and A76 with methionine residues could improve VDW interactions by 1.35 and 2.68 kcal•mol^{-1}, respectively. A598 is located at the β-turn of the C domain. Empirically, alanine residue is not a good turn-maker. According to Ramachandran plot, it is found that the dihedrals for the backbone of A598 are suitable for the proline residue. Because of its special side-chain,

proline residues are prone to locate at turns. Properly substituting a residue with a proline residue can notably decrease the conformational entropy of the unfolded state, and thereby stabilizing the protein.

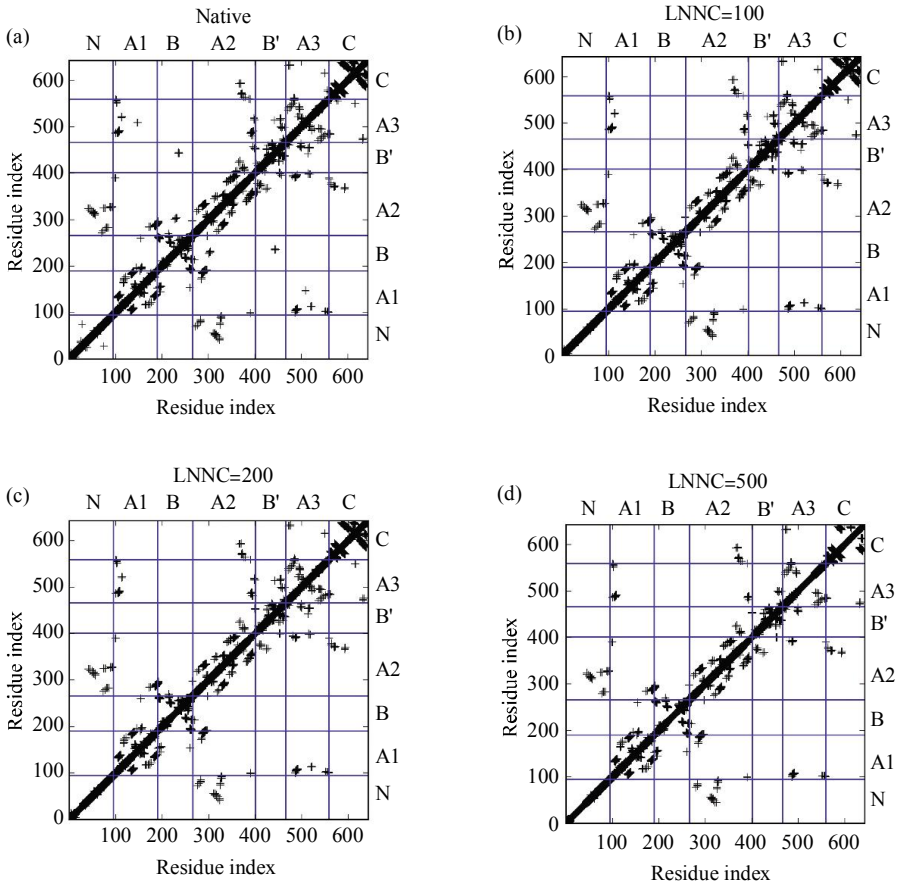

Fig. 7　The contact maps for the native (a), LNNC= 100 (b), 200 (c) and 500 (d) states of AcAS

Each contact is represented by a "+" mark.

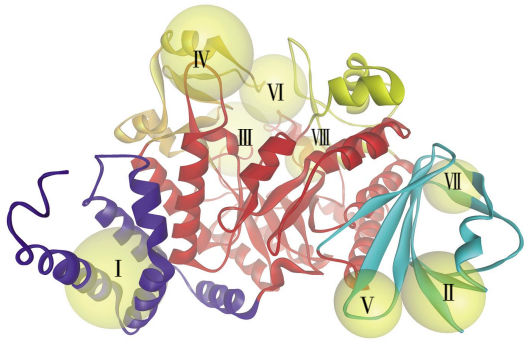

Fig. 8　Weak regions revealed by the iterative GNM for AcAS

Weak regions are represented by yellow spheres. The center of each sphere locates at the geological center of each weak region and the diameter of the sphere roughly reflects the size of the weak region. The used color code: catalytic domain A (red), domain N (blue), domain C (cyan), domain B (orange) and domain B' (yellow).

According to the contact map of the native state of AcAS (Figure 7a), it is found that there are only a few connections between the B and B' domain. The involving residues are I235, F236 and P237 in the B domain as well as F443, D445 in the B' domain (region Ⅵ). These connections were broken at the beginning of the unfolding process as well (around LNNC = 80). Actually, these residues constitutes the important geometrical door of the main active site of AcAS, therefore, the contacts between these residues are likely critical for the functions of the enzyme. Since these critical connections break at a very early stage, AcAS may become inactive at the beginning of the unfolding process. This may partly account for the reason why AcAS is notoriously susceptible to heat. Another weak region of domain B (region Ⅳ) is

located around the interface with the A domain (A2 part), which is composed of an α-helix in B domain and a loop in A-domain. The two weak regions of B domain explain why the unfolding curve for inter contacts of B domain drop sharply during the first 80 steps (Figure 6). It is known that the domain B and B' are very important for the activity and specificity of AS[13], therefore the connections among the B, B' and A domains are likely crucial for the enzyme functions and the weak regions Ⅳ and Ⅴ could be an ideal place for engineering to enhance the stability of AcAS. The rest two weak regions are located in the domain A:E147～P509(region Ⅲ) and S398～H519(region Ⅷ). These two regions are also important to the functions of AcAS. Therefore, these regions could also be suitable regions for engineering.

3 Conclusions

Compared with DgAS, AcAS exert its functions as a monomer. This can be mainly attributed to its lack of the extended loop (i.e. ^{341}RAHHG^{345} in DgAS) and two pivotal salt-bridges. In addition, the three residues longer loop (^{582}PPE^{584}) in AcAS may be harmful to the dimer formation due to the introduction of steric hindrance. On the whole, most functionally important residues of AcAS are identical to those of NpAS. This fact implicates the functions of AcAS are similar with those of NpAS. For example, with the conserved R233 in active site, the isomerization activity of AcAS may be similar with that of NpAS. On the other hand, owing to the weaker oligosaccharide binding at OB2 site and the defectiveness of OB3 site, the elongation activity of AcAS may be quite different with that of NpAS. This just account for the lower yield of insoluble glucan and increased yield of soluble maltooligosaccharides catalyzed by AcAS. According to the results of GNM, the structure of AcAS can be roughly divided into two regions. The motions of the two regions are negatively correlated, which suggests they move along different directions. These functional motions may be important to the binding of the receptor molecule and the glucose. Finally, through thoroughly investigating the unfolding pathway of AcAS, we discovered eight kinetically weak regions which could be suitable regions for engineering. In particular, the weak regions Ⅳ and Ⅴ, which locate at the interface between B and B'/A domain, are proposed as the ideal place for engineering to enhance the stability of AcAS.

In conclusion, the 3D structure of AcAS was modeled using the crystal structures of NpAS and DgAS as templates. The functions and dynamics properties of AcAS were then thoroughly investigated. The working mechanism of AcAS is similar with that of NpAS however with some distinct differences. The unfolding properties of AcAS were also carefully studied to reveal novel insight into rational design on thermostability. Based on the results of the iterative GNM, some kinetically weak regions of AcAS were identified. It was suggested these regions should be ideal places for stability engineering. Hopefully these discoveries can be useful to the industrial developments of AS.

References

[1] Henrissat B. A classification of glycosyl hydrolases based on amino-acid-sequence similarities. Biochem J, 1991, **280**(Pt2): 309−316

[2] Henrissat B, Davies G. Structural and sequence-based classification of glycoside hydrolases. Curr Opin Struct Biol, 1997, **7**(5): 637−644

[3] Cantarel B L, Coutinho P M, Rancurel C, *et al*. The carbohydrate-active enZymes database (CAZy): an expert resource for Glycogenomics. Nucl Acids Res, 2009, **37**(Database issue): D233−D238

[4] Hehre E J, Hamilton D M. Bacterial synthesis of an amylopectin-like polysaccharide from sucrose. J Biol Chem, 1946, **166**(2): 777−778

[5] MacKenzie C R, McDonald I J, Johnson K G. Glycogen metabolism in the genus *Neisseria*: synthesis from sucrose by amylosucrase. Can J Microbiol, 1978, **24**(4): 357−362

[6] Riou J Y, Guibourdenche M, Popoff M Y. A new taxon in the genus *Neisseria*. Ann Microbiol (Paris), 1983, **134B**(2): 257−267

[7] Potocki de Montalk G, Remaud-Simeon M, Willemot R M, *et al*. Amylosucrase from *Neisseria polysaccharea*: novel catalytic properties. FEBS Lett, 2000, **471**(2-3): 219−223

[8] Pizzut-Serin S, Potocki-Veronese G, van der Veen B A, *et al*. Characterisation of a novel amylosucrase from *Deinococcus* radiodurans. FEBS Lett, 2005, **579**(6): 1405−1410

[9] Emond S, Mondeil S, Jaziri K, *et al*. Cloning, purification and characterization of a thermostable amylosucrase from *Deinococcus geothermalis*. FEMS Microbiol Lett, 2008, **285**(1): 25−32

[10] Ha S J, Seo D H, Jung J H, *et al*. Molecular cloning and functional expression of a new amylosucrase from *Alteromonas* macleodii. Biosci Biotechnol Biochem, 2009, **73**(7): 1505−1512

[11] Seo D H, Jung J H, Choi H C, *et al*. Functional expression of amylosucrase, a glucan-synthesizing enzyme, from *Arthrobacter chlorophenolicus* A6. J Microbiol Biotechnol, 2012, **22**(9): 1253−1257

[12] Skov L K, Mirza O, Henriksen A, *et al*. Crystallization and

preliminary X-ray studies of recombinant amylosucrase from *Neisseria polysaccharea*. Acta Crystallogr D Biol Crystallogr, 2000, **56**(Pt 2): 203-205

[13] Skov L K, Mirza O, Henriksen A, *et al*. Amylosucrase, a glucan-synthesizing enzyme from the alpha-amylase family. J Biol Chem, 2001, **276**(27): 25273-25278

[14] Skov L K, Mirza O, Sprogoe D, *et al*. Oligosaccharide and sucrose complexes of amylosucrase. Structural implications for the polymerase activity. J Biol Chem, 2002, **277**(49): 47741-47747

[15] Jensen M H, Mirza O, Albenne C, *et al*. Crystal structure of the covalent intermediate of amylosucrase from *Neisseria polysaccharea*. Biochemistry, 2004, **43**(11): 3104-3110

[16] Hehre E J. Synthesis of a polysaccharide of the starch-glycogen class from sucrose by a cell free, bacterial enzyme system. J Biol Chem, 1949, **177**(1): 267-279

[17] Potocki de Montalk G, Remaud-Simeon M, Willemot R M, *et al*. Characterisation of the activator effect of glycogen on amylosucrase from *Neisseria polysaccharea*. FEMS Microbiol Lett, 2000, **186**(1): 103-108

[18] Preiss J, Ozbun J L, Hawker J S, *et al*. ADPG synthetase and ADPG-glucan 4-glucosyl transferase: enzymes involved in bacterial glycogen and plant starch synthesis. Ann N Y Acad Sci, 1973, **210**: 265-278

[19] De Montalk G P, Remaud-Simeon M, Willemot R M, *et al*. Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. J Bacteriol, 1999, **181**(2): 375-381

[20] Emond S, Andre I, Jaziri K, *et al*. Combinatorial engineering to enhance thermostability of amylosucrase. Protein Sci, 2008, **17**(6): 967-976

[21] Liu M, Wang S, Sun T G, *et al*. Insight into the structure, dynamics and the unfolding property of *Amylosucrases*: implications of rational engineering on thermostability. Plos One, 2012, **7** (7): e40441

[22] Larkin M A, Blackshields G, Brown N P, *et al*. Clustal W and clustal X version 2.0. Bioinformatics, 2007, **23**(21): 2947-2948

[23] Eswar N, Webb B, Marti-Renom M A, *et al*. Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci, 2007, **50**: 2.9.1-2.9.31

[24] Phillips J C, Braun R, Wang W, *et al*. Scalable molecular dynamics with NAMD. J Comput Chem, 2005, **26**(16): 1781-1802

[25] Laskowski R A, Macarthur M W, Moss D S, *et al*. Procheck - a program to check the stereochemical quality of protein structures. J Appl Cryst, 1993, **26**: 283-291

[26] Bahar I, Atilgan A R, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. Folding & Design, 1997, **2**(3): 173-181

[27] Miyazawa S, Jernigan R L. Estimation of effective interresidue contact energies from protein crystal-structures - quasi-chemical approximation. Macromolecules, 1985, **18**(3): 534-552

[28] Bahar I, Jernigan R L. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. J Mol Biol, 1997, **266**(1): 195-214

[29] Kundu S, Melton J S, Sorensen D C, *et al*. Dynamics of proteins in crystals: Comparison of experiment with simple models. Biophys J, 2002, **83**(2): 723-732

[30] Su J G, Li C H, Hao R, *et al*. Protein unfolding behavior studied by elastic network model. Biophys J, 2008, **94**(12): 4586-4596

[31] Su J G, Xu X J, Li C H, *et al*. An analysis of the influence of protein intrinsic dynamical properties on its thermal unfolding behavior. J Biomol Struct Dyn, 2011, **29**(1): 105-121

[32] De Montalk G P, Remaud-Simeon M, Willemot R M, *et al*. Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. J Bacteriol, 1999, **181**(2): 375-381

[33] Sarçabal P, Remaud-Simeon M, Willemot R-M, *et al*. Identification of key amino acid residues in *Neisseria polysaccharea* amylosucrase. FEBS Letters, 2000, **474**(1): 33-37

[34] Guerin F, Barbe S, Pizzut-Serin S, *et al*. Structural investigation of the thermostability and product specificity of amylosucrase from the bacterium *Deinococcus geothermalis*. J Biol Chem, 2012, **287**(9): 6642-6654

[35] Albenne C, Skov L K, Tran V, *et al*. Towards the molecular understanding of glycogen elongation by amylosucrase. Proteins-Structure Function and Bioinformatics, 2007, **66**(1): 118-126

[36] Jung J H, Seo D H, Ha S J, *et al*. Enzymatic synthesis of salicin glycosides through transglycosylation catalyzed by amylosucrases from *Deinococcus geothermalis* and *Neisseria polysaccharea*. Carbohydrate Res, 2009, **344**(13): 1612-1619

[37] Brooks B, Karplus M. Harmonic dynamics of proteins - normal-modes and fluctuations in bovine pancreatic trypsin-inhibitor. Proc Natl Acad Sci USA, 1983, **80**(21): 6571- 6575

[38] Keskin O, Durell S R, Bahar I, *et al*. Relating molecular flexibility to function: A case study of tubulin. Biophys J, 2002, **83**(2): 663-680

[39] Xu C Y, Tobi D, Bahar I. Allosteric changes in protein structure computed by a simple mechanical model: Hemoglobin T <-> R2 transition. J Mol Biol, 2003, **333**(1): 153-168

[40] Sluis-Cremer N, Temiz N A, Bahar I. Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. Curr Hiv Res, 2004, **2**(4): 323-332

[41] Loris R, Langhorst U, De Vos S, *et al*. Conserved water molecules in a large family of microbial ribonucleases. Proteins-Structure Function and Genetics, 1999, **36**(1): 117-134

[42] Demirel M C, Atilgan A R, Jernigan R L, *et al*. Identification of kinetically hot residues in proteins. Protein Sci, 1998, **7**(12): 2522-2532

淀粉蔗糖酶 AcAS 的工作机理与去折叠性质研究

刘 明[1] 李 平[2] 刘 斌[3] 苏计国[4]* 王存新[3]*

([1] 中国医学科学院北京协和医学院药物研究所，北京 100050；[2] 中国科学院过程工程研究所，北京 100190；
[3] 北京工业大学生命科学与生物工程学院，北京 100124；[4] 燕山大学理学院，秦皇岛 066004)

摘要 淀粉蔗糖酶(amylosucrase, AS)是一种葡萄糖基转移酶(E.C. 2.4.1.4)，隶属于糖苷水解酶 13 家族．当体系中存在葡聚糖时，AS 可以蔗糖为唯一底物催化合成直链葡聚糖．与 AS 相比，其他拥有合成直链淀粉样多糖的酶都需要利用昂贵的核苷酸激活糖来进行合成．这使得 AS 成为一种具有工业应用潜力的工业酶．尽管具有较大应用潜力，但是较弱的稳定性严重影响其在工业上的应用．本工作中，以 *NpAS* 和 *DgAS* 的晶体结构为模板，对氯酚节杆菌(*Arthrobacter chlorophenolicus*)中发现的一种新型 AS(AcAS)结构进行了模建．通过对 AcAS 与其他 AS 结构进行详细的比较，推断出 AcAS 的工作机理．随后，利用高斯网络模型(GNM)对其功能型运动进行了分析．根据 GNM 的结果，发现 AcAS 可分为两个运动方向相异的部分．最后，利用迭代高斯网络对 AcAS 的去折叠性质进行了研究．这些结果对随后的淀粉蔗糖酶的工业开发有一定帮助．

关键词 淀粉蔗糖酶，同源模建，去折叠，高斯网络模型，迭代高斯网络模型

学科分类号 Q6，Q51 **DOI**: 10.3724/SP.J.1206.2013.00044

* 通讯联系人.

苏计国. Tel: 0335-8057027, E-mail: jiguosu@ysu.edu.cn

王存新. Tel: 010-62757456, E-mail: cxwang@bjut.edu.cn

收稿日期：2013-02-01，接受日期：2013-04-01