Predicted Unfolding Order of the 13 α-Helices in the Catalytic Domain of Glucoamylase from *Aspergillus awamori* var. X100 by Molecular Dynamics Simulations

Hsuan-Liang Liu* and Wen-Chi Wang

Department of Chemical Engineering, National Taipei University of Technology, No. 1 s. 3 Chung-Hsiao E. Rd., Taipei, Taiwan

The unfolding mechanism of the 13 α-helices in the catalytic domain of *Aspergillus awamori* var. X100 glucoamylase was investigated by 200 ps molecular dynamics simulations in explicit water with temperature jump technique. Rather than a simultaneous event, the unfolding of these 13 α-helices followed a random ordered mechanism as α8→α1→α11→α7→α10→α3→α12→α13→α4→α5→α9→α6→α2. No significant relationships were found between the unfolding order and the length and the hydrophobicity of the helix. α-Helix 8 located in the inner region of the catalytic domain was predicted to be the first helix to unfold, indicating that the destruction of the secondary structure motif was initiated from the inner region of the catalytic domain. The dynamic behavior of these α-helices induced by increased kinetic energy during the unfolding process is considered to be similar to the expansion and compression of a series of springs under the influence of mechanical stress.

Introduction

The large number of interatomic interactions in force fields that have been taken into account for a realistic description of a protein in solution has made molecular dynamics (MD) technique a promising tool in examining the early events in protein folding/unfolding. Unfortunately, the fs time-step for the integration of the equation of motion limits the MD simulations to an extremely small time-span (e.g., ns or μs), which is relatively short compared to the second time-scale of the real protein folding/unfolding processes. It indicates that it is certainly still a long way to go to reach the purpose of simulating the folding/unfolding processes of a full-length protein "in silico". Instead of directly predicting protein folding processes, a lot of attention has been paid to unfolding simulations, in which the unfolding of a protein in its nativelike structure is usually initiated by raising the temperature or induced by changing the solvent (*1*−*4*). The idea of this approach is based on the assumption that protein folding and unfolding processes are reversible (*5*).

Although there is no guarantee that thermal denaturation is similar to the protein unfolding process under physiological conditions, temperature jump technique has still been widely applied to simulate the unfolding pathways for relatively small peptides and protein fragments in a rather small time-scale (*6*−*8*). According to the Arrhenius equation, where the reaction rate is about 2-fold faster when the temperature is increased by 10 °C, one may expect the rate of unfolding to be approximately 2^{10} (= 10^3)-, 10^6-, 10^9-, and 10^{12}-fold faster than is observed experimentally when the temperature is increased by 100, 200, 300, and 400 °C, respectively. For example, if a protein unfolded in a second time scale in vivo or in vitro, a MD simulation of protein unfolding in silico can be conducted within nanosecond time scale by increasing the protein denaturation temperature by 400 °C. This hypothesis makes temperature jump technique a powerful tool in simulating protein unfolding mechanism. A recent study of peptide folding was successfully carried out by Daura et al. (*9*), in which three 50 ns MD simulations were executed for a β-heptapeptide in a box containing 962 methanol molecules. This simulation indicates that the helical fraction is strongly temperature-dependent, meaning that the stability of the secondary structure is dramatically reduced when the temperature is raised from 298 to 350 K.

Glucoamylase (1,4-α-D-glucan glucohydrolase, EC 3.2.1.3, GA) catalyzes β-D-glucose release from the nonreducing ends of starch and related oligo- and polysaccharides by hydrolysis of the α-1,4-glucosidic linkages (*10*). GA is an industrially important, multidomain, N- and O-glycosylated enzyme (*10*). Both *A. awamori* and *A. niger* GAs, sharing the identical amino acid sequence, consist of three functional domains: (1) a globular N-terminal catalytic domain (CD) from residue 1 to 470; (2) a linear and extended linker region from residue 471 to 508; and (3) a globular C-terminal starch-binding domain (SBD) from residue 509 to 616 (*11*). The 3D structure of the CD from *A. awamori* var. X100 GA has been solved by X-ray crystallography in its native state at different pH values and in complexes with different inhibitors (*12*−*17*). There are 13 α-helices with limited regions of antiparallel β-strands in the vicinity of the active site (Figure 1) (*12*). Twelve of these α-helices (α-helix 11 excluded) fold into an unusual $(\alpha/\alpha)_6$-barrel, with the core of the barrel defining a pocket containing the general catalytic acid, Glu179 (*18, 19*), and the catalytic base, Glu400 (*12, 14, 16*). α-Helices 1, 3, 5, 7, 9, and 12 are located on the surface of the $(\alpha/\alpha)_6$-barrel, whereas

* To whom correspondence should be addressed. Phone: +886-2-2771-2171 ext. 2542. Fax: +886-2-2731711. E-mail: f10894@ntut.edu.tw.

10.1021/bp034045q CCC: $25.00
Published on Web 07/04/2003

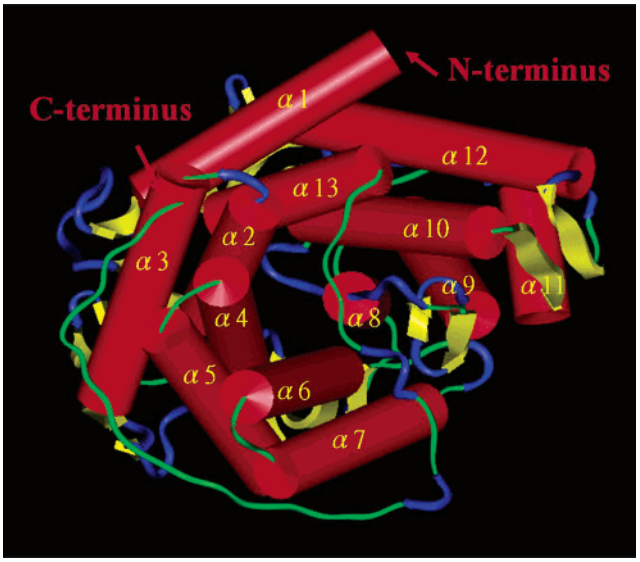

Figure 1. X-ray crystallographic structure of the CD from *A. awamori* var. X100 (*12, 13*) visualized by Insight II. The N- and C-termini are indicated. α-Helices 1 to 13 shown in red cylinders are numbered in the sequential order from the N- to the C-terminus. All α-helices except α-helix 11 form the $(\alpha/\alpha)_6$-barrel with the core of the barrel defining a pocket where the catalytic center is located. *β*-Strands are presented as yellow arrows pointing from the N- to the C-terminus. The polypeptide backbones belonging to the turn and random coil regions are shown in blue and green, respectively.

α-helices 2, 4, 6, 8, 10, and 13 are located in the interior region of the $(\alpha/\alpha)_6$-barrel. It is noteworthy that α-helix 11 does not belong to the $(\alpha/\alpha)_6$-barrel motif.

Many efforts have been made to improve the substrate selectivity (*20, 21*) and thermostability (*22*−*26*) of GA for better industrial applications. The CD of GA undergoes irreversible thermoinactivation at moderately acidic pH and 70 °C (*26*). Neither deamidation (*27, 28*) nor Asp-X peptide bond hydrolysis (*29*) determines the thermoinactivation rate. In addition, mismatched disulfide bonds are not observed at 70 °C and pH 3.5 and 4.5. These results indicate that some other mechanism must account for the thermoinactivation of the CD at high temperatures. The mechanism appears to be a change in the GA's secondary or tertiary structure, such as unfolding, that destroys the integrity of the active site (*26, 30*). In the present study, 200 ps MD simulations were conducted to predict the unfolding order of the 13 α-helices in the CD of GA from *A. awamori* var. X100 by temperature jump technique.

Materials and Methods

The initial structure of the CD of GA, consisting of 471 amino acid residues, from *A. awamori* var. X100 was taken from the X-ray crystallographic structure from the Protein Data Bank (PDB entry 1GLM) (*12*). The crystallographic CD structure was modeled in the SGI O200 workstation (Silicon Graphics, Inc., Mountain View, CA)−Insight II (Accelyrs, San Diego, CA) system with the force field Discover CVFF (consistent valence force field) (*31*−*33*). The force field parameters in CVFF, coverage of macromolecules including proteins, nucleic acids, carbohydrates, and lipids, were developed by computing the properties of nearly 2000 different molecules resulting in over two million quantum mechanically computed energies and energy derivatives. The crystallographic CD structure was subjected to energy minimization calculations by steepest descent method with 3,000 iterations followed by Newton−Raphson method with 5,000 iterations to be used as the initial lowest energy structure for further structural comparison. The energy-minimized CD was then submitted to 200 ps MD simulations in explicit water system with a 5-Å water layer after equilibrating for about 10 ps at pH 4.5 and 300, 400, 600, and 800 K, respectively, using the Discover module of the Insight II program. The temperature and pressure were maintained for each MD simulation by weak coupling the system to a heat bath and an external pressure bath at 1 atm according to the method described by Berendsen et al. (*34*). The time-step of the MD simulations was 1 fs. The trajectories and coordinates of the catalytic domain were saved every 2 ps for further analysis. All simulations were performed on SGI O200 workstation with 64-bit HIPS RISC R12000 270 MHz CPU and PMC-Sierra RM7000A 350 MHz processor. About 30 s of CPU time was required for one 1 fs simulation, leading to a total running time of 70 CPU days for each MD simulation.

For each MD simulation, the root-mean-square deviations (RMSDs) of the trajectories recorded every 2 ps interval were calculated for the backbone C_α atom of the CD during the course of 200 ps MD simulations at different temperature with reference to the starting structure according to the following equation (*35*):

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(||x(i) - y(i)||^2)}$$

where N is the number of total atoms in the list of equivalence, and x(i) and y(i) are the ith atomic coordinates in proteins A and B, respectively. The RMSDs were calculated after optimal superimposition of the coordinates to remove translational and rotational motion (*36*). The starting structure of the MD simulations, which was generated from about 10 ps equilibration, generally had a folding structure very similar to that of the native crystallographic structure with RMSD < 0.5 Å (data not shown). The secondary structure (helix, sheet, turn, and random coil) was predicted on the basis of the Kabsch−Sander algorithm (*37*), in which pattern recognition of hydrogen-bonded was correlated to the geometrical features. The default hydrogen bonding energy criterion of −0.5 kcal/mol was used. The helicity, $H(t)$, for each α-helix was defined as the ratio of the number of the residual cooperative H-bonds at time t to the number of the cooperative H-bonds in the starting structure. The average helicity $\langle H \rangle$ was then defined by $1/t_f \int_0^{t_f} H(t)\, dt$, with t_f as the total length of the MD simulation. The melting temperature, T_m, for each α-helix, calculated according to Arrhenius equation, was defined as the temperature at which the α-helix retained 50% average helicity $\langle H \rangle$. The N→C distance for each α-helix was defined as the linear distance between the two C_α of the N- and C-end residues of the α-helix.

Results and Discussion

The starting structure was obtained from the X-ray crystallographic structure of the CD from *A. awamori* var. X100 followed by energy minimization. By superimposition of the starting and the crystallographic structures, we found that the RMSD of the backbone C_α is within 0.5 Å (data not shown), indicating that these two structures are very similar and that the choice of the energy minimized structure as the starting structure is reasonable. The total energy of the stating structure

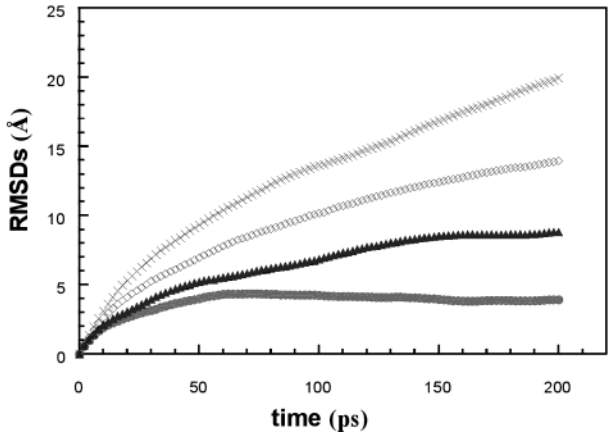

Figure 2. The root-mean-square deviations (RMSDs) of the entire backbone C_α with reference to the starting structure of CD during the 200 ps MD simulations at 300 (●), 400 (▲), 600 (◇), and 800 K (×).

decreased from 15,000 kcal/mol for the crystallographic structure to about 9,000 kcal/mol, mainly as a result of loss of the nonbonded energy, which accounts for most of the interactions between water molecules in the crystallographic environment and the polar or charged groups in the CD. Figure 2 shows the RMSDs of the entire backbone C_α with reference to the starting structure in the course of 200 ps MD simulations at different temperatures. The RMSDs increased from 0 to about 4 Å at 300 K within 60 ps and remained at 4 Å afterward. It indicates that the dynamic property made the CD keep sampling different conformations in its folded state and finally reached a steady state conformation, which is slightly different from its starting structure. It also indicates that the increased kinetic energy obtained by raising temperature compensated for the increased entropy, leading to spontaneous structural fluctuations, even at low temperature such as 300 K. In contrast, the RMSDs increased dramatically during the course of the 200 ps MD simulations, with the final RMSDs reached 8, 14, and 20 Å at 400, 600, and 800 K, respectively. Higher temperature provides more kinetic energy to overcome the energy barrier between the folded and the transition states, leading to faster unfolding rate. It indicates that the strategy to induce the unfolding of the CD by temperature jump technique is successful.

The 13 α-helices were defined by a series of at least four residues, the ϕ and φ angles of which are within ±30° of values typical for an α-helix ($\phi = -64°$, $\varphi = -40°$) as in the crystallographic structure of the CD (*12, 13*). The C-terminal regions of these helices in the CD are often distorted toward a 3_{10}-conformation. Some residues belonging to the distorted C-termini of helices were classified as part of the helices, whereas some of them were defined as reverse turns according to Aleshin et al. (*13*). In general, α-helices have negatively charged residues at their N-termini and positively charged ones at their C-termini (*38*−*40*). Therefore, one may expect that the N- and C-termini of these α-helices will unfold faster than the middle region. α-Helix 11 is the only helix not participating in the $(\alpha/\alpha)_6$-barrel, so its absence in some GA families suggests that it plays no major structural or functional role (*40*). The inner set of α-helices (labeled α2, α4, α6, α8, α10, and α13 in Figure 1 and Table 1) in the $(\alpha/\alpha)_6$-barrel are approximately parallel to one another, while the outer set of helices taken as a group (labeled α1, α3, α5, α7, α9, and α12 in Figure 1 and Table 1) are approximately antiparallel to the inner set of helices (*13*). Such well-organized secondary structure motif should allow the CD to be resistant to thermal denaturation at elevated temperature, as found for the extremely thermostable α-amylase, which folds into the $(\alpha/\beta)_8$-barrel. In contrast, CD was found to undergo unexpected thermal denaturation at moderate temperature. It makes the investigation of the unfolding process of the 13 α-helices in the CD induced by temperature jump more significant.

Table 1 presents the total number of residues; the hydrophobicity, *H*, calculated on the basis of the hydrophobicity index (*41*); the melting temperature, T_m, obtained by extrapolating the obtained Arrhenius equation; the predicted unfolding order; and the corresponding amino acid sequence of each α-helix. The unfolding process follows the order of α8→α1→α11→α7→α10→α3→α12→α13→α4→α5→α9→α6→α2. It is obvious that the unfolding order is not necessarily correlated to the number of residues and hydrophobicity of the α-helix. Previous studies have shown that an amino acid with higher hydrophobicity has the higher propensity to form an ideal α-helix (*42*−*44*) in solution. In the present study, we do not find the tendency that an α-helix with a higher hydrophobicity is more resistant to unfolding. It suggests that the factors influencing the stability of the secondary structure in a protein are different from that for a short peptide or a protein fragment in solution.

Although α-helix 8 is located in the inner set of the $(\alpha/\alpha)_6$-barrel, it was found to be the first α-helix to unfold at extremely low temperature, probably due to its short helical length (i.e., both α-helices 8 and 11 consist of only 10 amino acid residues). In addition, the hydrophobicity of α-helix 8 is 2.1 (*41*), which is relatively low compared to the thermally stable α-helix 6, whose hydrophobicity is 21. Table 1 also shows that α-helix 1 unfolded very quickly with $T_m = 308$ K. It is not surprising since it is located at the N-terminus of the CD. Previous studies have shown that the N-terminal region of the CD is extremely diversified. For example, Svensson et al. (*45*), discovered that about 50% of the GAs from *A. niger* lack the three residues from the N-terminus. Aleshin et al. (*12, 13*) found that the partial proteolysis of the first three residues from the N-terminus may be largely responsible for the high thermal parameters that persist over the first three to four residues from the N-terminus. In, addition, the hydrophobicity of α-helix 1 is only 0.7, which is relatively low compared to the other thermostable helices. α-Helix 1 is located on the surface of the CD and is subsequently connected by a thermal labile region between residues 20 and 30 (*24*), leading to low resistance to thermal unfolding. This result also indicates that the stability of the α-helix attributes mainly to its location and local environment in the CD. α-Helix 11 was predicted to be the third helix to unfold in the present study probably because it is one of the two shortest α-helices in the CD and that it is the only α-helix not located in the $(\alpha/\alpha)_6$-barrel and is not protected by the well-defined secondary structure motif from thermal unfolding.

Three intrinsic disulfide bonds exist between the following pairs of Cys residues: (1) 222 and 469; (2) 210 and 213; and (3) 262 and 270 in the CD (*12, 13*). The dihedral angles observed for these disulfide bonds are near the typical values of −90° and +90° (*45*). In addition, the distances between the C_α of these Cys residues involved in disulfide linkages range from 4.83 to 5.45 Å, which are typical for pairs of disulfide-linked Cys residues (*45*). This evidence indicates that the three disulfide linkages may contribute to the structural stability of the CD to a great extent. Table 1 shows that two of the above

Table 1. Number of Amino Acid Residues, Hydrophobicity, Predicted Melting Temperature, Predicted Unfolding Order, and Amino Acid Sequence of Each of the 13 α-Helices in the CD of GA from *A. awamori* var. X100[e]

α-helix[a]	number of residues	H[b]	T_m (K)[c]	U/O[d]	amino acid sequence
α1	21	0.7	308	2	1-ATLDSWLSNEATVARTAILNN-21
α2	16	4.6	619	13	53-TRDSGLVIKTLVDLFR-68
α3	19	4.7	403	6	72-TDLLSTIEHYISSQAIIQG-90
α4	20	3.1	459	9	126-DGPALRATAMIGFGQWLLDN-145
α5	22	3.5	461	10	148-TSAATEIVWPLVRNDLSYVAQY-169
α6	21	21.0	505	12	186-FFTIAVQHRALVEGSAFATAV-206
α7	17	1.5	350	4	211-SWC*DSQAPQILC*YLQSF-227
α8	10	2.1	281	1	246-SNTLLGSIHT-255
α9	16	-11.2	461	11	272-PRALANHKQVVDSFRS-287
α10	21	-1.4	358	5	318-FLCTLAAAEQLYDALYQWDKQ-338
α11	10	7.5	311	3	345-DVSLDFFKAL-354
α12	24	13.1	410	7	368-STYSSIVSAVKTFADGFVSIVETH-391
α13	15	-10.9	450	8	416-TTSYAALLTAANRRN-430

[a] The numbering of the α-helices is from the N− to the C-terminus according to Aleshin et al. (*12*). [b] H = hydrophobicity of the helix calculated by the sum of the hydrophobicity index (*41*) for each amino acid. [c] T_m = melting temperature of the helix, at which the average helicity over the 200 ps MD simulation is 50%, predicted by extrapolating the obtained Arrhenius plot. [d] U/O = The unfolding order is predicted according to the predicted T_m, with lower T_m corresponding to faster unfolding order. [e] α-Helice 1, 3, 5, 7, 9, and 12 are located on the outer region of the $(\alpha/\alpha)_6$-barrel (unshaded rows), whereas α-helice 2, 4, 6, 8, 10, and 13 are located in the inner region of the $(\alpha/\alpha)_6$-barrel (light shaded rows). α-Helix 11 is not part of the $(\alpha/\alpha)_6$-barrel (dark shaded row). (*) The two Cys residues Cys213 and Cys222, located in α-helix 7, form two disulfide linkages with Cys210 and Cys469, respectively, in the crystallographic structure of the CD.

six Cys residues, Cys213 and Cys222, are located in α-helix 7. One may expect that α-helix 7 is more thermostable than the other helices because of the covalent disulfide linkages. In contrast, α-helix 7 was predicted to be the fourth helix to unfold when induced by temperature jump. In order for Cys213 and Cys222 located in α-helix 7 to maintain their disulfide linkages with Cys210 and Cys469, respectively, α-helix 7 has to adjust its conformation as compensation when the backbone of the CD keeps expanding as a result of the increased kinetic energy at elevated temperatures. Protein seems to have the tendency to maintain the proper disulfide linkages through the loss of part of its secondary structures to obtain a higher structural stability. At first glance, one may expect that the outer set of α-helices in the $(\alpha/\alpha)_6$-barrel unfold faster than the inner set of α-helices. In contrast, we found that the location of each α-helix in the $(\alpha/\alpha)_6$-barrel does not necessarily determine its unfolding order (Table 1). Therefore, we propose here that the unfolding process of the 13 α-helices in the CD follows a random ordered mechanism.

Figure 3 shows the dynamics of the N→C distances, the RMSDs, and the helicities of α-helices 1, 2, 8, and 11 at 300, 400, 600, and 800 K during the 200 ps MD simulations. As can be seen, the N→C distance fluctuated slightly at 300 K, whereas it fluctuated more significantly when the temperature is higher, especially at 600 and 800 K. It indicates that lengths of these α-helices were constantly expanded or compressed during the MD simulations because of the destruction of the H bonds in the secondary structures. The behavior of these α-helices' response to the increased temperatures is believed to be similar to a series of springs responding to mechanical stress. The increased kinetic energy by raising the temperature results in destabilization of the well organized parallel hydrogen bonds in an α-helix. The expansion or compression of each α-helix depends mainly on its local environment and on its interactions with other portions of the polypeptide chain. The RMSD for α-helix 1 increased dramatically at the temperatures of 400 K and higher, whereas it remained lower than 2 Å for α-helix 2 between 300 and 600 K. It indicates that α-helix

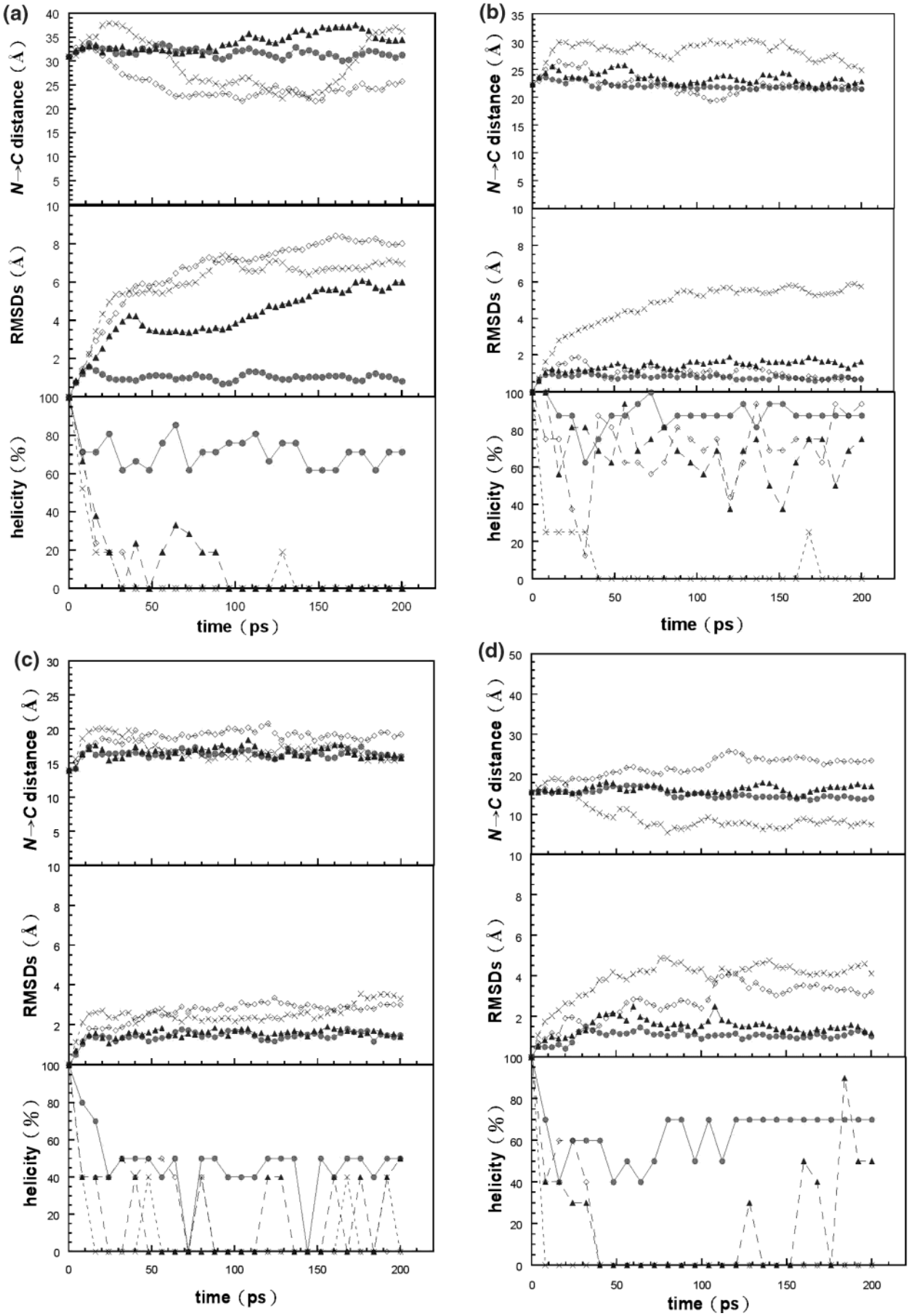

Figure 3. The N→C distances (upper panel), RMSDs (middle panel), and helicities (lower panel) of (a) α-helix 1; (b) α-helix 2; (c) α-helix 8; and (d) α-helix 11 at 300 (●), 400 (▲), 600 (◇), and 800 K (×) during the 200 ps MD simulations. The N→C distance, the RMSD, and the helicity of each α-helix were defined in the text.

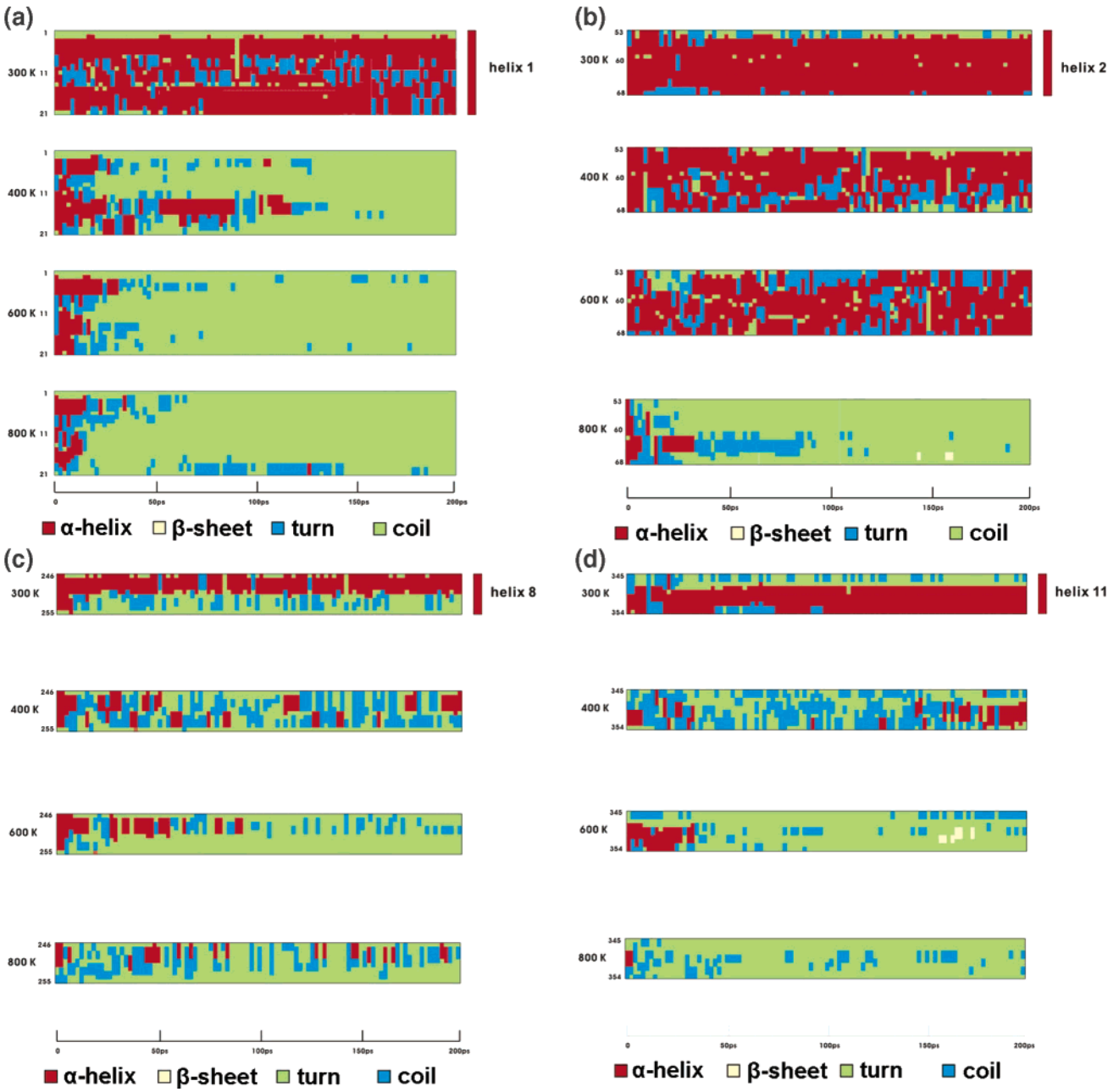

Figure 4. Secondary structures as a function of MD simulation time for (a) α-helix 1; (b) α-helix 2; (c) α-helix 8; and (d) α-helix 11 at 300, 400, 600, and 800 K. The numbers shown at the uppermost and the lowest end points on the *y*-axis indicate the N- and C-end residues of the corresponding α-helix. α-Helix, *β*-sheet, turn, and coil estimated according to the Kabsch and Sander algorithm (*37*) are shown in red, light yellow, blue, and green, respectively.

1 undergoes a significant conformation change at low temperature, whereas α-helix 2 maintains the stable conformation at its folded state at elevated temperatures. The RMSDs for both α-helice 8 and 11 were not as large as α-helix 1, probably because of their short lengths. The helicity decreased sharply for α-helices 1, 8, and 11 at high temperatures, whereas it remained above 60% for α-helix 2 between 300 and 600 K. This evidence again suggests that α-helix 2 is one of the most stable helices in the CD.

Figure 4 shows the dynamics of the α-helices 1, 2, 8, and 11 predicted on the basis of the method described by Kabsch and Sander (*37*) at 300, 400, 600, and 800 K. The type of secondary structure for three N-terminal residues in α-helix 1 has switched from helix as defined in the crystallographic structure (*12, 13*) to random coil even at 300 K. This result is consistent with the previous findings that these residues are very unstable (*12, 13*). α-Helix 2 maintained a well organized helix structure at temperatures up to 600 K and subsequently turned to random coil at 800 K. Both α-helices 8 and 11 have 10 residues, but the unfolding directions were opposite. The unfolding of α-helices 8 and 11 began from the C- and N-end residues, respectively. As can be seen from Table 1, the side chains of residues Ser252, His254, and Thr255 located at the C-terminal half of α-helix 8 are either polar or charged; therefore it is not favorable for them to reside in the apolar helix configuration. Similarly, the charged and polar side chains of Asp345 and Ser347 located near the N-terminal result in the destabilization of the helix conformation.

In conclusion, the unfolding of the 13 α-helices in the CD of *A. awamori* var. X100 GA were predicted to follow the random ordered mechanism with the unfolding sequence of α-helices 8, 1, 11, 7, 10, 3, 12, 13, 4, 5, 9, 6, and 2. It indicates that whether an α-helix is located in

the inner or the outer sets of the $(\alpha/\alpha)_6$-barrel seems to have little effect on its structural stability. No significant correlation between unfolding order and the number of residues and the hydrophobicity for the α-helix was found. α-Helix 8, located in the inner region of the catalytic domain, was predicted to be the first helix to unfold, indicating that the destruction of the secondary structure motif was initiated from the inner region of the catalytic domain. The dynamic behavior of these α-helices induced by increased kinetic energy during the unfolding process is considered to be similar to the expansion and compression of a series of springs under the influence of mechanical stress. The strategy to predict the unfolding order of the 13 α-helices of the catalytic domain from *A. awamori* var. X100 GA in this work can be applied to the other proteins or protein domains in order to obtain the structural/functional relationships.

Acknowledgment

The authors thank National Science Council of Taiwan for financial support (project NSC-89-2214-E−027-006). We also thank Fu-Yuan Cheng, Chia-Ming Hsu, Chia-Yuen Hu, Ming-Yi Hwang, Kauo-Cheng Liu, and Yen-Chi Su of NTUT for data analysis.

References and Notes

(1) Mark, A. E.; van Gunsteren, W. F. Simulation of the thermal denaturation of HEW-lysozyme: Trapping the *molten globule* state. *Biochemistry* **1992**, *31*, 7745−7748.

(2) Caflisch, A.; Karplus, M. Molecular-dynamics simulation of protein denaturation−Solvation of the hydrophobic cores and secondary structure of barnase. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1746−1750.

(3) Caflisch, A.; Karplus, M. Acid and thermal denaturation of barnase investigated by molecular dynamics simulations. *J. Mol. Biol.* **1995**, *252*, 672−708.

(4) Li, A.; Daggett, V. Identification and characterization of the unfolding transition state of chymotrypsin inhibitor 2 by molecular dynamics simulations. *J. Mol. Biol.* **1996**, *257*, 412−429.

(5) Finkelstein, A. V. Can protein unfolding simulate protein folding? *Protein Eng.* **1997**, *10*, 843−845.

(6) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. Fast events in protein-folding: helix melting and formation in a small peptide. *Biochemistry* **1996**, *35*, 691−697.

(7) Muñoz, V.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. A statistical-mechanical model for beta-hairpin kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5872−5879.

(8) Thompson, P. A.; Eaton, W. A.; Hofrichter, J. Laser temperature-jump study of the helic-reversible-arrow-coil kinetics of an alanine peptide interpreted with a kinetic zipper model. *Biochemistry* **1997**, *36*, 9200−9210.

(9) Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Reversible peptide folding in solution by molecular dynamics simulations. *J. Mol. Biol.* **1998**, *280*, 925−932.

(10) Liu, H.-L. Structural/functional relationships and protein engineering of glucoamylase. *Recent Res. Devel. Protein Eng.* **2001**, *1*, 75−106.

(11) Coutinho, P. M.; Reilly, P. J. Glucoamylase structural, functional, and evolutionary relationships. *Proteins* **1997**, *29*, 334−347.

(12) Aleshin, A.; Golubev, A.; Firsov, L. M.; Honzatko, R. B. Crystal structure of glucoamylase from *Aspergillus awamori* var. *X100* to a 2.2-Å resolution. *J. Biol. Chem.* **1992**, *267*, 19291−19298.

(13) Aleshin, A. E.; Hoffman, C.; Firsov, L. M.; Honzatko, R. B. Refined crystal structures of glucoamylase from *Aspergillus awamori* var. *X100. J. Mol. Biol.* **1994**, *238*, 575−591.

(14) Aleshin, A. E.; Firsov, L. M.; Honzatko, R. B. Refined structure for the complex of acarbose with glucoamylase from *Aspergillus awamori* var. *X100* to 2.4-Å resolution. *J. Biol. Chem.* **1994**, *269*, 15631−15639.

(15) Aleshin, A. E.; Stoffer, B.; Firsov, L. M.; Svensson, B.; Honzatko, R. B. Crystallographic complexes of glucoamylase with maltooligosaccharide analogues: Relationship of stereochemical distortions at the nonreducing end to the catalytic mechanism. *Biochemistry* **1996**, *35*, 8319−8328.

(16) Harris, E. M. S.; Aleshin, A. E.; Firsov, L. M.; Honzatko, R. B. Refined structure for the 1-deoxynojirimycin with glucoamylase from *Aspergillus awamori* var. *X100* to 2.4-Å resolution. *Biochemistry* **1993**, *32*, 1618−1626.

(17) Stoffer, B.; Aleshin, A. E.; Firsov, L. M.; Svensson, B.; Honzatko, R. B. Refined structure for the D-*gluco*-dihydroacarbose with glucoamylase from *Aspergillus awamori* var. *X100* to 2.2-Å resolution: dual conformations for extended inhibitors bound to the active site of glucoamylase. *FEBS Lett.* **1995**, *358*, 57−61.

(18) Sierk, M. R.; Ford, C.; Reilly, P. J.; Svensson, B. Catalytic mechanisn of fungal glucoamylas as defined by mutagenesis of Asp176, Glu179 and Glu180 in the enzyme from *Aspergillus awamori*. *Protein Eng.* **1990**, *3*, 193−198.

(19) Svensson, B.; Clarke, A. J.; Svendsen, I.; Møller, H. M. Identification of carboxylic acid residues in glucoamylase G2 from *Aspergillus niger* that participate in catalysis and substrate binding. *Eur. J. Biochem.* **1990**, *188*, 29−38.

(20) Liu, H.-L.; Coutinho, P. M.; Ford, C.; Reilly, P. J. Mutations to alter *Aspergillus awamori* glucoamylase selectivity. III. Asn20→Cys/Ala27→Cys, Ala27→Pro, Ser30→Pro, Lys108→Arg, Lys108→Met, Gly137→Ala, 311−314Loop, Tyr312→Trp and Ser436→Pro. *Protein Eng.* **1998**, *11*, 389−398.

(21) Liu, H.-L.; Ford, C.; Reilly, P. J. Mutations to alter *Aspergillus awamori* glucoamylase selectivity. IV. Combinations of Asn20→Cys/Ala27→Cys, Ser30→Pro, Gly137→Ala, 311−314Loop, Ser411→Ala and Ser436→Pro. *Protein Eng.* **1999**, *12*, 163−172.

(22) Chen, H.-M.; Li, Y.; Panda, T.; Buehler, F. U.; Ford, C.; Reilly, P. J. Effect of replacing helical glycine residues with alanines on the thermostability and production of *Aspergillus awamori* glucoamylase. *Protein Eng.* **1996**, *9*, 499−505.

(23) Li, Y.; Reilly, P. J.; Ford, C. Effect of introducing proline residues on the stability of *Aspergillus awamori* glucoamylase. *Protein Eng.* **1997**, *10*, 1199−1204.

(24) Li, Y.; Coutinho, P. M.; Ford, C. Effect on thermostability and catalytic activity of introducing disulfide bonds into *Aspergillus awamori* glucoamylase. *Protein Eng.* **1998**, *11*, 661−667.

(25) Allen, M. J.; Coutinho, P. M.; Ford, C. F. Stabilization of *Aspergillus awamori* glucoamylase by praline substitution and combining stabilizing mutations. *Protein Eng.* **1998**, *11*, 783−788.

(26) Liu, H.-L.; Doleyres, Y.; Coutinho, P. M.; Ford, C.; Reilly, P. J. Replacement and deletion mutations in the catalytic domain and belt region of *Aspergillus awamori* glucoamylase to enhance thermostability. *Protein Eng.* **2000**, *13*, 655−659.

(27) Chen, H.-M.; Ford, C.; Reilly, P. J. Substitution of asparagine residues in *Aspergillus awamori* glucoamylase by site-directed mutagenesis to eliminate N-glycosylation and inactivation by deamidation. *Biochem. J.* **1994**, *301*, 275−281.

(28) Chen, H.-M.; Bakir, U.; Reilly, P. J.; Ford, C. Increased thermostability Asn182→Ala mutant *Aspergillus awamori* glucoamylase. *Biotechnol. Bioeng.* **1994**, *43*, 101−106.

(29) Chen, H.-M.; Ford, C.; Reilly, P. J. Identification and elimination by site- directed mutagenesis of thermolabile aspartyl bonds in *Aspergillus awamori* glucoamylase. *Protein Eng.* **1995**, *8*, 575−582.

(30) Williamson, G.; Belshaw, N. J.; Noel, T. R.; Ring, S. G.; Williamson, M. P. *O*-glycosylation and stability. *Eur. J. Biochem.* **1992**, *207*, 661−670.

(31) Peng, Z.; Ewig, C. S.; Hwang, M.-J.; Waldman, M.; Hagler, A. T. Derivation of class II force fields. 4. van der Waals parameters of alkali metal cations and halide anions. *J. Phys. Chem. A* **1997**, *101*, 7243−7252.

(32) Maple, J. R.; Hwang, M.-J.; Jalkanen, K. J.; Stockfisch, T. P.; Hagler, A. T. Derivation of class II force fields: V. Quantum force field for amides, peptides, and related compunds. *J. Comput. Chem.* **1998**, *19*, 430−458.

(33) Hwang, M.-J.; Ni, X.; Waldman, M.; Ewig, C. S.; Hagler, A. T. Derivation of class II force fields. VI. carbohydrate compounds and anomeric effects. *Biopolymers* **1998**, *45*, 435−468.
(34) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Comput. Phys.* **1984**, *81*, 3684−3690.
(35) Koehi, P. Protein structure similarities. *Curr. Opin. Struct. Biol.* **2001**, *11*, 348−353.
(36) Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A* **1976**, *32*, 922−923.
(37) Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577−2637.
(38) Argos, P.; Palau, J. Amino acid distribution in protein secondary structures. *Int. J. Peptide Protein Res.* **1982**, *19*, 380−393.
(39) Richardson, J. S.; Richardson, D. C. Amino acid preferences for specific locations at the ends of alpha-helices. *Science* **1988**, *240*, 1648−1652.
(40) Coutinho, P. M.; Reilly, P. J. Structural similarities in glucoamylase by hydrophobic cluster analysis. *Protein Eng.* **1994**, *7*, 749−760.
(41) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105−132.
(42) Jasanoff, A.; Fersht, A. R. Quantitative determination of helical propensities from trifluoroethanol titration curves. *Biochemistry* **1994**, *33*, 2129−2135.
(43) Chakrabartty, A.; Baldwin, R. L. Stability of α-helices. *Adv. Protein Chem.* **1995**, *46*, 141−176.
(44) Hamada, D.; Kuroda, Y.; Tanaka, T.; Goto, Y. High helical propensity of the peptide fragments derived from *β*-lactoglobulin, a predominantly *β*-sheet protein. *J. Mol. Biol.* **1995**, *254*, 737−746.
(45) Svensson, B.; Pedersen, T. G.; Svendsen, I.; Sakai, T.; Ottesen, M. Characterization of two forms of glucoamylase from *Aspergillus niger*. *Carlsberg Res. Commun.* **1982**, *47*, 55−69.

Accepted for publication June 3, 2003.

BP034045Q