Comparison of Protein Internal Motion by Inter-helical Motional Correlations and Hydrogen Bond Ratio

Byoungkook Kim & Chang No Yoon*
Biocomputing Lab., Center for Biotransformation and Biocatalysis,
Korea Institute of Science and Technology
39-1 Hawolgok-dong, Seongbuk-gu, Seoul, 136-79, Republic of Korea
** Correspondance Email : cody@kist.re.kr, Tel : +82-2-958-5068, Fax : +82-958-5059*

ABSTRACT: Internal motion of the protein has been described in many papers with C_α correlation coefficients to find motional correlation and functional characteristics. To describe the secondary structural motion and stability in protein, we have studied molecular dynamics (MD) simulations on FADD Death Domain and FADD Death Effector Domain which have a similar structure but have different functional characteristics. After 10ns MD simulations, the inter-helical motional correlations and the hydrogen bond ratios were compared between the two domains. From these data we could distinctly compare the internal motions of them and could explain the differences in experimental thermodynamic melting behaviors at molecular level.

1 INTRODUCTION

Atomic motions and their correlated character in proteins are important to understanding of their biological function [1]. Although difficult to obtain experimentally, this information may easily be extracted from molecular dynamics (MD) simulations. To identify the groups of atoms moving in concert and conformational changes occurring upon thermal unfolding of a protein, atomic positional fluctuation and cross-correlation analysis techniques has been used for during past 20 years [2]. However the fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations, the driving force of the correlated internal motions was not explicitly explained. In this study we used inter-helical motional correlation and hydrogen bond dynamics during MD simulation to compare the difference in protein melting behavior.

Two protein domains from Fas-associated death domain protein (FADD) which have similar structures and different melting behaviors are selected for the MD simulation and analysis [3]. The C-terminal death domain (FADD-DD) and N-terminal death effector domain (FADD-DED) in FADD belong to death domain superfamily which plays a pivotal role in signaling events that regulate apoptosis [4]. Although the two domains, FADD-DD and FADD-DED, have a similar structural motif which consists of six helices, their functional characteristics in the apoptosis signal transduction are very different. In a Fas-mediated apoptosis pathway, a death signal from Fas-receptor is transferred to caspase-8, which activates a subsequent effector caspase such as caspase-3, via an adaptor protein FADD through the death domain (DD) interactions of Fas and FADD [4] and death effector domain (DED) interactions of caspase-8 and FADD [5]. The thermal melting experiments showed that FADD-DED have structurally more stable than FADD-DD [6].

We have performed molecular dynamics (MD) simulation on the two domains and compared the motional characteristics. From the difference in the inter-helical stabilities and the inter-helical motional correlations of the two domains, we could explain the difference in thermal stability of the two domains. The strong correlation between inter-residue correlation coefficients and internal hydrogen bond ratio (or probability) during MD simulation in the two domains could explain the driving force of the inter-residue motional correlations and melting behaviors at the atomic or molecular level.

2 MATERIAL AND METHODS

2.1 Molecular Dynamics Simulation

We constructed the mouse FADD-DED(F25Y) structure, which has the residue numbers from 1 to 83, and the wild type mouse FADD-DD structure, which has the residue numbers from 92 to 182, with NAMD psfgen [7] and Gromacs-3.3beta [8]. From the NMR structure of human FADD-DED(F25Y) (PDB accession ID: 1A1W) [9], 15 residues in human FADD-DED(F25Y) were replaced by corresponding residues in mouse FADD-DED(F25Y). The wild type mouse FADD-DD structure was constructed by the point mutation in Tyr-96 to Asp from the NMR structure of mouse FADD-DD(D96Y) (PDB accession ID: 1FAD) [10] (Figure 1). (In this paper, FADD-DED will represent the FADD-DED(F25Y) unless explicitly mentioned.)

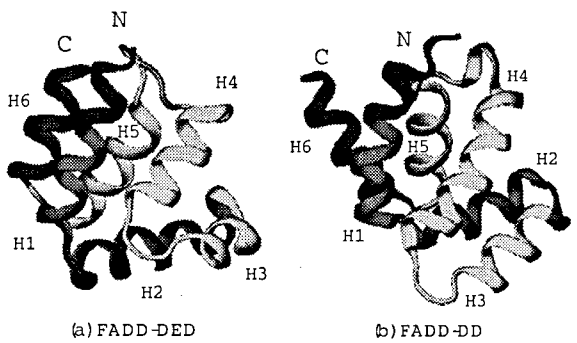


Figure 1: Structure of FADD death effecter domain (a) and death domain (b). Molecular dynamics simulation was achieved in 150mM NaCl and SPC explicit solvent water.

The protonated states of all charged residues were determined by the module in the gromacs, pdb2gmx with oplsaa force field [11] for pH 7.0 condition. After steepest descent algorithm were used for energy minimization with 100 steps, the protein domains were solvated with 10 Å water shell in a truncated octahedron box and ionized with 150mM NaCl both for the ionic strength of the solution and for the system neutralization. Total number of SPC water molecules was 5737 for FADD-DED solvation and 6106 for FADD-DD.

First pre-MD was performed with fixing protein coordinates during 10ps with isothermal and isobaric ensemble (NPT) in 1 atm and 300K. Second pre-MD was performed without positional fixing constraint during 500ps in 1 atm and 300K. Temperature coupling with Berendsen-thermostat was used to a bath with reference temperature 300K, with time constant 0.1ps [12]. For pressure coupling Berendsen exponential relaxation pressure coupling was used with time constant 0.5 ps.

Then 10.5 ns canonical ensemble (NVT) simulation was performed. The short range electrostatic and van der Waals energy was calculated in every 1 fs time intervals within 10 Å. For long range electrostatic energy, PME calculation was used with 1 Å fourier-spacing, and for long range van der Waals energy, the energy of atoms within 14 Å was recalculated in every 10 steps. Coordinates, velocities and energies of the system were sampled every 200 fs during 11ns MD simulation.

2.2 Analysis of MD Simulation

To compare the difference in motional characteristics between the two domains, variations in helix length and inter-helical motional correlations were investigated. At first, the secondary structures of the two domains were monitored by DSSP program [13]. Then helix axes and cross angles between inter-helices were calculated from the 'rotational fit' methods [14]. The variance of helix length was considered as one of criteria in helix stability.

2.2.1 Correlation Coefficients

Time correlation function of the center of helices was calculated for each domain (Figure 4) and the correlation coefficients of the center of helix axes at 100ps was compared. To find out the characteristic motional correlation between residues, the time correlation function of the inter-residue C_α atom coordinates and the cross-correlation coefficients of the C_α atom coordinates at τ = 100ps were observed (Figure 5). Both the time correlation function of the center of helices and that of the inter-residue C_α atom coordinates were defined in equation (1);

$$C_{ij}(t)_\tau = \sum_{t}^{10ns-k} \frac{(r_{t+\tau} - r_t)_i \bullet (r_{t+k} - r_t)_j}{\sqrt{(r_{t+\tau} - r_t)_i^2 (r_{t+k} - r_t)_j^2}} \qquad (1)$$

where $C_{ij}(t)_\tau$ is a ensemble average of the correlation coefficient between two residues i, and j after sampling time τ from time t during 10ns. r_{t+k} is the vector of the coordinate that may be a helical center or a C_α atom at time $(t + \tau)$. $C_{ij}(t)_\tau$ ranges between (-1) and (+1). Total 15 cross-correlation coefficients between the inter-helices were calculated with maximum τ = 5ns during 10ns. Specific correlation maps for the τ = 100ps were selected for the comparison of the inter-helical correlations and the inter-residue correlations.

2.2.2 Internal Hydrogen Bond Ratio and Life Time

Total internal hydrogen bond dynamics were monitored. Each hydrogen bond donor and acceptor pair was checked for their hydrogen bonding existence at every 0.2ps samples from 10ns MD trajectory. The criteria for the hydrogen bonding is the angle (<30°) and the distance (< 3.5 Å) between the donor and acceptor atoms. Internal hydrogen bond ratio r (or probability) during 10ns MD simulation was defined as the equation (2).

$$r = \frac{\text{Sum of hydrogen bonding time}}{\text{Total simulation time}} = \frac{\text{hydrogen bonding life time}}{\text{hydrogen bonding life time} + \text{hydrogen bond dead time}} \qquad (2)$$

The hydrogen bond life time and the hydrogen bond dead time is defined as the average length of the continuous hydrogen bonding time and the average length of the continuous hydrogen nonbonding time. The hydrogen bond ratio means the probability of the existence of the hydrogen bonding pair during MD simulation. For example, the value of r = 1 means that the hydrogen bonding is always found during MD simulation. The r = 0.5 means that the probability of the hydrogen bonding during MD simulation is 0.5.

3 RESULTS

3.1 Validation of Equilibrium

After 1ns equilibrium, total 50,000 frames were analyzed for the protein motions with 0.2ps sampled space from the 10ns production MD trajectories. The root mean square displacement (RMSD) of backbone and the root mean square fluctuation (RMSF) of C_α atoms are usually considered as the reference of the system equilibrium. The backbone RMSD of mouse FADD-DD was more fluctuated than that of mouse FADD-DED. Average backbone RMSD of the FADD-DED was about 1.8 Å and that of the FADD-DD was 2.3 Å for the PDB structure as the reference fitting (Figure 2a). After 1ns equilibrium, the average value of RMSD in backbone for FADD-DED and for FADD-DD during 10ns are decreased to about 1.0 Å and 1.6 Å respectively for 1ns MD equilibrium structure as the reference fitting. From these results, even though the sequence of mouse FADD-DED is different from that of human FADD-DED in 15 residues and the sequence of mouse FADD-DD is different from that of mouse FADD-DD(D96Y) only in a residue, it could be clearly seen that the mouse FADD-DED structure was more similar to human FADD-DED than mouse FADD-DD structure to FADD-DD(D96Y). RMSF of C_α in each residue during 10ns MD simulation was slightly larger than that in NMR structure. This large value in RMSF is due to the 10ns time averaging widows [2]. Radius of gyration of FADD-DED is about 12.29 ± 0.068 Å and that of FADD-DD is about 12.86 ± 0.085 Å during 10ns MD simulation. Consequently we concluded that FADD-DD have slightly larger backbone

fluctuation intrinsically than FADD-DED at 300K

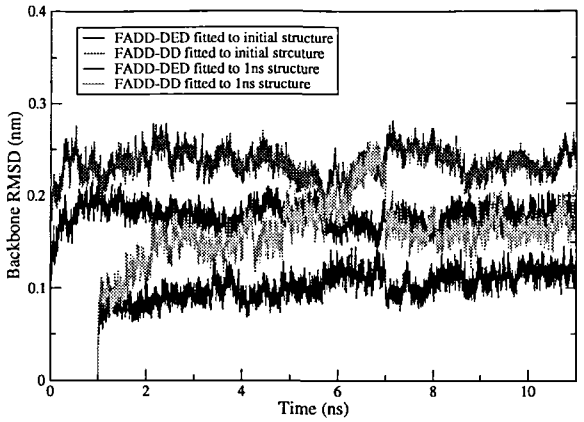


Figure 2(a): Backbone RMSD of mouse FADD-DED and FADD-DD during 10 ns MD simulation

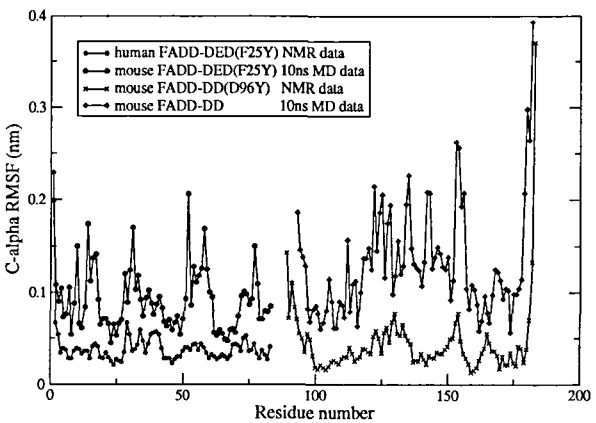


Figure 2(b): C_α RMSF of FADD-DED and FADD-DD from NMR data and 10ns MD results.

3.2 Stability and Motional Correlation of Helix

Protein primary sequence has divided into secondary structural elements by the DSSP program from the MD trajectory, then the helix lengths of the each six helical axes are monitored (Figure 3). The fluctuations (standard deviations) of all helix lengths in FADD-DED are in a range 0.047 ~ 0.085 nm except helix 3 which has large fluctuation value 0.387 nm in helix length. Helix 3 in FADD-DED is composed of 4 residues (residue number 33 to 36) and naturally very short. During the MD simulation the dihedral angles of these four residues are on the edge of the α-helix definition, so the helix length is more fluctuated from 0 to 1.65 nm (Figure 3(a)). In FADD-DD the helix 3 is composed of 9 residues (residues 124 to 132), and the length of helix 3 is not as fluctuated as that of FADD-DED. The fluctuations of all helix lengths in FADD-DD are in a range 0.073 ~ 0.156 nm. So the fluctuations of the length in all six helices from FADD-DD are larger than those in all five helices from FADD-DED (Figure 3(b)).

To investigate the inter-helical motional correlation, we calculated the cross-correlation coefficients of the center of helix axes and the cross angle of inter-helices about z axis after least square fitting of 10ns trajectory onto the 1ns protein structure. The result of the cross correlation coefficients in the center of helix axes is similar to that in the cross angle of inter-helices (data not shown).

Inter-helix motional correlations of the two domains were different in helix numbers and correlation times (Figure 4). The largest negative cross-correlation value (-0.3) were found between helix 5 and helix 6 (H5: H6) in FADD-DED. The cross-correlation values between helix 3 and helix 4 (H3:H4) in FADD-DD were more negative value (-0.75) than any others in FADD-DED.

Overall patterns of inter-helical motional correlations of the two domains were very different. In FADD-DD the inter-helical motional correlation coefficients had more negative values than those in FADD-DED. The difference in these inter-helical cross correlation showed that if a protein has large RMSD in backbone (or C_α) like FADD-DD, then the absolute value of its motional correlation coefficients in the secondary structural elements is large. This means that if a protein has a larger value in RMSD with maintaining its' globular fold, then the protein

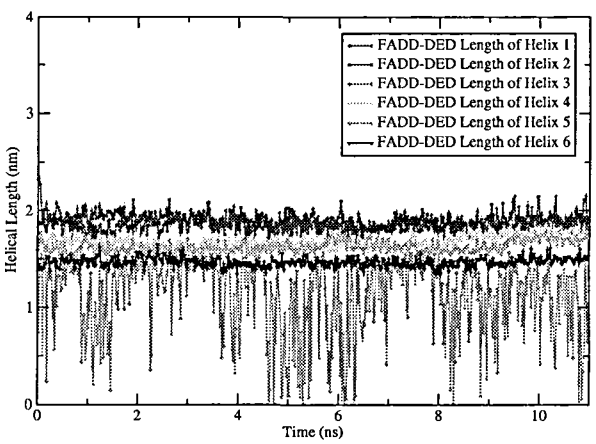


Figure 3.(a): Helix length of FADD-DED during 11ns. Helix 1 is located in Residue 1~15, Helix 2 in 16~28, Helix 3 in 29 ~38, Helix 4 in 39~54, Helix 5 in 55~72, and Helix 6 in 73~83.

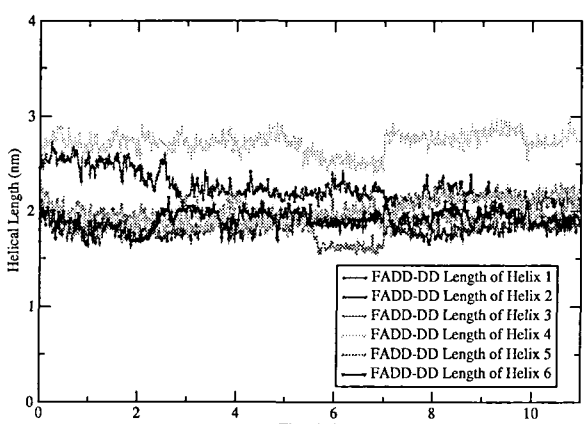


Figure 3(b): Helix length of FADD-DD during 7ns. Helix 1 is located in Residue 94~108, Helix 2 in 109~28, Helix 3 in 29 ~38, Helix 4 in 39~54, Helix 5 in 55~72, and Helix 6 in 73~83.

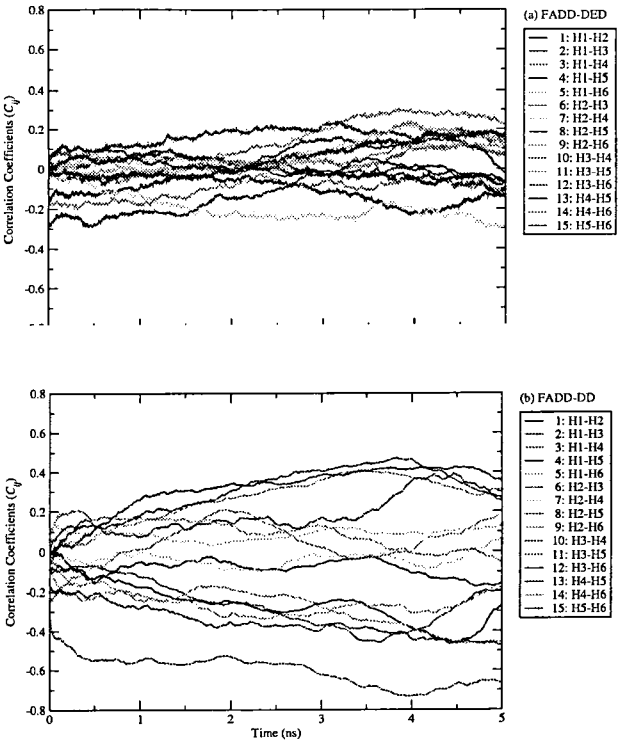


Figure 4: Correlation coefficients of the center of helical axes in (a) FADD-DED and (b) FADD-DD during 5ns after 2ns equilibrium. Each of 'H1' to 'H6' represents 'Helix 1' to 'Helix 6' respectively.

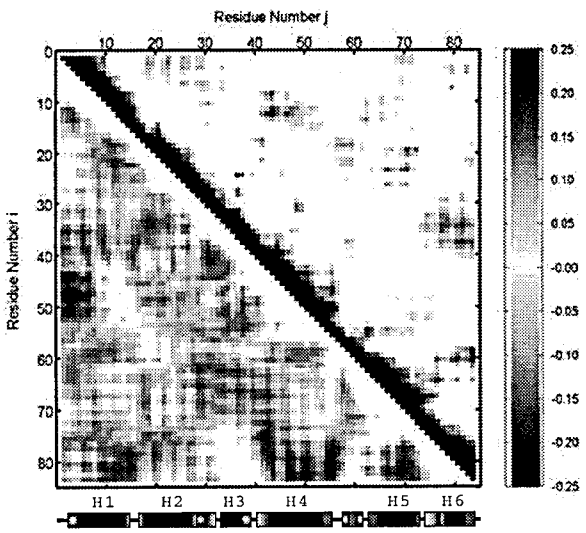
3.3 Hydrogen bond Ratio

The number of the hydrogen bond donors and acceptors in FADD-DED is 131 and 251 but that in FADD-DD is 138 and 271 respectively. The average number of hydrogen bond per frame in FADD-DED is 57.671 out of the possible number of hydrogen bonding pairs 16315 but that in FADD-DD is 63.117 out of 18699 possible. Even though the absolute number of hydrogen bonds per frame in FADD-DED is smaller than that in FADD-DED, the ratio in the hydrogen bonds in FADD-DED, 57.651/ 16315 (= 0.00353) is larger than that in FADD-DD, 63.117/18699 (= 0.00337). The physical meaning of the ratio in hydrogen bonds per time per possible hydrogen numbers is easily imaginable. If a protein has more fluctuation motion with more hydrogen bonding donors and accepters which do not participated in internal hydrogen bonding than other proteins, it could be thought that the protein have more possibility to make a hydrogen bonding with solvent water. So the hydrogen bonding donors and accepters, which do not participate in internal hydrogen bonding in FADD-DD, could frequently interact with solvent water by hydrogen bond than those in FADD-DED.

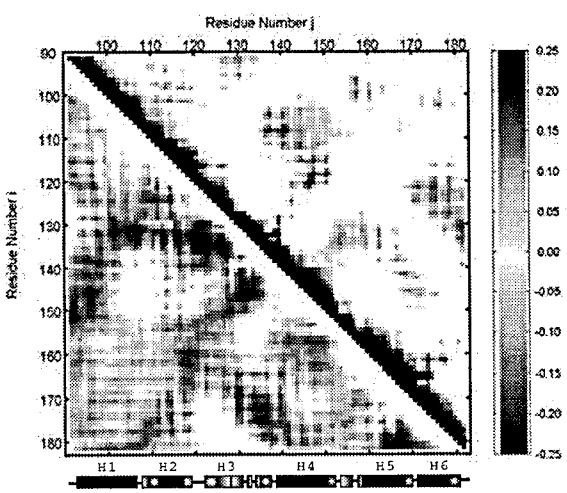
Also to compare the motional characteristics with hydrogen bond in residue basis, inter-residue correlation coefficients maps from Equation (1) at τ = 100ps (Figure 5) and inter-residue hydrogen bonding maps were plotted (Figure 6). The number of internal hydrogen bonding in the two domains respectively and the hydrogen bond life time and unbound life time were monitored in MD trajectory. Then the hydrogen bond ratio defined in equation (2) calculated and residue based donor to acceptor plot were drawn (Figure 6).

3.3.1 Inter-helical hydrogen bond

The inter-helical hydrogen bond is important for the stability in tertiary structure of protein. To compare the tertiary stability of the two domains, the numbers of inter-helical hydrogen bond are monitored from the long range (>= n---n+6) off-diagonal elements in inter-residue hydrogen bonding map (Figure 6). In FADD-DED inter-helical hydrogen bonds were in H1:H3, H1:H4, H1:H6, H2:H3, H2:H5, H3:H4, and H5:H6, but in FADD-DD those were in H1:H3~H4 loop, H1~H2loop:H3, H1:H6, H2:H3, H3:H4, H5:H6 from cross peaks in the figure 6. Consequently FADD-DED has more inter-helical hydrogen bond than FADD-DD.

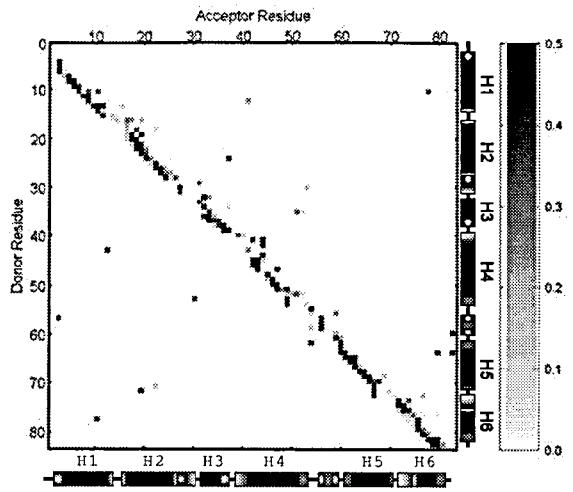


(a) FADD-DED

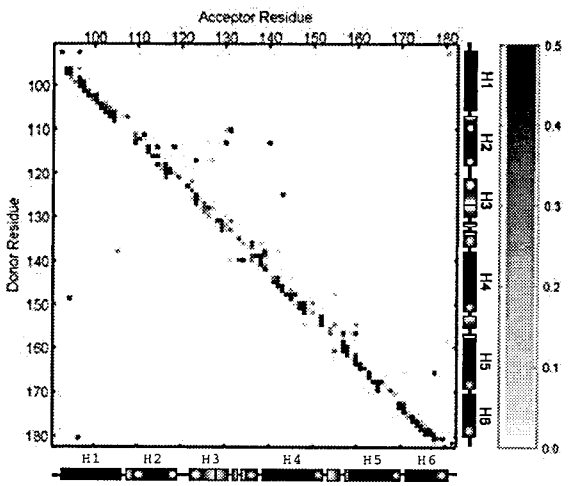


(b) FADD-DD

Figure 5: Inter-residue correlation coefficient map in (a) FADD-DED and (b) FADD-DD during 10ns. Upper off-diagonal elements in the maps represent the positive correlation and lower off-diagonal elements in the maps represent the negative correlation. Each value represents the correlation coefficients in τ_c =100ps.



(a) FADD-DED



(b) FADD-DD

Figure 6: Inter-residue hydrogen bond interaction map in (a) FADD-DED and (b) FADD-DD during 10ns Each plotted value represents the hydrogen bond ratio (or probability) during MD simulation which defined in equation (2).

3.3.2 Intra-helical hydrogen bond

The intra-helical hydrogen bond in the residue n---n+3, n---n+4, or n---n+5, is the main driving force of the conformation of helix [15]. The average numbers of the intra-helical hydrogen bond are similar in the two domains. But the intra-helical hydrogen bond ratios are different. In FADD-DED the length of helix 3 is shorter than other five helices from inter-residue hydrogen bond interaction map, but the inter-residue hydrogen bonding ratio *r* is similar to other five helices. So the stability of the helix 3 is similar to others. However in FADD-DD the length of helix 3 is longer than that in FADD-DED, but intra-helical hydrogen bond ratio *r* is lower than other helices of FADD-DD. This makes that the helix 3 had more positional fluctuation.

4 CONCLUSION

Helical stability can be monitored with the helical lengths and the inter-residue hydrogen bond interaction map during MD simulation. Variations of helical length in FADD-DED and in FADD-DD showed that helixes in FADD-DED were more stable than FADD-DD except in helix 3 because of the critical number of helix formation.

Inter-helical positional correlation could be described by the cross correlation coefficients in the center of helical axes. Large negative values in inter-helical correlation coefficients, or inter –residue correlation coefficients found in FADD-DED mean that the protein have more fluctuated in those inter-helical regions or inter-residue regions.

Positive inter-residue correlation coefficients also strongly related with inter-residue hydrogen bonding during MD simulations. The most positively correlated residues are the hydrogen bonding residues which have the large hydrogen bonding ratio (or probability) during MD simulation. In FADD-DED and FADD-DD, the stability of the protein are distinctly related with the number and the intensity of long range (>= n---n+6) inter-residue hydrogen bonding and the intensity of intra-helical (n---n+3, n---n+4, n---n+5) hydrogen bonding during MD simulation.

Motional correlation coefficients analysis and internal hydrogen bond dynamics could explain the difference in the melting temperature of the two domains, FADD-DED and FADD-DD. Also the internal hydrogen bond ratio could explain the positively correlated inter-residue interactions.

Acknowledgements

This work was supported by the Korea Research Foundation Grant funded by Korea Government (National Research Laboratory Fund), (M1-0104-00-0142), and the Korean Systems Biology Research Grant, (M1-0309-00-0012).

REFERENCES

[1] R. Huber & W. S. Bennett, Jr., Fuctional significance of flexibility in proteins. Biopolymers. 22:261-279, 1983.

[2] P. H. Hünenberger, A. E. Mark, and W. F. van Gunsteren., Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. J. Mol. Biol. 252:492-503, 1995.

[3] C. H. Weber, and C. Vincenz, The death domain superfamily: a tail of two interfaces? Trends Biochem Sci. 26:475-481, 2001

[4] H. Duan, A. M. Chinnaiyan, P. L. Hudson, J. P. Wing, W. W. He, and V. M. Dixit, ICE-LAP3, a novel mammalian homologue of the *Caenorhabditis elegans* cell death protein Ced-3 is activated during Fas- and tumor necrosis factor-induced apoptosis. J. Biol. Chem. 271: 1621-1625, 1996.

[5] M. Muzio, A. M. Chinnaiyan, F. C. Kischkel, K. O'Rourke, A. Shevchenko, J. Ni, C. Scaffidi, J. D. Bretz, M. Zhang, R. Gentz, M. Mann, P. H. Krammer, M. E. Peter, and V. M. Dixit. FLICE, a novel FADD-homologous ICE/CED-3-like protease, is recruited to the CD95 (Fas/APO-1) death--inducing signaling

complex. Cell 85: 817-827, 1996.
[6] K.-S. Kim, BioNMR Laboraroty, Biomedical Research Center, Korea Institute of Science and Technology, Seoul, Republic of Korea , *Personal Communications*
[7] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, NAMD2: Greater scalability for parallel molecular dynamics, J. Comp. Phy. 151: 283-312, 1999.
[8] E. Lindahl, Hess, B., and D. van der Spoel, Gromacs 3.0: A package for molecular simulation and trajectory analsyis. J. Mol. Model. 7: 306-317, 2001.
[9] M. Eberstadt, B. Huang, Z. Chen, R. P. Meadows, S. C. Ng, L. Zheng, M. J. Lenardo, and S. W. Fesik, NMR structure and mutagenesis of the FADD (Mort1) death-effector domain. Nature 392: 941-945.1998.
[10] E. J. Jeong, S. Bang, T. H. Lee, Y. I. Park, W. S. Sim, and K. S. Kim, The solution structure of FADD death domain. Structural basis of death domain interactions of Fas and FADD, J Biol Chem 274: 16337-16342, 1999.
[11] W. L. Jorgensen, Maxwell, D. S., and Tirado-Rivers, J., Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118: 11225~11236, 1998.
[12] D. van der Spoel, Lindahl, E., Hess, B., van Buuren, A. R., Apol, E., Meulenhoff, P. J., Teleman, D. P., Sijbers, A. L. T. M., Feenstra, K. A., van Druen, R., and Berendsen, H. J. C., Gromacs User Manual version 3.2, www.gromacs.org, 2004.
[13] W. Kabsch, and Sander, C., Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. Biopolymers, 22: 2577~2637, 1983.
[14] M. F. Lensink, A. M. Haapalainen, J. K. Hiltunen, T. Glumoff, and A. H. Juffer, Response of SCP-2L domain of human MFE-2 to ligand removal: binding site closure and burial of peroxisomal targeting signal. J. Mol. Biol. 323: 99-113, 2002.
[15] A. V. Finkelstein and O. B. Ptitsyn, Protein Physics, A Course of Lectures. Academic Press, Lecture 7, 2002.