Correlation Between The Helix-Forming Propensity Of Peptides Eptides Obtained By NS MD Simulation and The Chou-Fasman Parameters

Shinichi Murayama[a], Toru Mori,[a] and Koichi Nishigaki[a, b*]

[a] Department of Functional Materials Science, Saitama University
255 Shimo-okubo, Sakura-ku, Saitama, Saitama 338-8570, Japan
[b] REDS (Rational Evolutionary Design of Advanced Biomolecules, Saitama Prefecture)
Collaboration of Regional Entities for the Advancement of Technological Excellence, JST
**e-mail: koichi@fms.saitama-u.ac.jp*

ABSTRACT: Molecular dynamics of nanosecond timescale could provide a significant result for the structure analysis of oligopeptides. The most surprising was the fact that Chou-Fasman parameters which have no direct relationship with nanosecond molecular dynamics could implicate the result to a certain extent. A novel parameter termed X %-stickiness, which is another measure of compactness of a molecule, was first introduced effectively.

1 INTRODUCTION

Peptide folding has been studied both experimentally and theoretically during these decades [1]. Recently, femtosecond laser analysis enabled to detect the events occurring in the time scale of nanoseconds or less [2], thus allowing ones to compare the experimental results with those obtained by computer simulation. Studies on peptide folding are also active due to the idea that the whole protein folding can be constructed from such peptide fragments or in a hierarchical manner [3]. This is partly because the current computer resources are not sufficiently enough to obtain the final solution of protein folding. In addition, there is another viewpoint which makes molecular dynamic simulation on peptides precious: rough estimation of the final structure of peptides may be sufficient to sort a library of peptides into suitable or not-suitable for a particular purpose. There are several axes of estimation such as *rapid folding or not, compact or not, α-helix forming or not, α-sheet forming or not*, and others. This is especially the case with evolutionary molecular engineering since it has to deal with a huge number (e.g., 10^{12} to 10^{15}) of different molecules and needs to select a preferable subset out of the whole molecules.

Lately, nanosecond solution structures of dodecanucleotides were suggested to be correlated to their macroscopic behavior of gel-electrophoretic mobilities [4]. Namely, the compactness of oligonucleotides obtained as the radius of gyration by nanosecond molecular dynamics (MD), so not yet reached its equilibrium, had a strong correlation with the actual mobility in gel electrophoresis, suggesting that the early stage structures in folding of such molecules as single-stranded DNA and polypeptide can provide partially the feature of their folding-matured structures. If true, this nature is very useful for evolutionary molecular engineering where fuzzy properties can be utilized temporarily and confirmed finally.

In this vein, we investigated ns MD of oligopeptides such as octamers and hexadecamers, resulting in an interesting observation that Chou-Fasman relationship may be tied with determining the early stage structures in folding of oligopeptides.

2 MATERIALS AND METHODS

2.1 MD simulation

MD simulations were performed using AMBER 7 molecular simulation package. AMBER 99 force field was adopted in all simulations. GB/SA continuum solvation model was used with the salt concentration set at 0.2 M. ONYX-3400 supercomputer (SGI) at the Information Technology Center of Saitama University or VPP5000 (Fujitsu) at the Information Technology Center of Nagoya University.

The initial conformations of peptides generated by our home-made software, *IniConf*, were subjected to the energy minimization (500 steps of steepest descent followed by 500 steps of conjugate gradient energy minimization) and then 10 ps equilibration with the temperature gradually raising from 0 K to 300 K. SHAKE was applied to all bonds involving hydrogen atoms.

2.2 Oligopeptides and Chou-Fasman parameters

The sequences of hexadecapeptides were randomly generated with or without bias of occurrence of amino acids which have α-helix-/β-sheet-forming tendency in proportion to Chou-Fasman estimation value [5]. For this purpose, we used the following formula:

$$P = (\Sigma_i C_i)/(\Sigma_i 1) \qquad (1)$$

where C_i stands for Chou-Fasman parameter related to α-helix-forming tendency of i-th amino acid and the denominator represents the number of amino acids. We calculated this by using Internet service of ExPASy (http://expasy.org/tools/pscale/) which was built based on the theory developed by Chou and Fasman [6]. Sixteen sequences thus generated are shown in Fig. 1, using PRS representation [7].The remaining sequences generated are: i) α-helix biased; EKEQQCCLDHKLMCEL CELCMMMKLKALLMLK CMHEQDEHQMCHHAKE KHKQLQCHLHGCLAHH QMLLGKAYKLHCDKQL ECYQLEIEKCECQQMM KQKLSHCACHQCSMQA HHGCEKQHQEEHQCAA QCMHECEAKQQHMKAE

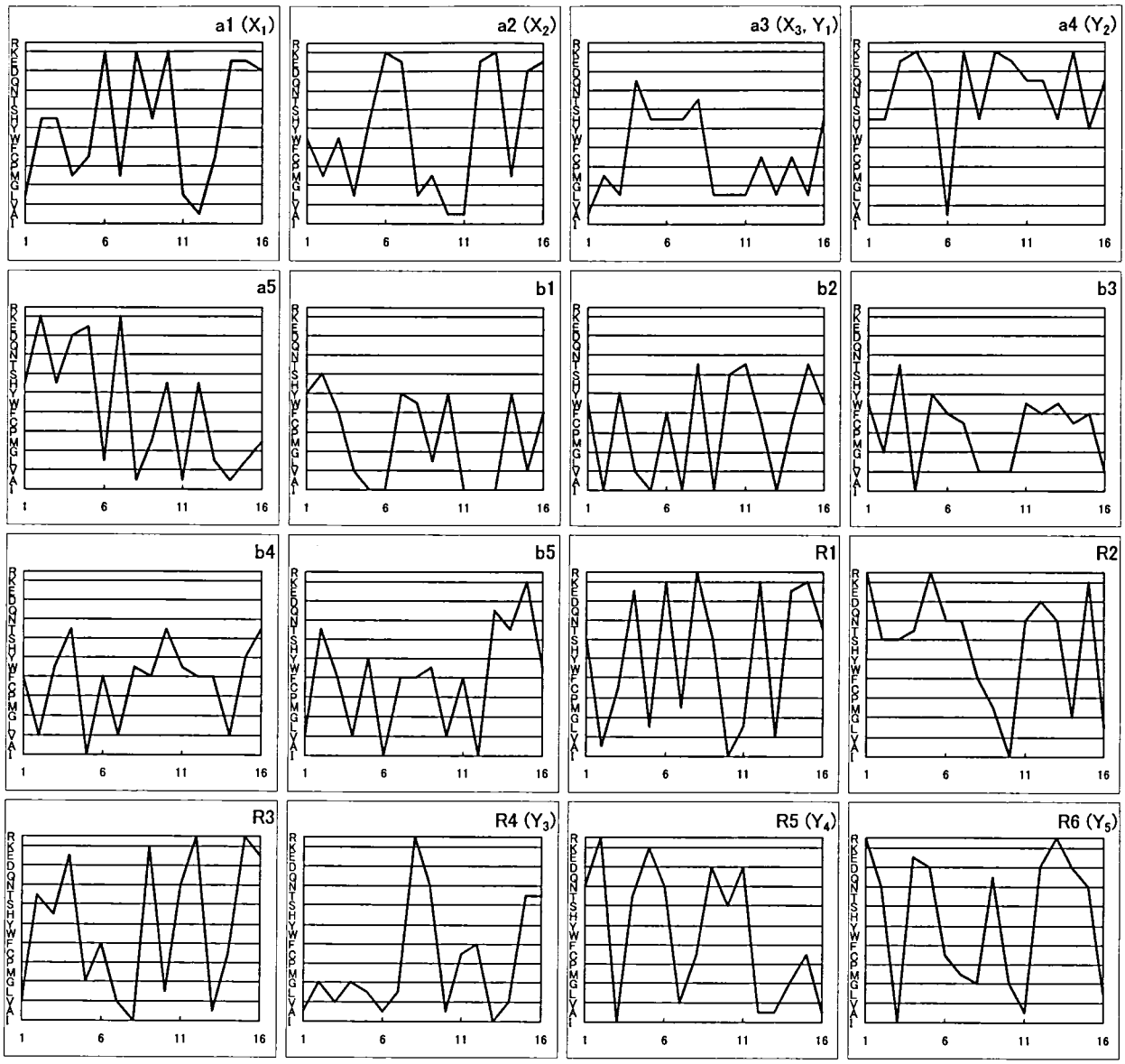

Fig.1 Some of hexadecapeptides used in this simulation are shown in planar representation of sequences (PRS). Arbitrarily selected 16 peptides are shown in PRS, which has the ordinate of amino acid species (shown by one letter abbreviation) and the abscissa of the sequence of peptides. Amino acids are arrayed hydrophobic (bottom) to hydrophilic (top). The letters X1 to X3 and Y1 to Y5 represent the points appearing Figures 3 and 4.

KGEKQKCWEAMMELLL
LQEEALLQYEMELKGQ
EEMCLGMQMHHKYQAH
LEHHKMKHCWDMKKAH
EHMKADMKLKFHMLMI
MEKKELLMHHCDLELK
YIIVYYYYVFWYIYHT
VWVWWVTWTWAVTFWF
GIFYFDIVTTWWWVYW
TVFWTVTWVWWTFFVT
VFVYWVYYVWTIVYIV
FFWYVWFYHFFVYYVW
WSVIVTTTTTFFFVFI
VYNYTTTVFWYIFTVY
YYYFYYVWVIVWCFII
IFYYWTWNTVVYTFWW

KALECMQQEKSIMCAE
HMMEFMMCLECEAACK
KMQHCQELQQMEHLLK
TLCQAELAESCQQLHE
LDAHAEEKWAHLMMMI

ii) β-sheet-biased;

WVTRIFTVTWIIFWWF
WVTWYIIYFTWTIMVY
YIITSIIYFVIFTYFG
WAVYWFFVHVWCGVWV
IETVYWWIFVIWYVTE
VTTTIFWFYYVWIYVT
ITFTVYFFVWFITIVI
STFVFWVVIIRVWITC
FWYVYWIFETIVYWVT
VCWWIVFVVGCFCIVW

iii) Non-biased;

ERNKMLLMETNKGNII
LRSNRVDVERSIVDGT
ANLHSGYETNNQGDNL
LTTFLKWDIRGLENNL
RMNNELDQIKRSTYYL
GYWTDRRLTMTYRNLI
GGVGAFNMFYIQILKL
LGKDRFLAVEVLHVTI
LKASSDYREAVLDFRT
KVNGNREDGFRRTYDK
MFKQDFACSMKGSVVC
AADIDSGLDYDRATNN

IYAFCGRRARAVGIDK
LGKGAIRFVLTETLRR
MANSGGLRLISEIFAR
ASKIKVAIECAMEGMF
RDCNSGFRKWTDIRNN
QYGGDRMDDWKDKELL
TMKRRTQENVRFIENS
WDNTLLGCIHFSCVQK
LRTERILCTFRFKGAG
GRGFEYQNTGSRVRYY
MEKVCAVMLILFLLLC
RNTVRGINRLHKTKVF

2.3 Parameter introduced: X % stickiness

Here, we introduced a parameter which can depict the compactness and stability of a molecule termed as X%-stickiness. This parameter expresses how many atom-to-atom neighborings occurred which have an accumulated life span of more than X % of the whole simulation time period.

3 RESULTS AND DISCUSSION

3.1 *Simple octapeptides*

Each of the homopolymers of octamer (19 amino acid species except proline) was subjected to MD analysis of 1 ns simulation. The R_g (radius of gyration), α-helix content and 20 %-stickiness of those molecules were calculated from their snapshots of 1 ns time period and summarized in Table 1. The octamer of Ala is shown to be the most compact among these molecules and those molecules which are composed of charged amino acids except $(Asp)_8$ are rather stretched due to the electrostatic repulsion of the elementary amino acid residues. Thus, Table 1 can be regarded as a list of properties of simple octapeptides. The values of α-helix content are also intriguing: being natural for $(Gly)_8$ but unfamiliar for $(Thr)_8$. What is the most interesting is shown in Fig. 2, where α-helix content of each octapeptide obtained by ns MD is plotted against its α-helix forming propensity calculated based on Chou-Fasman parameter. Octapeptides can be grouped into 4 categories here: *a*) those composed of non-aromatic hydrophobic amino acids (4 members in the filled circles), *b*) hydrophilic but non-charged (6 members in diamonds), *c*) charged (5 members in star) and *d*) the others (4 members in filled triangles). All of these groups can be separately analyzed to have a correlation between experimentally-obtained α-helix content and that calculated based on the Chou-Fasman parameter, except the case of charged group (Fig. 2). Among all, the group *a* has a highest correlation. Though intriguing, this fact must be conservatively translated since this phenomenon is supported only by a less sufficient amount of data, appearing here, and the peptide grouping was rather arbitrarily made.

#	Sequence	Mean R_g (10^2 pm)	α-helix content (%)	β-sheet content (%)	20% stickiness average (pairs/molecule)
1	$(Arg)_8$	7.03	10.8	0.5	1.34
2	$(Trp)_8$	6.78	8.25	0	2.08
3	$(His)_8$	6.67	10.5	0	0.72
4	$(Tyr)_8$	6.58	14.3	0.5	2.16
5	$(Lys)_8$	6.49	25.8	0	0.46
6	$(Phe)_8$	6.29	19.5	0.5	1.58
7	$(Glu)_8$	6.20	26.0	0	1.60
8	$(Gln)_8$	5.93	31.8	0.5	1.72
9	$(Asp)_8$	5.82	25.3	0	1.32
10	$(Met)_8$	5.79	25.3	0.5	2.86
11	$(Ile)_8$	5.73	29.3	0	0.90
12	$(Asn)_8$	5.62	32.0	0	1.82
13	$(Val)_8$	5.54	23.5	0	0.78
14	$(Gly)_8$	5.52	0	1.0	0.66
15	$(Leu)_8$	5.49	39.3	0	1.54
16	$(Cys)_8$	5.41	28.0	0	5.96
17	$(Thr)_8$	5.28	19.8	0	4.72
18	$(Ser)_8$	5.24	23.3	0	5.14
19	$(Ala)_8$	4.96	35.3	0	4.50

Table 1 Properties of simple octapeptides obtained by ns molecular dynamics.

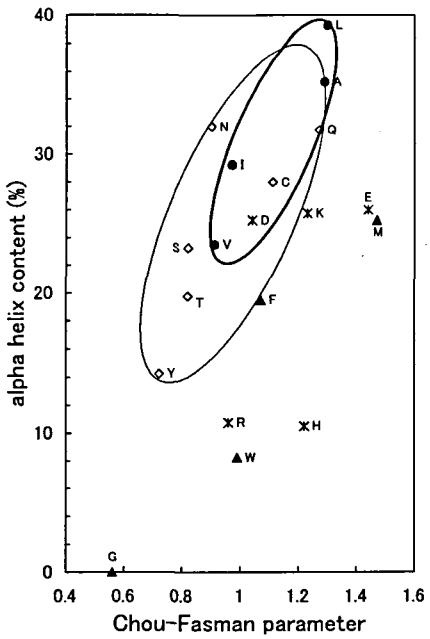

Fig.2 Correlations between α-helix contents and Chou-Fasman parameters of simple octapeptides for each cluster of peptides. The peptides were classified to be composed of hydrophobic (●, alkane; ▲, non-alkane) and hydrophilic (◊, non-charged; *, charged) amino acids (one letter abbreviation is used to show each component amino acid). Tyrosine(Y) is tentatively classified to be hydrophilic due to its hydroxide moiety.

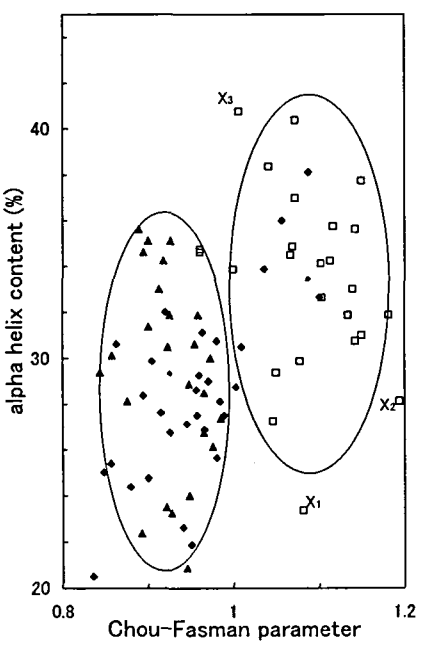

Fig.3 Correlations between α-helix content and Chou-Fasman parameters of hexadecapeptides of three categories. The symbols represent α-helix-biased (□), β-sheet-biased (▲), and non-biased (♦). The ellipsoids are drawn to focus the area where most of the components are located. The centers of mass for α-helix/β-sheet biased categories are shown by a circle (●).

Categories	Number of samples	Mean R_g (10^2pm)	α-helix content (%)	β-sheet content (%)	20% stickiness average (pairs/molecule)
Non-biased*	30	7.81	28.4	0.36	11.9
α-helix-biased†	25	8.06	33.4	0.28	10.6
β-sheet-biased§	25	8.10	29.3	0.46	7.74

Table 2 Properties of hexadecapeptides obtained by ns-MD
*Occurrence frequencies of amino acids were adjusted to those of lysozyme (a typical protein).
†Those amino acids which have a high propensity of forming a-helix, according to Chou-Fasman, were preferentially selected to generate the sequence of 16mer.
§The amino acids of a high b-sheet-generating propensity were preferentially selected.

3.2 Hexadecapeptides

Three categories of peptides were arbitrarily generated based on Chou-Fasman parameters: i) peptides composed of amino acids biased to forming α-helix, ii) those biased to forming β-sheet, and iii) those of non-bias. Figure 3 shows that there is a significant difference in the amount of α-helix content between the two categories (α-helix-biased and β-sheet-biased) , which is consistent with the result theoretically expected (The α-helix content of the centers of mass of these are 1.09 and 0.93, respectively). The clear separation of two categories in the horizontal direction means that the stochastic process is working as expected. It is also reasonable that those peptides of non-biased category are spread over the areas of both categories. Those points far separated from the remaining are labeled with X_1 to X_3 but seem to be less abnormal in their sequence (see Fig. 1).

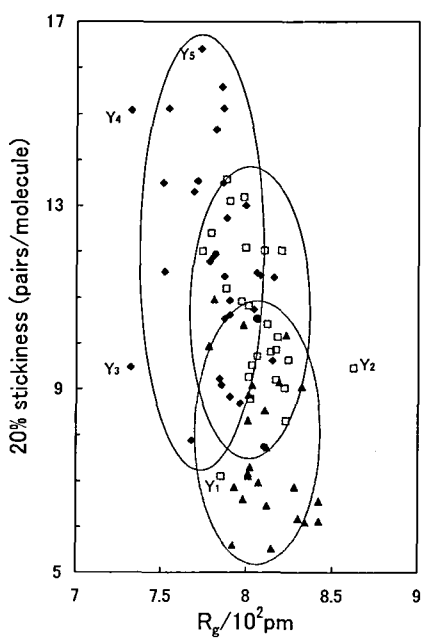

Fig.4 Correlations between 20%-stickiness and R_g of each category of hexadecapeptides. The same symbols are used in Fig.3. The center of mass for α-helix-biased, β-sheet-biased, and non-biased peptides are (8.06, 10.6), (8.10, 7.74) and (7.81, 11.9) on the coordinated. Far-deviated points are also designated as Y_1~Y_5 for analysis.

3.3 Novel finding with 20 %-stickiness analysis

The radius of gyration of the peptides analyzed here is evidently correlated with α-helix-forming propensity as can be confirmed by the horizontal separation of the relevant two categories (Fig. 4). To our surprise, these two categories are also separated in 20 %-stickiness axis more clearly. Since the stickiness is a measure of compactness of a molecule as explained in Materials and Methods, this result is consistent with our knowledge that α-helix is more compact than β-sheet in general.

These facts evidently show that Chou-Fasman parameter which had been obtained by analyzing crystallographic data of proteins [5] and have, a priori, no direct relation with ns structures of peptides was found to be related with it by this simulation. We have no rational interpretation for this phenomenon at this moment. However, Chou-Fasman parameters must be thought to have some implication for MD analysis.

4 CONCLUSION

Nanosecond molecular dynamics simulation of oligopeptides evidently showed that the propensity of α-helix formation can be observed to some extent even in such a limited duration, suggesting the utility of ns simulation for such a purpose as evolutionary molecular engineering [8]. Although it is yet mysterious, Chou-Fasman parameters can be used to predict the MD results of oligopeptides to some extent. Therefore, nanosecond analysis of solution structure dynamics of peptides may be able to depict the actual behavior of those molecules.

This work was supported by the Grant donated by the Rational Evolutionary Design of Biomolecules (REDS) Project, Saitama Prefecture Collaboration of Regional Entities for the Advancement of Technological Excellence supported by JST.

REFERENCES

[1] S. Gnanakaran, H. Nymeyer, J. Portman, K. Y. Sanbonmatsu and A. E. Garcia. Peptide folding simulations. Current Opinion in Structural Biology, 13:168--174, 2003

[2] S. Spörlein, H. Carstens, H. Statzger, C. Renner, R.Behrendt, L. Moroder, P. Tavan, W. Zinth, and J. Wachtveitl. Ultrafast spectroscopy reveals subnanosecond peptide conformational dynamics and validates molecular dynamics simulation. Proceedings of the National Academy of Sciences of United States of America, 99(12):7998--8002, 2002

[3] R. L. Baldwin and G. D. Rose. Is protein folding hierarchic? I. Local structure and peptide folding. Trends in Biochemical Sciences, 24(1):26--33, 1999

[4] M. Biyani and K. Nishigaki. Single-strand conformation polymorphism (SSCP) of oligonucleotides: An insight into solution dynamics of DNAs provided by gel electrophoresis and molecular dynamics simulations. Journal of Biochemistry (in press)

[5] P. Y. Chou and G. D. Fasman. Empirical Prediction of Protein Conformation. Annual Review of Biochemistry, 47:251--76, 1978

[6] P. Y. Chou and G. D. Fasman. Prediction of the secondary structure of proteins from their amino acid sequence. Advances in Enzymology and Related Areas of Molecular Biology, 47:45--148, 1978

[7] K. Kitamura, C. Yoshida and K. Nishigaki. GFPs of insertion mutation generated by molecular size-altering block shuffling. FEBS Letters, 555(3):483-488, 2003

[8] K. Kitamura, Y. Kinoshita, S. Narasaki, N. Nemoto, Y. Husimi and K. Nishigaki. Construction of block-shuffled libraries of DNA for evolutionary protein engineering: Y-ligation-based block shuffling. Protein Engineering, 15(10):843--853, 2002