

ナノ秒のMDシミュレーションから求めたペプチド二次構造の確率予測 三角マップ表示した%スティッキネス法の有用性

村山 真一, 吉田 昼也, 青山 崇, 浦田 賢, 西垣 功一*

埼玉大学 工学部 機能材料工学科, 〒 338-8570 さいたま市桜区下大久保 255

*e-mail: koichi@fms.saitama-u.ac.jp

(Received: May 10, 2006; Accepted for publication: July 7, 2006; Published on Web: September 4, 2006)

ペプチドのナノ秒 (ns) 分子動力学計算 (MD) では、安定構造を得るのには遙かに不十分な計算時間であることが分かっている。にもかかわらず、新たに導入した%スティッキネスの三角マップ表示 (TMR-% σ) によって、安定構造に含まれる二次構造 (特に α ヘリックス) のおぼろげな像が得られることが示された。%スティッキネスとは、生体高分子の動的なコンフォメーション変化を記述する目的で以前に導入された指標である。TMR-% σ 表示において、ペプチド形成傾向は、マップ上に α ヘリックス固有の相互作用を示すシフトした対角線ラインとして現れた。ns-MD の結果から計算されたペプチドの% σ と 旋回半径 (R_g) の座標プロットから、ペプチドの二次構造予測に関する確率マトリクスを導出できることが示された。

キーワード: ns-MD, Chou-Fasman パラメータ, α ヘリックス, スティッキネス, 三角マップ表示, 二次構造予測

1 序文

アミノ酸配列 (一次構造) 情報からタンパク質の立体構造 (三次構造) を予測することはフォールディング問題として、多くの研究がなされてきたが、未だ十分に実用的な解法は得られていない [1–3]。ポストゲノム計画として立ち上がってきた「タンパク 3000 プロジェクト」は、現存するタンパク質のフォールド全体を X 線や NMR で解析し、それらを一般的構造予測の資源としようとする壮大な試みである。この場合、一次構造を三次構造に直接結び付ける手がかりを探すことになる。一方、一次構造から二次構造さらに三次構造へと順次構造を組み上げていくアプローチもあり、多くの試みがなされてきた [4]。この一つに分子動力学 (MD) 的アプローチがあり、まだ膨大な計算時間を必要とし実用的ではないものの、対象がペプチドレベルであれば、分子動学的に平衡構造を求めることが可能となってきている。例えば、ペプチドがヘ

リックスを形成するには、約 300ns 要することが示されてきた [5]。

この状況で、最近筆者らは、ナノ秒の計算で求めたオリゴペプチドやオリゴヌクレオチドの構造が実験時間 (秒 ~ 分) での現象と関係付けられることを発見した [6, 7]。特に、ペプチドにおいては二次構造形成傾向がナノ秒段階で、ある程度窺い知ることができるといえる、いわば「梅檀は双葉より芳しい」ともいえる現象を見出した [7]。このことは、確率論的ではあるものの、ナノ秒段階で将来の構造を予測しうることであり、現在興隆しつつある膨大な分子ライブラリー ($10^{12\sim 15}$ 分子) を扱う進化分子工学には、「*in silico* 淘汰」の基礎になる重要な事実といえる。とりわけ、“モジュール”や“ドメイン”などのタンパク質部分構造をシャッフルして組合せ化学的に分子構築を行う“ブロックシャフリング法” [8, 9] においては、その出発ブロック材料の設計問題が大きなテーマであるが、このための有力なツールを提供する可能性がある。先の

論文 [7] で不思議な現象として、結晶構造解析データに基づき経験則として作り上げられたはずの「タンパク質におけるアミノ酸の二次構造形成性向」に関する Chou-Fasman パラメータ [10] が、ペプチドのナノ秒構造における α ヘリックス含量の予測に用いるという事実を報告した。本論文ではこのことをより一層鮮明に捉えるために、三角マップ法の導入が有効であることやその後、追加補強したデータに基づきながら新たに「(確率的) 2 次構造簡易判別法」を作成することを報告する。

2 方法

(i) シミュレーション試料

シミュレーションには 16 残基のペプチド配列を用いた (一部文献 [7] に) 配列は、Chou-Fasman の経験則を用いながら一定の方式で 3 つのカテゴリ (α -helix-biased (α 型)、 β -sheet-biased (β 型)、Non-biased (非バイアス型)) に分類されるものを合計 100 種 (内、新規 20 種) 作成した。

(ii) MD シミュレーション

MD 計算は既報の文献 [7] に従った。すなわちまず、そ

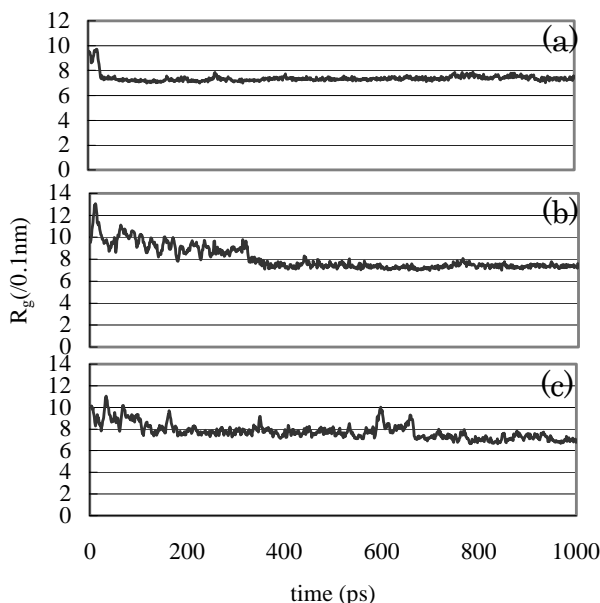


Figure 1. R_g trajectories of three representative peptides (a) α -helix forming biased peptide (P α -a05_46), (b) β -sheet forming biased (P β -b29_13), and (c) Non-biased (P0-018_44), which have the sequences : QMLL-GKAYKLHCDKQL, VFYVEIWWYCTFFYQW, and GGVGAFNMFYIQLKL, respectively.

れぞれのペプチド配列についてランダムに初期構造を作成し、その後、エネルギー極小化と系の平衡化を行い、1 ナノ秒の MD シミュレーションを行った。計算は、分子動力学計算パッケージ AMBER 7 を用い、力場は AMBER 99 の下で、溶媒モデルとして GB/SA を採用した。

(詳細は、Appendix および <http://gp.fms.saitama-u.ac.jp/sequence.htm> を参照)

3 結果と考察

出力された原子座標から求めた旋回半径 (R_g) の一例とアミノ酸配列を Figure 1 に示す。これらの結果に対して統計処理したものが Table 1 にまとめられている。ここで、X% スティッキネス ($X\%\sigma$) とは、文献 [7] で導入された動的構造記述パラメータであり、分子内の 2 つの原子が計測時間 (ここでは 1 ns) 中の X% (ここでは 20%) 以上の時間、定義された原子間距離 (ここでは 0.5 nm) に存在したときに、その 2 つの原子は $X\%\sigma$ にあると呼び、構造安定性の指標となる。

Table 1 から明らかなように α 型は β 型に比べて、1 ns 時点での平均 α ヘリックス含量が高い。同時に 20% σ にある原子の割合が有意に高いが、これは β 型より α 型の方が (ns 以下の寿命の話ではあるが) 安定した構造にある原子の割合が多いことを示していると解される。一方、非バイアス型は配列生成時点においてはバイアスをかけていないが、生成したペプチドは結果的に α 偏向や β 偏向した配列を含むことになる。有限分

Table 1. Some properties of hexadecapeptides obtained after ns-MD simulation

Categories	Mean R_g (10^2 pm)	α -helix content	β -sheet content	20%stickiness* /molecule
α -biased	8.05 (0.23)	33.1	0.29	10.4 (1.87)
β -biased	8.07 (0.17)	29.1	0.46	8.31 (1.76)
Non-biased	7.81 (0.21)	28.4	0.36	11.9 (2.26)

* The number of atomic pairs having 20%-stickiness per molecule (see Ref.[7]).

† Standard deviations are shown in parentheses.

子集団 (30 試料) に対しての結果であるが、非バイアス型は各種のパラメータ全体では α 型とも β 型とも異なる挙動を示しているのが注目される。

3.1 %スティッキネスの三角マップ表記

今回、%スティッキネスを視覚化し有益な情報をえる目的で、三角マップ表示を導入した (Figure 2)。すなわち、ペプチド分子の 1 ns - MDシミュレーションの間にそれぞれの原子対で 20% σ が成立したものを数え上げた (実際の表示はそれぞれのペプチド分子に関して異なる初期構造 (50 種) から出発したシミュレーション結果の総和となっている)。

Figure 2 の代表例に見られるように、 α 型では、三角マップの対角線から 3 残基分だけ平行シフトしたライン (“対角線シフト”) が現れ、ちょうど α ヘリックスが形成したときに期待されるものと対応している。

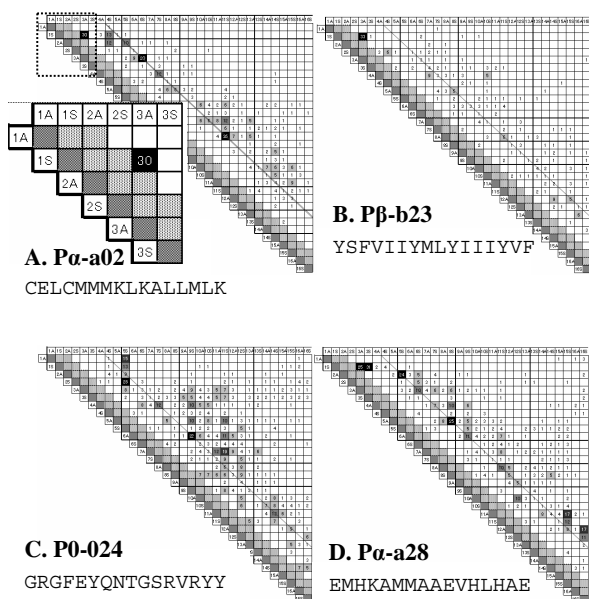


Figure 2. Triangle map representation of 20%-stickiness of a peptide. The C_{α} atoms (iA) and the mass center of the side chain (iS) are aligned on the top and at the side in a numerical order to provide the %stickiness result (summation over 50 different simulations) at the intersect for the relevant atomic pairs, which is effectively monitoring $NH \cdots C=O$ pairs. The higher the number at the cross is, the more the %stickiness is. The peptides shown here are (A) α -helix forming biased ($P\alpha$ -a02), (B) β -sheet biased ($P\beta$ -b23), (C) non-biased (P0-024), and (D) the highest α -helix content monitored by snapshot analysis.

実際には、全体の一部 (20%) の寿命であり、それぞれの原子対のフェーズが必ずしも揃って現れるものではないことを考えれば、この “対角線シフト” の出現は α ヘリックスの形成を積極的に意味するものではないが、それにもかかわらず、それを示唆するものとして興味深い。ここにも 「梅檀は双葉より芳しい」 性質が現れている可能性がある。

Table 2 に、 α 型 (35 分子)、 β 型 (35 分子)、非バイアス型 (30 分子) 合計 100 種のヘキサデカペプチドに関してこの処理をしてえたものを統計処理した結果を示している。

Table 2 から、平均値としては、この処理においても α 型が β 型よりも高い %スティッキネスを示すことがわかり、Table 1 の結果と符合している。ただし、標準偏差がやや大きいためにこの解釈の信頼性はやや低いものとなっていることに留意する必要がある。

ここで、以上のような統計的解析を離れて、Chou-Fasman 経験則から α ヘリックス形成能が最も高いと考えられる配列のペプチド (すなわち、 $P\alpha$ -a28) の三角マップを描いたものが Figure 2D であり、他のペプチドに対するもの (Figure 2B, C) とは異なり、かなり有意に対角線シフトが現れている。このことから、この三角マップ表記により、ナノ秒レベルの観察で α ヘリックス形成能の推測が可能になったと考えられる。

Table 2. %stickiness pairs appearing on the whole or on the off-diagonal line of the triangle map.

Categories	Whole		Off-diagonal line	
	Total 20% σ	Average [†]	Average [‡]	Average [¶] over threshold
α -biased	18137	518.2	112.5 (30.1)	80.2 (34.0)
β -biased	14546	415.6	94.6 (29.6)	52.6 (32.2)
Non-biased	17914	597.1	99.1 (21.5)	53.8 (26.8)

* The summation was taken over all the different peptides of non-biased (30 species), α -helix biased (35), and β -sheet biased (35).

[†] (the number of %stickiness atomic pairs on the triangle map) / (the number of molecules).

[‡] (the number of %stickiness atomic pairs on the off-diagonal line) / (the number of molecules).

[¶] The number of %stickiness pairs which is larger than or equal to the threshold value was subjected to the average calculation

3.2 確率的二次構造簡易判定方式

最後に、Figure 3 に示す MD シミュレーションで得られた“ペプチドの 20% σ と 旋回半径 (R_g) の相関”を元にして、これまでの結果を踏まえて、ペプチドが形成する 2 次構造の判定方式として次のものを試みに提唱する。

$$P = S(R_g/0.5, \sigma/2) \quad (1)$$

ここにおいて、 R_g および σ はそれぞれ、試料についてえられた旋回半径および 20% スティッキネスの値を示し、 S は $R_g - \sigma$ 平面の区画座標の関数を表している。 P はその時の確率であり、経験的に Table 3 で与えられるものとする。Table 3 は現在得られている 100 種の異なるヘキサデカペプチドに関するものであり、統計的に信頼性のあるものにはなっていないが、今後、試料数を増やすことによりこのテーブルの信頼性を高めていくことが出来ると考えられる。この表はペプチドに関する計算から得られているがタンパク質の部分構造としての二次構造予測にも当然使いうると期待される。

4 結論

本研究を通じて、ヘキサデカペプチド 100 種に関して ns - MD シミュレーションを行ったが、その結果えられたペプチドの R_g や % スティッキネス解析 (動的

構造解析) などからペプチドの α ヘリックス形成能が確率的に予測可能となった。一方、% スティッキネスの三角マップ解析を導入することで α ヘリックス形成傾向が可視化されたと考えられる。

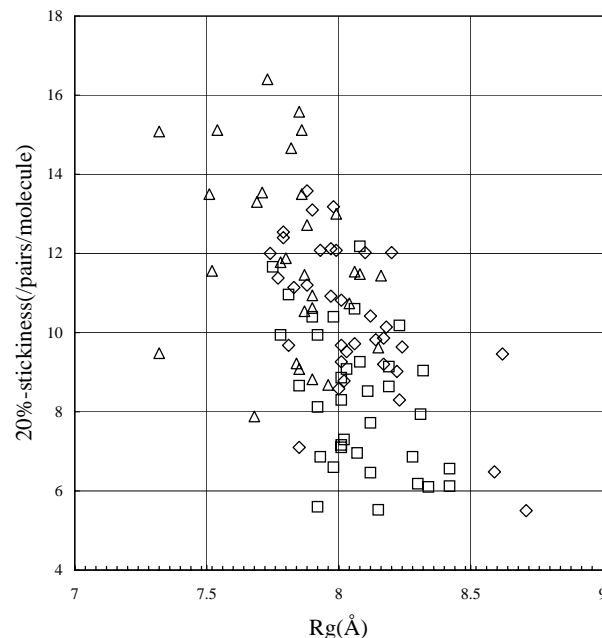


Figure 3. Correlations between 20%-stickiness and R_g of each category of hexadecapeptides. The symbols, \diamond , \square , \triangle , represent α -helix forming biased, β -sheet forming biased, and non-biased, respectively. The number of each symbol per block was counted to obtain the probability.

Table 3. Probability matrix of secondary structure formation obtained by MD simulation.

Stickiness \ R_g -interval		15*				16				17				18			
		7.0~7.5				7.5~8.0				8.0~8.5				8.5~9.0			
3 [†]	4.0~6.0	-	-	-	-	‡0.01	0	1	0	0.01	0	1	0	0.01	1	0	0
4	6.0~8.0	-	-	-	-	0.04	1	2	1	0.12	0	12	0	0.01	1	0	0
5	8.0~10.0	0.01	0	0	1	0.09	1	4	4	0.21	12	8	1	0.01	1	0	0
6	10.0~12.0	-	-	-	-	0.16	5	4	7	0.09	3	2	4	-	-	-	-
7	12.0~14.0	-	-	-	-	0.14	8	0	6	0.03	2	1	0	-	-	-	-
8	14.0~16.0	0.01	0	0	1	0.04	0	0	4	-	-	-	-	-	-	-	-
9	16.0~18.0	-	-	-	-	0.01	0	0	1	-	-	-	-	-	-	-	-

* The serial number for R_g -Interval. [†] The serial number for %stickiness Interval. [‡] Figures within a section represent the ratio of the number of points within a section, the total number of points, and the number of points belonging to α -helix-biased, β -sheet-biased, and non-biased actually found.

本研究の一部は、JSTの支援による埼玉県地域結集型共同研究事業「埼玉バイオプロジェクト (REDS)」の補助を受けて行われたものである。

参考文献

- [1] Gnanakaran S., Nymeyer H., Portman J., Sanbonmatsu K. Y., and Garcia A. E., *Curr. Opin. Struct. Biol.*, **13**, 168-174 (2003).
- [2] Snow C. D., Sorin E. J., Rhee Y. M., and Pande V. S., *Annu. Rev. Biophys. Biomol. Struct.*, **34**, 43-69 (2005).
- [3] Fetrow J. S., Giammona A., Kolinski A., Skolnick J., *Curr. Pharm. Biotechnol.*, **3**, 329-47 (2002).
- [4] Baldwin R. L. and Rose G. D., *Trends Biochem. Sci.*, **24**, 26-33 (1999).
- [5] Spörlein S., Carstens H., Statzger H., Renner C., Behrendt R., Moroder L., Tavan P., Zinth W., and Wachtveit J., *Proc. Natl. Acad. Sci. USA.*, **99**, 7998-8002 (2002).
- [6] Biyani M. and Nishigaki K., *J. Biochem.*, **138**, 363-373 (2005).
- [7] Murayama S., Mori T., and Nishigaki K., *BIOINFO2005*, 300-04 (2005).
- [8] Gilbert W., *Nature*, **271**, 501 (1978).
- [9] Kitamura K., Kinoshita Y., Narasaki S., Nemoto N., Husimi Y., and Nishigaki K., *Protein Eng.*, **15**, 843-853 (2002).
- [10] Chou P. Y. and Fasman G. D., *Annu. Rev. Biochem.*, **47**, 251-76 (1978).

Stochastic Prediction of Peptide Secondary Structure Based on Molecular Dynamics Simulation Performed for Nano-second—Effectiveness of the %stickiness Approach Represented by Triangle Map Mode.

Shinichi MURAYAMA, Chuya YOSHIDA, Takashi AOYAMA, Satoshi URATA
and Koichi NISHIGAKI*

Department of Functional Materials Science, Saitama University
255 Shimo-okubo, Sakura-ku, Saitama, Saitama 338-8570, Japan
*e-mail: koichi@fms.saitama-u.ac.jp

Nano-second (ns) MD simulations of peptides, which are known to be insufficient to obtain stable structures, were confirmed to provide faint secondary structure (especially, α -helix) images through a novel approach termed triangle map representation of %stickiness (TMR-% σ). %stickiness is a measure formerly introduced to describe dynamic conformation changes of biopolymers. In TMR-% σ representation, the α -helix forming propensity of a peptide was expressed as an off-diagonal dotted line, indicating α -helix intrinsic interactions. The coordinate information derived from % σ and radius of gyration (R_g) of a peptide calculated from ns-MD results could be converted to a probability matrix of secondary peptide structure formation.

Keywords: ns-MD, Chou-Fasman parameter, α -helix, %stickiness, Triangle map representation, Secondary structure prediction

Appendix

Table A1. シミュレーションに用いたペプチド配列

α ヘリックス形成傾向バイアス		β シート形成傾向バイアス		非バイアス	
Code	Sequence	Code	Sequence	Code	Sequence
P α -a01	EKEQQCCLDHKLMLCEL	P β -b01	YIIIVYYYYVFWYIYHT	P0-001	IYAFCGRRRARAVGIDK
P α -a02	CELCMMMMLKALLMLK	P β -b02	WVTRIFVTWIIFFWWF	P0-002	AGVGLALRNACFIVTT
P α -a03	CMHEQDEHQMCHHAKE	P β -b03	VWVWVWVWVWVWVWVWV	P0-003	ERNKMLLMETNKGNI I
P α -a04	KHKQLQCHLHGCLAHH	P β -b04	WVTWYIIYFTWTIMVY	P0-004	LGKGAIRFVLTETLRR
P α -a05	QMLLGKAYKLHCDKQL	P β -b05	GIFYFDIVTTWWVWVY	P0-005	LRSNRVDVERSIVDGT
P α -a06	ECYQLEIEKCECQMM	P β -b06	YIITSIIYFVIIFTYFG	P0-006	RSSTRNFMINDNGKL
P α -a07	KQKLSHCACHQCSMQA	P β -b07	TVFWTVTWVWVWVWVWV	P0-007	MANSGLRLISEIFAR
P α -a08	HMGCEKQHQEEHQCAA	P β -b08	WAVYWFVHVWVWVWVWV	P0-008	ANLHSGYETNNQGDNL
P α -a09	AMLQHHHTLLCLCLH	P β -b09	VFVYVWVYVWVWVWVWV	P0-009	ASKIKVAIECAMEGMF
P α -a10	QCMHECEAKQQHMKAE	P β -b10	IETVYWWIFVIWVYVTE	P0-010	LTTFKWDIRGLENNL
P α -a11	LHHMCKMKHKLACEED	P β -b11	FFWYVWFYHFFVYVWV	P0-011	RDCNSGFRKWTDIRNN
P α -a12	KGEKQKCEWAMMELLL	P β -b12	VTTTTIFWYVWVWVWV	P0-012	RMNNELDQIKRSTYYL
P α -a13	KALECMQQEKSIMCAE	P β -b13	WSVIVTTTTTFFVFI	P0-013	QYGGDRMDDWDKKELL
P α -a14	LQEEALLQYEMELKGQ	P β -b14	ITFTVYFFVWFITIVI	P0-014	GYWTDRLTMTYRNLI
P α -a15	HMMEFMMCLECEAACK	P β -b15	VYNYTTTVFWYIFTVY	P0-015	SACELKMRSILKVEKT
P α -a16	EEMCLGMQMHKYQAH	P β -b16	STFVFWVVIIRVWITC	P0-016	NRITKNVCDSDAAGCA
P α -a17	KMQHCQELQQMEHLLK	P β -b17	VTFVYIFFWVFIQTW	P0-017	TMKRRTQENVRFIENS
P α -a18	LEHHKMKHCWDMKKAH	P β -b18	FVWTFVWVWVWVWVWV	P0-018	GGVGFNMFYIQILKL
P α -a19	TLQQAELAESCQQLHE	P β -b19	YFFVYVWVWVWVWVWV	P0-019	WDNTLLGCIHFSCVQK
P α -a20	HHEKQAKHKEQQHXYQ	P β -b20	FWYVWVWVWVWVWVWV	P0-020	RNIEDCMGQADRDNL
P α -a21	EHMKADMMLKFHMLMI	P β -b21	WIYVIFITISTCICTW	P0-021	LGKDRFLAVEVLHVTI
P α -a22	HKHDELKAMHAHLALM	P β -b22	WGTYFVWVWVWVWVWV	P0-022	LRTERILCTFRFKGAG
P α -a23	LDAHAEKWAHLMMMI	P β -b23	YSFVYVWVWVWVWVWV	P0-023	LKASSDYREAVLDFRT
P α -a24	WMWLHKELMAAEKMD	P β -b24	IFYYWVWVWVWVWVWV	P0-024	GRGFYQNTGSRVRY
P α -a25	MEKKELMHHCDELK	P β -b25	VCWVWVWVWVWVWVWV	P0-025	KVNGNREDGFRRTYDK
P α -a26	HAEADDKALHDALDAK	P β -b26	CIFCIYFVFCGSCSC	P0-026	MEKVCVAVMLILFLLC
P α -a27	AKHKMELKHLLELADH	P β -b27	CGYFTWVYTWNTTII	P0-027	MFKQDFACSMKGSVVC
P α -a28	EMHKAMMAAEVHLHAE	P β -b28	TCCWVHVICVCVCIC	P0-028	RNTVRGINRLHKTQVF
P α -a29	MLAKKEDEDMHMLTA	P β -b29	VFYVEIWWYCTFFYQW	P0-029	AADIDSGLDYDRATNN
P α -a30	EDEADDADHDAEDDLA	P β -b30	YIWFYVWVWVWVWVWV	P0-030	VTHEGFVIKLNACRE
P α -a31	CAMQAVKEMIKQFDAQ	P β -b31	YMYMVIQICIVTYTCV		
P α -a32	WVWLCVWVWVWVWVWV	P β -b32	MIMTYLMWVWVWVWV		
P α -a33	CFQARMHIEALMMMV	P β -b33	LLWLLCVWVWVWVWV		
P α -a34	WKKHELHWHKHFHKK	P β -b34	LFVWYFVWVWVWVWV		
P α -a35	QKWMWVKHSWLFMMFQ	P β -b35	AWWVWVWVWVWVWV		

*16 残基の各位置に 9 割の確率で、Chou-Fasman の経験則において α ヘリックス形成傾向の高いアミノ酸が入るように作成した (α ヘリックス形成傾向バイアス: P α -a01 ~ a35) 同様に β シートに対しても行った (β シート形成傾向バイアス: P β -b01 ~ b32) 一方、 α ヘリックス、 β シートどちらにも偏向をかけない非バイアス (P0-001 ~ 030) を作成した。

Table A2. MD シミュレーション条件

非結合による相互作用のカット	OFF
GB/SA モデル	ON
ステップ数	500,000
時間刻み	2 fs
温度	300K で一定
塩濃度	0.2 M