

Generalized Born energy の検証：
レプリカ交換 MD 法による様々なペプチドの構造サンプリング

亀田倫史

産業技術総合研究所 生命情報科学研究センター(CBRC)

kameda-tomoshi@aist.go.jp

生体高分子のシミュレーションをする際に、溶媒である水の効果をきちんと考慮することが重要である。当然、水分子を陽に含めた系を用いれば良いわけであるが、原子数が多いため、計算には現在の計算機資源をもってしても非常に時間がかかってしまう。この問題に対処するためのアプローチは大きく言って2つある。一つは水分子を含んだ系のエネルギーを多極展開などの近似を用いて高速に求める方法で、もう一つは水分子を陽には含めず、溶媒効果を近似ポテンシャルで表現し高速化する方法である。これまで双方ともに様々な手法が提案されてきた。

2000年に入り、近似ポテンシャルの一つである Generalized Born(GB)エネルギーを用いた MD によって、20 残基程度のペプチドを伸びきった構造から天然構造へと巻き戻らせることに成功したと複数のグループが報告し、その有用性に注目が集まっている。しかし一方で、水分子を陽に含んだ場合はうまく実験結果を再現するのに、GB を用いた場合は大きく結果が異なってしまう、との報告もいくつかある。

今回はこの GB エネルギーを用いて、様々な構造のペプチド (α 、 β 、poly-proline II helix(PPII helix、この構造を知らない人は Discussion を参照))を計算し、実験結果と比較することにより、その有用性を検証した。

Methods

力場は CHARMM の param19 を用いた。構造空間に無数存在するエネルギー極小点に留まることなく空間を広く探索し、低温での構造アンサンブルを正しく得るために、従来の MD 法ではなくレプリカ交換 MD 法を用いて計算した。Time step は 1fs、水素原子を含む結合に SHAKE を適用し bond 長を固定した。熱浴は Nose-Hoover 法を用いた。

ここで GB エネルギーについて少し述べておく。GB エネルギーは Still らによって Born エネルギーを拡張する形で導入された。以下の式で表現される。

$$G_{pol} = -166.0 \left(1 - \frac{1}{\epsilon}\right) \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{\sqrt{r_{ij}^2 + b_i b_j} \exp\left(-\frac{r_{ij}^2}{4b_i b_j}\right)}$$

$\{b_i\}$ は Born 半径と呼ばれており、「原子がどれだけペプチド内部に埋もれているか」を表現する量である。粗い表現をすれば「各原子の埋もれ度 $\{b_i\}$ 」に応じて、各原子ペアごとの誘電率を定義し、静電ポテンシャルを求めることで溶媒環境を表現するのが GB エネルギーである。様々な Born 半径の求め方が提案されているが、今回は Dominy らが CHARMM 用に開発したものを使用した²。

さらに、レプリカ交換 MD 法についても簡単に説明しておく。低温から高温までの MD (レプリカと呼ぶ) を同時に

多数行い、定期的に隣り合うレプリカ間の温度を交換し、計算を続けていくのがレプリカ交換MD法である³ (MDではなくモンテカルロ法で行うことも可能である。実際、はじめにレプリカ交換法を提案した福島ら⁴はモンテカルロを使用している) 温度交換をメトロポリス判定によって行えば、交換の前後で詳細釣り合いを満たすので、計算結果は正しい熱平衡に達する。蛋白質はスピニングラスと同様にフラストレーションが存在する系であり、エネルギー空間上に局所的最小が多数存在する。このため、従来のMDを低温で行うと局所的最小に系が捕まってしまう、構造が変化しなくなってしまうことがあったが、レプリカ交換MD法を用いればその構造が高温に移ることで局所的最小から抜け出すことができ、効率のよいサンプリングを行うことができる。

実際に計算する前に、レプリカ交換MD法を行うためのパラメータ (レプリカ数、温度幅) 決定を行う必要がある。レプリカ間の交換を促進し、準安定状態に捕らわれないようにするにはできるだけレプリカ数を多くすべきである。また、様々な構造をサンプリングするには広い温度範囲をとる必要がある。適切なパラメータを用い、シミュレーションが様々な構造をきちんとサンプリングしているならば、異なった初期構造から始めた計算結果はすべて同じふるまいを記述するはずである。しかし、レプリカ数が大きな計算はそのぶん計算機リソースを大きく消費するので、実際に計算する際にはできるだけ減らしたい。そこで、生理的条件下で α helixを形成するペプチドAc-AAAAA(AAARA)₃A-NH₂を用いたテスト計算を行い、適切なパラメータ設定を調べた。具体的にいうと、初期構造として伸びきった構造 (extend) を全レプリカに振り分けた計算と、 α helix構造を全レプリカに振り分けた計算を行い、計算結果が一致するか比較した。Fig.1にレプリカ数を32、温度幅を270~500Kとして行った二つの計算結果を示す。2つの計算結果は非常に似たものとなり、このパラメータを用いた計算をすれば、十分にサンプリングできることがわかった。

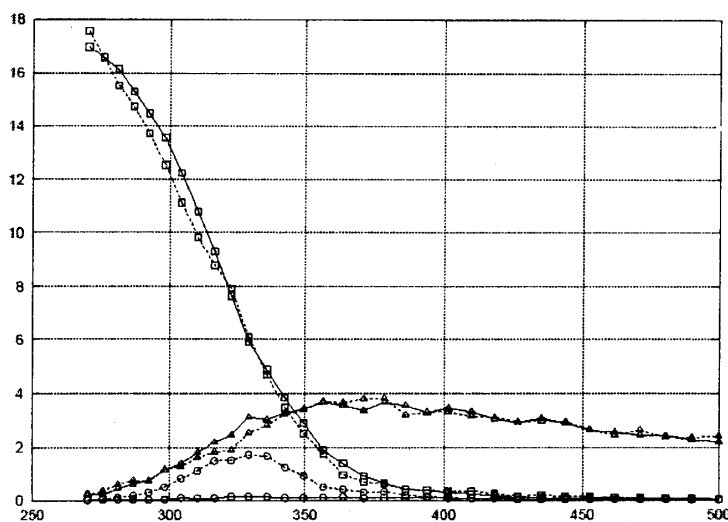


Fig.1 各温度 (レプリカ) に対するAc-AAAAA(AAARA)₃A-NH₂の二次構造形成度
横軸は温度、縦軸は二次構造を形成している平均アミノ残基数を示す。実線は α helixを初期構造とした計算、点線はextend構造を初期構造とした計算結果である。□は α helix、○は β 構造、△はPPII helixを示す。二次構造形成度が収束するまで計算を行い、後半4ns間のデータの平均値を用いた。具体的には α helixから始めた計算は12.5ns、extendから始めた計算は20ns行った。また、温度交換は5psごとに温度が隣り合うレプリカ間でメトロポリス判定することによって行った。

Results

以下、計算結果を示す。初期構造としてextend構造を用いた計算を、各ペプチドに対して一回ずつ行った。計算に用いた条件などはすべてテスト計算と同一のものを使用した。これは、テスト計算で得られたパラメータをそのまま使用すれば十分なサンプリングをすることが期待できるからである。なぜなら、十分なサンプリングに必要なレプリカ数は自由度の0.5乗に比例することがすでに示されているが⁵、テスト計算に用いたペプチドは21残基、今回計算するペプチドはすべて20残基以下の大きさであるから、多くとも32レプリカ用意すれば十分と考えられるからである。

はじめに、生理的条件下で α helixを形成するペプチドAc-AAAAA(AAARA)₃A-NH₂ (Fig.1)、C-peptide(アミノ酸配列:KETAAAKFERQHS、Fig.2)の計算結果を示す。

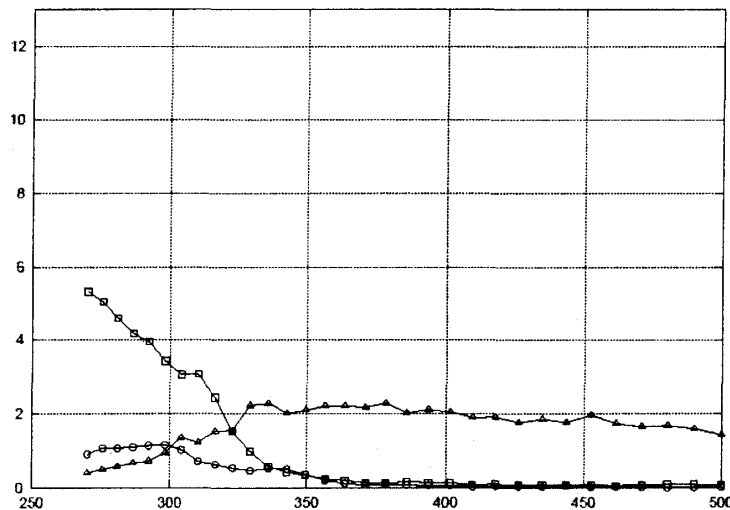


Fig.2 各温度 (レプリカ) に対する C-Peptide の二次構造形成度。

横軸は温度、縦軸は二次構造を形成している平均アミノ残基数を示す。

□は α helix、○は β 構造、△はPPII helixを示す。

Ac-AAAAA(AAARA)₃A-NH₂ (Fig.1)、C-peptide(Fig.2)ともに、低温で α helixが大きな割合で存在しており、実験に合致する。さらに実験によって270Kにおいて、Ac-AAAAA(AAARA)₃A-NH₂の α helix形成度が約80%⁶、C-peptide⁷では約30%存在することが示されているが、今回の計算結果はそれぞれ α =83.1%、40.9%となり定量的にもよい一致を示している。また、Thompsonらによって転移温度(α helixの存在率が50%になる温度)がAc-AAAAA(AAARA)₃A-NH₂は308Kであることがわかっているが⁶、今回の計算結果は315Kを示し、これもよく一致している。

次に、 β sheetを形成するペプチド β -3s(TWIQNGSTKWYQNGSTKYIT、Fig.3,4)、

betanovaFLM(RGWSFQNGKYTLNGKTMEGR)の計算結果を示す。

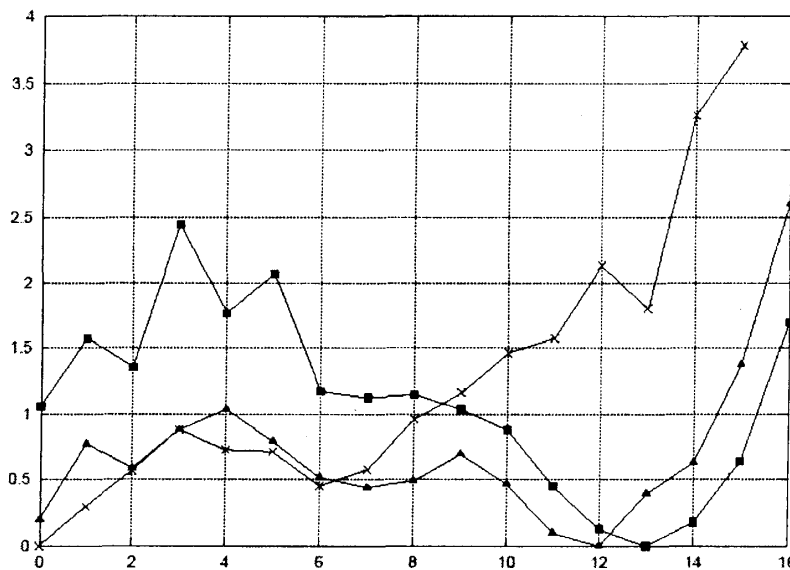


Fig.3 各温度 (レプリカ) に対する beta3s の構造形成度。

横軸は NativeContact 数 (天然構造内で形成されている水素結合、疎水結合などを NativeContact と呼ぶ。この NativeContact が計算で得られた構造中にどのくらい形成しているか数えたのが NativeContact 数である。この値は、大きければその構造は天然構造に近く、小さければ変性状態での構造に近いことを示すため、立体構造形成の反応座標として広く使われている) 縦軸は相対自由エネルギー。各温度での自由エネルギーの最小値を基準 (F=0) とした。■は 270K、▲は 325K、×は 375K での結果を表す。

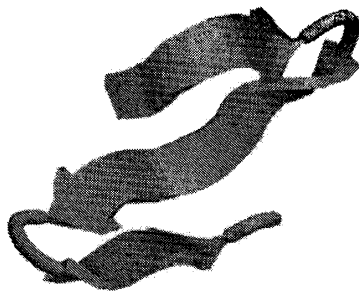


Fig.4 (右図) 270K における代表構造 (NativeContact 数=13)

Fig.3 は beta3s の各温度での構造形成度を示している。低温(270K)では天然構造によく似た構造 (Fig.4) が支配的で、高温 (375K) では、ほどけた変性構造が多数を占めていることがわかる。また、天然構造、変性構造の存在率が等しくなる温度を転移温度と定義すると、305K となり概ね妥当である。また、betanovaFLM についても (詳細は割愛する) 同様の結果が得られている。

最後にpoly-proline II helix (PPII helix)を形成するペプチドについての結果を示す。PPII helixは名前の通りproline (プロリン)のみからなるポリペプチドが形成するhelix構造のことであるが、prolineを含まないペプチドでもPPII helixを形成するとの報告が、2000年頃から多数なされている。例えばShiら⁹はalanine 7merが主にPPII helixをとることをNMRを用いた測定で明らかにしている。また、McCollら⁸は酸性下でのpoly-lysine、アルカリ性下でのpoly-glutamateが、主にPPII helixをとることをラマン散乱によって示している。

今回はそのうちpoly-lysine (lysine 8mer と 21mer)を計算した。すると270Kにおいて8merはPPII helixが70.4%、21merでは46.7%とかなりの高率で存在し、かつ他の二次構造 (α helix、 β sheet) はまったく存在しないことをしめし、実験に合致する結果を得た。

また、従来変性状態ではペプチドの構造はランダムコイルであると考えられてきたが、PPII helixがかなり存在していることを示唆する報告も出てきた。例えばAsherら¹⁰はAc-AAAAA(AAARA)₃A-NH₂を高温にすると、構造が α helixからPPII helixに変化することを示している。ここで改めてFig.1をみると、温度が上昇するにつれて α helixは減少しPPII helixが増加するようすを示しておりAsherらの報告と合致する。

Discussion

ここで、PPII helixについていくつか補足説明をしておく (PPIIについての研究が盛んになってきたのはここ数年のことで、知らない方が多いと思われるので。) PPII helix はらせん一周期当り3残基と α helixに比べ長細い形をしている (Fig.5)

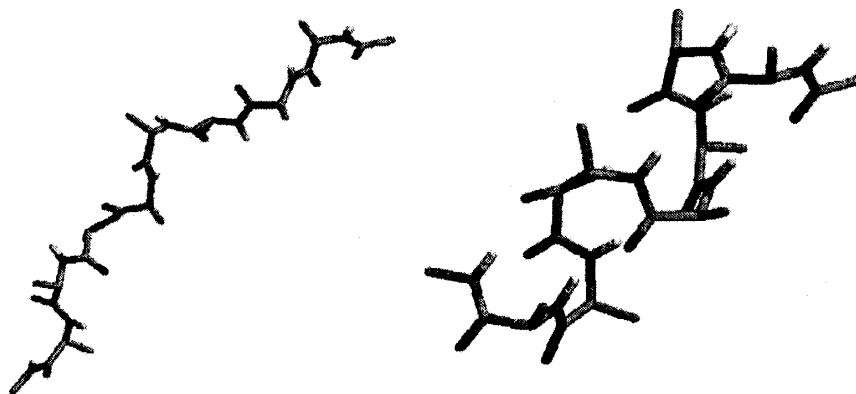


Fig.5 poly-proline II helix(左図)と α helix(右図)

ともに alanine 8mer からなる。縮尺の関係でどちらも同じ大きさに見えるが、PPII helix はらせん1ピッチ当り3残基、 α helixは3.6残基なので、PPII helixの方が長細い。

α helixは主鎖のNH基とCO基の間に形成される水素結合によって安定化されるが、PPII helixは細長い形状のためそのような主鎖原子間の結合は存在しない。そのかわり、主鎖のNH、CO基が水 (溶媒) によく面しているので、それらが水分子との水素結合ネットワークを形成することで安定化されている。このことは水分子をexplicitに入れたシミュレーションで示されている。例えば、Sreeramaら¹¹はMD計算を用いてpoly-alanineのNH基、CO基が水分子1~2個を介した水素結合ネットワークで結ばれている様子を示しているし、Mezeiら¹²は第一水和層に属する水分子とペプチ

ド間の相互作用がPPII helixを安定化させていることを示すなど、様々な研究が水分子との相互作用の重要性を強調している。しかし今回示した結果は、水分子を陽に含めずGBエネルギーを用いた計算であるにもかかわらず、PPII helixの存在を再現している。つまり「implicit solvent でもPPII helixを再現できる」ことを示したわけで、驚くべきことである。また今回示した結果を、「GBエネルギーは水和に関するエネルギーを再現するので、PPII helix形成を再現できた」と言い換えることもできるが、GBエネルギーがPPII helixを安定化させていることも確認している¹³。

Conclusion

以上より、GBエネルギーとレプリカ交換MD法を組み合わせれば、蛋白質の天然構造を正しく予測でき、かつ、その転移温度もかなりの精度で予測することがわかった。

References

1. Still WC et al.(1990)J. Am. Chem. Soc., 112:6127-6129
2. Dominy BN and Brooks CL(1999)J. Phys. Chem. B, 103:3765-3777
3. Sugita Y and Okamoto Y.(1999)Chem. Phys. Lett., 314:141-151
4. Hukushima K and Nemoto K(1996)J. Phys. Soc. Jpn, 65(6):1604-1608
5. Fukunishi H et al.(2002)J. Chem. Phys., 116(20):9058-9067
6. Thomson PA et al.(1997)Biochemistry, 36(30):9200-9210
7. Bierzynski A et al.(1982)Proc. Natl. Acad. Sci. USA, 79(8):2470-2474
8. Shi Z et al.(2002) Proc. Natl. Acad. Sci. USA, 99(14):9190-9195
9. McColl IH et al.(2003) J. Am. Chem. Soc., 125:10019-10026
10. Asher SA et al.(2004) J. Am. Chem. Soc., 126:8433-8440
11. Sreerama N and Woody RW(1999) Proteins, 36:400-406
12. Mezei M et al.(2004) Proteins, 55:502-507
13. Kameda T and Takada S, prepared