

AMBER を用いたタンパク質の自己組織化の機構解析

佐々 和洋^{1*}, 宮本 貴也¹, 後反 克典¹, 宇野 健², 林 治尚³

¹ 福井工業高等専門学校 物質工学科 〒916-8507 福井県鯖江市下司町

² 県立広島大学 地域創生学部 地域創生学科 〒734-8558 広島県広島市南区宇品東一丁目 1-71

³ 兵庫県立大学 学術総合情報センター 〒671-2280 兵庫県姫路市書写 2167

*sasa@fukui-nct.ac.jp

Mechanism Analysis of Protein Folding with AMBER

Kazuhiro SASA,^{1*} Takaya MIYAMOTO,¹ Katsunori GOTAN,¹

Takeshi UNO,² Haruhisa HAYASHI³

¹Department of Chemistry and Biology, National Institute of Technology, Fukui College
(Geshi, Sabae, Fukui 916-8507, Japan)

²Department of Regional Developments, Prefectural University of Hiroshima
(1-1-71 Ujinahigashi, Minami-ku, Hiroshima, Hiroshima 734-8558, JAPAN)

³Library and Academic Information Center, University of Hyogo
(2167 Shosha, Himeji, Hyogo 671-2280, JAPAN)

(Received April 12, 2022; Accepted May 9, 2022)

Abstract

Protein forms a biologically functional steric structure by folding a linear polypeptide consisting of amino acid residues into a three-dimensional structure peculiar to an amino acid sequence. The process of protein self-assembly is called folding, which is said to be assisted by chaperone molecules. However, details have not been confirmed yet. In this study, we performed molecular dynamics simulations of the folding process of tryptophan cage protein (Trp-cage protein), which is a small protein consisting of 20 amino acid residues. Data analysis suggests that the formation of a protein secondary structure is the first step in the folding process, and chaperones may provide energy to overcome the steric energy barrier.

Key words : molecular dynamics simulation, protein folding, chaperone, Trp-cage

1. はじめに

動植物の細胞内で DNA を鋳型として作製されるタンパク質は、長い鎖状の分子が非常に複雑に折りたたまれた立体構造を有している。その構造を取ることにより、タンパク質は非常に特異的な機能を持つことができる。そして、生体内で発現する複雑な生命活動も、このタンパク質の機能により実現していることはよく知られるところである。この重要性から、生体内の様々な反応を考察するために、タンパク質の構造を知る様々な研究が行われてきた。

1961 年、アメリカの Christian Anfinsen は、タンパク質を変性させ、1 次構造（直鎖状のポリペプチド）にした後、しばらくすると再び元の構造と同じように折りたたまれるという機構（アンフィンセンのドグマ）[1]を発見した。つまり、DNA から転写された m-RNA をリボゾームが翻訳することにより、m-RNA に記されたアミノ酸残基の順に直鎖状のポリペプチドが生成されるのであるが、その後ポリペプチドは天文学的な数ほどある構造のうち、ただひとつの構造に折りたたまれると言うことを Anfinsen は説明している[2, 3]。この、ポリペプチドが再現性のある方法で生物学的に機能する立体構造を獲得する物理的な過程のことは、タンパク質の自己組織化（フォールディング）と呼ばれている[4]。

タンパク質がフォールディングを起こす原動力となっているものは疎水性効果[5, 6]とシャペロン[7]であると言われている。疎水性効果によってタンパク質は一般的に、疎水基を側鎖に持つアミノ酸残基が内側に、親水基を側鎖に持つアミノ酸残基が外側になるように折りたたまれる。これは、タンパク質が水溶媒の中に存在する場合が多いため起こる現象である。これがフォールディングの原則とされており、ほとんどのタンパク質がこれに当てはまる。また、フォールディングにはシャペロン分子が関与していることも示唆されている。シャペロンとは、リボゾームから翻訳された直後のポリペプチドと結合し、ポリペプチドが他のタンパク質や細胞質などの影響を受けないようにしながら、正しい構造にフォールディングできるよう環境や機会を提供するタンパク質分子の総称である。シャペロン

の中には、複雑に絡み合った異常タンパク質を直鎖のポリペプチドにほどくような働きを見せるものもあると言われ、アンフォールディングしたものはこのシャペロンの働きにより再び正しいフォールディングへと導いているのではないかと考えられている[8]。

このような微小な細胞内で起こる反応は、化学実験では非常に解析しづらく、現在でも化学実験によるフォールディングの機構の解析には段階的にアンフォールディング[9]や蛍光分光法[10, 11]などを用いた構造変化の観察に依存している。そのため、計算化学的研究がフォールディングにはよく用いられている[12]。本研究では、タンパク質構造をコンピュータ内に再現し、シミュレーションソフト AMBER[13]を用いた分子動力学（Molecular Dynamics: MD）法によるコンピュータシミュレーションでタンパク質のフォールディングを再現し機構を解明することとともにシャペロンの働きについて予測することを目的とする。

2. 方法

生体内のタンパク質は一般的に原子数が非常に多く、分子内における相互作用も複雑に絡み合っている。そこで、本研究では出来る限り小さな分子かつタンパク質の特徴を所持している TC5B (PDB ID: 1L2Y) [14]をシミュレーションの対象とした。PDB (Protein Data Bank) に登録されている TC5B の構造（以降、天然構造と称する）を図 1 に示す。TC5B は 20 アミノ酸残基からなる小さな人工タンパク質であり、第 1 番目の残基から第 12 番目の残基までで 1 つの α -ヘリックス構造を形成する。そして、 α -ヘリックス内の 6 番目残基のトリプトファン側鎖であるインドール環を、タンパク質全体が取り囲む、Trp-cage 構造をしていることが特徴的である。本研究ではこの TC5B に対して、総じて 7 種類の分子動力学計算を行った。

実験 1 として、TC5B の周りに水分子を配置せず、周囲に水の影響のみを考慮する仮想溶媒中の条件下で全原子について 1 ステップ 1 フェムト秒で 2000 万ステップ(計

20 ナノ秒)の MD 計算を行い、実験 2 として、TC5B の周りに水分子を配置する実溶媒中の条件下で全原子について 1 ステップ 1 フェムト秒で 500 万ステップ(計 5 ナノ秒)の MD 計算を実行した。



図 1 TC5B の天然構造

実験 3 として、TC5B の周りに水分子を配置せずに、仮想溶媒中の条件下で N 末端から順に拘束を外しながら 1 ステップ 1 フェムト秒で 500 万ステップ(計 5 ナノ秒)の MD 計算を行った。さらに、実験 4 として、TC5B の周りに水分子を配置し、N 末端から順に拘束を外しながら 1 ステップ 1 フェムト秒で 25 万ステップ(計 0.25 ナノ秒)の MD 計算(実験 4)を実行した。

実験 5 として、タンパク質の周りに水分子を配置せずに、仮想溶媒中の条件下で 2 次構造の部分のみを先に 1 ステップ 1 フェムト秒で 300 万ステップ(計 3 ナノ秒)の MD 計算を行った。また、実験 6 として、タンパク質の周りに水分子を配置し、2 次構造の部分のみを先に 1 ステップ 1 フェムト秒で 500 万ステップ(計 5 ナノ秒)の MD 計算を行い、次に全原子において 1 ステップ 1 フェムト秒で 500 万ステップ(5 ナノ秒)の MD 計算を行った。

実験 7 として、2 次構造の部分のみを作製し、タンパク質の周りに水分子を配置せずに、仮想溶媒中の条件下でその分子において 1 ステップ 1 フェムト秒で 200 万ステップ(計 2 ナノ秒)の MD 計算。その後、残りの残基を追加し、全原子において 1 ステップ 1 フェムト秒で 1200 万ステップ(計 12 ナノ秒)MD 計算を実行した。

実験 3 と実験 4 については、原子を初期構造の座標に強い力で拘束し、N 末端からその拘束を順に外しながら MD

計算を行っていくことで、リボゾームによる、翻訳と同時に端から順にフォールディングが行われるという、より現実に近い状況を再現した。

実験 5、実験 6、実験 7 については、初めに 2 次構造以外の部分を形成させることで、実験 I の条件と比べフォールディングにどのような相違点が生まれるかを確認した。

今回の MD 計算では、1 ステップを 1 フェムト秒としておよそ 1 万~1000 万ステップ(10 ピコ秒~10 ナノ秒)程度の計算を実行した。

設定温度は、生体内を仮定するために、280K~320K の温度条件下でそれぞれ MD 計算を行った。また、気圧はすべて 1atm の定圧条件下で行った。

3. 結果と考察

仮想溶媒中にほぼ直線状の構造から開始した実験 1 において、天然構造との 1000 ステップ (1 ピコ秒) ごとの類似性の指標として RMSD (Root Mean Square Deviation) の算出結果を図 2 に示す。RMSD は対応する 2 点の距離それぞれを二乗し、その相加平均の平方根として定義され、値が小さいほど類似性が高いことを表す。およそ 150 万ステップ(1.5 ナノ秒)で直鎖状のポリペプチドが折りたたまれ、しばらくその形を保つが、300 万ステップ(3 ナノ秒)で構造変化がおき、さらに 800 万ステップ(8 ナノ秒)で大きな構造変化が見られた。このような結果となった理由として、150 万ステップ (1.5 ナノ秒) までは、第 6 番目の残基であるトリプトファンの側鎖を取り囲むようにして折りたたまれていた構造(トリプトファンケージ構造)が、300 万ステップ (3 ナノ秒) でトリプトファンの側鎖が外側に出されたことにより安定しなくなったためであると考えられる。

また、天然構造では構造の内部に取り囲まれたままのトリプトファンの側鎖がこのシミュレーション中で外部に出されてしまったのは、天然構造では形成されている α -ヘリックスがこのシミュレーション中では形成されず、系が不安定になったためであると考えられる。

これらのことから、タンパク質の 2 次構造を先に形成し

た後に全体の構造を形成するという順でフォールディングを行うと、より天然構造に近づくのではないかと予測される。このことについては、後述の実験3以降から検証した。

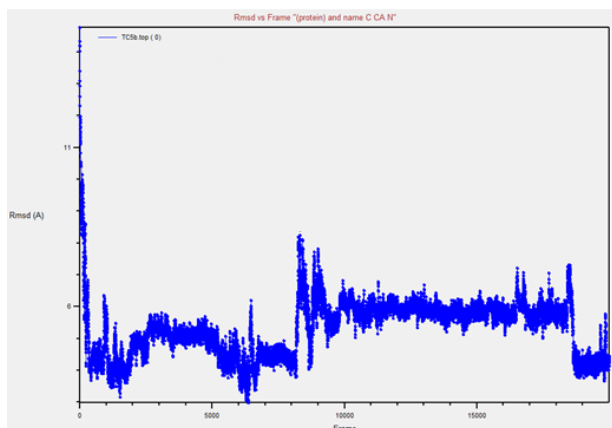


図2 実験1における、シミュレーション結果と天然構造とのRMSD算出値

実溶媒中に実験1と同様の構造を配置した実験2では、500万ステップ(5ナノ秒)のシミュレーションにおいて実験1で確認されたトリプトファンケージ構造は見られず、実験1においての17万ステップ目(170ピコ秒目)と同じような構造でとどまった。

仮想溶媒を使用した実験1と比べて構造変化が緩やかになった理由として、水分子によるタンパク質への水素結合の形成(水和)や、水分子の立体障害により、タンパク質構造がある程度拘束されてしまったと考えられる。また、水分子を配置したことによる仮想溶媒の結果と大きく異なる構造変化についても見られなかった。つまり、仮想溶媒のものと同じような構造変化となっている。このことから、短いシミュレーション時間でフォールディングを再現するためには、仮想溶媒を用いることが有効であると示された。

仮想溶媒中でタンパク質の一部の原子を動かないように拘束し、徐々に拘束を外すシミュレーションを試みた実験3では、計算途中で原子の速度が大きくなりすぎるといった現象が生じ、計算途中でエラーを起こした。エラーを起こした原因を調査すると、拘束している原子と拘束してい

ない原子間の結合距離が大きすぎる値をとり、構造が崩れてしまっていることが確認された。

実溶媒中でタンパク質の一部を拘束し、徐々に拘束を外すシミュレーションを試みた実験4では、実験3のように分子の崩壊は起こらなかったが、分子はほとんど構造変化を起こさず、初期構造に近い直線状の構造を保った。実験3と比べて分子の崩壊が起こらなかった理由としては、水分子が周りに配置されていることによる立体障害や水和により、拘束されている原子とされていない原子の境目に生じる加速度が、比較的大きくなりにくかったためと考えられる。しかし、構造変化がほとんど見られなかったのは、ステップ数が非常に少なかったためと、少しずつ拘束を外していったために、初めのうちは自由に動ける原子が少なく構造変化がしにくかったためと考えられる。よって、よりステップ数を大きくし、拘束を解除する範囲を大きくすると、さらに良いデータが得られるのではないかと予想される。

仮想溶媒中でタンパク質のループ部分を拘束し、先に α -ヘリックス形成を試みた実験5においても、実験3と同じく原子の速度超過によるエラーが生じた。第13残基目のセリン以降のループ部を形成する残基全てを拘束したが、第12残基目とのつなぎ目の部分の原子に速度超過が確認された。実験3および実験5の結果から、仮想溶媒中で分子を拘束するのは得策でなく、2次構造を形成する部分だけ先に計算を行うには、この方法では難しいということが確認された。この改善策として、2次構造を形成する部分とそうでない部分の2つにタンパク質を分け、別の分子として2次構造の部分のみ計算を実施し2次構造を形成させてから、残りの分子と結合させて1つの分子とし、再度タンパク質全体を計算するという方法が考えられる。この方法については後述の実験7にて検証した。

実溶媒中でループ部分を拘束した実験6では、実験5のような分子構造の乱れによるエラーは起こらなかった。また、実験4とは異なり、計算ステップ数も増やし、自由に運動させる部分を天然構造では α -ヘリックス構造が形成される(第1~12残基)と、比較的大きく取った。その結果、実験4と比べ構造変化は大きなものとなったが、それでもま

だ2次構造を形成するには至らなかった。ループ部分を拘束した最初の500万ステップ(5ナノ秒)において、エラーが起こらなかった理由としては実験4と同様に、周囲の水実分子による影響が大きいと考えられる。実験4と比べ非常に長いステップ数と、自由な残基を第1~12残基と広く設定したことにより、構造変化の様子が見られたが、目的とする α -ヘリックスの形成には至らなかった。この理由として挙げられるのはやはり、計算ステップ数がまだ足りなかったためと考えられる。さらに、 α -ヘリックスが形成されていない状態でループ部分の拘束を解除してシミュレーションを続けたが、 α -ヘリックスの形成もトリプトファンページ構造もとることは無かった。 α -ヘリックスを形成させるためには計算ステップ数をさらに増やす必要があるが、これには極めて多くの計算コストが必要となる。これを仮想溶媒として計算するにも、エラーが生じる結果となった。そこで、第13残基以降を削除し、 α -ヘリックス部分(第1~12残基のみ)を初期構造とするシミュレーションを実行する実験7を行った。

実験7では、仮想溶媒中においてタンパク質の一部を拘束することによる原子の速度超過を防ぐために、まずは α -ヘリックス構造の部分(タンパク質の第1~12残基)の構造のみを作製し、その構造に対して計算することでまずは2次構造の形成を試み、200万ステップ(2ナノ秒)のMD計算を行った。シミュレーション結果と天然構造の第1~12残基とのRMSDを算出した結果、徐々に α -ヘリックス構造に近づき、200万ステップ(2ナノ秒)後には2.5Åとなり、わずかに α -ヘリックス構造を形作っていることが確認された。

この実験7は、第1~12残基で構成される2次構造部分のみを作製し200万ステップ(2ナノ秒)のシミュレーションを行ったが、この構造にさらに第13~20残基部分も追加してTC5B全体のシミュレーションも1200万ステップ(12ナノ秒)行った。この方法を採用することで、実験3や実験5のような、分子が崩壊するエラーを起こすことなく2次構造のみを先に形成させ、その後ループ部分を含むシミュレーションの実行を試みた。2次構造のみを先に形成させたタンパク質を初期構造とした1200万ス

テップ(12ナノ秒)のMD計算結果と、タンパク質の天然構造とのRMSDを算出した結果を図3に示す。なお、この計算では構造変化を激しくするために、途中の600万ステップ目(6ナノ秒目)からは温度を380Kと高い温度に設定して行った。

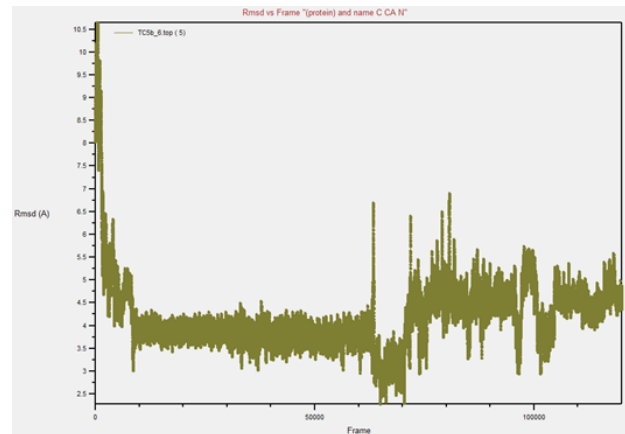


図3 実験7における、シミュレーション結果と天然構造とのRMSD算出値

図3から、100万ステップ(1ナノ秒)以降RMSD値で3.7Å程度の天然構造に近い構造で安定していることが確認された。さらに温度を上げた600万ステップ(6ナノ秒)以降では、620万ステップ(6.2ナノ秒)付近でループ部分が一度開きRMSD値が増加した後、650万ステップ(6.5ナノ秒)付近でRMSD値2.0Åと最も天然構造に近い構造となった(図4a参照)。天然構造に最も近い構造とタンパク質の天然構造とを重ねて表わしたものを図4bに示す。

図4から、この条件で行った計算結果は極めて天然構造と近い構造となっていることが確認できた。しかし、最も天然構造に近い構造をとったにも関わらず、700万ステップ(7ナノ秒)後には他の構造へと構造変化を起こした。これは高エネルギー状態のままシミュレーションを続けたため更なる構造変化を起こしたものと考えられる。構造変化前に温度を常温に戻すことで天然構造に近い構造は維持されたと思われる。

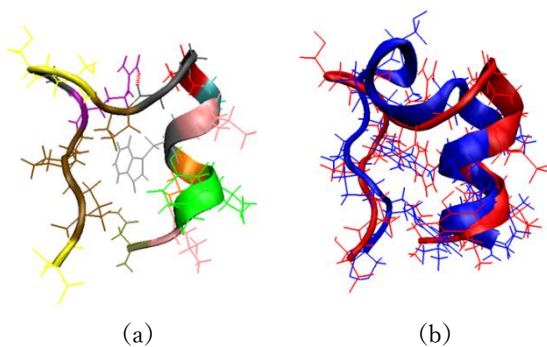


図4 実験7における、(a) 最も天然構造に近いシミュレーション結果 (b) 構造aと天然構造との重ね合わせ。赤が構造a, 青が天然構造

よって、 α -ヘリックスを先に形成させることで天然構造に近い構造へのフォールディングが再現され、フォールディングされた構造は安定性の高い構造をとることがわかった。さらに、加温することで各原子の運動エネルギーを増加させると、さらなるエネルギー障壁となっている構造を経てほぼ天然構造と同じ構造をとることが明らかとなった。これらのことから、 α -ヘリックスなどの2次構造を先に形成させることが、タンパク質のフォールディングを再現するためには有効な手段であり、実際のタンパク質のフォールディングにおいても2次構造の形成が先んじて行われているのではないかと示唆された。

4. 結論

今回の研究においては、天然構造と全く同じ構造にフォールディングするものは見られなかったが、実験1や実験7などにみられるように、非常に近い構造へのフォールディングは再現された。しかし、最も天然構造に近い構造にフォールディングが行われるためには、常温以上のエネルギーが必要であることも明らかとなった。さらに、常温以上の温度では天然構造に近い構造をとったとしても、すぐに他の準安定構造へと構造変化してしまった。このことから、タンパク質は単純な熱力学的分子運動のみで天然構造へのフォールディングは行われず、構造的なエネルギー障

壁により天然構造になる前の準安定な構造でトラップしてしまうことを表している。

しかし、今回の実験では長くても20ナノ秒のシミュレーションであり、20残基程度のタンパク質が実際のフォールディングにかかる時間はミリ秒程度であるという報告[15]と比べるとときわめて少ない時間であるため、本シミュレーションでは天然構造までたどり着かなかっただけであるということも考えられる。

また、実溶媒と仮想溶媒で行ったシミュレーションでは分子の挙動においてほとんど違いは見られなかった。ただし、実溶媒では配置される水分子の影響により分子の構造の変化が非常に遅いものであった。しかし、仮想溶媒では空間中に分子の一部を拘束して計算を行うと分子構造が壊れることなど、実溶媒でしか行えないような計算も存在することも明らかとなった。

このような中、実験7に見られるように、 α -ヘリックス構造などの2次構造の部分みのMD計算ならば単純な熱力学的分子運動のみで短時間で形成されることがわかる。この理由としては、分子内での立体障害や水素結合の影響が少なくなるため、2次構造の形成に必要な水素結合が形成されやすいからと考えられる。そして、2次構造が先に形成されるとより安定にフォールディングが行われることが確認された。

2次構造を先に形成させることで完全なフォールディングの再現には至らなかったが、本研究の中では最も天然構造に近くなった。これより、例えば生体の細胞内に存在するシャペロンの内部には、タンパク質の2次構造をとる部分を判別できる部位があり、先に2次構造を形成させてから3次構造をとらせるようにしていることなど、生体内でもこれに似た機構が存在するのではないかと考えられる。さらには、シャペロンはフォールディング中のタンパク質へエネルギーを与え、天然構造に達すればエネルギーを奪う挙動をしているのではないかとすることも考えられる。そして、シャペロンの内部ではATPなどのエネルギーが貯蓄されていることが報告されている[16]。よって、シャペロンは自身に貯蓄したエネルギーを与えることで、タンパク質の各原子の運動エネルギーを上げることに

より、エネルギー障壁を突破させて天然構造へと導いてい
るのではないかと予測される。

References

- [1] C B Anfinsen, E Haber. *J. Biol. Chem.* 236: 1361–1363. (1961) PMID 13683523
- [2] C B Anfinsen. *Biochem J.* 128(4): 737–749. (1972) DOI:10.1042/bj1280737. PMC: 1173893. PMID 4565129
- [3] C B Anfinsen. *Science.* 181(4096):223-30. (1973) DOI: 10.1126/science.181.4096.223. PMID: 4124164
- [4] Alberts Bruce, Johnson Alexander, Lewis Julian, Raff Martin, Roberts Keith, Walters Peter. “The Shape and Structure of Proteins”. *Molecular Biology of the Cell*; Fourth Edition. New York and London: Garland Science. (2002)
- [5] C N Pace, B A Shirley, M McNutt, K Gajiwala. *FASEB J.* 10(1): 75-83. DOI: 10.1096/fasebj.10.1.8566551. PMID: 8566551
- [6] Di Cui, Shuching Ou, Sandeep Patel. *Proteins.* 82(12): 3312-26. (2014) DOI: 10.1002/prot.24683. PMID: 25204743.
- [7] R A Laskey, B M Honda, A D Mills, J T Finch. *Nature.* 275(5679): 416-20. (1978) DOI: 10.1038/275416a0. PMID: 692721.
- [8] Hartl, F.U., Bracher, A., & Hayer-Hartl, M. (2011) *Nature*, 475, 324–332.
- [9] J K Myers, C N Pace, J M Scholtz. *Protein Sci.* 4(10), 2138–2148 (1995)
- [10] Hugues Bedouelle. *Biochimie.* 121: 29-37. (2016) DOI: 10.1016/j.biochi.2015.11.013. PMID: 26607240.
- [11] Elodie Monsellier, Hugues Bedouelle. *Protein Eng Des Sel.* 18(9): 445-56. (2005) DOI: 10.1093/protein/gzi046. PMID: 16087653.
- [12] Mario Compiani, Emidio Capriotti. *Biochemistry.* 52(48): 8601-24. (2013) DOI: 10.1021/bi4001529. PMID: 24187909
- [13] Weiner, S. J., Kollman, P. A., Nguyen, D. T. and Case, D.A., *J. Comput. Chem.*, 7, 230 (1986)
- [14] Neidigh, J. W., Fesinmeyer, R. M., Andersen, N. H. Designing a 20-residue protein. *Nat. Struct. Biol.*, 9: 425-430, (2002) PMID: 11979279 DOI: 10.1038/nsb798
- [15] *Current Opinion in Structural Biology* 14 (1): 76–88. (February 2004). DOI:10.1016/j.sbi.2004.01.013. PMID 15102453.
- [16] *Nature* 381 (6583): 571–9. (June 1996). DOI:10.1038/381571a0. PMID 8637592.