

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3669704号  
(P3669704)

(45) 発行日 平成17年7月13日(2005.7.13)

(24) 登録日 平成17年4月22日(2005.4.22)

(51) Int.Cl.<sup>7</sup>

G06F 17/50

F I

G06F 17/50 638

請求項の数 4 (全 19 頁)

|               |                         |           |                              |
|---------------|-------------------------|-----------|------------------------------|
| (21) 出願番号     | 特願平8-514452             | (73) 特許権者 | 板井 昭子                        |
| (86) (22) 出願日 | 平成7年10月30日(1995.10.30)  |           | 東京都文京区本郷5-16-6               |
| (86) 国際出願番号   | PCT/JP1995/002219       | (74) 代理人  | 特許業務法人特許事務所サイクス 代表者          |
| (87) 国際公開番号   | W01996/013785           |           | 今村 正純                        |
| (87) 国際公開日    | 平成8年5月9日(1996.5.9)      | (72) 発明者  | 板井 昭子                        |
| 審査請求日         | 平成14年10月29日(2002.10.29) |           | 東京都文京区本郷5-16-6               |
| (31) 優先権主張番号  | 特願平6-267688             | (72) 発明者  | 水谷 実穂                        |
| (32) 優先日      | 平成6年10月31日(1994.10.31)  |           | 東京都文京区小石川5-31-4 マンションかねぎ203号 |
| (33) 優先権主張国   | 日本国(JP)                 |           |                              |
|               |                         | 審査官       | 田中 幸雄                        |

最終頁に続く

(54) 【発明の名称】 三次元構造データベースから新規リガンド化合物を検索する方法

## (57) 【特許請求の範囲】

## 【請求項1】

三次元構造データベースから生体高分子に対するリガンド化合物を検索する方法であって、以下の工程：

- (i) 2種以上の試行化合物の三次元座標に対し、水素結合性カテゴリー番号、分子力場計算用の情報および配座作成用の情報を用意する第1工程；
- (ii) 生体高分子の三次元座標から、水素結合性官能基の水素結合の相手となり得るヘテロ原子の位置にダミー原子を用意する第2工程；
- (iii) 試行化合物の一つを選択し、第1工程で用意した試行化合物の三次元座標に対する水素結合性カテゴリー番号、分子力場計算用の情報および配座作成用の情報、並びに、第2工程で用意したダミー原子を用いて、試行化合物の配座を変化させながらダミー原子間の距離と試行化合物中の水素結合性ヘテロ原子間の距離とを比較することにより、生体高分子と試行化合物をドッキングさせて複合体の構造を作成し、生体高分子と試行化合物との間の相互作用を評価し、最安定の複合体構造を探索する第3工程；
- (iv) 第3工程で探索した最安定複合体構造における生体高分子と試行化合物との相互作用エネルギーの値に基づいて、該試行化合物をリガンド候補化合物として選択するか否かを判定する第4工程；および
- (v) すべての試行化合物について第3工程および第4工程を繰り返す第5工程を含む方法。

## 【請求項2】

10

20

上記第 3 工程が以下の工程：

(1)生体高分子中の水素結合性官能基の水素結合の相手となり得るヘテロ原子の位置に設定したダミー原子と試行化合物中の水素結合性ヘテロ原子との対応づけを組合せ的に網羅することにより、生体高分子 - 試行化合物間の結合様式を網羅する工程 a；

(2)試行化合物の配座を系統的に変化させながら、前記のダミー原子間の距離と前記の水素結合性ヘテロ原子間の距離を比較することにより、生体高分子 - 試行化合物間の可能な水素結合様式及び試行化合物の水素結合性部分の配座を同時に推定する工程 b；及び(3)工程 b で得られた水素結合様式と配座毎に、試行化合物中の水素結合性ヘテロ原子とダミー原子との対応関係に基づいて、試行化合物の全原子の座標を生体高分子の座標系に置き換えることにより、生体高分子 - 試行化合物の複合体構造を得る工程 c を含む、請求項 1 に記載の方法。

10

【請求項 3】

上記第 3 工程が以下の工程：

(1)生体高分子中の水素結合性官能基の水素結合の相手となり得るヘテロ原子の位置に設定したダミー原子と試行化合物の部分構造中の水素結合性ヘテロ原子との対応づけを組合せ的に網羅することにより、生体高分子 - 試行化合物間の結合様式を網羅する工程 a；

(2)試行化合物の配座を系統的に変化させながら、前記のダミー原子間の距離と前記の水素結合性ヘテロ原子間の距離を比較することにより、生体高分子 - 試行化合物間の水素結合様式及び試行化合物の部分構造の配座を同時に推定する工程 b；

(3)工程 b で得られた水素結合様式と配座に基づいて、前記の試行化合物の部分構造中で不可能な水素結合様式を与えるダミー原子と水素結合性ヘテロ原子の組合せ、及びダミー原子と水素結合を形成し得ない水素結合性ヘテロ原子を保存する工程 c；

20

(4)工程 c で保存された水素結合性ヘテロ原子を含む組合せ、及び工程 c で保存されたダミー原子と水素結合性ヘテロ原子の組合せを除いて、ダミー原子と試行化合物全体の水素結合性ヘテロ原子との対応づけを組合せ的に網羅することにより、生体高分子 - 試行化合物間の結合様式を網羅する工程 d；

(5)試行化合物の配座を系統的に変化させながら、前記のダミー原子間の距離と前記の水素結合性ヘテロ原子間の距離を比較することにより、生体高分子 - 試行化合物間の可能な水素結合様式及び試行化合物の水素結合性部分の配座を同時に推定する工程 e；及び(6)工程 e で得られた水素結合様式と配座毎に、試行化合物中の水素結合性ヘテロ原子とダミー原子との対応関係に基づいて、試行化合物の全原子の座標を生体高分子の座標系に置き換えることにより、生体高分子 - リガンド分子の複合体構造を得る工程 f を含む、請求項 1 に記載の方法。

30

【請求項 4】

上記第 3 工程が以下の工程：

(1)生体高分子中の水素結合性官能基の水素結合の相手となり得るヘテロ原子の位置に設定したダミー原子と試行化合物中の水素結合性ヘテロ原子との対応づけを組合せ的に網羅することにより、生体高分子 - 試行化合物間の結合様式を網羅する工程 a；

(2)試行化合物の配座を系統的に変化させながら、前記のダミー原子間の距離と前記の水素結合性ヘテロ原子間の距離を比較することにより、生体高分子 - 試行化合物間の可能な水素結合様式及び試行化合物の水素結合性部分の配座を同時に推定する工程 b；

40

(3)工程 b で得られた水素結合様式を保持しつつ、ダミー原子と試行化合物中の水素結合性ヘテロ原子の位置が一致するように試行化合物の配座を最適化し、次いで、分子内エネルギーの高い試行化合物の配座を除去する工程 c；

(4)工程 c で除去されなかった配座毎に、試行化合物中の水素結合性ヘテロ原子とダミー原子との対応関係に基づいて、試行化合物の全原子の座標を生体高分子の座標系に置き換えることにより、生体高分子 - リガンド分子の複合体構造を得る工程 d；

(5)工程 d で得られた複合体構造から、試行化合物中の水素結合性部分の分子内エネルギー及び生体高分子 - 該試行化合物中の水素結合性部分の分子間相互作用エネルギーの高い複合体構造を除去し、次いで、残った複合体構造について構造の最適化を行う工程 e；

50

(6)工程 e で得られた複合体構造毎に、試行化合物の非水素結合性部分の配座を発生させて新たな複合体構造を得る工程 f ; 及び(7)工程 f で得られた複合体構造から、試行化合物全体の分子内エネルギー及び生体高分子 - 試行化合物の分子間相互作用エネルギーの高い複合体構造を除去し、次いで、残った複合体構造について構造の最適化を行う工程 g を含む請求項 1 に記載の方法。

#### 【発明の詳細な説明】

##### 技術分野

本発明は、医薬、農薬、その他の生理活性化合物の構造設計に利用可能な三次元構造データベースの検索方法に関する。

##### 背景技術

一般に、薬物は標的となる生体高分子と強い相互作用をすることによってその薬理作用を発現する。近年、X線結晶解析やNMRの技術の進歩により、生体内で重要な働きをする生体高分子の立体構造が次々と解明されている。このような研究の進展にともない、従来もっぱらランダムスクリーニングや偶然の発見によってなされてきた新規薬物のリード創製についても、生体高分子の三次元構造の情報に基づいた論理的なアプローチの方法論や成功例が報告されるようになってきた。現在、このようなアプローチの重要性はますます強く認識されており、世界中で様々な角度から研究が行われている。

これらの研究の目標の1つは、リガンドとなり得る構造をコンピューターに構築させる方法を提供することであり、他の目標は、既存化合物の三次元構造データベースからリガンド構造を検索する方法を提供することである。前者(構造自動構築法)は、可能なりガンド構造を構造既知・未知に関係なく広く示唆できる点で優れてる。一方、後者(データベース法)は、既存化合物の構造を蓄積したデータベースから新しい生理活性化合物を見出すのに用いられるが、特に、自社内に保存されていたり入手可能な市販化合物についてのデータベースを対象として検索を行うことにより、検索条件を満たした化合物(ヒット化合物)を合成的労力なしに入手して、標的生体高分子への結合定数や生物活性を測定することができる点で優れている。

新しい骨格構造を有する化合物の合成には、一般に専門家でも数カ月を要するので、数多くの有望な構造を合成して活性を測定するには長い年月と多くの労力を要するが、入手可能な化合物であれば数十もの有望なりガンド候補化合物について結合定数や生物活性を測定するのは容易である。また、ある程度強い結合や活性が認められた化合物の構造を基にすれば、より結合定数や生物活性の強い化合物を設計し、合成することによって効率的にリード創製を行うことが可能である。このような理由から、近年、既知化合物の三次元構造データベースを検索して、新しい生理活性化合物を発見するアプローチに関心が集まっている。

もっとも、この方法の問題点は、どのような検索条件で検索を行うかということにある。一般には、鑄型となる代表的な生理活性化合物が有する原子団や官能基のうち、活性に必須であると推定される原子団や官能基を有しているか否か、あるいはそれらの位置関係が鑄型化合物のものと似ているかどうかを検索条件にして、三次元構造データベースを検索する方法が採用されている。標的生体高分子の構造が不明な場合には基本的にはこの手法に頼らざるを得ないが、このような検索条件は推定や仮説に基づいているために、ヒット化合物が鑄型化合物と同じ活性を有しない場合も多い。構造の特徴が異なる2個以上の分子が同種活性を有している場合、それらの分子の構造に関する情報を用いても、活性に必須な官能基やそれらの位置関係についてよりもっともらしい作業仮説を構築することが可能になるにすぎない。

三次元構造データベースを検索する上で最も困難な問題は、試行化合物の配座の自由度の取扱いである。生体高分子に結合する(最安定複合体を形成する)際のリガンドの配座は、その分子自身の結晶中・溶液中における構造や、エネルギー計算による最安定エネルギーの構造のいずれとも一致する必然性はなく、同じリガンド分子でも相手の生体高分子によって異なる配座で最安定複合体を形成することができることが知られている。一般に、データベース中には、1つの化合物については存在し得る多数の配座のうちの1つの配座

10

20

30

40

50

の三次元構造の原子座標が収められていることが多い。また、結晶構造由来のデータベースを除けば、各化合物の三次元構造の情報は、二次元に入力された構造を計算的に三次元構造にしたものを基にしている。このような三次元構造は、その分子自体のとりうる局所安定構造の1つであることが多い。

そこで、データベースに収められている化合物の配座のみを基にして検索条件を満たすかどうかを調べると、他の配座を考えればヒットすべき化合物が適切に選択されない場合が生じる。考慮すべき配座の数は、結合回転可能な結合数によって変化し、また、どの程度きめ細かく配座を考慮するかなどの要因によっても変わるが、中程度にフレキシブルな回転が可能な結合を3～6個含む分子では一般に数十個から数十万個の配座を扱う必要がある。これらの可能な配座を考慮するには、予め有望な配座を選んでデータベース側に入力しておくか、あるいは検索時に作り出して調べるかのいずれかの方法しかない。いずれにしても、膨大な計算機資源（記憶容量）と計算時間とが必要となる。

最近、Merck社のKearsleyらは、データベース側で1化合物につき最大20個の配座を収めたデータベースを作成し、官能基間の位置関係を主とした検索条件で検索するシステムFLOGを発表した(M. D. Miller, S. K. Kearsley, D. J. Underwood, and R. P. Sheridan: Journal of Computer-Aided Molecular Design, 8, 1994, pp.153-174, FLOG: A system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure.)。FLOGは10万化合物の200万配座を持つデータベースを検索するのに、CRAYというスーパーコンピュータを用いて約1週間を要するという。各化合物について平均20個の配座は、予め各化合物についての配座解析を行ってエネルギーの安定な局所安定構造を選んで収めてあるが、そのための時間と労力は莫大なものである。しかしながら、1化合物当たり20個の配座では十分とはいえない。

また、検索条件としては、鑄型となる活性化合物中で活性に必須と推定される官能基の有無やそれらの間の距離や角度・方向を採用しているが、これらの条件は考えられる条件のうち最も幼稚なものであり、アルゴリズムが易しく計算時間もかからないという利点はあるものの、ヒット化合物が実際に所望の活性を有する確率は高くはない。この理由は、官能基だけが望ましい位置関係にあったとしても、分子全体の形状や大きさが適当でない分子は活性を発揮し得ないからである。標的生体高分子の立体構造が既知の場合にはその情報を用いるのが最も効果的であり、ヒット化合物が所望の活性を有する確率が高い。

例えば、Dupont Merck社のEyermannらは、生体高分子との複合体としてX線結晶解析されたりリガンド分子中の官能基の位置関係に基づいて、データベースを検索した(P. Y. S. Lam, P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong, C. H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharp, and S. Erickson-Viitanen: Science, 263, 1994, pp.380-384, Rational design of Potent, Bioavailable, nonpeptide cyclic Ureas as HIV Protease Inhibitors.)。HIVプロテアーゼの系では、ペプチド性リガンド分子中の2個のベンゼン環の中心でプロテアーゼとリガンドとを水素結合で結んでいる水分子中の酸素原子の3点が同じような位置関係にあるような分子を検索した。ヒット化合物の中からリガンド結合部位にうまく嵌合し、検索条件に入れた酸素以外にも水素結合が形成されている化合物を1つ選び出した。この化合物の構造を基にして、酵素阻害活性を測定しつつ適宜構造を修正した化合物を合成する作業を繰り返した結果、非常に活性の高い化合物を見いだしたと報告している。しかしながら、検索法についての詳しい記述がなく、配座をどのように扱っているか不明である。

ヒット化合物が所望の活性を有する確率が最も高くなる検索条件は、標的生体高分子のリガンド結合部位に化合物が安定に結合できるか否かという検索条件である。もっとも、この条件が満たされた場合においても物性面（水溶性など）が原因となって活性が発現されない場合もあるが、この条件を満足することは活性発現のための最小限の条件である。一方、特定の官能基の有無や官能基の位置関係は必ずしも活性発現にとって必須ではなく、化合物系が異なり分子骨格が相違しても、標的生体高分子との間で同程度に安定な複合体を形成できれば同様な生物活性を発揮できる可能性がある。従って、リガンド結合部位に

10

20

30

40

50

安定に結合できるという条件を満足する化合物は、実際にリガンドとなり得る化合物である可能性が高い。つまり、標的生体高分子のリガンド結合部位へのフィットという検索条件を用いることにより、所望の生物活性を有するより広汎な構造の化合物をデータベースから選び出すことが可能になる。

標的生体高分子に対してある特定のリガンド分子が結合できるか否かを判断するためには、可能な複合体構造のうち最も安定な複合体構造を探索し、その複合体がどの程度に安定かを知ることがあり、与えられた生体高分子とリガンド分子との間の最も安定な複合体構造を探索するには、すべての可能な結合様式（一方の分子を固定して、他方の分子を回転・併進させることに対応する）と可能なリガンド配座を網羅して安定性を評価する必要がある。しかしながら、この作業は膨大な計算を必要とするために、人間がグラフィック・ディスプレイに向かって会話的に行うのは困難であり、これ作業を自動的かつ効率的に行う自動ドッキング法の開発が待ち望まれていた。

自動ドッキングの方法として、Kuntzらはリガンド結合部位の形状を数個～数十個の内接する球を用いて表現し、それらの球の中心がつくるベクトル群とリガンド化合物の分子内の原子間ベクトル群とを比較することにより、可能な結合様式を推定する方法を開発した（R. L. Desjarlais, R. P. Sheridan, G. L. Seibel, J. S. Dixon, I. D. Kuntz and R. Venkataraghavan : J. Med. Chem. 31(1989) 722-729. Using Shape Complementarity as an Initial Screen in Designing Ligands for a Receptor-Binding Site of Known 3-Dimensional Structure.）。最近の改良法では、ベクトルの一致度を示すスコアの他に、分子間エネルギーも計算できるように変更されている。

しかしながら、この方法の欠点は、結合様式の網羅だけについても相当の時間を要することである。改良法では、1つの化合物のドッキングについて配座を変化させることができるように変更されたものもあるが、すべてのドッキングのプロセスを自動的に進むように設計されているのではない。また、エネルギー評価はヒットした化合物についてはできても、結果を得るまでの過程では簡略なスコアしか計算できないし、水素結合等も考慮できない。また、精度にも問題があり、結晶として観測された正しい複合体構造が最安定の結果を与えず、かなり低い順位に見積もられることもあるなど問題点が多い。しかし、致命的な欠陥は基本的に配座を固定して行う方法であり、配座の自由度が取り扱えないことである。特に、多数の化合物を自動的に扱う必要がある三次元データベース検索では、上記の欠点を人間の手仕事で補うことができないため、実用性がきわめて低い。

本発明者らは、生体高分子とリガンド分子間の最安定複合体構造を自動的に、かつ先入観を入れずに推定する方法としてADAMを開発し、上記の問題をすべて解決することに成功した（PCT/JP93/0365, PCT国際公開WO 93/20525; M. Yamada and A. Itai, Chem. Pharm. Bull., 41, p.1200, 1993; M. Yamada and A. Itai, Chem. Pharm. Bull., 41, p.1203, 1993; and M. Yamada Mizutani, N. Tomioka and A. Itai, J. Mol. Biol., 243, pp. 310-326, 1994）。

結晶解析により複合体構造が知られているいくつかの酵素 - 阻害剤分子の系にADAMを適用してみたところ、すべての場合にエネルギー最低の複合体モデルが結晶中の複合体構造を非常によく再現しており、リガンド分子中の結合回転できる結合のねじれ角や水素結合様式なども極めて良く再現されていた。ADAMの高い精度は、全部で3回行われる構造最適化（エネルギー最小化）計算に起因している。所要時間はコンピュータの性能やダミー原子数、ヘテロ原子数、回転可能な結合数などに依存するので一概にはいえないが、広く普及しているワークステーション（R4400）を用いた場合、普通の薬物分子の構造で初期複合体モデルを得るのに数分から1時間弱程度である。この解析速度は、配座固定でなく全部の可能な配座を考慮している方法としては、現在発表されている最速の方法と比べても数十倍速い。

しかしながら、ADAMは一つの生体高分子と一つのリガンド分子間との最安定複合体構造を探索する目的には適しているものの、三次元構造データベース中に収められている多種多様な化合物の中から新規なリガンド化合物を探索するのに利用するには適しておらず、さらなる改良が求められていた。

従って本発明は、従来技術の問題点を解決し、多種多様な化合物の中から新規なリガンド化合物を探索するために適した方法を提供することを目的とする。

#### 発明の開示

本発明者らは鋭意努力した結果、試行化合物の配座を変化させながら生体高分子と試行化合物とをドッキングさせて複合体の構造を作成し、生体高分子と試行化合物分子の間の相互作用エネルギー（例えば、水素結合、静電相互作用、及びファンデルワールス力）を評価し、構造最適化を繰り返すことにより前記複合体の最安定結合様式を探索し、得られた最安定複合体における生体高分子と試行化合物との相互作用エネルギーの値に基づいて試行化合物群の中からリガンド候補化合物を選択することにより、膨大な数の試行化合物から多数の有望なリガンド化合物候補を選択する方法を開発することに成功し、本発明を完

10

すなわち、本発明は、

(i) 2 種以上の試行化合物の三次元座標に対し、水素結合性カテゴリー番号、分子力場計算用の情報および配座作成用の情報を用意する第 1 工程；

(ii) 生体高分子の三次元座標から、リガンド結合領域の物理化学的情報およびダミー原子を用意する第 2 工程；

(iii) 第 1 工程で用意した試行化合物の三次元座標に対する水素結合性カテゴリー番号、分子力場計算用の情報および配座作成用の情報、並びに、第 2 工程で用意したリガンド結合領域の物理化学的情報を基にして、試行化合物の配座を変化させながら生体高分子と試行化合物をドッキングさせて複合体の構造を作成し、生体高分子と試行化合物分子の間の相互作用を評価することにより前記複合体の最安定構造を探索する第 3 工程；

20

(iv) 第 3 工程で探索した最安定複合体構造における生体高分子と試行化合物の相互作用エネルギーの値に基づいて、該試行化合物をリガンド候補化合物として選択するか否かを判定する第 4 工程；および

(v) 全ての試行化合物について第 3 工程および第 4 工程を繰り返す第 5 工程；

を含む、三次元構造データベースから生体高分子に対するリガンド化合物を検索する方法を提供するものである。

さらに、本発明の方法は、生体高分子と形成しうる水素結合の数および／または生体高分子との相互作用エネルギーの値に基づいて、第 4 工程で選択されたリガンド候補化合物を絞り込む第 6 工程を含んでいてもよい。

30

#### 【図面の簡単な説明】

第 1 図は、本発明の方法の構成の概念図として示した図である。

第 2 図は、本発明の方法の一態様についてのフローチャートを示す図である。図中、S は各ステップを示す。

第 3 図は、本発明の方法によりリガンド候補化合物として得られたジヒドロ葉酸還元酵素の基質および既知阻害剤アナログの一例を示す図である。

第 4 図は、本発明の方法により得られたジヒドロ葉酸還元酵素の新規な構造をもつリガンド候補化合物の一例を示す図である。

#### 発明を実施するための最良の形態

本発明の方法において、試行化合物は構造が既知の化合物であれば特に限定されず、例えば、現存するデータベースに収納されている種々の化合物を用いることができる。水素結合性カテゴリー番号とは、水素結合を形成しうる原子団についての識別用番号であり、その原子団中で直接水素結合に関与するヘテロ原子に付与される。その番号により、その原子団の幾何学的構造及び水素結合の性質、並びにその相手となる水素結合性原子（ダミー原子）の位置が即座に作り出される。分子力場計算用の情報とは、分子力場を用いて分子内・分子間の相互作用エネルギーを計算するために各原子に与えられる番号や電子の状態のことであり、原子タイプ番号や原子電荷を含むものとして用いる。また、配座作成用の情報とは、結合回転可能な結合のねじれ角を変化させて異なる配座を系統的に作成するために用いるそれらのねじれ角の初期値・終値及び何度ずつ回転させるかの増分値のことであり、1 個の結合についてねじれ角を構成する 4 つの原子の入力順番号とねじれ角の初期

40

50

値、終値、及び回転角が含まれる。

生体高分子とは、生体に見出される高分子の他、生体に見出される高分子を模擬した分子をも含む概念である。リガンド結合領域の物理化学的情報とは、リガンドが結合できる生体高分子の凹んだ領域の内部において生体高分子の全原子の及ぼすポテンシャル的影響をいい、リガンド結合領域内の三次元格子点の水素結合性（水素受容性又は水素供与性）、その三次元格子点にプローブ原子を置いたときに生体高分子とプローブ原子との間に働くファンデルワールス相互作用エネルギーおよび静電相互作用エネルギーが含まれる。ドッキングとは、あるリガンド化合物の分子とある生体高分子との複合体を形成させること、および両者の安定複合体構造を探ることを含む概念として用いる。ドッキングは、一般的には、分子模型を用いる方法、コンピューターとコンピューター・グラフィックスを用いたシュミレーション手法、又は自動ドッキング法等を用いて行うことができる。

10

生体高分子と試行化合物との相互作用とは、生体高分子と試行化合物との間に働く力、例えば、水素結合、静電相互作用、ファンデルワールス相互作用等などのことである。リガンドあるいはリガンド化合物とは、生体高分子に結合する低分子量の化合物をいい、その例としては、薬物、酵素の基質・阻害剤・補酵素、その他の生理活性物質を挙げることができる。

本発明の方法の一態様によれば、上記の第3工程は、以下の工程：

- (1) 生体高分子中の水素結合性官能基の水素結合の相手となり得るヘテロ原子の位置に設定したダミー原子と試行化合物中の水素結合性ヘテロ原子との対応づけを組合せ的に網羅することにより、生体高分子 - 試行化合物間の有利な結合様式を網羅する工程 a ；
- (2) 試行化合物の配座を系統的に変化させながら、前記のダミー原子間の距離と前記の水素結合性ヘテロ原子間の距離を比較することにより、生体高分子 - 試行化合物間の可能な水素結合様式及び試行化合物の水素結合性部分の配座を同時に推定する工程 b ；及び
- (3) 工程 b で得られた水素結合様式と配座毎に、試行化合物中の水素結合性ヘテロ原子とダミー原子との対応関係に基づいて、試行化合物の全原子の座標を生体高分子の座標系に置き換えることにより、生体高分子 - 試行化合物の複合体構造を得る工程 c

20

を含んでいてもよい。

また、本発明の方法の別の態様に従えば、上記の第3工程は、以下の工程：

- (1) 生体高分子中の水素結合性官能基の水素結合の相手となり得るヘテロ原子の位置に設定したダミー原子と試行化合物の部分構造中の水素結合性ヘテロ原子との対応づけを組合せ的に網羅することにより、生体高分子 - 試行化合物間の有利な結合様式を網羅する工程 a ；
- (2) 試行化合物の配座を系統的に変化させながら、前記のダミー原子間の距離と前記の水素結合性ヘテロ原子間の距離を比較することにより、生体高分子 - 試行化合物間の水素結合様式及び試行化合物の部分構造の配座を同時に推定する工程 b ；
- (3) 工程 b で得られた水素結合様式と配座に基づいて、前記の試行化合物の部分構造中で不可能な水素結合様式を与えるダミー原子と水素結合性ヘテロ原子の組合せ、及びダミー原子と水素結合を形成し得ない水素結合性ヘテロ原子を保持する工程 c ；
- (4) 工程 c で保存された水素結合性ヘテロ原子を含む組合せ、及び工程 c で保存されたダミー原子と水素結合性ヘテロ原子の組合せを除いて、ダミー原子と試行化合物全体の水素結合性ヘテロ原子との対応づけを組合せ的に網羅することにより、生体高分子 - 試行化合物間の有利な結合様式を網羅する工程 d ；
- (5) 試行化合物の配座を系統的に変化させながら、前記のダミー原子間の距離と前記の水素結合性ヘテロ原子間の距離を比較することにより、生体高分子 - 試行化合物間の可能な水素結合様式及び試行化合物の水素結合性部分の配座を同時に推定する工程 e ；及び
- (6) 工程 e で得られた水素結合様式と配座毎に、試行化合物中の水素結合性ヘテロ原子とダミー原子との対応関係に基づいて、試行化合物の全原子の座標を生体高分子の座標系に置き換えることにより、生体高分子 - 試行化合物の複合体構造を得る工程 f ；

30

40

を含んでいてもよい。

第3工程が上記のようなサブ工程を含むことにより、生体高分子 - 試行化合物の安定複合

50

体構造の探索の高速化が可能になるとともに、生体高分子及び／又は試行化合物が複雑な構造を有する場合においても、最安定を含む安定な複合体の構造を短時間に探索することが可能となる。

本発明のさらに別の態様によれば、上記の第3工程は、以下の工程：

- (1) 生体高分子中の水素結合性官能基の水素結合の相手となり得るヘテロ原子の位置に設定したダミー原子と試行化合物中の水素結合性ヘテロ原子との対応づけを組合せ的に網羅することにより、生体高分子 - 試行化合物間の有利な結合様式を網羅する工程 a ；
  - (2) 試行化合物の配座を系統的に変化させながら、前記のダミー原子間の距離と前記の水素結合性ヘテロ原子間の距離を比較することにより、生体高分子 - 試行化合物間の可能な水素結合様式及び試行化合物の水素結合性部分の配座を同時に推定する工程 b ；
  - (3) 工程 b で得られた水素結合様式を保持しつつ、ダミー原子と試行化合物中の水素結合性ヘテロ原子の位置が一致するように試行化合物の配座を最適化し、次いで、分子内エネルギーの高い試行化合物の配座を除去する工程 c ；
  - (4) 工程 c で除去されなかった配座毎に、試行化合物中の水素結合性ヘテロ原子とダミー原子との対応関係に基づいて、試行化合物の全原子の座標を生体高分子の座標系に置き換えることにより、生体高分子 - 試行化合物の複合体構造を得る工程 d ；
  - (5) 工程 d で得られた水素結合性部分の複合体構造から、試行化合物中の水素結合性部分の分子内エネルギー及び生体高分子 - 該試行化合物中の水素結合性部分の分子間相互作用エネルギーの高い複合体構造を除去し、次いで、残った複合体構造について構造の最適化を行う工程 e ；
  - (6) 工程 e で得られた複合体構造毎に、試行化合物の非水素結合性部分の配座を発生させて新たな試行化合物全体を含む複合体構造を得る工程 f ；及び
  - (7) 工程 f で得られた複合体構造から試行化合物全体の分子内エネルギー及び生体高分子 - 試行化合物の分子間相互作用エネルギーの高い複合体構造を除去し、次いで、残った複合体構造について構造の最適化を行う工程 g ；
- を含んでいてもよい。

第3工程が上記のようなサブ工程を含むことにより、生成させる配座の数を減少させても適切な複合体構造を選択することが可能になり、精度のよい安定複合体が得られる。

上記方法において、水素結合性官能基には、水素結合に関与することのできる官能基及び原子が含まれる。また、水素結合性ヘテロ原子とは、試行化合物中に存在する水素結合性官能基を構成するヘテロ原子をいう。水素結合性部分とは、試行化合物の構造のうち、ダミー原子と対応づけられる水素結合性ヘテロ原子を含む構造部分をいい、非水素結合性部分とは、水素結合性部分以外の構造部分をいう。

本発明のさらに別の態様によれば、試行化合物の三次元座標を読み込んで、必要な場合には水素原子の三次元座標を付加し、原子タイプ番号、水素結合性カテゴリー番号、原子電荷および結合回転の様式を付与することにより、三次元データベースを作成する方法が提供される。

以下、第2図のフローチャートを参照しつつ、本発明の好ましい態様について説明する。第2図中のSは各ステップを示しており、ADAMとは、国際公開W093/20525号明細書（国際公開日1993年10月14日）に開示された生体高分子 - リガンド分子の安定複合体構造の探索方法のことを示す。なお、PCT国際公開W0 93/20525号明細書の開示を本明細書の開示に含める。

最初に、2種以上の試行化合物のADAM仕様三次元データベースを作成する（S1）。上記のデータベースは、例えば以下のようにして作成することができる。まず、試行化合物の三次元座標を入力する。試行化合物の三次元座標としては、二次元構造データベースから得られる二次元座標を力場計算的に三次元座標に変換した座標、単独の結晶の構造解析、結晶データベースやエネルギー計算を基にしたモデルビルディングより得られた立体構造等から得た座標などを用いることができる。試行化合物の原子座標が表す立体構造については、原子間の結合距離、結合角等の幾何学的な量が正確であれば、基本的には配座を問題とする必要はない。もっとも、試行化合物を構成する各原子の電荷をできるだけ

10

20

30

40

50



正確に算出するためには、ある程度単独で安定な立体構造を与える原子座標を入力することが好ましい。

上記の試行化合物の三次元座標を入力した後、必要な場合には水素原子を付加し、さらに各原子の原子タイプ番号、水素結合性ヘテロ原子の水素結合性カテゴリー番号、各原子の原子電荷、結合回転の様式（どの結合を回転させるか、あるいは何度から回転を始めて何度おきに回転させるか等）を自動的に付加することにより、三次元データベースを作成することができる。

試行化合物を構成する原子の原子タイプ番号は、例えばWeinerら (Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., and Weiner, P., J. Am. Chem. Soc., 106, 1984, pp.765-784)の行った番号付けを用いることができる。試行化合物中に存在する水素結合性ヘテロ原子の水素結合性カテゴリー番号は、下記の表1に従って付加することができる。試行化合物を構成する各原子の電荷は、例えばGasteiger法、MOPACプログラム中のMNDO法、AM1法など用いた分子軌道法計算により算出することができる。結合回転の様式については、ねじれ回転可能な単結合を  $60^{\circ} \sim 120^{\circ}$  の一定の回転角で系統的に回転させる回転様式を指定することが好ましい。

表 1. 水素結合性官能基を構成する原子の種類と水素結合性カテゴリー番号

| 水素結合性カテゴリー番号 | 水素結合性官能基を構成する原子の種類    |    |
|--------------|-----------------------|----|
| 1            | 第一アミンの $sp^2$ の N     | 10 |
| 2            | 第一アミンの $sp^3$ の N     |    |
| 3            | アンモニウムイオンの $sp^3$ の N |    |
| 4            | アミドの $sp^2$ の N       |    |
| 5            | 第二アミンの $sp^3$ の N     | 20 |
| 6            | 芳香族の N                |    |
| 7            | プロトン化した芳香族の N         |    |
| 8            | 第三アミンの N              |    |
| 9            | プロトン化した第三アミンの N       | 30 |
| 10           | C-O結合の回転可能なヒドロキシルの O  |    |
| 11           | エーテルの O               |    |
| 12           | カルボニルの O              |    |
| 13           | カルボシキレートアニオンの O       |    |
| 14           | カルボン酸の O              |    |
| 15           | ホスフェートの O             |    |
| 16           | 水分子の O                |    |
| 17           | メルカプトの S              |    |
| 18           | チオエーテルの S             |    |
| 19           | 水素の位置が固定したヒドロキシの O    |    |

40

つぎに、生体高分子を構成する原子の番号と原子座標（但し、水素原子の原子座標は除く。）を入力する（S2）。生体高分子の原子座標は、X線結晶解析やNMR解析により得られた立体構造の情報、蛋白質結晶データベース等から入手可能な情報、あるいはこれらの情報を基にして構築された生体高分子モデルの原子座標などを用いることができる。生体高分子の原子座標は三次元座標系で表されることが好ましい。また、生体高分子に結合して構造的にも機能的にも重要な役割を果たす補因子を生体高分子の一部とみなして、該補因子を結合した状態の生体高分子の原子座標を入力して以後のステップを行ってもよい。このような補因子としては、補酵素、水分子、又は金属イオン等を挙げることができる。

上記ステップに続いて、生体高分子中に存在するアミノ酸残基を構成する水素原子の原子

50

座標を算出する (S3)。一般に、生体高分子の水素原子の位置を X 線結晶解析や NMR 解析等の実験的手法で求めることは困難であり、また、蛋白質結晶データベース等からも水素原子に関する情報は入手できない。従って、生体高分子中に存在するアミノ酸残基の構造に基づいて、該アミノ酸残基を構成する水素原子の原子座標を算出する必要がある。回転可能なアミノ酸残基に結合しているためにその原子座標を一義的に決定できない水素原子は、トランス位に存在すると仮定して原子座標を算出することが好ましい。

その後、生体高分子中に存在するアミノ酸残基を構成する原子に電荷を付加し (S4)、生体高分子中に存在する水素結合性官能基を構成するヘテロ原子について水素結合性カテゴリ番号を付加する (S5)。電荷の値としては、各アミノ酸について算出されている文献値、例えば、Weiner の値 (Weiner, S. J., Kollman, P. A. Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr. and Weiner, P., J. Am. Chem. Soc., 106, 1984, pp.765-784) などを用いることができる。水素結合性カテゴリ番号は、上記の表 1 に従って付加すればよい。

つぎに、リガンド結合領域を指定する (S6)。リガンド結合領域としては、生体高分子の任意の部位を含む領域、好ましくは直方体の領域を選択することができる。目的に応じてリガンド結合ポケットやその周辺の生体高分子の部位を選択してもよく、必要に応じてエフェクター等の他の分子の結合する生体高分子の部位を含む領域を選択することも可能である。なお、リガンド結合ポケットとは、生体高分子の凹んだ分子表面にあり、基質や阻害剤などのリガンド分子が結合する内孔をいう。リガンド結合領域の範囲の指定には、例えばプログラム GREEN (Tomioaka, N., Itai, A., and Itaka, Y., Journal of Computer Aided Molecular Design, vol.1, 1987, pp.197-210) の機能の一部を利用することができる。

S6 で指定した領域内に三次元格子点を発生させ、各三次元格子点につき番号付けと格子点情報の算出を行う (S7)。三次元格子点とは、生体高分子のリガンド結合領域内の一定間隔毎に発生させる三次元格子上の点をいい、格子点情報とは、各三次元格子点上にプローブとなる原子があると仮定して算出した生体高分子とプローブ原子の間に働くファンデルワールス相互作用エネルギー及び静電相互作用エネルギー、並びに水素結合性等のリガンド結合領域の局所的な物理化学的情報を含む概念である。この三次元格子点の格子点情報を利用することにより、以後のステップで行う生体高分子と試行化合物の分子間相互作用エネルギーの近似計算の高速化が可能となり、また、以後のステップで設定するダミー原子の位置を合理的に決定することが可能となる。その結果、生体高分子と試行化合物のドッキングモデルを短時間に網羅的に探索できるようになる。

三次元格子点は、0.3 ~ 2.0 オングストローム、好ましくは、0.3 ~ 0.5 オングストロームの一定間隔で S6 で指定した領域内に発生させることができる。プローブ原子としては、リガンドとなる可能性のある化合物に含まれるすべての原子種を採用することが好ましい。各三次元格子点に配置した各プローブ原子と生体高分子との間に働くファンデルワールス相互作用エネルギーは、経験的なポテンシャル関数を用いて常法の原子対計算により算出することができる。経験的なポテンシャル関数としては、以下の式で表されるような Lennard-Jones type 関数を用いることができる。

$$G_{vdw,i} = \sum_j (A/r_{ij}^{12} - B/r_{ij}^6)$$

式中、i はプローブ原子の位置を示す番号を、j は生体高分子の原子の番号を表す。A 及び B は極小のポテンシャルの位置と大きさを決定するパラメーターを表す。 $r_{ij}$  は、i 番目に配置したプローブ原子と生体高分子の j 番目の原子との間の距離を表す。なお、上記の A 及び B のパラメーターとしては、例えば Weiner らの値 (Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr. and Weiner, P., J. Am. Chem. Soc., 106, 1984, pp.765-784) を用いることができる。

試行化合物をこの三次元格子点上に配置した際の試行化合物 - 生体高分子間のファンデルワールス相互作用エネルギーは、以下の式 (式中、m は試行化合物中の原子の番号であり、N は試行化合物中の原子の数であり、 $m_0$  は m 番目の原子に最も近い三次元格子点の番

10

20

30

40

50

号である) から算出することができる。

$$E_{vdw} = \sum_{m=1}^N G_{vdw, mo}$$

各三次元格子点に配置した各プローブ原子と生体高分子との間に働く静電相互作用エネルギーは、以下の式(式中、 $i$ 、 $j$ 、及び $r_{ij}$ は前記定義のとおりであり、 $q_j$ は生体高分子の $j$ 番目の原子の電荷を示し、 $K$ はエネルギー単位を変換するための定数を示し、 $\epsilon$ は誘電率を示す)を用いて算出することができる。上記の誘電率としては一定の値を用いてもよいが、Warshelらが提案しているような $r_{ij}$ に依存した値(Warshel, A., J. Phys. Chem., 83, 1979, pp.1640-1652)を用いることが好ましい。

$$G_{elc, i} = \sum_j K q_j / \epsilon r_{ij}$$

10

試行化合物をこの三次元格子点上に配置した際の試行化合物 - 生体高分子間の静電相互作用エネルギーは、以下の式(式中、 $m$ 、 $N$ 、及び $m_0$ は上で定義したとおりである)から算出することができる。

$$E_{elc} = \sum_{m=1}^N q_m G_{elc, mo}$$

水素結合性格子点情報とは、水素供与性の原子又は水素受容性の原子のうちいずれの原子がその三次元格子点に配置されれば生体高分子の水素結合性官能基と水素結合を形成し得るか、あるいはどちらの原子がその三次元格子点に配置されてもその原子が生体高分子の水素結合性官能基と水素結合を形成できる(または水素結合を形成できない)ことを示す情報のことである。例えば、水素受容性である三次元格子点に1、水素供与性である三次元格子点に2、両方の性質をもつ三次元格子点に3、水素結合性のない三次元格子点に0の番号を付加することにより、三次元格子点の水素結合性を表すことができる。

20

三次元格子点の水素結合性は以下のようにして求めることができる。ある三次元格子点Pと生体高分子中に存在するある水素供与性原子Dとの間の距離DPが水素結合を形成し得る距離(例えば、2.5~3.1オングストローム)であり、かつ、三次元格子点P、水素H、及び水素供与性原子Dがなす角度DHPが水素結合を形成し得る角度(例えば、30°以上)であれば、この三次元格子点は水素受容性であると判定される。同様に、ある三次元格子点Pと生体高分子中に存在するある水素受容性原子Aとの間の距離PAが水素結合を形成し得る距離であり、かつ、三次元格子点P、孤立電子対L、及び水素受容性原子Aがなす角度ALPが水素結合を形成し得る角度であれば、この三次元格子点は水素供与性であると判定できる。なお、ある三次元格子点が水素受容性でも水素供与性でもない場合には、水素結合性なしとして取り扱う。

30

つぎに、生体高分子中に存在する水素結合性官能基のうちS6で指定した領域内の存在する水素結合性官能基の中から、試行化合物と水素結合を形成すると予想される水素結合性官能基を選択する(S8)。多数の水素結合性官能基が存在する場合には、その重要度に応じて選択することができる。さらに、S7で算出した三次元格子点の格子点情報に基づいて、S8で選択した各水素結合性官能基に対してダミー原子を設定する(S9)。

このステップでは、S7で算出した三次元格子点の水素結合性に基づいて、S8で選択した各水素結合性官能基と水素結合を形成し得る領域(水素結合性領域)を決定した後、適当な数(例えば5~20個)の三次元格子点を含む水素結合性領域を構成する三次元格子点の中心にダミー原子を配置する。ダミー原子は、該水素結合性領域内で他の原子のファンデルワールス半径外に配置される。水素結合性領域は、同じ水素結合性を有しており、かつ互いに隣接する一群の三次元格子点から構成される領域である。1個の水素結合性官能基から2個以上のダミー原子が設定される場合があり、逆に、1個のダミー原子も設定されない場合があることに留意すべきである。設定されたダミー原子の数が多い場合には、リガンド結合領域の内孔の奥深くにあって重要と思われるダミー原子のみを残すようにして、10個以下、好ましくは5~10個に絞り込むことが好ましい。なお、ダミー原子には、その中心に該ダミー原子を配置した三次元格子点と同じ水素結合性を与えておく。

40

50

つぎに、試行化合物の一つを選択し ( S 1 0 )、選択した試行化合物の三次元座標、水素結合性カテゴリ番号、分子力場計算用の情報および配座作成用の情報を S 1 で作成した A D A M 仕様データベースから読み込む ( S 1 1 )。その後、ダミー原子と試行化合物中の水素結合性ヘテロ原子との間に水素結合対応関係を付与する ( S 1 2 )。ダミー原子の数を m 個、試行化合物中に存在する水素結合性ヘテロ原子の数を n 個とすると、i 個の水素結合を形成する対応関係 ( 組合せ ) の数 N ( i ) は、 ${}_m P_i \times {}_n C_i$  ( 式中、P は順列、C は組合せを表す ) 個となる。ダミー原子と、試行化合物中の水素結合性ヘテロ原子の可能な組合せをすべて選択することが好ましい。

その後、 $l_{min} \leq i \leq l_{max}$  の関係を満たすすべての i について、S 1 2 ~ S 2 9 のステップを繰り返す。その結果、ダミー原子と試行化合物中に存在する水素結合性ヘテロ原子との間で、

$$\sum_{i=l_{min}}^{l_{max}} N(i)$$

個の組合せがすべて選択されることとなる。この操作によって、生体高分子と試行化合物とにより形成される水素結合の全組合せが選択されることになり、生体高分子と試行化合物との結合様式を系統的かつ効率的に探索できるようになる。 $l_{min}$ としてはダミー原子と水素結合性ヘテロ原子の数のうち小さい方、 $l_{max}$ としてはダミー原子と水素結合性ヘテロ原子の数のうち大きい方の数を採用することが好ましい。

より具体的に説明すると、S 1 2 で対応関係をつけた、ダミー原子と試行化合物中の水素結合性ヘテロ原子との組合せの一つを選択する ( S 1 3 )。この際、ダミー原子の水素結合性と試行化合物中の水素結合性ヘテロ原子の水素結合性とが一致しない組合せは選択しない。つぎに、S 1 3 で選択した組合せに含まれるダミー原子について、各ダミー原子間の距離を算出する ( S 1 4 )。なお、ダミー原子の数及び試行化合物中の水素結合性ヘテロ原子の数の両方が 1 個の場合は、S 1 4 ~ S 2 1 のステップを行わずに S 2 2 のステップへ飛び、次いで、S 2 3 ~ S 2 5 のステップを行わずに S 2 6 のステップへ飛ぶ。また、ダミー原子の数又は試行化合物中に存在する水素結合性ヘテロ原子の数のどちらか一方が 1 個の場合は、S 1 4 ~ S 2 1 のステップを行わずに S 2 2 のステップへ飛ぶ。

試行化合物を水素結合性部分と非水素結合性部分とに分割し ( S 1 5 )、S 1 5 で分割した試行化合物の水素結合性部分についてねじれ回転が可能な結合とその回転の様式を選択する ( S 1 6 )。その後、S 1 6 で選択した結合を、S 1 で作成した A D A M 仕様三次元データベースに入力されている結合回転様式に従って回転させ、試行化合物の配座を順次生成させる ( S 1 7 )。そして、生成した配座毎に以後の S 1 8 ~ S 2 9 のステップを繰り返す。

これらのステップの詳細は以下のとおりである。S 1 7 で生成した配座について、S 1 3 で選択した組合せに含まれる試行化合物中の各水素結合性ヘテロ原子の原子間距離を算出する ( S 1 8 )。S 1 4 で得られたダミー原子の原子間距離と S 1 8 で得られた対応する試行化合物中の水素結合性ヘテロ原子の原子間距離との差の 2 乗の和である F の値が一定の範囲以上となる水素結合組合せと配座を除去する ( S 1 9 )。すなわち、

$$F = \sum_j^{\frac{n(n-1)}{2}} (r_{li} - r_{di})^2$$

とした場合、 $k_a - F \leq k_a'$  ( 式中、n は水素結合の数を示し、 $r_{di}$  は i 番目のダミー原子間距離を示し、 $r_{li}$  は試行化合物中の水素結合性ヘテロ原子の i 番目の原子間距離を示し、 $k_a$  は F の下限値を示し、 $k_a'$  は F の上限値を示す ) である水素結合組合せ以外のものを除去する。上記  $k_a$  としては、0.6 ~ 0.8、 $k_a'$  としては、1.2 ~ 1.4 が好ましい。このステップにより、生体高分子と試行化合物の水素結合様式及びリガンド分子の配座の可能性を効率的に網羅することができる。

その後、S 1 9 で除去されなかった水素結合組合せと配座について、前記の F の値が最小になるように試行化合物の水素結合性部分の配座を最適化する ( S 2 0 )。このステッ

10

20

30

40

50

ブにより、生体高分子と試行化合物が安定な水素結合を形成するように、試行化合物の配座が修正される。試行化合物の水素結合性部分に存在するねじれ回転可能な結合のねじれ角を変数として、前記の関数  $F$  の値をFletcher-Powell法(Fletcher, R., Powell, M.J.D., Computer J., 6, 1968, p.163)を用いて最小化することにより、試行化合物の水素結合性部分の配座を最適化することができる。

ついで、S 2 0のステップで最適化した試行化合物の水素結合性部分の分子内エネルギーを算出して、該分子内エネルギーの値が一定値以上の配座を除去する(S 2 1)。例えば、AMBER 4.0の分子力場を用いて試行化合物の水素結合性部分の分子内エネルギーを算出した場合、該分子内エネルギーが100 Kcal/mol以上となった配座を除去することが好ましい。その後、上記のS 2 1までのステップで得られた配座を有する試行化合物の水素結合性ヘテロ原子の座標と対応するダミー原子の座標とが一致するように、試行化合物の原子座標を生体高分子の座標系に置き換える(S 2 2)。このステップには、Kabshの最小二乗法(Kabsh, W., Acta Cryst., A32, 1976, p.922; Kabsh, W., Acta Cryst., A34, 1978, p.827)を用いることができる。この操作により、可能な水素結合様式と試行化合物の水素結合性部分の配座とを同時に粗く推定することが可能になる。

上記ステップに続いて、生体高分子と試行化合物の水素結合性部分との分子間相互作用エネルギー(ファンデルワールス相互作用エネルギー及び静電相互作用エネルギーの和)、及び試行化合物の水素結合性部分についての分子内エネルギーを算出する(S 2 3)。生体高分子と試行化合物の水素結合性部分との分子間相互作用エネルギー  $E_{inter}$  は、試行化合物中の各原子  $k$  に最も近い三次元格子点の格子点情報のファンデルワールス相互作用エネルギー  $G_{vdw}(k)$  及び静電相互作用エネルギー  $G_{elec}(k)$  を用いて、以下の式(式中、 $q_k$  は原子  $k$  の電荷を表す)から算出することができる。

$$E_{inter} = \sum_k \{G_{vdw}(k) + G_{elec}(k) \cdot q_k\}$$

また、試行化合物の水素結合性部分の分子内エネルギー  $E_{intra}$  は、既知の方法に従って算出することができる。例えば、 $E_{intra}$  はAMBER4.0等の公開された力場を用いて、以下の式(式中、 $V_n$  はねじれ角を構成する4つの原子の原子タイプの並びに対して与えられている定数を示し、 $n$  はねじれ角回転に対するポテンシャルの対称性を表す定数を示し、 $\phi$  はねじれ角を示し、 $\gamma$  はねじれ角回転に対するポテンシャルの位相(ねじれ角を構成する4つの原子の原子タイプの並びに対して与えられている)を示し、 $A_{ij}$  と  $B_{ij}$  は試行化合物分子中の  $i$  番目と  $j$  番目の原子の原子タイプの対に対して設定された定数を示し、 $R_{ij}$  は  $i$  番目と  $j$  番目の原子間の距離を示し、 $q_i$  は試行化合物中の  $i$  番目の原子の電荷を示し、 $q_j$  は試行化合物中の  $j$  番目の原子の電荷を示し、 $\epsilon$  は誘電率を示す)により算出できる。

はねじれ角を示し、 $\gamma$  はねじれ角回転に対するポテンシャルの位相(ねじれ角を構成する4つの原子の原子タイプの並びに対して与えられている)を示し、 $A_{ij}$  と  $B_{ij}$  は試行化合物分子中の  $i$  番目と  $j$  番目の原子の原子タイプの対に対して設定された定数を示し、 $R_{ij}$  は  $i$  番目と  $j$  番目の原子間の距離を示し、 $q_i$  は試行化合物中の  $i$  番目の原子の電荷を示し、 $q_j$  は試行化合物中の  $j$  番目の原子の電荷を示し、 $\epsilon$  は誘電率を示す)により算出できる。

$$E_{intra} = \sum_{\text{Dihedrals}} \sum_n V_n / 2 \{1 + \cos(n\phi - \gamma)\} + \sum_{i < j} (A_{ij} / R_{ij}^{12} - B_{ij} / R_{ij}^6) + \sum_{i < j} (q_i q_j / \epsilon R_{ij})$$

S 2 3で算出された分子間相互作用エネルギーと分子内エネルギーとの和が一定値以上となった試行化合物の配座を除去する(S 2 4)。例えば、上記のエネルギーの和が分子量100当たり10kcal/mol以上となった配座を除去することが好ましい。その後、S 2 4で除去されなかった配座を有する試行化合物の水素結合性部分の構造を最適化する(S 2 5)。構造最適化計算を行い試行化合物の水素結合性部分のねじれ角、並びに試行化合物の位置及び方向を最適化することにより、試行化合物の水素結合性部分の構造を最適化することができる。

例えば、以下の式：

$$E_{total} = E_{inter} + E_{intra} + Whb \cdot Nhb \cdot Chb$$

(式中、 $E_{inter}$  及び  $E_{intra}$  は前記のとおりであり、 $Whb$  は重み、 $Nhb$  は水素結合の数、及び  $Chb$  は1個の水素結合による安定化エネルギー(例えば、2.5kcal/mol)を示す)から算

10

20

30

40

50

出される全エネルギー  $E_{total}$  が最小となるように、Powell法 (Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., Numerical Recipes in C, Cambridge University Press, Cambridge, 1989)等を用いて試行化合物の水素結合性部分の構造を最適化することができる。

ついで、試行化合物の非水素結合性部分の結合回転が可能な結合を S 1 で作成した A D A M 仕様三次元データベースに入力されている結合回転様式に従って回転させて試行化合物の配座を順次生成させ (S 2 6)、生成した配座毎に以後の S 2 7 ~ S 2 9 のステップを繰り返す。具体的には、生体高分子と試行化合物との分子間相互作用エネルギー、及び試行化合物の分子内エネルギーを算出し (S 2 7)、S 2 7 で算出された分子間相互作用エネルギーと分子内エネルギーとの和が一定値以上となった試行化合物の配座を除去する (S 2 8)。その後、S 2 8 で除去されなかった配座を有する試行化合物分子全体の構造を最適化し (S 2 9)、S 2 9 までのステップで得られた生体高分子 - 試行化合物の複合体構造のエネルギーを算出して該エネルギーが最小である最安定の複合体構造を選択する (S 3 0)。

10

分子間相互作用エネルギー及び分子内エネルギーは S 2 3 と同様の方法で算出できる (ただし、S 2 3 では試行化合物分子の水素結合性部分についてのみ計算を行うが、S 2 7 では試行化合物分子全体について計算を行う必要がある)。また、上記のエネルギーの和の上限値としては、分子量 100 あたり 10 kcal / mol が好ましい。このステップにより、生体高分子と試行化合物の安定な複合体、及び試行化合物の活性配座が得られる。試行化合物分子の構造の最適化は S 2 5 のステップと同様の方法で行うことができる (ただし、S 2 5 では試行化合物分子の水素結合性部分の構造のみ最適化を行うが、S 2 9 では試行化合物分子分子全体について構造の最適化を行う必要がある)。

20

上記のような生体高分子と試行化合物のドッキング操作は、自動ドッキングプログラム A D A M (PCT国際公開 W093/20525; Yamada, M. and Itai, A., Chem. Pharm. Bull., 41, 1993, p.1200; Yamada, M. and Itai, A., Chem. Pharm. Bull., 41, 1993, p.1203; Yamada Mizutani, M., Tomioka, N., and Itai, A., J. Mol. Biol., 243, 1994, pp.310-326)を用いて行うことができる。

S 3 0 で得られた最安定の複合体構造における生体高分子と試行化合物との分子間相互作用エネルギー、試行化合物の分子内エネルギーの値、水素結合数、及び環の数などの基準のすべてあるいはいくつかに基づき、該試行化合物をリガンド候補化合物として選択するか否かを判定する (S 3 1)。分子間相互作用エネルギーとしては、静電相互作用エネルギー、ファンデルワールス相互作用、水素結合エネルギー、及びそれらの和を挙げることができる。例えば、静電相互作用エネルギーとファンデルワールス相互作用との和が分子量 1 0 0 あたり - 2 kcal / mol 以下の試行化合物分子をリガンド候補化合物として選択することが好ましい。その後、すべての試行化合物を選択したか否かを判定し (S 3 2)、すべての試行化合物を選択していない場合には S 1 0 のステップに戻って S 1 0 ~ S 3 1 のステップを繰り返し、選択した場合には S 3 3 のステップに進む。

30

S 3 1 で選択されたりガンド候補化合物の中からより有望なりガンド候補化合物への絞り込みを行う (S 3 3)。この絞り込みに先立って、リガンド候補化合物の情報 (例えば、生体高分子と試行化合物の分子間相互作用エネルギー、試行化合物の分子内エネルギーの値、水素結合数、及び環の数等) のリストを出力して、絞り込みの基準を決定すると便利である。例えば、生体高分子とリガンド候補化合物との間で形成されうる水素結合の数、生体高分子とリガンド候補化合物の分子間相互作用エネルギーの値、及び原子数等を基準として上記の絞り込みを行うことが好ましい。

40

分子間相互作用エネルギーとしては、静電相互作用エネルギー、ファンデルワールス相互作用、及びそれらの和を挙げることができる。例えば、静電相互作用エネルギーとファンデルワールス相互作用の和が分子量 1 0 0 あたり - 8 kcal / mol 以下の試行化合物分子を有望なりガンド候補化合物として絞り込むことが好ましい。水素結合の数は生体高分子の種類により任意の数を設定できるが、例えば 2 以上、好ましくは 3 以上を設定することが望ましい。原子数も任意の数を設定できるが、例えば 2 0 個以上、好ましくは 2 0 ~ 4 0

50

個の範囲を設定することができる。

生体高分子から非常に多くのダミー原子が作成される場合、あるいは試行化合物がかなり大きな配座の自由度を有するか又は多くのヘテロ原子を有する場合には、本発明のより好ましい実施態様に従って、S 9のステップで試行化合物に関する情報を入力した後、試行化合物の部分構造についてS 13～S 25までのステップを行い、その部分構造中で不可能な水素結合様式を与えるダミー原子と試行化合物中の水素結合性ヘテロ原子との組合せ、及びダミー原子と水素結合を形成し得ない試行化合物中の水素結合性ヘテロ原子についての情報を保存する。試行化合物の部分構造はいかなる構造的な制約を受けることなく任意に決めることができるが、3個以上の水素結合性官能基を含む構造部分が好ましい。なお、その構造部分の配座の自由度の有無は無視してもよい。

上記ステップに続いて試行化合物の全体構造についてS 13～S 33のステップを行うが、前記ステップで保存された水素結合性ヘテロ原子を含む組合せ及びダミー原子と水素結合性ヘテロ原子との組合せをS 12のステップで作成される対応関係の組合せから除く。この結果、試行する組合せ及び配座の数を減じることができ、生体高分子-試行化合物の安定複合体の構造探索に要する時間が大幅に短縮される(上記方法をPre-Pruning法:P P法と呼ぶ場合がある)。P P法は、試行化合物の部分構造において不可能な水素結合様式やリガンド配座は試行化合物の全体構造においても不可能であるという仮説に基づいている。このP P法を採用することにより、得られる生体高分子-試行化合物の複合体構造の精度や信頼性に全く影響を与えることなく探索時間を大幅に短縮することが可能である。

#### 実施例

本発明の方法の有効性を検証するために、ジヒドロ葉酸還元酵素を標的生体高分子として、基質や既知阻害剤、あるいはそれらのアナログがリガンド候補化合物として選択されるか否かをテストした。

ジヒドロ葉酸還元酵素(以下、「DHFR」と記す。)の原子座標として、プロテインデータベースエントリー4DFRから入手できる大腸菌のDHFR-メトトレキセートの二元複合体の原子座標から得られる値を用いた。DHFR-葉酸-NADP<sup>+</sup>の三元複合体の結晶構造に基づき、補因子NADP<sup>+</sup>分子の原子座標を前記の二元複合体の原子座標に付加した。この際、非常に多くDHFRに結合して化学的に除去できない2個の水分子以外の全ての水分子は、前記の二元複合体の原子座標から除去した。除去しなかった2個の水分子のうちの一つは、DHFRの21位のトリプロファン及び26位のアスパラギン酸に、他の一つは114位のロイシン及び116位のトレオニンに水素結合している。

試行化合物のデータベースとして、約110,000個の分子のデータを含むThe Available Chemicals Directory(ACD-3D、MDL社)を用い、上記データベースに収納された分子の中から、原子数が5個以下の分子、H、C、Siおよびハロゲンだけで構成されている分子、H、C、N、O、S、Pおよびハロゲン以外の元素を含む分子、6本以上の回転可能な結合を持つ分子、および環を持たない分子を除いた10,017個の分子を解析対象とし、水素原子の付加と、原子タイプ番号、水素結合性カテゴリー番号、原子電荷、及び結合回転の様式(120°の回転角で回転)の付与とを行った。リガンド候補化合物として選択する判定基準は、標的生体高分子(酵素)との相互作用エネルギーの値が分子量100あたり-7 Kcal/mol以下で分子間水素結合数が3本以上であり、かつ原子数が25個以上とした。

本発明の方法によって、最終的に約1400個の試行化合物がリガンド候補化合物として選択され、生体高分子の三次元座標入力からリガンド候補化合物の選択までの所要時間は約80時間であった。それらのリガンド候補化合物には、第3図に示した化合物をはじめ、多数の基質及び既知阻害剤アナログが含まれていた。阻害剤メトトレキセートと基質ジヒドロ葉酸とは、回転できる結合数の制限により今回の探索からは除外されていたが、トリメプリームはリガンド候補化合物として選択されていた。また、本発明の方法により提示されたリガンド候補化合物の結合様式および配座は、基質や阻害剤アナログと酵素との複合体結晶構造から推測される結合様式および配座とよく合致していた。

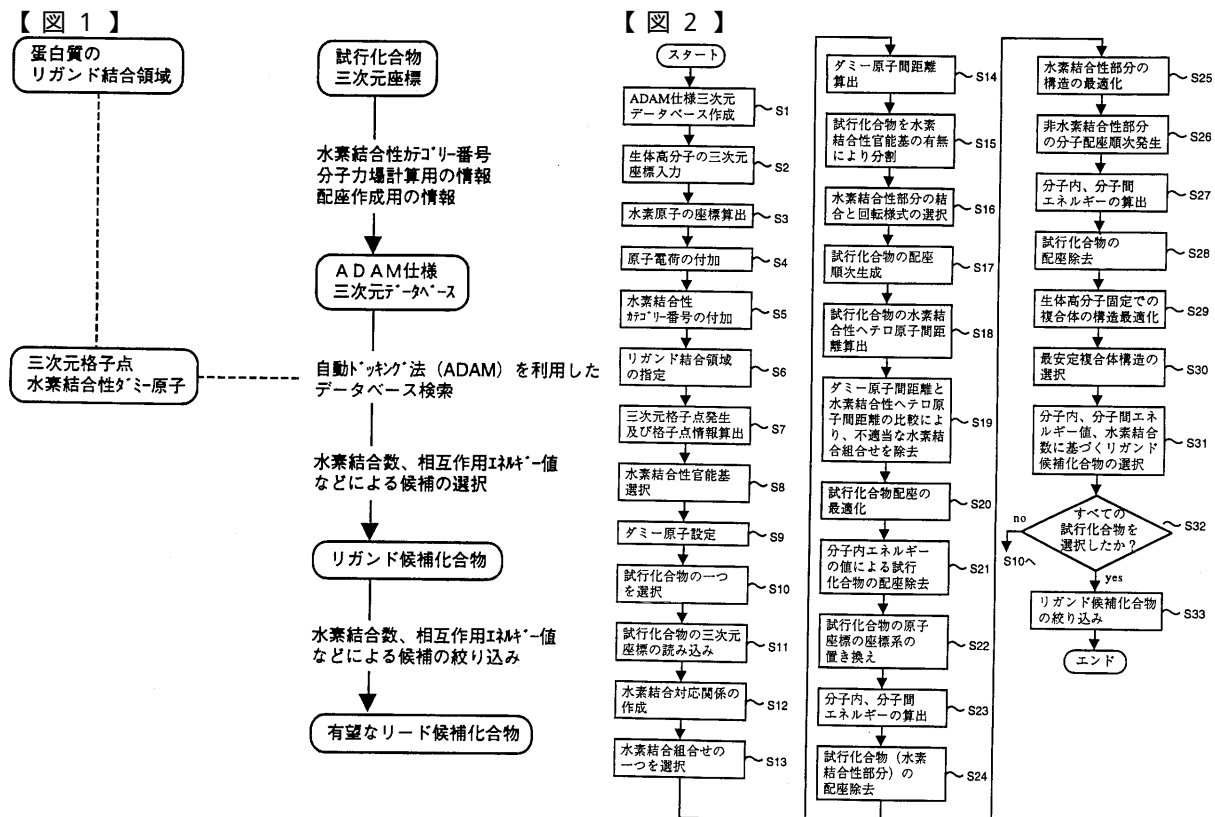


また、本発明の方法により、基質や既知阻害剤の構造と類似性の少ない新規構造をもつリガンド候補化合物も多数選択された。その例をそのドッキングモデルとともに第4図に示す。図中、太線はリガンド分子の骨格、点線は水素結合、細い実線は生体高分子の一部、破線で書いた鳥カゴはリガンドの原子の中心に許される領域を示す。これらのドッキングモデルを見ると、水素結合が何本か形成されており、酵素のキャビティとの形状の相補性もかなり良いことがわかる。従って、これらのリガンド候補化合物は、ジヒドロ葉酸還元酵素の阻害剤のリード化合物として期待される化合物である。

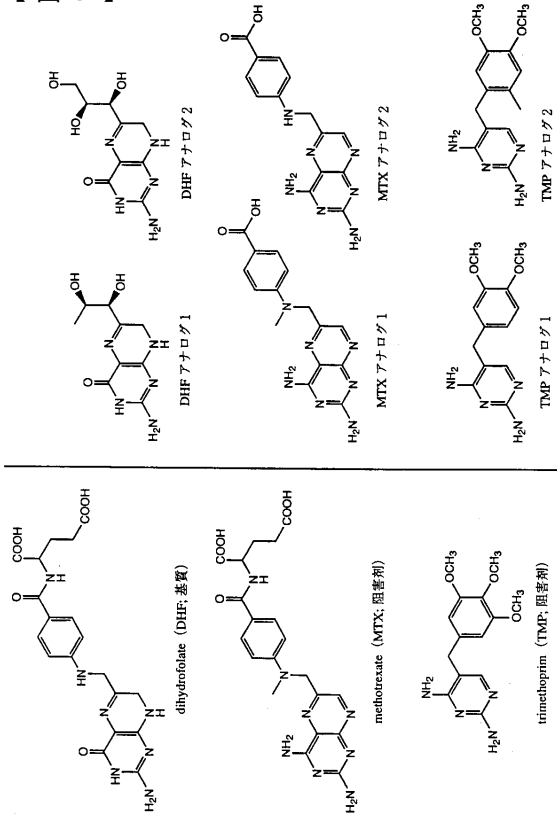
#### 産業上の利用可能性

本発明の方法によれば、三次元構造データベースから構造既知の標的生体高分子に対するリガンド化合物を配座の自由度を考慮しながら短時間で検索することができるうえ、極めて信頼性の高い検索結果を容易に入手できる。従って、本発明の方法は、医学及び薬学などの分野において生体高分子に対して阻害剤やアゴニスト又はアンタゴニストとして作用するリード化合物創製などに有用である。

10



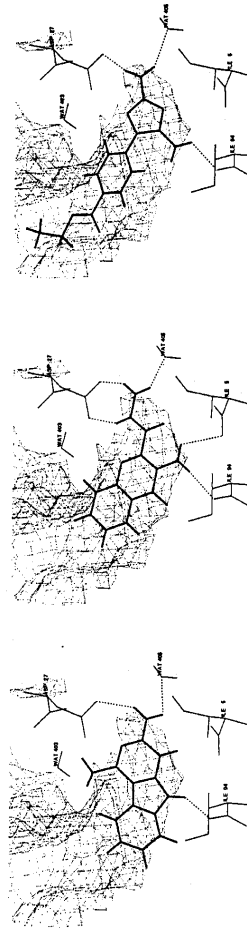
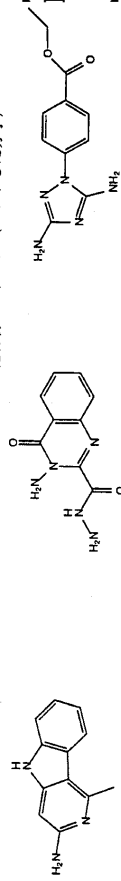
【図 3】



【図 4】

左の化合物のアナログ (ヒットした分子)

基質、既知阻害剤



本発明の方法によりヒットした新規化合物の一例

---

フロントページの続き

## (56)参考文献 国際公開第93/020525(WO, A1)

Simon K. Kearsley et al., Flexibases: A way to enhance the use of molecular docking methods, Journal of Computer-Aided Molecular Design, NL, ESCOM Science Publisher B. V., 1994年10月, vol. 8, no. 5, p565-582

Joseph B. Moon et al., Computer Design of Bioactive Molecules: A Method for Receptor-Based de Novo Ligand Design, PROTEINS: Structure, Function, and Genetics, 米国, WILEY-LISS INC., 1991年, vol. 11, no. 4, p314-328

(58)調査した分野(Int.Cl.<sup>7</sup>, DB名)

G06F 17/50