

【公報種別】特許法第17条の2の規定による補正の掲載
 【部門区分】第6部門第3区分
 【発行日】平成18年1月12日(2006.1.12)

【国際公開番号】W02003/060765
 【年通号数】公開・登録公報2005-019
 【出願番号】特願2003-560791(P2003-560791)
 【国際特許分類】

G 0 6 F 17/30 (2006.01)

【F I】

G 0 6 F 17/30 1 7 0 F

【手続補正書】
 【提出日】平成17年11月2日(2005.11.2)

【手続補正1】
 【補正対象書類名】明細書
 【補正対象項目名】特許請求の範囲
 【補正方法】変更
 【補正の内容】
 【特許請求の範囲】

【請求項1】

与えられた任意のアミノ酸配列を有する単数鎖より成るタンパク質の立体構造を構築する経験的なモデリング法を拡張して(拡張型モデリング法)、与えられた任意のアミノ酸配列を有する複数鎖より成るタンパク質の立体構造を構築する方法であって、コンピュータソフトウェアにおける当該複数鎖の入力ファイル形式を擬似単数鎖としての形態を有するように修正し(配列アライメントの修正)、コンピュータソフトウェアにおけるポテンシャル計算式の計算では複数鎖を有するものとして当該モデリング法によりその立体構造を構築することを特徴とする複数鎖を有するタンパク質の立体構造構築方法。

【請求項2】

当該経験的なモデリング法がホモロジーモデリング法又はスレッディング法である請求の範囲1記載の方法。

【請求項3】

全自動的又は手動的な構築方法である請求の範囲1記載の方法。

【請求項4】

当該修正が、目的タンパク質と同一本数の複数のポリペプチド鎖から成る参照タンパク質のアミノ酸配列を選択し、ポリペプチド鎖のアミノ酸配列のC末端部に区切り文字を付加することを含み、単数鎖のタンパク質として取り扱うことができるようにした(配列アライメントの修正)請求の範囲1~3何れか記載の方法。

【請求項5】

参照タンパク質が立体構造データベースから検索され、参照タンパク質のアミノ酸配列と目的配列との配列アライメントを行う方法を含む請求の範囲4記載の方法。

【請求項6】

参照タンパク質を検索しアライメントを出力するソフトウェアがFAMS、FASTA、PSI-BLAST、LIBRA、RBS-BLAST、IMPALA、ClustalW、HMMER、及びBIOCESの少なくとも1種である請求の範囲5記載の方法。

【請求項7】

当該修正された配列アライメントが、各ポリペプチド鎖のアミノ酸配列のC末端部に区切り文字を有するように配置した書式を有し、区切り文字で区切られたアライメント区間ごとに参照タンパク質IDを指定するファイルを使用することで、同種又は異種の参照タンパク質のアミノ酸配列を記入したマルチプルアライメントに対処可能にしてそれ等の重

ね合わせにより平均構造が得られるようにした請求の範囲 4 又は 5 記載の方法。

【請求項 8】

修正後の当該擬似単数鎖から各タンパク質鎖の末端残基番号を判定し、その境界における化学結合ポテンシャル及び化学結合角ポテンシャルの切断、並びに当該境界における原子間相互作用ポテンシャルの付加を行うことで、立体構造データベース及び/又は当該立体構造データベースを類似構造の重複を回避するように加工したデータベースに基づき C 原子座標及び主鎖原子座標の構築を行い、シミュレーテッドアニーリング法、分子力学計算、及びモンテカルロ法の少なくとも 1 種により疑似エネルギー値を表す目的関数の最小化（最適化）を行うことを含む請求の範囲 1 ~ 7 何れか記載の方法。

【請求項 9】

目的タンパク質を構成する複数鎖のうち少なくとも二つの鎖がポリペプチド鎖であり、タンパク質 - タンパク質相互の認識部位に位置するアミノ酸残基のそれぞれについて 20 種類のアミノ酸の可能な組合せに基いて、ポテンシャルエネルギー値を指標として、適合性の優れた、又は劣った改変体のデータセットを作成して、当該少なくとも二つのポリペプチド鎖の各々のタンパク質の機能を上昇させた、又は低下させた立体構造を構築できるようにした請求の範囲 1 ~ 8 何れか記載の方法。

【請求項 10】

目的タンパク質を構成する複数鎖のうち少なくとも一つの鎖がアミノ酸誘導体又はペプチド誘導体（ペプチド性リガンド）であり、かつ参照タンパク質において該当するリガンド分子と類似した化学構造を持つものである場合に、目的タンパク質の当該誘導体を新たな残基名及び 1 文字コードで定義し、かつ参照タンパク質の当該リガンドを更に新たな残基名及び 1 文字コードで定義したアライメントを手動的又は自動的に作成して、当該リガンドの配列を構成するそれぞれの残基について 20 種類のアミノ酸及びその誘導体の可能な組み合わせに基づいて、ポテンシャルエネルギー値が低い順にランク付けした場合の上位複数を、受容体タンパク質の結合部位への適合性の優れた改変体として格納したアミノ酸誘導体又はペプチド誘導体リガンドモデルデータセットを構築できるようにした請求の範囲 1 ~ 8 何れか記載の方法。

【請求項 11】

目的タンパク質を構成する複数鎖の成分の少なくとも一つがペプチド性リガンドであり、当該リガンドのアミノ酸配列を固定し、当該リガンドを認識する部位に位置するアミノ酸残基のそれぞれについて 20 種類のアミノ酸の可能な組み合わせに基づき、ポテンシャルエネルギー値を指標として、上位複数の受容体タンパク質の結合部位への適合性の優れた改変体のデータセットを作成して、当該リガンドに結合し得る多様なリガンド受容体タンパク質の立体構造を構築できるようにした請求の範囲 1 ~ 9 何れか記載の方法。

【請求項 12】

当該複数鎖が単数鎖ペプチドを領域や部品に切り分けたものであり、擬似単数鎖に復元できるようにした請求の範囲 1 ~ 11 何れか記載の方法。

【請求項 13】

当該目的タンパク質が、通常のアミノ酸及び当該通常のアミノ酸が複数結合したペプチドの何れにも含まれない物質であって、PDB 又はそれを加工したデータベースに登録されている物質を含む請求の範囲 1 ~ 12 何れか記載の方法。

【請求項 14】

立体構造データベースの中から、目的配列に適した参照タンパク質を検索し、検索した複数の参照タンパク質のアミノ酸配列との配列アライメントを行うステップ；目的配列に対して E - v a l u e の小さい参照タンパク質のアミノ酸配列を選択するステップ；及び参照タンパク質に含まれる鎖の各アミノ酸配列の末端部に区切り文字を付加し、目的配列の対応する位置にも区切り文字を付加する（配列アライメントの修正）ステップを含む請求の範囲 1、4 又は 8 記載の方法。

【請求項 15】

更に、アライメントの情報に基いて、目的配列のアミノ酸中の構成原子の一つである C

原子について、当該参照タンパク質のアミノ酸配列を選択するステップで定められた参照構造から座標を取得し、C の原子座標を最適化するステップ；得られたC 座標に主鎖原子座標を立体構造データベースより付加し、主鎖原子座標を最適化するステップ；及び得られた主鎖原子座標に側鎖原子座標を立体構造データベースより付加し、側鎖原子座標を最適化するステップを含む請求の範囲14記載の方法。

【請求項16】

当該ポテンシャル計算式が下記の内容を含む請求の範囲1記載の方法：

全鎖数 = M のときのポテンシャル計算式については、N がタンパク質鎖の番号を、 k_N が N 番目のタンパク質鎖における C 末端残基の通し番号をそれぞれ表し、 $i = 1, \dots, M - 1$ を簡略化して、 $i = 1, M - 1$ と表記するとき、

(A) C 原子座標の構築と最適化プロセスにおける計算では、擬似化学結合ポテンシャルの $i = k_N (N = 1, M - 1)$ の場合を含めず、擬似化学結合角ポテンシャルの $i = k_N (N = 1, M - 1)$ 、 $i = k_N (N = 1, M - 1) + 1$ の場合を含めない。また、原子間相互作用ポテンシャルの場合は、 $i = k_N - 1$ なら $j > i + 1$ 、 $i = k_N$ なら $j > i$ を各々追加する。

(B) 主鎖原子座標の構築と最適化プロセスにおける計算では、化学結合ポテンシャルには $i = k_N (N = 1, M - 1)$ のときの C_i と N_{i+1} 間の結合を含めず、化学結合角ポテンシャルには、C、O をそれぞれカルボニルの炭素原子と酸素原子、C を炭素原子、N を窒素原子とするとき、 $i = k_N (N = 1, M - 1)$ の角度 C_i - C_i - N_{i+1}、O_i - C_i - N_{i+1} 及び C_i - N_{i+1} - C_{i+1} を含めず、更に、化学結合ねじれ角ポテンシャルには、 $i = k_N (N = 1, M - 1)$ のときの角度 N_i - C_i - C_i - N_{i+1}、C_i - C_i - N_{i+1} - C_{i+1}、及び C_i - N_{i+1} - C_{i+1} - C_{i+1} を含めない。また、原子間相互作用ポテンシャルには、原子間距離を r とするとき、 $i = k_N (N = 1, M - 1)$ の、 $r_{ij} \{ r_{Ni, Ni+1}; r_{Ci, Ni+1}; r_{Ci, Ci+1}; r_{Ci, Ni+1}; r_{Ci, Ci+1}; r_{Ci, Ci+1}; r_{Ci, Ci+1}; r_{O_i, Ni+1}; r_{O_i, Ci+1} \}$ について r_{ij} 規定値の場合を追加する。

【請求項17】

請求の範囲1～16何れか記載の方法で構築されたことを特徴とするタンパク質の立体構造モデル。

【請求項18】

請求の範囲1～16何れか記載の方法で構築されたタンパク質の立体構造モデル、請求の範囲10において使用するリガンドモデル及び請求の範囲11において使用するリガンド受容体タンパク質の立体構造から成るデータを固定化し、組み合わせたことを特徴とする拡張型モデリング法に使用可能なデータベース。

【請求項19】

請求の範囲1～16何れか記載の方法で構築されたタンパク質の立体構造モデル、請求の範囲10において使用するリガンドモデル及び請求の範囲11において使用するリガンド受容体タンパク質の立体構造のデータを、コンピュータで閲覧又は検索できるようにデータベース化したことを特徴とするデータベース。

【請求項20】

下記内容をコンピュータで閲覧又は検索できるようにしたことに特徴を有するデータベース構造：

複数鎖より成る目的タンパク質の遺伝子同定コード又はタンパク質同定コード、1行程度の機能説明、目的アミノ酸配列、及び当該目的タンパク質の3次元立体構造の座標；参照タンパク質の遺伝子同定コード又はタンパク質同定コード、1行程度の機能説明、参照アミノ酸配列、及び当該参照タンパク質の3次元立体構造の座標；並びに目的配列と参照配列間のアライメント結果、ホモロジー値、及び E - v a l u e 。

【請求項21】

請求の範囲18～20何れか記載の、データベースの内容を閲覧又は検索でき、又はデ

データベース構造を利用できるようにしたことを特徴とするコンピュータソフトウェアプログラム又は当該プログラムを搭載したコンピュータ。

【請求項 2 2】

請求の範囲 1 ~ 1 6 何れか記載の方法で構築された立体構造データベースの中から、閲覧したいタンパク質を、生物種固有の任意の記号、タンパク質コード名、参照タンパク質名、1 行程度の機能説明の文字列との部分一致等の、論理積検索によって、目的とするタンパク質にたどり着けるように設計されたことを特徴とするインターフェース。

【請求項 2 3】

当該アミノ酸誘導体が、Asp及びGlu等の非天然アミノ酸又はその誘導体である請求の範囲 1 0 記載の方法。

【請求項 2 4】

請求の範囲 1 ~ 1 6 何れか記載の方法を含むことを特徴とするプログラム(new FAMS)、又はそれを載せたことを特徴とするコンピュータ。

【請求項 2 5】

複数鎖より成る目的タンパク質が、ポリペプチド鎖を一つ又は複数含む請求の範囲 1 記載の方法。

【請求項 2 6】

当該複数鎖のうち少なくとも一つの鎖が、天然若しくは非天然アミノ酸及びその誘導体等アミノ酸誘導体、ペプチド誘導体、薬剤成分、核酸類、糖類、有機金属化合物、金属酸化物及びそのイオン、並びに金属及びそのイオンから選択される請求の範囲 2 5 記載の方法。

【請求項 2 7】

目的タンパク質及び選択された参照タンパク質について、それぞれその中に含まれる複数鎖の各アミノ酸配列を、N末端部とC末端部とで順次結合した状態の単数鎖とみなした後、このように得られた擬似単数鎖の参照配列と擬似単数鎖の目的配列との配列アライメントを行い対応関係の確認を行った後、目的配列中のアミノ酸残基中の構成原子の一つであるC原子を配置し、その間をアミド結合でつなぎ、更に側鎖を付加してその他の構成原子についても座標を構築し、最適化を行い、当該モデリング法により当該目的タンパク質の立体構造を構築することを含む請求の範囲 1 記載の方法。

【請求項 2 8】

目的タンパク質の立体構造構築が、得られたアライメントの情報に基いて、目的タンパク質主鎖アミノ酸中のC原子について選択された参照タンパク質の立体構造から座標を取得し、目的関数を最小化するようにCの原子座標を最適化し、最適化されたCの原子座標に主鎖の他の原子座標(Cの原子座標を含む。)を付加して目的関数を最小化するように主鎖の原子座標を最適化し、最適化された主鎖の原子座標に側鎖の他の原子座標を付加し目的関数を最小化するように側鎖の原子座標を最適化することにより行われる請求の範囲 1 又は 2 7 記載の方法。