

平成 15 年度～平成 17 年度

科学研究費補助金(基盤研究 (c) (2))

研究成果報告書

研究課題名

タンパク質の基準振動解析データベースを利用した動的構造の比較研究

課題番号 15570139

研究組織

研究代表者 輪湖 博 (早稲田大学社会科学総合学術院・教授)

研究分担者 猿渡 茂 (北里大学理学部・助教授)

研究経費

平成 15 年度 1, 100 千円

平成 16 年度 1, 000 千円

平成 17 年度 1, 400 千円

研究発表

(1) 学会誌等

H. Wako, M. Kato and S. Endo, Improvements in *ProMode* (a database of normal mode analyses of proteins), *Genome Informatics* **14**:663-664 (2003).

H. Wako, M. Kato and S. Endo, *ProMode*: a database of normal mode analyses on protein molecules with a full-atom model, *Bioinformatics*, **20**:2035-2043 (2004).

(2) 学会報告

猿渡 茂、加藤雅樹、輪湖 博、*ProMode*: タンパク質の基準振動データベースの解析—動的ドメインの同定—、第3回日本蛋白質科学会年会、2003年6月

加藤雅樹、輪湖 博、木下賢吾、木寺詔紀、データベースを利用したタンパク質の分子認識機構の解析、第3回日本蛋白質科学会年会、2003年6月

輪湖 博、加藤雅樹、猿渡 茂、*ProMode* の開発と動的ドメインの解析、日本生物物理学会第41回年会、2003年9月

加藤雅樹、輪湖 博、木下賢吾、木寺詔紀、単体・複合体の立体構造変化の解析による分子認識機構解析、日本生物物理学会第41回年会、2003年9月

高橋卓也、猿渡 茂、輪湖 博、溶媒効果を取り込んだ蛋白質エネルギー極小化アルゴリズムの2面角系への拡張、第26回溶液化学シンポジウム、2003年10月

H. Wako, M. Kato and S. Endo, Improvements in *ProMode* (a database of normal mode analyses of proteins), GIW2003、2003年12月

H. Wako, M. Kato and S. Endo, *ProMode*: a collection and comparison of normal mode analysis results of protein molecules, The 1st Pacific-Rim International Conference on Protein Science, 2004年4月

T. Takahashi, H. Wako and S. Endo, Development of energy-minimization and molecular simulation algorithm considering solvation effects in dihedral angles space, The 1st Pacific-Rim International Conference on Protein Science, 2004年4月

輪湖 博、大塚元央、富澤裕樹、加藤雅樹、猿渡 茂、*ProMode*—知見情報の基準振動アニメーションへの反映、日本生物物理学会第 42 回年会、2004 年 12 月

猿渡 茂、多田美希子、輪湖 博、アミノ酸置換に伴うタンパク質の動的構造の変化—ヒトリゾチーム変異型における基準振動解析—、日本生物物理学会第 42 回年会、2004 年 12 月

H. Wako, M. Otsuka, Y. Tomizawa, M. Kato and S. Endo, Reflection of knowledge information in *ProMode* (a database of normal mode analyses on proteins), GIW2004, 2004 年 12 月

猿渡 茂、輪湖 博、相同タンパク質の動的構造 I. ヒトリゾチームの動的構造に対するアミノ酸置換の影響、第 5 回日本蛋白質科学会年会、2005 年 6 月

輪湖 博、大塚元央、富澤裕樹、加藤雅樹、猿渡 茂、相同タンパク質の動的構造 II. 同一スーパーファミリー内での比較、第 5 回日本蛋白質科学会年会、2005 年 6 月

猿渡 茂、輪湖 博、異種及び同種 2 量体タンパク質の基準振動解析、日本生物物理学会第 43 回年会、2005 年 11 月

輪湖 博、大塚元央、富澤裕樹、加藤雅樹、猿渡 茂、基準振動解析による相同なタンパク質の動的構造の比較、日本生物物理学会第 43 回年会、2005 年 11 月

H. Wako, M. Otsuka, Y. Tomizawa, M. Kato and S. Endo, Comparisons between dynamic properties of homologous protein structure in *ProMode* (Database of normal mode analyses on proteins), GIW2005, 2005 年 12 月

(3) 出版物

なし

研究の概要

(1) 研究の目的

タンパク質立体構造の構築原理を明らかにし、構造と機能の関係を理解することは、構造生物学、バイオインフォマティクスなどにおけるポスト・ゲノムの重要な研究課題の一つである。そのためには静的構造のみならず、動的構造を考慮した研究が不可欠であることが広く認識されているが、実験から空間的にも時間的にも高い精度の動的構造に関する情報を得ることは一般に難しいのが現状である。こうした中で、コンピュータによる理論的アプローチは、それを補完する重要な役割を担うことのできる手法であり、多くの有用な情報を提供してくれる。

コンピュータによるタンパク質の動的構造に関する研究には、大きく分けて2つの方法がある。一つは、PDB (Protein Data Bank) に登録されたデータの解析である。PDB には、同一のタンパク質について、たとえばモノマーのときと複数鎖のとき、あるいはリガンドが結合しているときと結合していないとき、といった複数の構造が含まれている場合があり、それらの構造比較からタンパク質の構造変化を調べ、動的構造に関する情報を引き出すことができる。しかし、そうしたデータがない場合、解析できないのが難点であり、実際、そうした解析ができるタンパク質はまだかなり限られている。また、PDB に温度因子として含まれる各原子のゆらぎも重要な情報であるが、原子間の相対的な動きを知ることはできないため、構造と機能の関係をダイナミクスの観点から論じるには不十分である。

これに対して、分子力学計算に基礎をおいた基準振動解析、モンテカルロ・シミュレーション (MC)、分子動力学シミュレーション (MD) は、静的構造という PDB データを時間軸の上に展開し、動的構造の情報を引き出すことのできる有力な方法である。ソフトの普及、コンピュータの高速化などによって、構造空間のサンプリング効率が最もよい MD 計算が多くの研究室で行われるようになり、多くのタンパク質に関する動的構造に関する情報が得られるようになってきた。しかし、高速化されたとはいえ、たとえば一つの研究室で行おうとすれば、特定の興味あるタンパク質についていくつか計算を実行することが現時点では精一杯であり、1,000 を超えるような多くのタンパク質について実行し、それらの比較から、あるいは統計的な解析手法を適用した研究を行うのは、現時点ではまだまだ難しいと言わざるを得ないであろう。

こうした状況の中で、多くのタンパク質の動的構造の特徴を比較したい、というのが本研究の動機である。そこでわれわれは、MD に比べると立体構造エネルギー近傍のゆらぎしか論じることができないが、計算時間が MD に比べたら短くてすむ基準振動解析によって、より多くのタンパク質についてその動的構造に関する情報を引き出し、それらの比較から、立体構造の構築原理、そして構造・機能相関の問題にアプローチすることを考えた。できるかぎり多くのタンパク質について基準振動解析を行い、その結果を比較する。量が質を凌駕するという言葉があるとおり、多くのタンパク質の動的構造に関するデータを集めることによって、これまで気がつかなかったタンパク質立体構造と機能の関係に関する描像、あるいは解析する視点が浮かび上がってくることを期待して研究を行った。

(2) 研究の特色

基準振動解析は、MC や MD と比べると探索される構造空間がきわめて狭く、これまであまり注目されてこなかった。しかし、MC や MD との比較研究がいくつか報告されるようになると、定性的には非常によく一致することが指摘された。これはおそらく、天然構造近傍でタンパク質自身が取ることのできる構造空間がそもそも限定されており、基準振動解析の近似がその適用範囲を多少超えても有効であるため、と考えられている。現在では、基準振動解析がタンパク質の動的構造をある程度の確に表現できるものと考えられるようになり、ここ 10 年、多くの研究結果が報告されるようになった。本報告書の末尾に、1980 年代以降の基準振動解析に関連する文献のリストを参考までに掲げたが、そうした状況をみてとることができよう。

ところで、基準振動解析がこれまであまり一般的に行われてこなかった理由の一つに、エネルギー極小化に関する問題があった。それは、基準振動解析を行うためには、十分な精度の（得られた極小点でのエネルギー関数の二次微分行列が正値となるような）エネルギー極小化を行う必要がある。固有値の平方根が解析過程で必要となるが、もし二次微分行列が正値でないと負の固有値が現れ、そこで解析が破綻するからである。しかし、そうした十分な精度でエネルギー極小構造を得ることは、タンパク質のような自由度の大きな系では決して容易ではない。実際、デカルト座標を独立変数とし、タンパク質のすべての原子を考慮した計算では、この条件を満たすことはきわめて難しいのが現状である。

この問題を解決するために、申請者らはかつて、京都大学郷信広教授と、二面角を独立変数にして変数の数をデカルト座標の 1/8 に減らし、かつ確実にエネルギー極小構造に達するプログラム FEDER を開発した。その FEDER を利用できることが、本研究遂行の最大の特色である。

実際、タンパク質の基準振動解析のデータベースがここ数年いくつか作られたが、われわれを除くすべてのデータベースはデカルト座標を独立変数としている。しかし、デカルト座標を独立変数としているデータベースで、すべての原子を考慮した系を扱っているものは 1 つだけであり(データベースというよりも、計算サーバで、利用者の要求によって基準振動解析を行なうサイトである)、その他はすべて、各アミノ酸を $C\alpha$ 原子で代表し、それら $C\alpha$ 原子間の相互作用を単純なバネ・ポテンシャルで表現する粗視化したモデル (Elastic Network Model) によるものである。これに対して、われわれの研究はすべての原子を考慮した解析である点も特徴的な点である。

ただ、Elastic Network Model では、PDB の構造をエネルギー最小構造とみなすため、エネルギー極小化を必要としない。そのため、タンパク質が大きくても計算が可能という利点がある。一方、われわれの計算ではエネルギー極小化がもっとも計算時間を必要とする行程であり、そのため大きなタンパク質については、計算時間の制約から実行できないという欠点がある。そうした意味からは、両者は補完的な関係にある。

もう一つの本研究の特色は、ダイナミクス計算がそれぞれの研究者の興味あるタンパク質に限られていて、大量のタンパク質について比較しようという発想そのものがこれまでなかったのに対して、PDB データから動的構造に関する情報を抽出し、比較研究したことである。基準振動解析そのものは決して新しいものではないが、それらの計算結果を蓄積し、タ

ンパク質のダイナミクスの比較を行うという研究はこれまでにはなく(特定のタンパク質に注目した研究はある)、タンパク質立体構造の構築原理解明という研究課題に新たな知見を提供することができると考えている。

(3) 研究の成果

アミノ酸置換は、いわゆる系に摂動を与え、その応答から系の性質を調べるという意味で、タンパク質研究の一つの常套手段であろう。PDBには、そうしたアミノ酸置換した変異型の立体構造が、野生型の立体構造とともにいくつか登録されており、それらの動的構造を比較することが研究の第1歩である。これに関して、BPTI とヒト・リゾチームについて行った結果を報告する。

また、互いに相同なタンパク質は、同じFoldをもちながらも、アミノ酸配列がときには大きく異なり、一部に挿入・欠失があることもあり、アミノ酸置換に比べ、より大きな変化が起こった系の比較材料を提供してくれる。そうした変化にもかかわらず保存されている動的構造と、それぞれのタンパク質に特異的な変化した動的構造などを明らかにしていくことは、非常に興味深い。そこでいくつかの相同タンパク質について比較解析した結果を報告する。

研究成果報告

1. データベース ProMode について

ProMode は、日本科学技術振興事業団の BIRD (Institute for Bioinformatics Research and Development) の支援を受けてわれわれが構築した、タンパク質立体構造の基準振動解析の結果を蓄積したデータベースである (Wako et al., 2004)。本研究課題は、ProMode に蓄積された個々のタンパク質の解析結果を利用して、以下に述べるように、野生型と変異型の動的構造の比較、あるいは相同タンパク質間の動的構造の比較などを行ったものであり、ProMode との関係は密接である。そこで、ProMode の概略をまずは述べることにする。

1.1. 基準振動解析計算

基準振動解析計算の手順は以下の通りである。

(1) Protein Data Bank (PDB) からタンパク質の立体構造データを選ぶ。

大きさは当面 300 残基以下とし、複数鎖からなるものは、それぞれの鎖を個別に扱う。リガンドやヘムなどの分子は無視する。現在、1,500 を越えるタンパク質のデータが ProMode に蓄積されている。

(2) PDB データの正規化と立体構造エネルギー極小化

われわれが開発したプログラム FEDER (コーネル大学 H.A. Scheraga グループが開発した ECEPP に由来する) によって、PDB データの正規化と立体構造エネルギー極小化を行う。正規化とは、FEDER が原子間の化学結合長、結合角を標準値に固定し、二面角を独立変数として立体構造を記述するため、その標準値のまわりで微妙にずれた値をとる PDB データから求めた二面角の値をそのまま使うことができない。そこで、FEDER の幾何学的パラメータを使って、PDB の立体構造を再現できるよう二面角の値を最適化する。この最適化は、立体構造エネルギーの極小化と同時に行い、最終的に、PDB に極めて近い構造をもつ、FEDER のパラメータでの立体構造エネルギー極小構造を求める。

(3) 基準振動解析計算

得られたエネルギー極小構造について、基準振動解析計算を行う。37°C のときの以下の物理量を求める。

時間平均および各モードについて(図 1-1 参照)

- ① 各原子の位置のゆらぎの大きさ。
- ② 各二面角のゆらぎの大きさ。
- ③ 原子の位置のゆらぎの相関。具体的には、2つの原子のゆらぎの単位ベクトルの内積。
(実際には、原子対の数が膨大であるので、C α 原子ペアについてのみ計算する)。

各モードについて

- ④ DynDom によって定義された動的ドメイン(図 1-3 参照)。

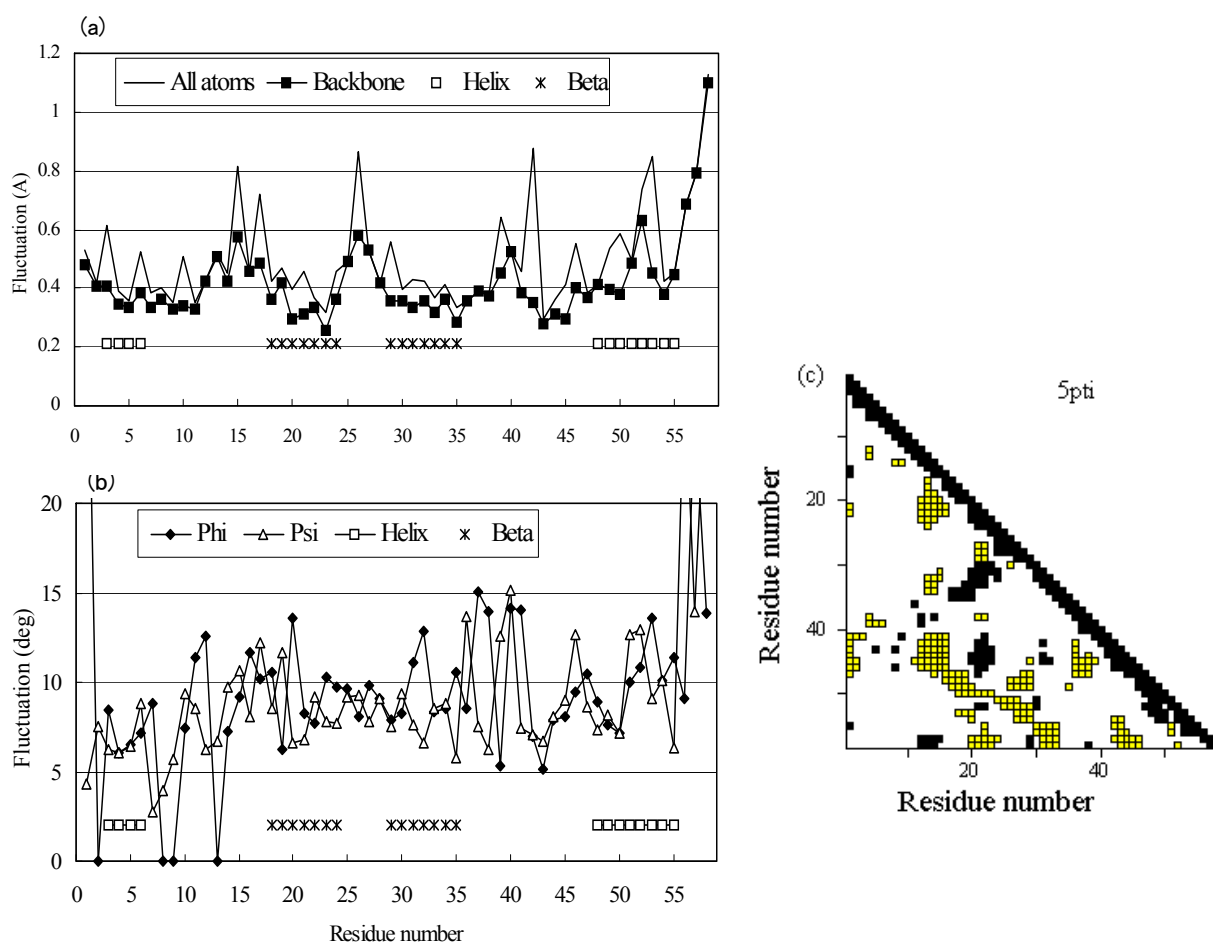


図 1-1 基準振動解析で求められた原子の位置のゆらぎ、二面角のゆらぎ、原子の位置のゆらぎの相関の例。BPTI (Bovine Pancreatic Trypsin Inhibitor; PDB コード 5PTI; 残基数 58) について行った結果を示した。(a) 原子の位置のゆらぎを、アミノ酸残基ごとに平均をとってプロットした。(All atoms)すべての原子の平均、(Backbone)主鎖の原子のみの平均。 α ヘリックスと β ストランドの位置も示した。(b) 主鎖の二面角 ϕ と ψ のゆらぎ。(c) $C\alpha$ 原子のゆらぎの相関。相関係数が0.2以上(黒■)と-0.2以下(黄■)の残基ペアを示した。

1.2 ProMode

基準振動解析の結果をデータベース化して蓄積し、ProModeと名づけ、WEB上に公開した(図 1.2 参照; <http://promode.socs.waseda.ac.jp/>)。ProModeの最大の特徴は、無料のソフトChimeやPDBjViewerを使って、タンパク質の立体構造がそれぞれの基準振動モードでゆらぐ様子をアニメーションで見られるようにしたことである。利用者は、特定の残基を指定して、分子の表示方法や色などを自由に操作することができる。また、基準振動解析から得られるさまざまな物理量も、グラフ等で表示している。

Wako, H., M. Kato and S. Endo, ProMode: a database of normal mode analyses on protein molecules with a full-atom model, *Bioinformatics*, **20**:2035-2043 (2004).

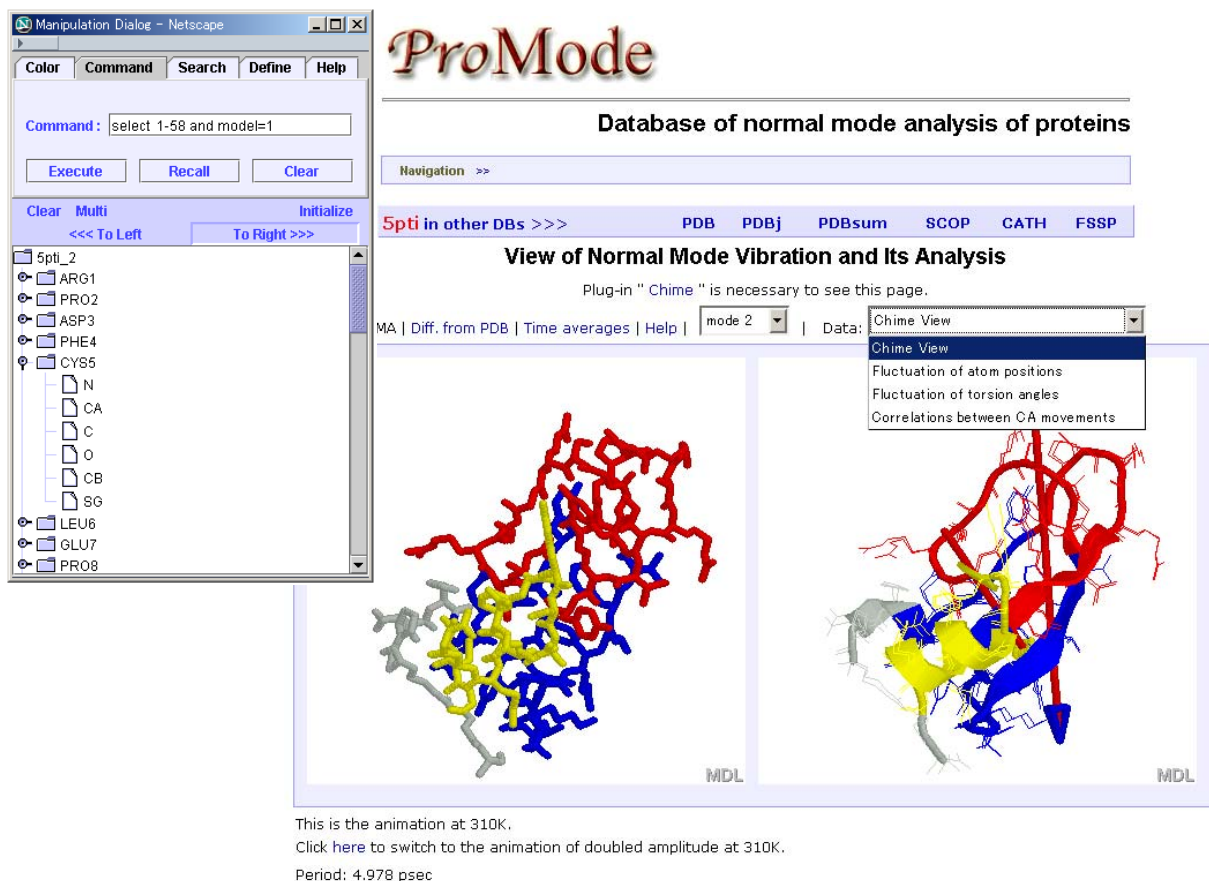


図 1-2 ProMode のスナップショット。 基準振動解析の結果は、例えば2つの Chime ウィンドウに表示される。左は指定されたモードで振動する様子をアニメーションで表示したものであり、右はその静的構造を示す。2つの図は利用者の操作（並進、回転、拡大・縮小等）に対して同期して反応する。表示されている色は、DynDom によって定義された動的ドメインを示す。また、矢印は、それらドメイン間の相対運動をらせん運動とみなしたときのらせん軸を示している。また、左にあるのは、Chime のポップアップ・メニューに含まれないコマンドを実行したり、特定の残基を指定し、その残基の表示法（線画、CPK モデル、リボンモデルなど）や色を変更したりすることができるようにわれわれが開発した Chime の操作用ダイアログである。

2. BPTI の基準振動解析—野生型と変異型の比較

多くの基準振動解析データが蓄積してくると、それらを比較することによって、タンパク質を個別に研究していたときにはわからなかった多くの情報を得ることができるであろう。そこでまずは BPTI(Bovine Pancreatic Trypsin Inhibitor)の立体構造について調べた結果を報告する。BPTI は α ヘリックスと β シートをとともを含み、残基数も少ないことから、タンパク質の理論的研究を行うとき、最も単純な系として古くから利用されてきたものである。幸い、野生型といくつかの変異型の立体構造が決定されている。ここでは、以下のデータを使った解析を報告する。

野生型	PDB コード	5pti			
変異型	PDB コード	1bpt	1bti	1fan	1nag
	アミノ酸置換部位	Tyr23Ala	Phe22Ala	Phe45Ala	Asn43Gly
	PDB コード	7pti	8pti		
	アミノ酸置換部位	Cys30Ala, Cys51Ala	Tyr35Gly		

図 1-1 に野生型 BPTI の原子の位置のゆらぎと二面角のゆらぎを示した。一般的に、二次構造 (α ヘリックス: 3-6、48-55; β ストランド: 18-24、29-35) の原子のゆらぎは比較的小さく、ループ領域(15-17、25-27、39-42)のゆらぎは大きい。ただし、 α ヘリックス内の残基 50-53 のゆらぎは例外的に大きい。アニメーションを観察すると、ループ 25-27 の大きなゆらぎが、S-S 結合 30-51 を通して、C 端の α ヘリックス、特に 50-53 のゆらぎに影響を与えているように見える。

一方、二面角のゆらぎは、ループ領域だけでなく、二次構造内の一部の残基についても大きい。二面角は、隣り合う ϕ_i と ϕ_{i+1} が逆方向に相補的に動くことによって、その両側にあるペプチド鎖の相対的な位置関係 (すなわち立体構造) をあまり変化させないようにすることがある程度可能であるため、大きい二面角のゆらぎが即大きな構造のゆらぎを意味しない。しかし、二面角の大きなゆらぎは、局所的な構造のゆらぎを意味していることは確かである。また、ループ 25-27 では、二面角のゆらぎは決して大きくはなく、 β 構造を形成している 2 つの β ストランドの二面角のゆらぎが、このループ全体のゆらぎを引き起こしているようである。一般に、ループ領域のゆらぎは大きい、それが局所構造を維持した剛体的なゆらぎと、ループ構造そのものが変化するゆらぎとがあることに注意する必要がある。

図 2-1、2-2 には DynDom によって定義された動的ドメインを野生型、6 つの変異型それぞれについて示した。1~4 番目の低振動モードに対する結果のみを示しているが、いくつかのモードではドメインが定義されなかった。定義されたものについてみると、おおよそ 2 つのドメインが定義されており、しかもモード間そしてタンパク質間でかなり共通していることがわかる (しかし、同じタンパク質の異なるモードについてアニメーションで観察すると、同じように定義されたドメインでも、それらの相対運動はモードごとに異なる)。

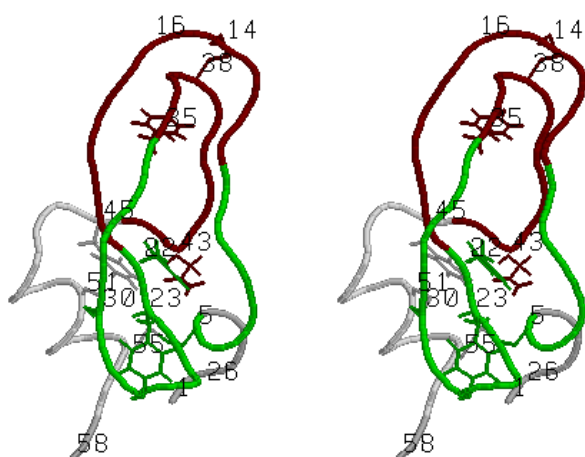


図 2-1 BPTI (5PTI) の立体構造 (立体視)。 2 番目の最低振動モードについて DynDom によって定義された 3 つの動的ドメインが色分けして示されている。また、置換されたアミノ酸残基と S-S 結合している Cys 残基の残基番号と側鎖が示されている。

1つのドメインは、ヘアピンループを含むβシートの半分(21-32あたり)とS-S結合5-55によって架橋されたN端およびC端領域からなる(図2-1の下半分)。モードによっては、C端領域が違うドメインとして定義されているものもある。もう一つのドメインは、βシートの残り半分とS-S結合14-38で架橋された2つの長いループからなる(図2-1の上半分)。どちらのドメインも一つながりではなく、複数の領域が空間的に接近して構成されている。

図2-3は、変異型BPTIにおける原子のゆらぎおよび二面角のゆらぎの野生型との差をプロットしたものである。ここで調べた変異型のアミノ酸置換は、すべて大きな側鎖から小さな側鎖への置換であり、したがって、置換部位周辺でのゆらぎが大きくなるだろうことが予想される。しかし、ゆらぎの変化は、必ずしもアミノ酸置換が起こった場所の周辺とは限らない。また、二面角のゆらぎの変化が大きい残基が、必ずしも原子のゆらぎの大きい残基ではない。偶然にも、ここで調べた変異型BPTIの置換部位22、23、35、43、45は、動的ドメインをつなぐ蝶番部位近くにあり、しかも多くが疎水コアを形成している。その結果、アミノ酸置換が、置換部位付近に限定されず、動的ドメイン、たとえば9-17、25-28、40-44といった領域の運動に影響を与えているようにみえる。

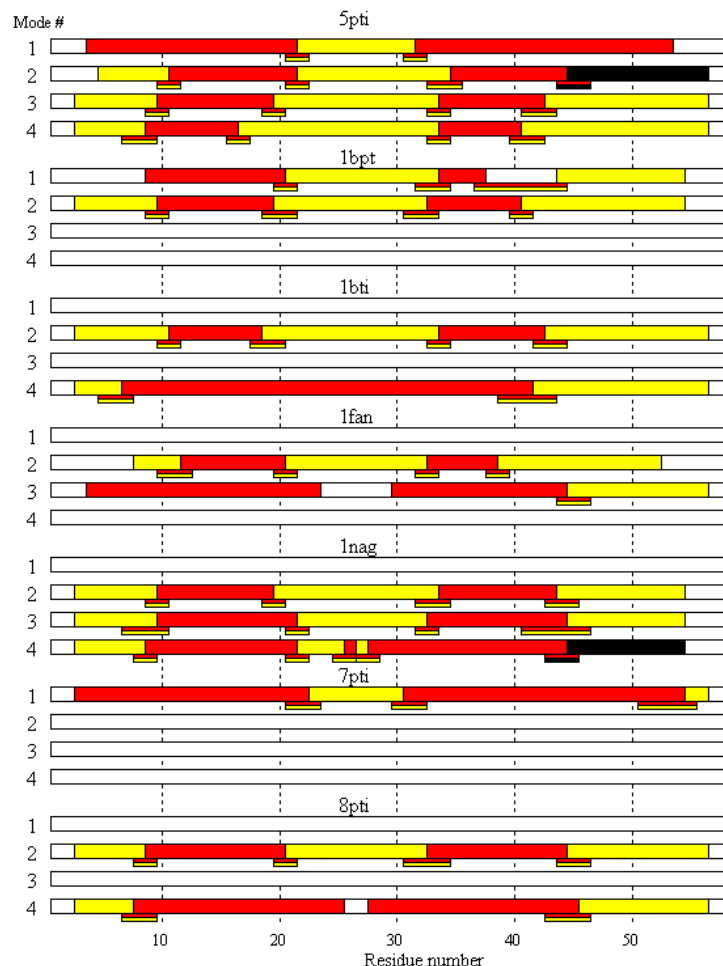


図 2-2 野生型および6つの変異型 BPTI の動的ドメイン。1～4番目の最低振動モードについて、DynDomによって定義されたドメインを示した。同じ色の残基は、同じドメインに属している。また、白い領域は、ドメインとして定義されなかった残基である。なお、色の境界の下にある短い横棒は、蝶番領域の残基を示している。

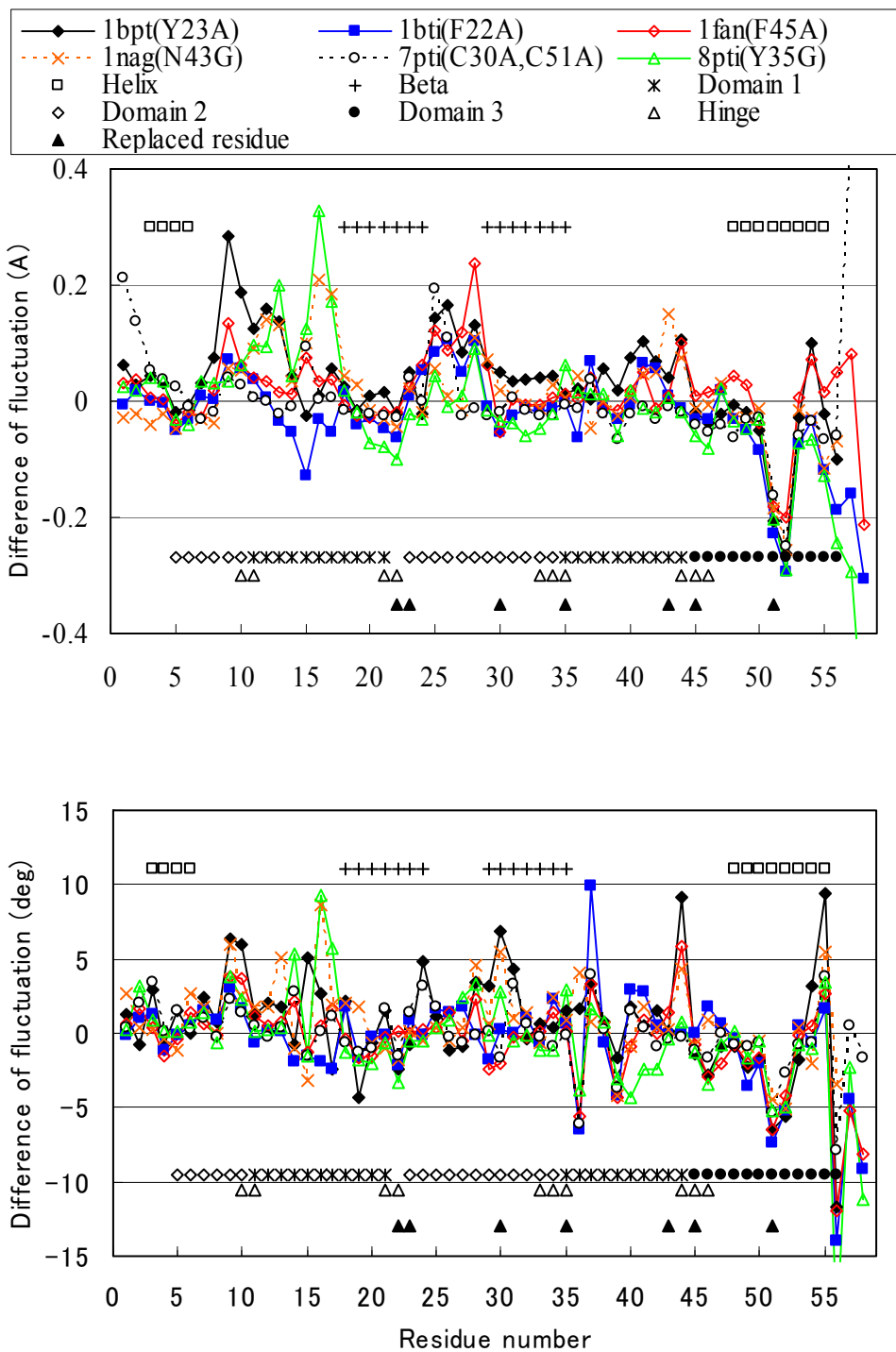


図 2-3 変異型 BPTI の原子の位置のゆらぎ（上図）および二面角のゆらぎ（下図）の野生型との差。6つの変異型について、置換された残基を凡例の PDB コードの後ろの括弧の中および図中の▲で示した。また、二次構造、2 番目の最低振動モードに対する動的ドメインと蝶番残基も図に示してある（凡例参照）。

図 2-3 をみると、その他に、51-52 での原子のゆらぎ、そして二面角のゆらぎがともに変異型で大きく抑えられているのが特徴的である。言い換えれば、野生型だけが大きくゆらいでいる。図 1-1(c)によれば、26-43 と 52-58 の 2つの領域は互いに負の相関をもって動いている。しかし、そうした相関は、図 2-4 でみられるように、変異型では弱くなっている。一般に、大きな正の相関あるいは負の相関を持っていた残基ペアは、変異型では野生型に比べ少なくなっている。互いに正の相関をもつ残基ペアから成り立つドメインとそれらドメイン間の負の相関という単純な図式が、野生型ほどすっきりしており、変異型ではそれが崩れているような印象をもつが、それを確証するには、もっと多くのタンパク質について調べることが必要であろう。

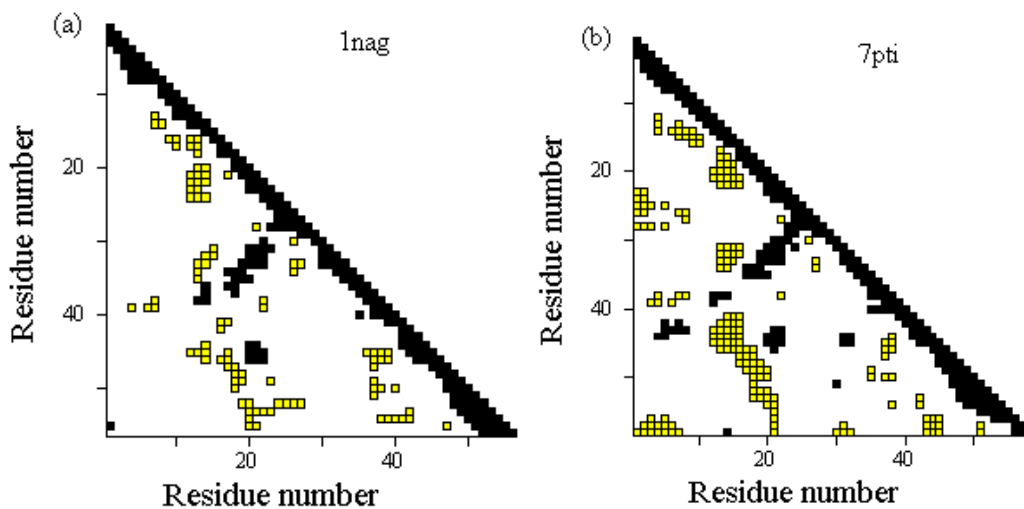


図 2-4 変異型 1nag と 7pti の C α 原子のゆらぎの相関図。図 1-1(c)と比較せよ。

3. ヒト・リゾチームの基準振動解析—野生型と変異型の比較

ヒト・リゾチームは、100 を超える変異型（1～3 個のアミノ酸置換体）の立体構造が解明されており、上で述べたような研究をより統計的に進めるのに適している。

図 3-1 に、ヒト・リゾチームの立体構造を示した。残基数 130 で、主として α ヘリックスからなる α ドメイン(1-40, 87-130)と β 構造からなる β ドメイン(41-86)の 2つのドメインで構成され、両ドメインにはさまれるように活性部位が形成されている。

二次構造の位置は以下の通りである（図 3-1 のマゼンタ色、黄色の部分）。

α -helix: 5-14, 20-22, 25-36, 81-85, 90-101, 105-108, 110-115, 122-124

β -sheet: 43-46, 51-54, and 59-60

4 本の S-S 結がある。6-128 と 30-116 は α ドメイン内に、65-81 は β ドメイン内にあり、77-95 が両ドメイン間を架橋している。

図 3-2 に、例として、4 種類の変異型リゾチーム(Y20F, T52A, V74G, V110G)について、原子位置のゆらぎの残基ごとの 2 乗平均値がアミノ酸置換に伴いどのように変化するかを、野生型の原子ゆらぎとの差をとって示した。どの変異型についても、アミノ酸置換の影響は置換した残基だけでなく分子全体に及んでいることがわかる。また、4 種類の変異型の置換

この点を調べるために、アミノ酸置換に伴う各残基のゆらぎの変化を野生型におけるゆらぎの大きさに対して、図 3-3 にプロットしてみた。野生型リゾチームにおいてあまりゆらがない残基では、残基置換によってゆらぎに変化が起こることは少ない。それは、たとえ置換部位であっても同じである。そして、野生型で大きくゆらいでいる部位で、残基置換に伴うアミノ酸残基のゆらぎが大きく変化できる、あるいは変化する傾向が見て取れる。

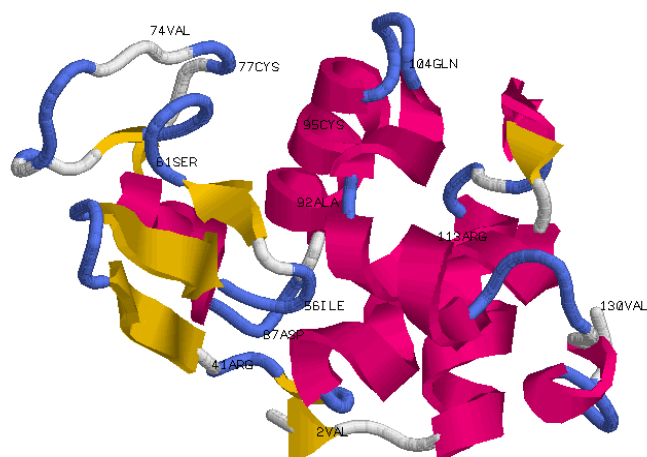


図 3-1 ヒト・リゾチームの立体構造

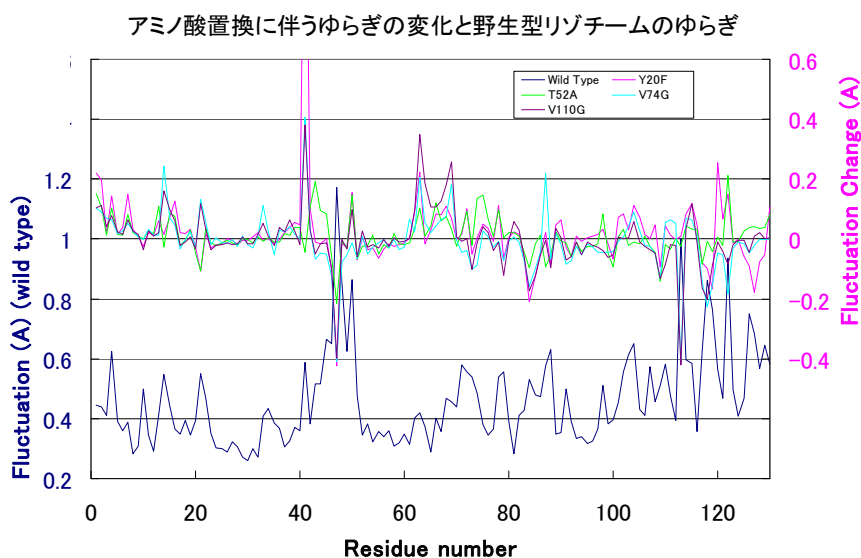


図 3-2 アミノ酸置換に伴う原子のゆらぎの変化(上、右目盛)と野生型リゾチームにおける原子のゆらぎ(下、左目盛)。凡例に示す、4つの変異型リゾチームについて示した。

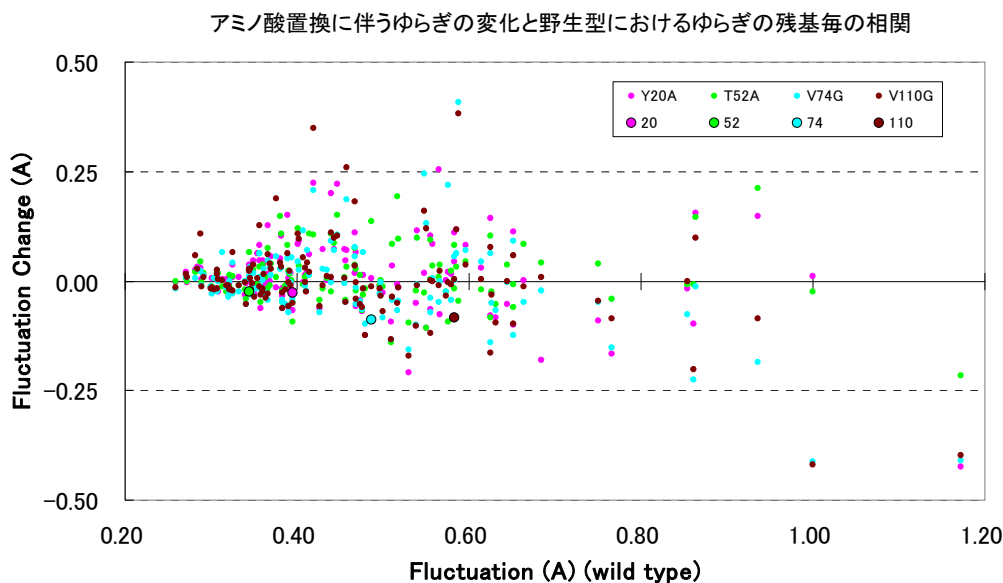


図 3-3 アミノ酸置換に伴う原子ゆらぎの変化と野生型における原子ゆらぎの残基ごとの相関

1つの点が4種類の変異型リゾチームにおける各アミノ酸残基に対応する。変異型ごとに色分けして表示している。黒で縁取りした円は置換されたアミノ酸残基のデータを示している。

図 3-4(a)は、アミノ酸置換による静的な立体構造の全体的変化と置換部位における変化との相関をみたものである。プロットした記号は、置換部位の二次構造を表している。やはり α リックスや β 構造に含まれる残基の変位は一般に小さく、大きく置換部位が変位しているのはループ領域である。中には主鎖が2 Å程度構造変化する場合もある。全体的変化と局所的变化には、局所的变化が大きくないとき、ゆるやかな相関があるようであるが、はっきりとしたものではない。

また、図 3-4(b)では、静的な立体構造でのアミノ酸置換部位の変化とそのゆらぎの変化との相関を調べてみた。置換部位のゆらぎの変化が大きい置換体で必ずしも静的構造が大きく変位しているわけではなく、ほとんど相関が見られない。

さらに図 3-5 には、アミノ酸置換に伴う動的構造の全体的変化と局所的变化の相関を示した。Val2、Val74 (ループ領域) や Val110 (α ヘリックスの先頭) で Arg、Met、Tyr 等の大きなアミノ酸に置換した場合に特徴的に置換部位のゆらぎが増大している。Val2、Val74、Val110 ではそれぞれ 10 種以上の置換体があるが、これら以外のアミノ酸への置換ではゆらぎが大きく変化することはない。よって、一般に、二次構造によらず、置換部位のゆらぎの変化は分子全体の平均的な変化に比べて大きくはなく、むしろあまり相関はない、といえよう。

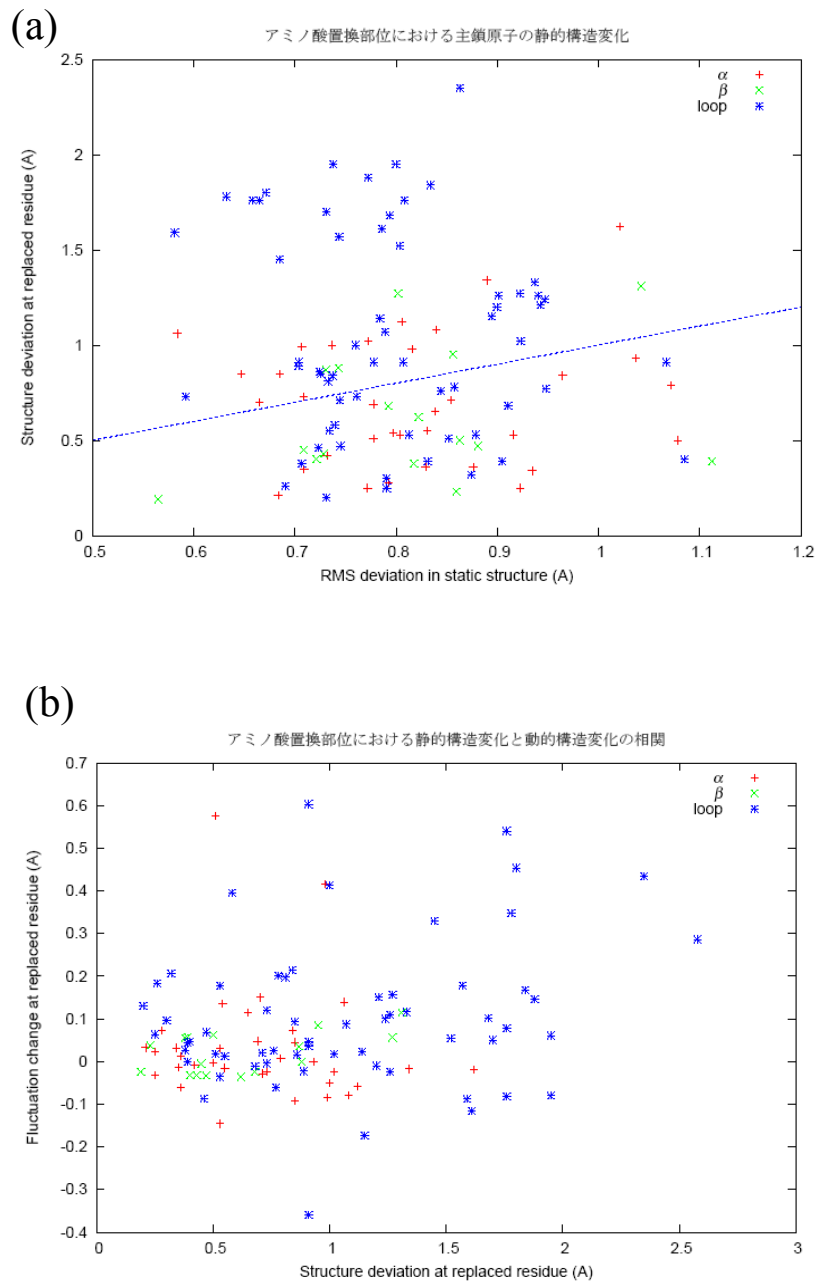


図3-4 静的立体構造の全体的変化と局所的变化の相関(a)および静的立体構造の局所的变化とゆらぎの変化との相関(b)。(a) 野生型とアミノ酸1残基置換体の立体構造を主鎖原子に関して重ね合わせ、原子位置の2乗変位を求めた。分子全体に対して平均した値を全体的変化、置換部位における変位を局所的变化として、その相関を示した。(b) 置換部位における変位(局所的变化)に対する置換部位のゆらぎの変化をプロットした。各アミノ酸残基は、それが属する二次構造によって異なった記号を用いてプロットしてある。

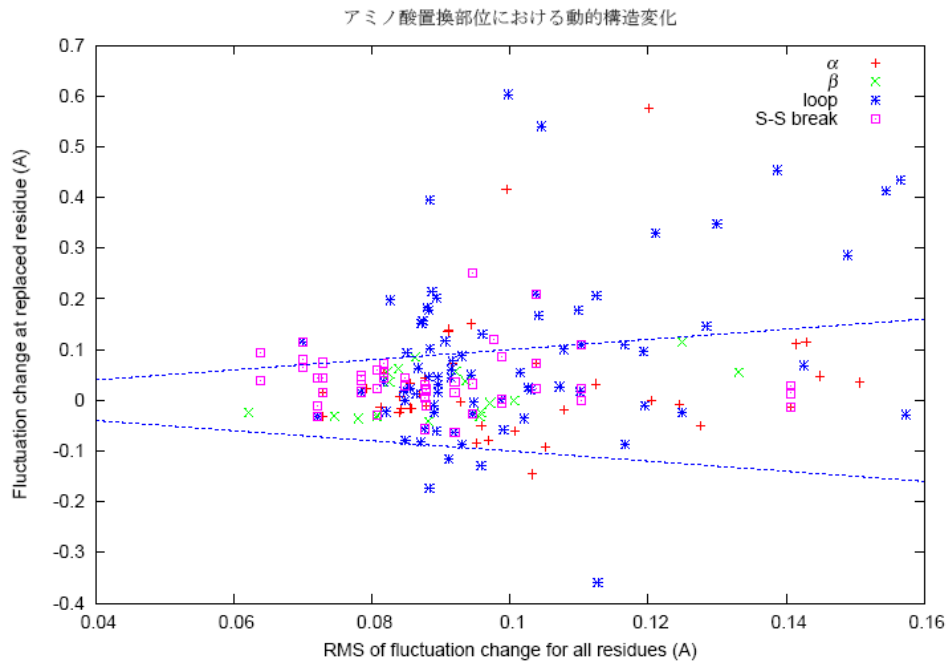


図 3-5 アミノ酸置換に伴うゆらぎの全体的変化と局所的变化の相関

各置換体について、野生型とのゆらぎの差を分子全体に対して平均した値(横軸)と置換部位におけるゆらぎの変化(縦軸)とをプロットした。1つの点が1つの置換体リゾチームに対応する。各置換対は、置換部位の二次構造によって記号を変えて表示している。

こうしてアミノ酸置換によるゆらぎの変化をみてくると、それが置換部位に顕著に起こるというわけではないことがわかる。むしろ、全体的な変化を引き起こしていると言えよう。そこで、全体的な変化がどんなものであるか明らかにするために、主成分分析を応用してみた。

図 3-6 は、各置換体リゾチームにおける 130 残基のゆらぎの変化量を 130 次元のベクトルデータとみなし、すべての置換体を標本として主成分分析を行った結果を示したものである。側鎖まで考慮した場合、ゆらぎの変化の主成分は表面のアミノ酸残基 (Arg41, Asp87, Arg113, Arg119, Arg122) の大きな変化に引きずられて孤立した鋭いピークとなっており、分子全体の変化が明確になっていない (a)。これに対して、主鎖原子のみを考慮したゆらぎの変化の主成分 (b) ではこうしたピークがなくなり、主鎖構造のゆらぎが大きいいくつかの部位がくっきりと現れている。具体的には、次のような部位が大きくゆらいでいる。

- A: β -rich ドメイン中の loop 領域 (62~76)
- B: α -helix をつなぐ loop 領域 (100~104)
- C: α -helix の後半+loop 領域 (113~116)

D: α -helix をつなぐ loop 領域 (117~121)

E: C 末端の loop 領域 (125~130)

そして、3つの主成分は、これら部位のゆらぎの大小の組み合わせで表現できる。

	A	B	C	D	E
<i>1st principal comp.:</i>	+	+	+	0	0
<i>2nd principal comp.:</i>	0	0	-	0	++
<i>3rd principal comp.:</i>	-	-	++	++	0

第1主成分は、 α ドメインと β ドメインの相対的な動きである。第2主成分は、C末端の大きなゆらぎである。第3主成分では、 β ドメインとそれに向かい合うループ(100~104)の動きが抑えられ、C端近くのループが大きくゆらぐ動きである。こうした解析は、アミノ酸置換によるゆらぎの変化が、変位部位周辺に起こるというよりも、第1~第3主成分で表されるような動きの組み合わせで基本的に表現できることを示唆している。

ところで、これらは1残基置換や3残基置換など、すべての置換体のデータに対する主成分分析である。3残基置換体は、77-95のS-S結合を切断する置換(Cys→Ala)に加え、他のもう一つの部位でアミノ酸置換したものである。S-S結合切断は、切断しない置換体とは大きく異なることが予想されたため、それらを除き、1残基置換体のデータだけを使って同じようにゆらぎの変化の主成分分析を行ってみた。しかし、大きな差異は見られなかった(図なし)。

図3-7には、アミノ酸置換体のゆらぎの変化の主成分値の分布を示した。各置換体の第1および第2主成分値をプロットした。野生型のパターンは、側鎖原子まで含めたゆらぎでは主成分値の分布の端に位置するような特異的なものであったが(a)、主鎖原子のゆらぎに関しては全標本の平均に近かった(b)。野生型と同じようなゆらぎを持っている mutant group A (X) においても、主鎖のゆらぎは分布の中央付近にあるのが見てとれる。しかし全原子のゆらぎに関して野生型以上に特異的な mutant group B (*) においても同様な傾向が見られるので、野生型に特徴的な性質ではないかもしれない。

図3-8には、原子ペアのゆらぎの相関の野生型と変異型の差を調べた結果を示した。1残基置換体と野生型の相関の差と変位部位のゆらぎとの相関を各残基に対して1次元にプロットした。相関の三角図(BPTIの図1-1(C)に対応する図)の非対角要素を、対角線の自己相関の値のルートで割って規格化したもの(いわゆる相関係数)の差(上図)と、規格化しない絶対値のままの値の差(下図)についてプロットした。変位部位を図の下の方にグラフと同じ色の点で示した。変異部位との相関が比較的大きいものを選んで示しており、これ以外の変異型では、相関はあまり大きくなかった。

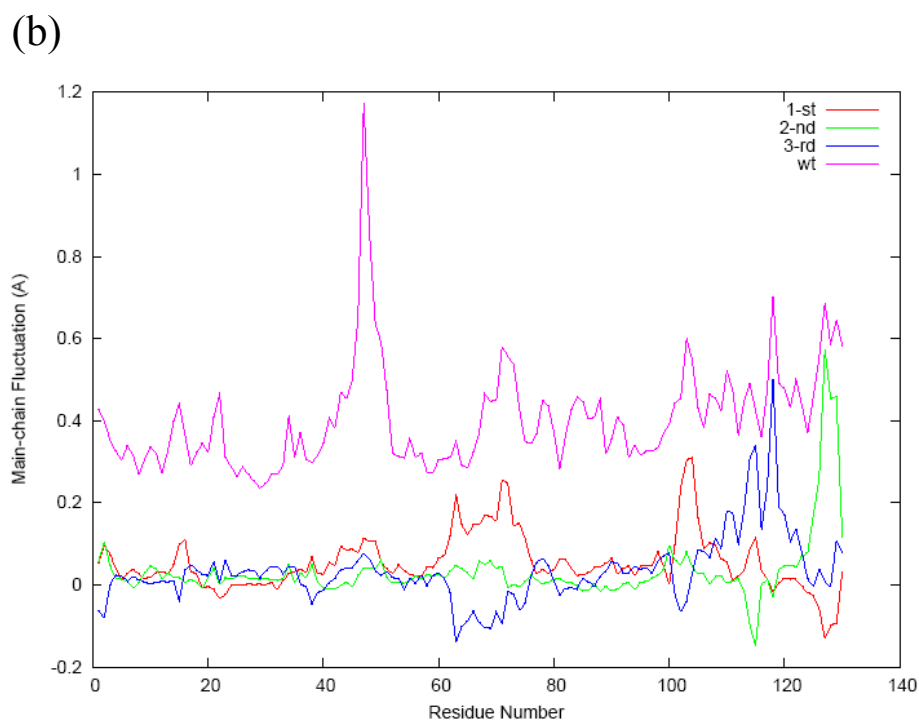
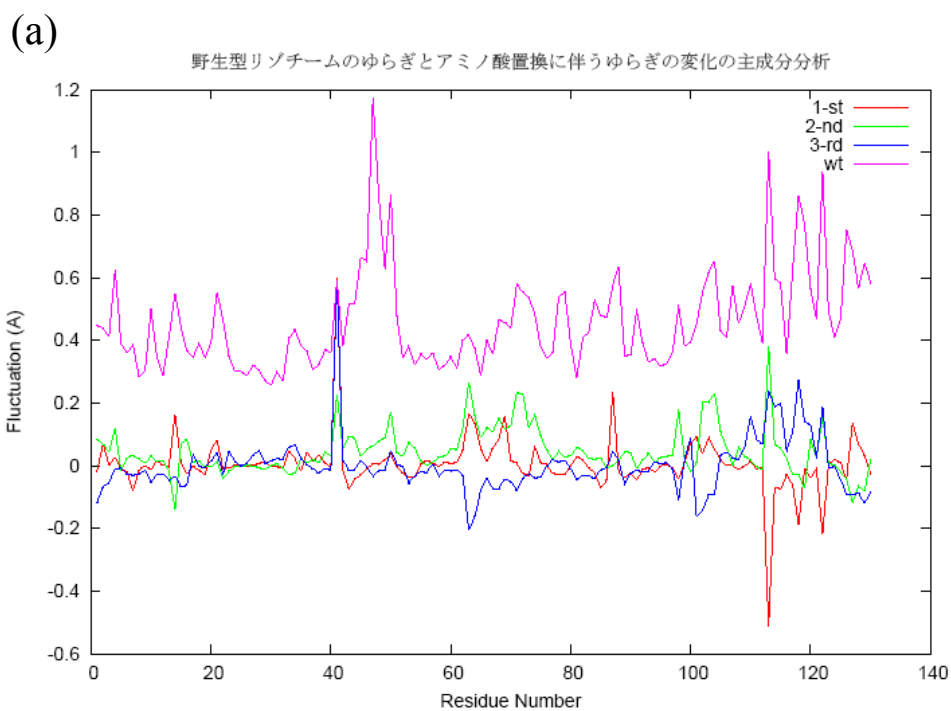


図 3-6 アミノ酸置換に伴うゆらぎの変化の主成分分析

分子全体 130 残基に対するゆらぎの変化量を 130 次元のベクトルデータとみなし、すべての変異体を標本として主成分分析を行った。第 1、第 2、第 3 主成分と、全主成分の総和を示している。(a) 全原子のゆらぎ、(b) 主鎖原子のみのゆらぎを考慮した分析結果。

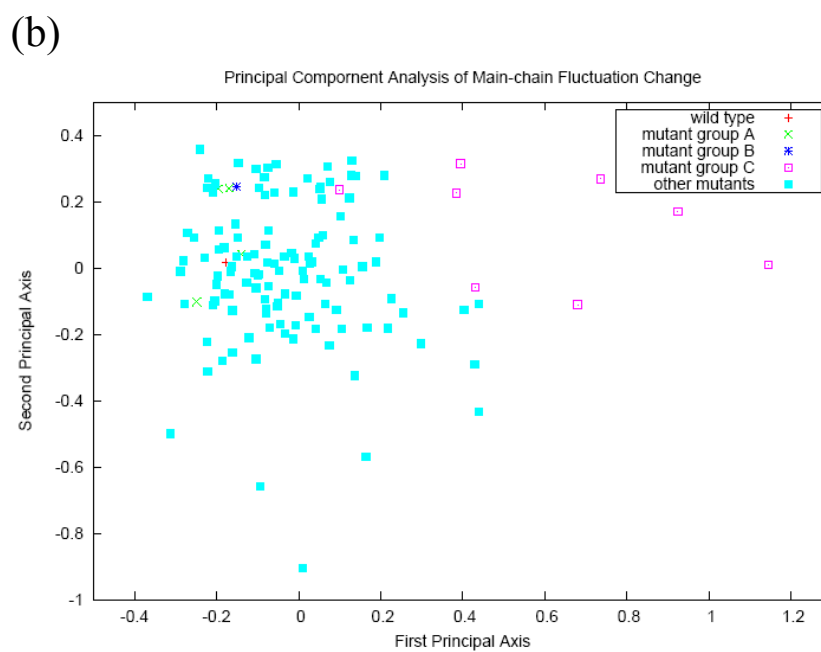
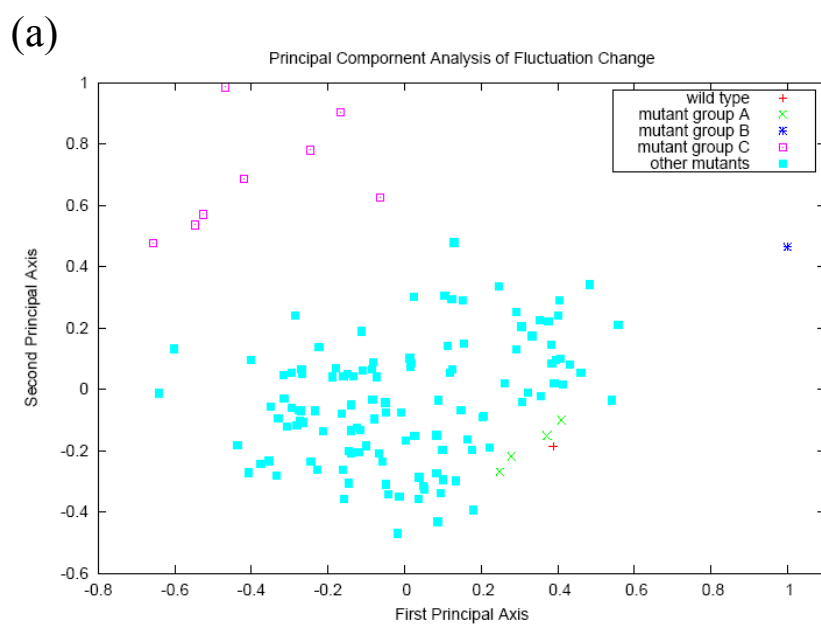


図3-7 アミノ酸置換体のゆらぎの変化の主成分値の分布

各置換体の第1および第2主成分値をプロットした。(a)側鎖原子まで含めたゆらぎの主成分値の分布、(b)主鎖原子のゆらぎの主成分値の分布。

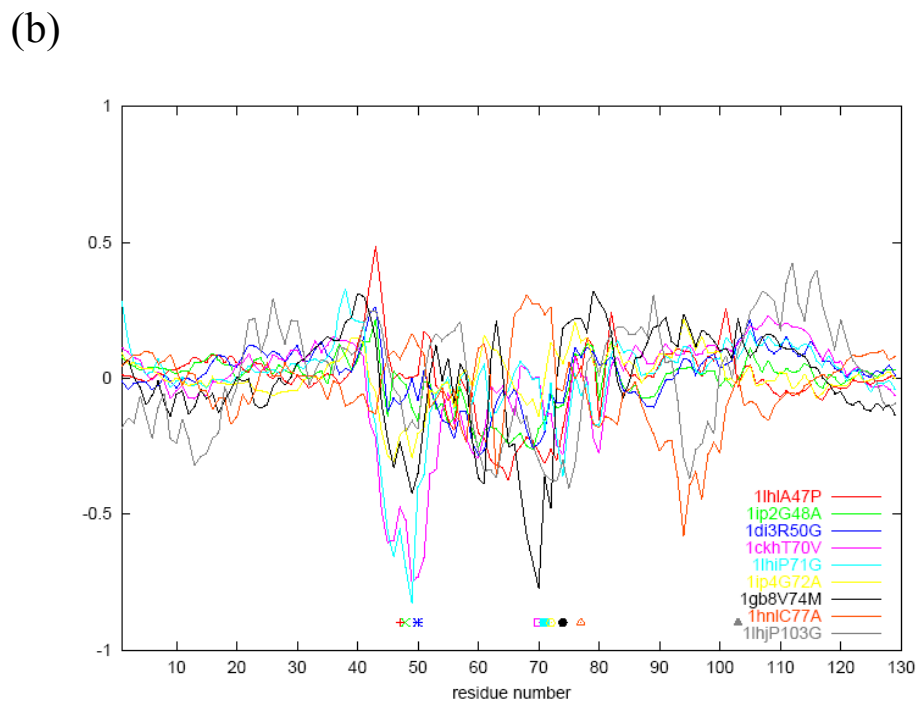
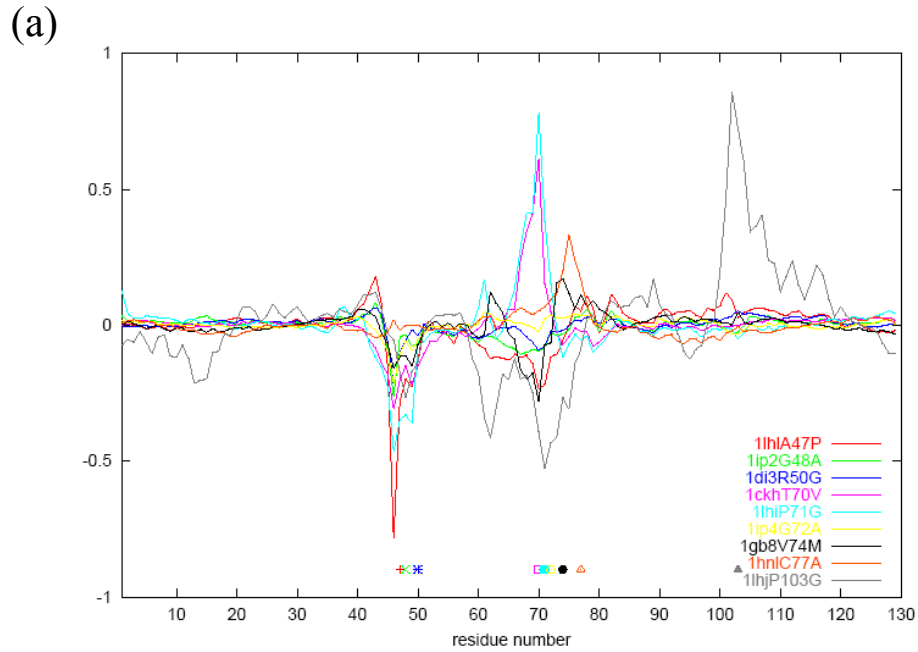


図 3-8 原子ペアのゆらぎの相関の野生型と変異型の差

1 残基置換体と野生型の相関の差と変位部位のゆらぎとの相関を各残基に対して 1 次元にプロットした。相関の三角関 (BPTI の図 1-1(C) に対応する図) の非対角要素を、対角線の自己相関の値のルートで割って規格化したもの (いわゆる相関係数) の差(上図)と、規格化しない絶対値のままの値の差(下図)についてプロットした。変位部位を図の下の方にグラフと同じ色の点で示した。変異部位との相関が比較的大きいものを選んで示しており、これ以外の変異型では、相関はあまり大きくなかった

4. 相同タンパク質間の比較

ProMode にデータが蓄積されてきて、互いに相同なタンパク質も含まれるようになってきた。そこで、*ProMode* では、SCOP で同じ superfamily に属するタンパク質を互いに相同なタンパク質とみなし、それらの比較することによって、タンパク質の立体構造の静的、動的特徴を調べることを試みた。相同タンパク質の比較では、上で述べたような 1 個から数個のアミノ酸置換とは異なり、全体的なフォールドは基本的に似ているものの、アミノ酸配列はかなり異なることがあり、また、アミノ酸残基の挿入・欠失もあることから、微細な変化を捉え、その原因を探るという研究ではなく、もう少し大局的な視点から、相同タンパク質間で変わらない性質や、個々のタンパク質に特異的な、他の相同なタンパク質とは異なる性質などに注目した解析が有効であろう。

図 4-1 は、相同タンパク質を比較した結果を表示した *ProMode* の画面の一例である。複数の互いに相同なタンパク質を ASH と呼ばれるソフトを使って立体構造アラインメントし、それに従って原子の位置のゆらぎ、および二面角のゆらぎを一つのグラフの中に表示し、比較できるようにした。従来の静的立体構造の比較ではなく、ゆらぎという動的特徴について、互いに似ている部分、異なる部分がわかるようにした。グラフの下には、アミノ酸配列と二次構造も示し、二次構造との関連、アミノ酸配列の違いとの関連などを利用者が調べることが出来るようにした。

また、4.1 から 4.5 には、互いに相同なタンパク質を比較した結果を 5 例列挙した。そこで、ここではまず、共通事項について述べ、それぞれの superfamily に固有の事項は、図の説明文の中で触れることにする。

解析方法

1) 相同タンパク質

SCOP の立体構造分類で同じ superfamily に属するものを相同なタンパク質とみなした。

2) 立体構造アラインメント

プログラム ASH(Toh, 1997; Daiyasu & Toh, 2000)によって、相同なタンパク質の立体構造マルチプル・アラインメントを行った。

Toh, H., Introduction of a distance cut-off approximation into structural alignment with the double dynamic programming algorithm. *CABIOS* **13** (1997) 387-396.

Daiyasu, H. & Toh, H. (2000) Molecular evolution of myeloperoxidase family. *J. Mol. Evol.* **51** (2000) 433-445.

3) 動的ドメインの同定

動的ドメインの同定は、各モードにおける構造のゆらぎに対しては DynDom (Hayward & Lee, 2002; Hayward et al., 1997)を用いた。また、すべてのモードの時間平均から得られる C α 原子ペアのゆらぎの相関をもとに、k-means 法によってアミノ酸残基をクラスター化し、それ

をドメインと定義することも行った(付録参照)。前者は *ProMode* で公開している動的ドメインであるが、各モードごとにドメインが定義されるため、複数のタンパク質を比較するにはあまり適さないと考え、後者の方法で定義されたドメインによって比較を行った。

S.Hayward, R.A.Lee, Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J Mol Graph Model*, **21** (2002) 181-183.

S.Hayward, A.Kitao, H.J.C.Berendsen, Model-Free Methods of Analyzing Domain Motions in Proteins from Simulation: A Comparison of Normal Mode Analysis and Molecular Dynamics Simulation of Lysozyme. *Proteins*, **27** (1997) 425.

解析結果

5つの superfamily について、それぞれ次の結果を示した。

1) 立体構造の重ね合わせ (図 4-2、4-8、4-14、4-20、4-26)

静的な構造比較のため、ASH を使った立体構造アラインメントの結果をもとに、立体構造の重ね合わせを行い、結果を図示した。それぞれのタンパク質を異なる色で表示している。また、SCOP での分類項目 Class と Fold の記述も示した。

2) アミノ酸配列、二次構造、残基の保存度など (図 4-3、4-9、4-15、4-21、4-27)

アミノ酸配列と二次構造を図式化した PDBsum (*)の図をコピーしたものを載せた。アミノ酸配列は、各アミノ酸残基の保存の程度によって色づけされており、保存される部位、変異しやすい部位がわかるようになっている。また、リガンド結合部位なども表示されている。さらに PROSITE のデータがあるものについては、別途そのデータも追加表示した。

* PDBsum <http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/>

3) 原子の位置のゆらぎ、二面角のゆらぎの比較 (図 4-4、4-10、4-16、4-22、4-28)

原子のゆらぎ、二面角のゆらぎのグラフを、立体構造アラインメントの結果に従い、対応する残基が同じ位置にくるように描画した。これによって、互いに相同なタンパク質の間でのゆらぎの様相の違いを見ることができる。また、比較のため、立体構造を重ね合わせたときのずれ(静的構造の違い)も表示した。

4) 動的ドメイン (図 4-5、4-11、4-17、4-23、4-29)

C α のゆらぎの相関をもとに、動的ドメインを求めた結果を、立体構造上に色分けして表示した。相同なタンパク質間で比較できるよう、できるだけ対応するドメインが同じ色になるように心がけた。しかし、定義されたドメインが相同タンパク質間で必ずしもぴったり一致するわけではないので、色の指定は試行錯誤で行った。しかし、構造を重ね合わせてみると、同じ領域の重なりから、共通性が見えてきた。

5) ゆらぎの相関マップと動的ドメイン (図 4-6、4-12、4-18、4-24、4-30)

$C\alpha$ のゆらぎの相関をもとに、動的ドメインが定義される様子を具体的に示すため、代表的なタンパク質を適当に一つ選び、相関マップと、それによって定義されたドメインの立体構造上での位置を色分けして示した。また、同じ **superfamily** 内のすべてのタンパク質に共通して正の相関をもつ残基ペア、あるいは負の相関をもつ残基ペアを強調して表示した。

6) 特定の基準振動モードでの動的ドメイン (図 4-7、4-13、4-19、4-25、4-31)

各タンパク質について、低振動モードから一つ選び、その振動モードに対して **DynDom** で定義されたドメインの例を、立体構造上に色分けして示した。基本的には最も低い振動モードを選んだが、**DynDom** ではドメインが定義されないことが多く (その場合は、全体が一つのドメインと定義されたとも考えられる)、その場合には、2 番目以降の低振動モードの例を示した。また、ドメイン間の相対運動を、らせん運動とみなしたときのらせん軸も表示した。ここで示した図は、**ProMode** 上で表示されている図と同じものである。

結果の概略

5 つの **superfamily** を概観して、以下の点が明らかになった。

- 1) 二次構造内の原子のゆらぎは概して小さいが、個々にみると大きい部位、小さい部位がある。 α ヘリックスは、二面角のゆらぎはすべての残基で小さいから、構造としては硬いと考えられ、部位によるゆらぎの大きさは、分子の中心に近い部位が比較的固定され、遠い部位がゆらぐような剛体運動を反映していると考えられる。一方、 β 構造は一般に二面角のゆらぎが大きく、やわらかい。その割には原子のゆらぎが小さいのは、二面角が互いに構造をあまり崩さないように相補的に動くためであろう。
- 2) 二次構造のゆらぎは、相同タンパク質間では、定性的にも定量的にもよく一致している。ループ部分にかなりの挿入・欠失がみられる場合でも、ゆらぎの大きさにあまり変化はない (たとえば、**flavoprotein** の **1qr2A**)。一方、ループ領域は、一般にゆらぎが大きい。しかしその大きさは、挿入・欠失を反映して、それぞれのタンパク質で大きく異なる。
- 3) $C\alpha$ 間のゆらぎの相関から定義されたドメインは、ドメインの境界は、相同タンパク質間でかなり異なるが、その主要部分はよく一致している。逆に、タンパク質によってかなり異なる領域もみられ、そうした部分はそれぞれのタンパク質の個性が反映されている部分と考えられる。ドメインを色分けして、構造を重ね合わせた図からこうしたことが読み取れるが、もう少し定量的な評価法を考え、客観的評価を行う必要がある。
- 4) 個々のモードでは、同じモード番号同士が対応するというわけではないので、比較することは難しい。しかし、同じようにドメインが定義され、同じようなドメイン間の相対運動を行うモードが相互にみられることも確かである。

Database of normal mode analysis of proteins

Comparison between proteins in the same superfamily

SuperFamily: **SH2 domain**

1:1ayaA 2:1csyA 3:1d4tA 4:1m1 5:2pldA 6:2pna

Fluctuation of atom positions Fluctuation of torsion angles

Execute SuperFamily List

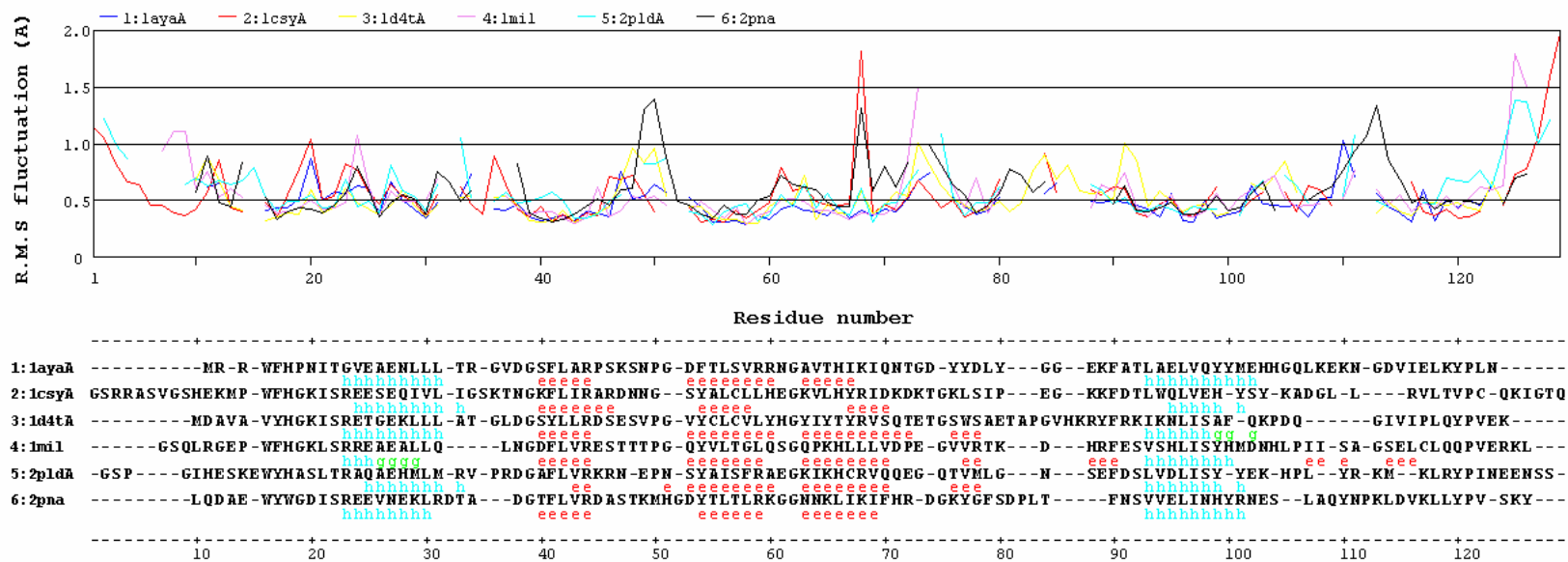


図 4-1 相同タンパク質の比較をした ProMode の例

SH2 domain superfamily に属する 6 つのタンパク質の原子のゆらぎを、立体構造アラインメントに即し、対応する残基が同じ残基位置にくるようにプロットしている。二面角についても同様の結果を表示することができる。

4.1 Superfamily: Cytochromes

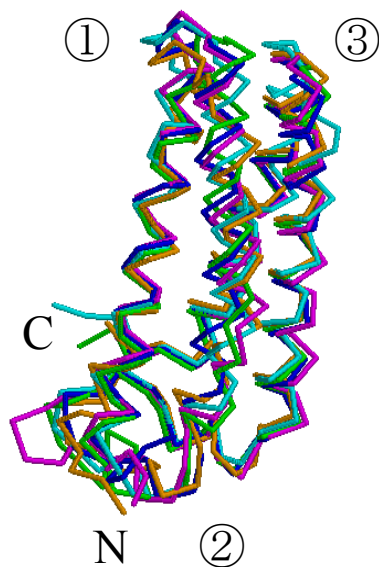


図 4-2 Cytochrome superfamily の立体構造の重ね合わせ。

PDB ID

1a7vA Cytochrome c' (Rhodopseudomonas palustris)

1jafA Cytochrome c' (Rhodocyclus gelatinos)

1bbhA Cytochrome c' (Chromatium vinosum)

1cpq Cytochrome c' (Rhodopseudomonas capsulate)

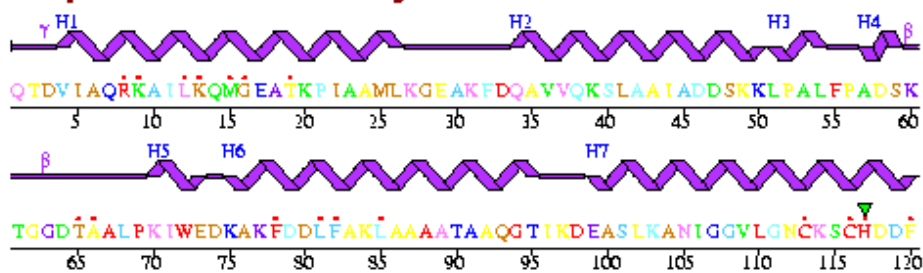
2ccyA Cytochrome c' (Rhodospirillum molischianum)

SCOP classification

Class: All alpha proteins

Fold: Four-helical up-and-down bundle

Sequence coloured by residue conservation:



Active site: - ▼ HEA Residue interactions: - • with ligand

Conservation colouring: Low 1 2 3 4 5 6 7 8 9 High

PROSITE: PS00190 - CYTOCHROME_C Cys113 to Asp118: C K S C H D

図 4-3 アミノ酸配列、二次構造、残基の保存性、PROSITE パターン

1a7vA について PDBsum からコピーした。

PDBsum: <http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/>

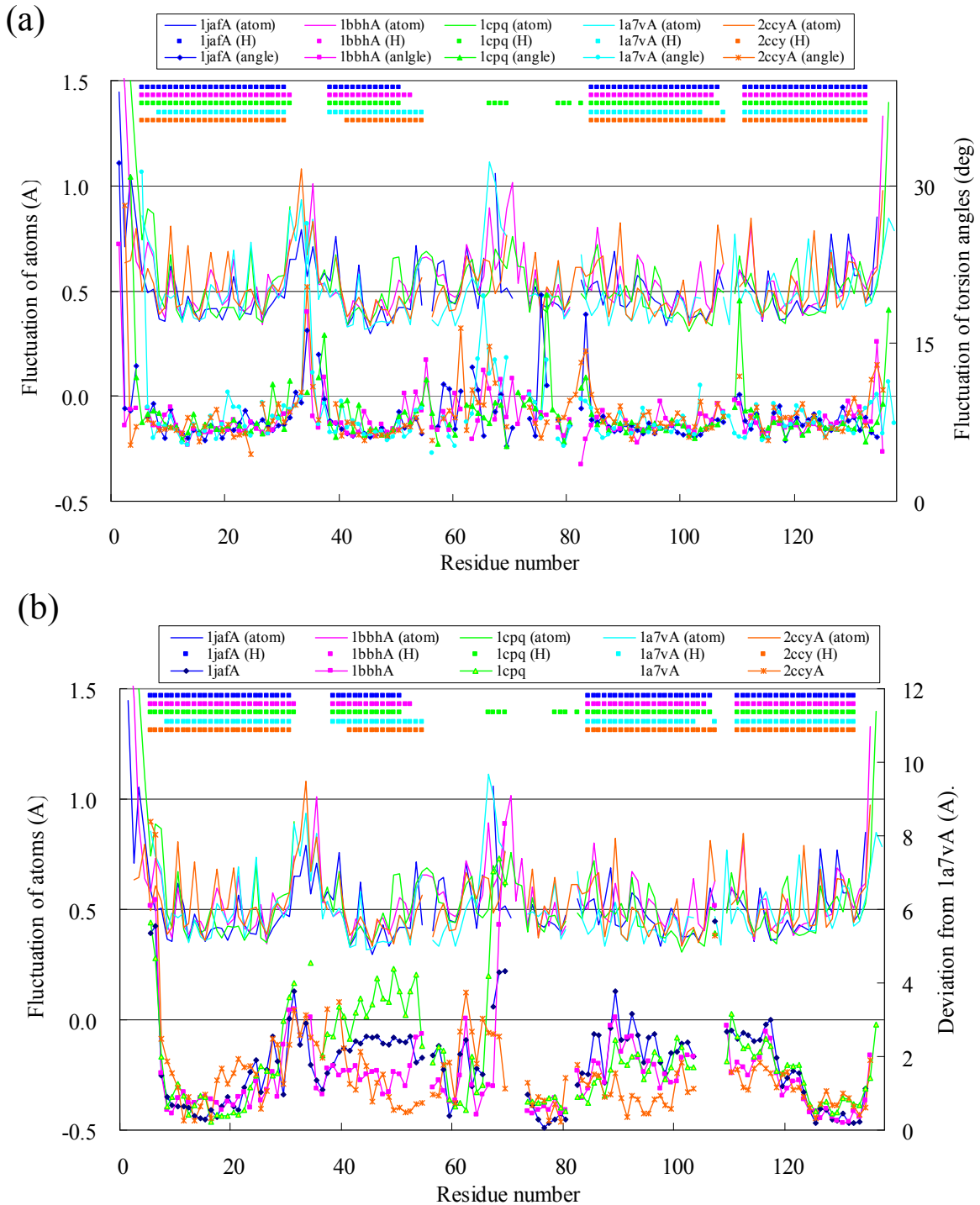


図 4-4 Cytochrome superfamily の原子の位置のゆらぎ、二面角のゆらぎの比較

(a) 上のグラフは原子の位置のゆらぎを、下のグラフは二面角 ϕ のゆらぎを表している。(b) 上のグラフは(a)と同じく原子の位置のゆらぎであり、下のグラフは、立体構造を重ね合わせたときの 1a7vA からのずれを表している。それぞれのグラフの上部にある ■ は α ヘリクスの位置を示している。また、色は、凡例に示すように、それぞれ特定のタンパク質に対応して用いられている。2 番目の α ヘリクスは、静的構造比較では大きなずれを生じているが、そのことが残基のゆらぎには影響を与えていない。

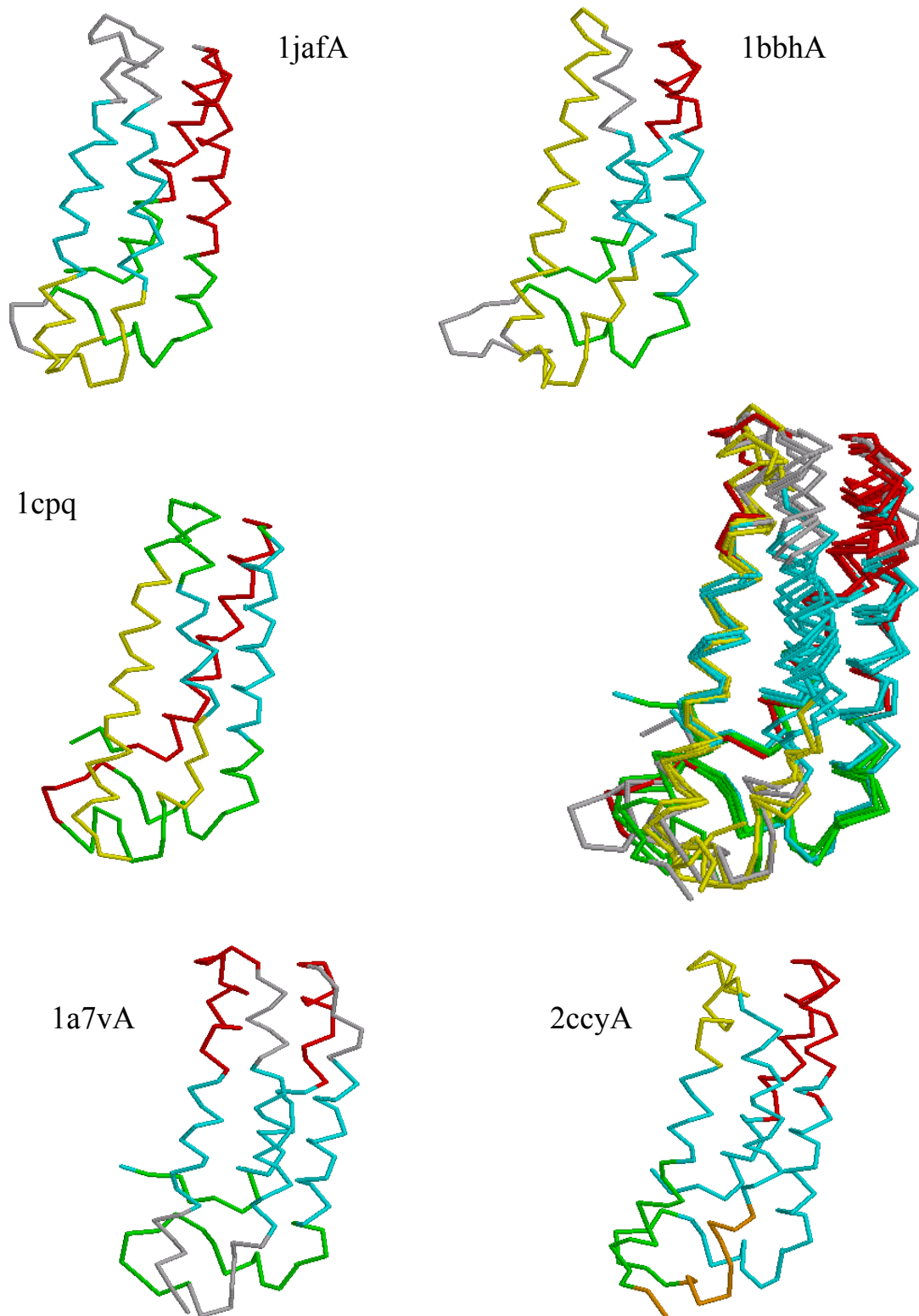
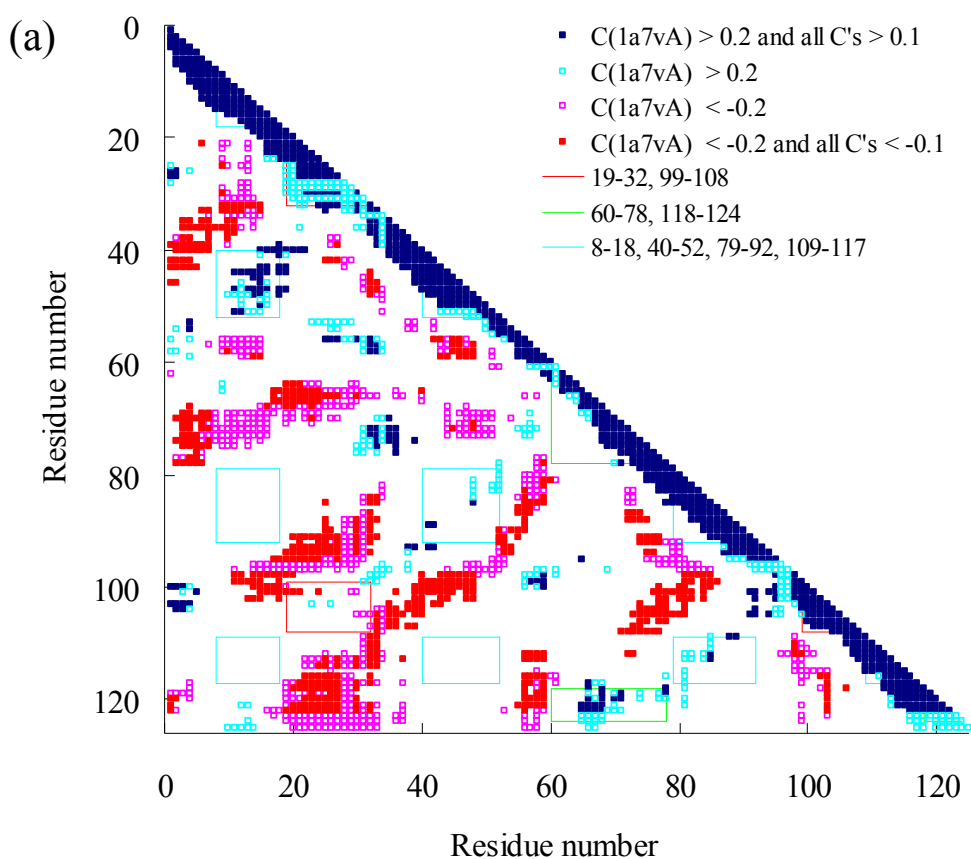


図 4-5 Cytochrome superfamily の動的ドメイン

C α のゆらぎの相関から定義した動的ドメインを、色分けして示した。中央右は、すべての構造を重ね合わせたもので、できるだけ同じ色のドメインが重なるように、色を指定した。4本の α ヘリックスの束からなる中央のドメイン、その両側のループ領域のドメインからなる。



(b)

Domain1	19-32, 99-108	(red)
Domain2	60-78, 118-124	(green)
Domain3	8-18, 40-52, 79-92	(cyan)

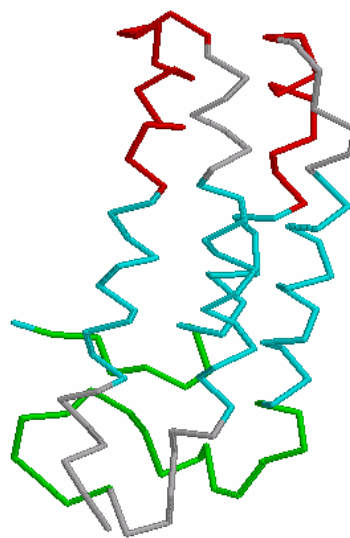


図 4-6 ゆらぎの相関マップと動的ドメイン

1a7vA を例に示している。(a) 1a7vA の相関係数 C が 0.2 以上の残基ペアを □ で、-0.2 以下の残基ペアを ■ で示した。さらに、他のタンパク質の対応する残基ペアもすべて 0.1 以上、あるいは -0.1 以下の場合、それぞれ ■ あるいは ■ で示している。また、定義された 3 つのドメインについて、それらに含まれる残基ペアをそれぞれ細線で囲んで示した。(b) (a) の三角マップによって定義されたドメインを立体構造上に色分けして示した。

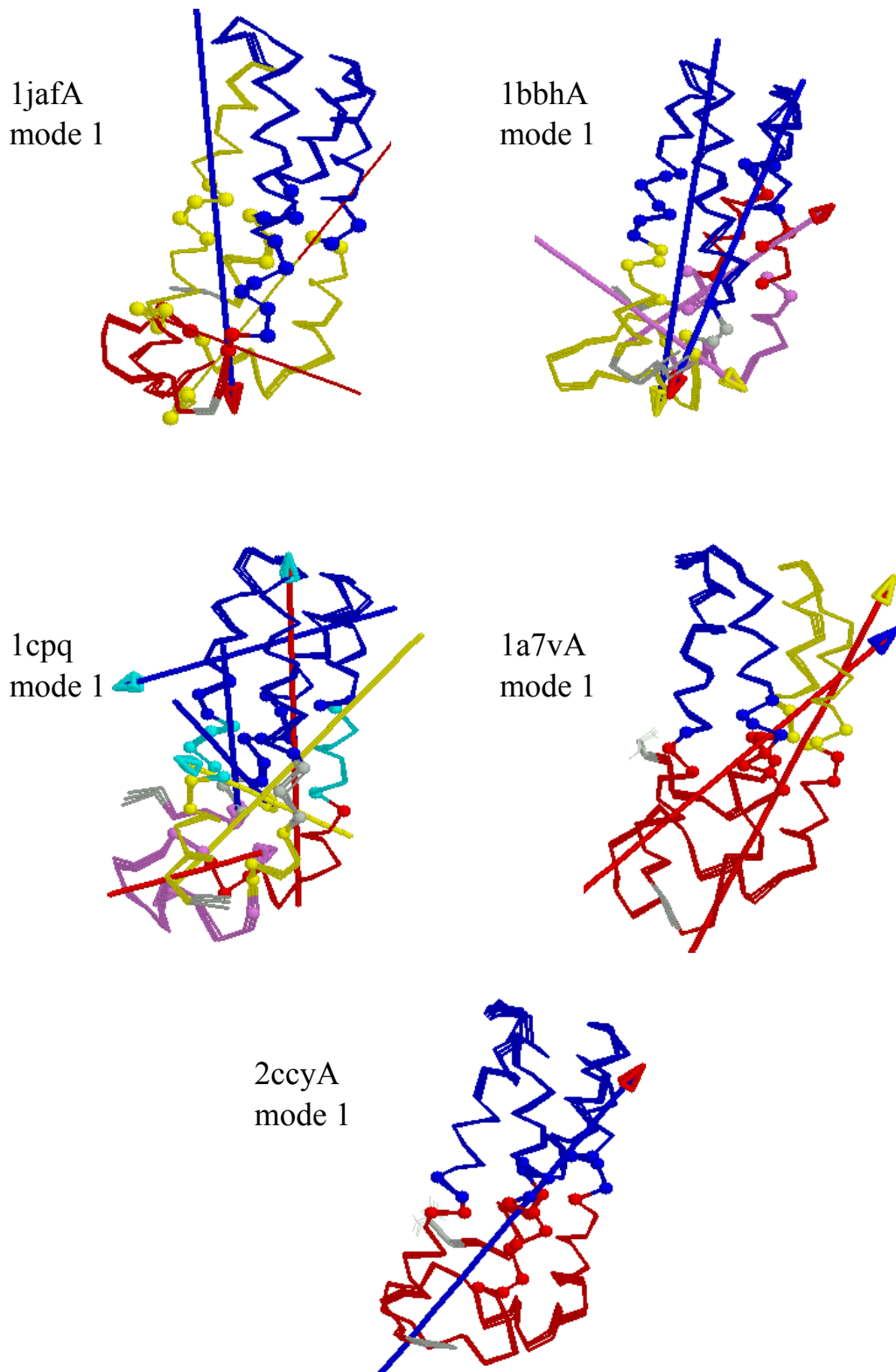
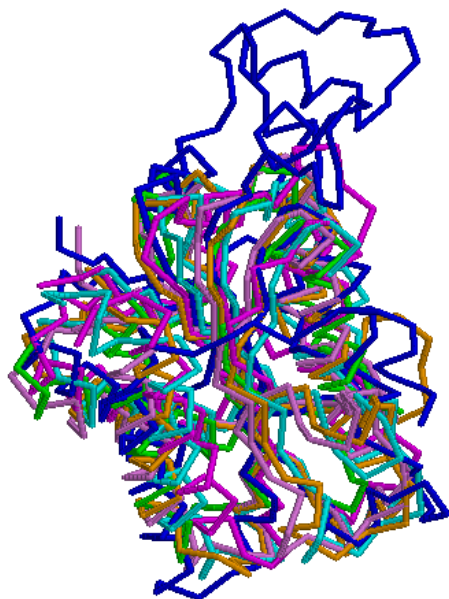


図 4-7 特定の基準振動モードに対して定義された動的ドメインの例

DynDom によって定義された動的ドメインを色分けして示した。ただし、灰色はドメインとして定義されなかった領域を意味している。矢印はそれらドメインの相対運動をらせん運動とみなしたときのらせん軸で、矢尻と軸との色が2つのドメインの色を示している。また、球は、蝶番領域のアミノ酸残基を表している。

4.2 Superfamily: Flavoproteins

図4-8 Flavoproteins superfamilyの立体構造の重ね合わせ。



PDB ID

- 1qr2A Quinone reductase type 2
- 1bvyF Cytochrome p450 bm-3
- 5nul Flavodoxin (Clostridium beijerinckii)
- 2fcr Flavodoxin (Chondrus crispus)
- 1ag9A Flavodoxin (Escherichia coli)
- 1rcf Flavodoxin (Anabaena)

SCOP classification

- Class: Alpha and beta proteins (a/b)
- Mainly parallel beta sheets (beta-alpha-beta units)
- Fold: Flavodoxin-like
- 3 layers, a/b/a; parallel beta-sheet of 5 strand

Sequence coloured by residue conservation:

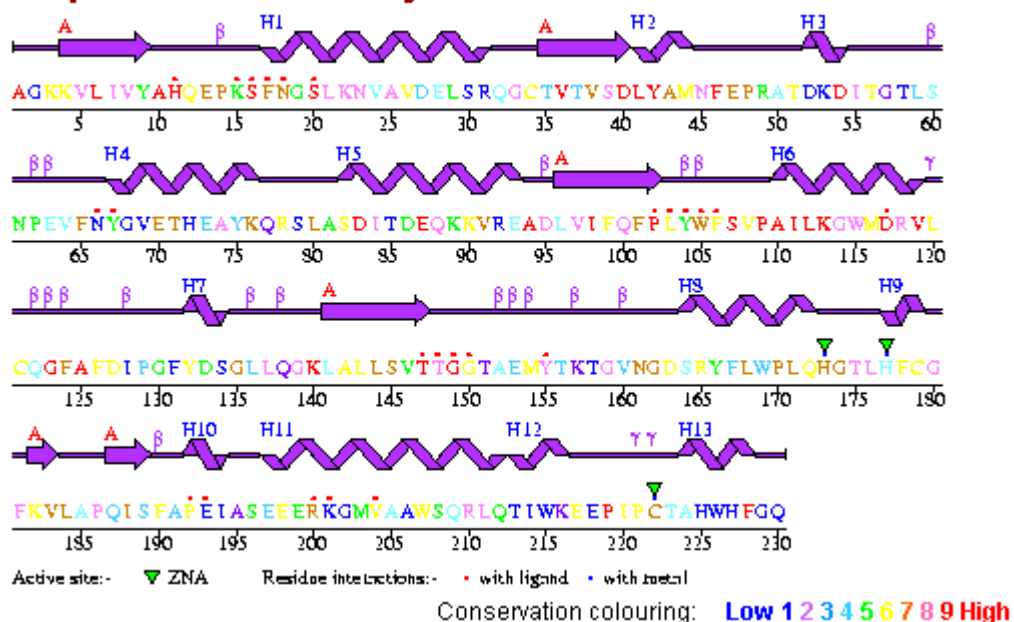


図 4-9 アミノ酸配列、二次構造、残基の保存性。

1qr2A について PDBsum からコピーした。

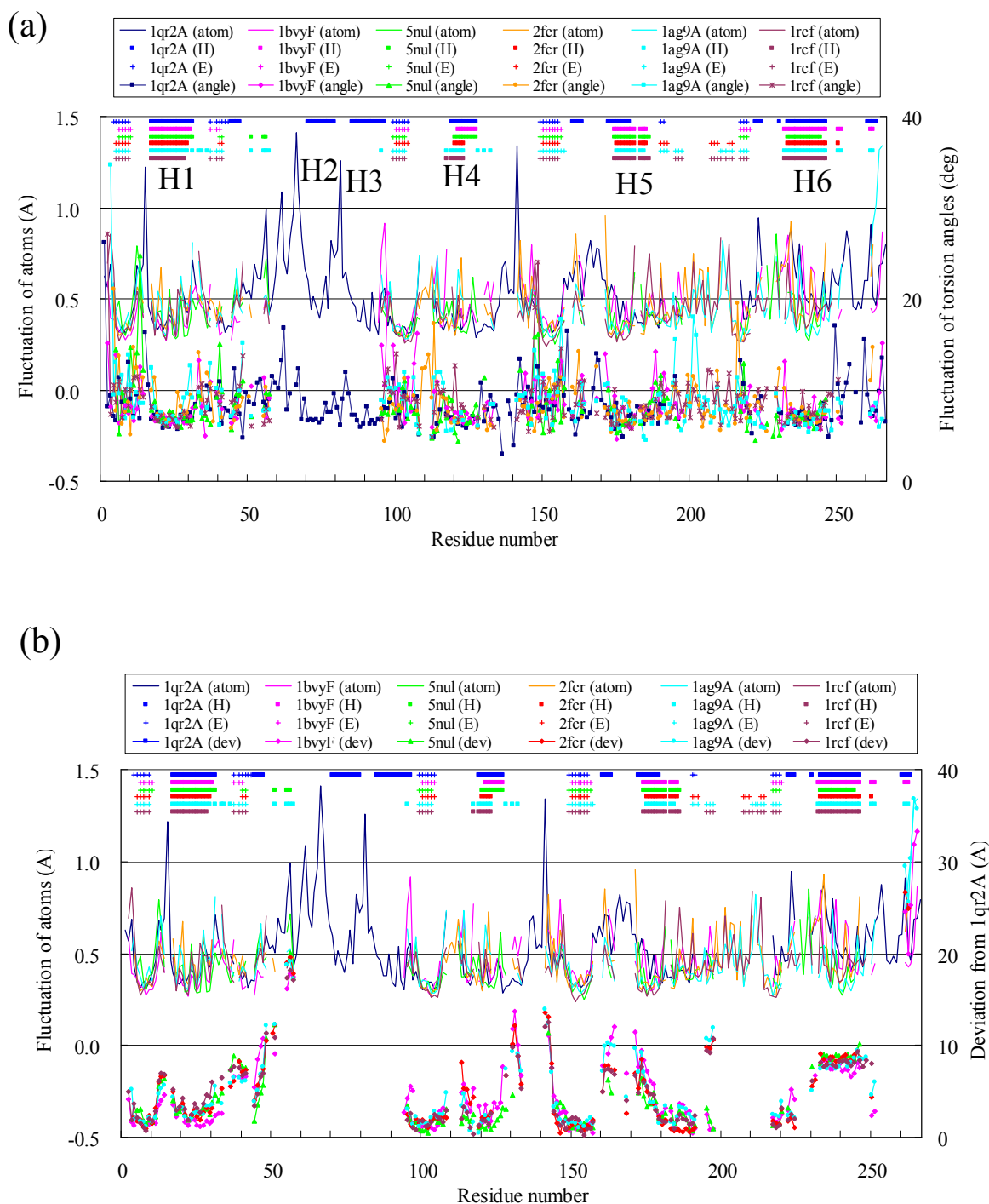


図 4-10 Flavoproteins superfamily の原子の位置のゆらぎ、二面角のゆらぎの比較

(a) 上のグラフは原子の位置のゆらぎを、下のグラフは二面角 ϕ のゆらぎを表している。
 (b) 上のグラフは(a)と同じく原子の位置のゆらぎであり、下のグラフは、立体構造を重ね合わせたときの 1qr2A からのずれを表している。それぞれのグラフの上部にある ■ と + はそれぞれ α ヘリックスと β ストランドの位置を示している。また、色は、凡例に示すように、それぞれ特定のタンパク質に対応して用いられている。1qr2A には大きな挿入があるにもかかわらず、細部にいたるまで、他のタンパク質とよく似ていることが注目される。

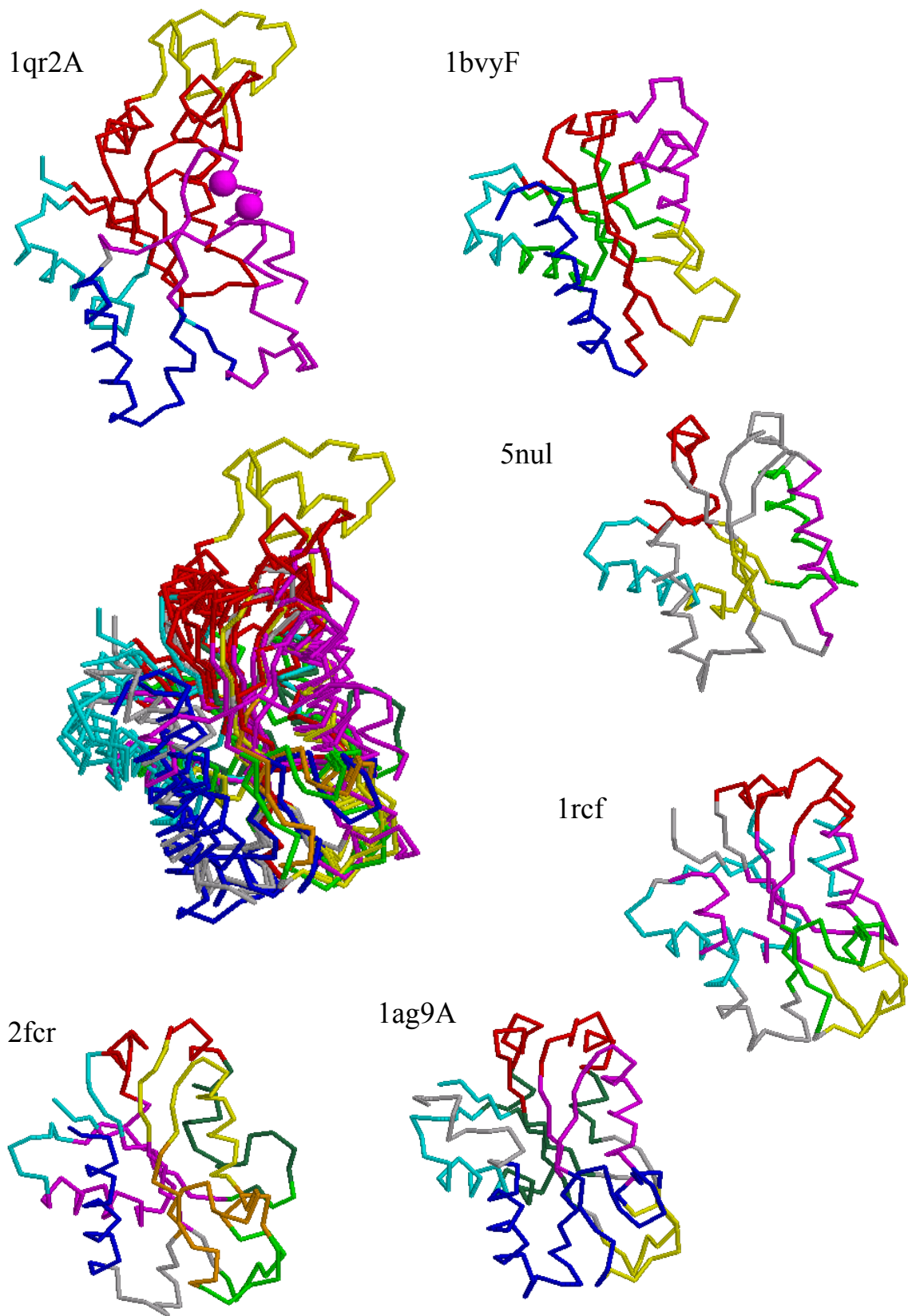
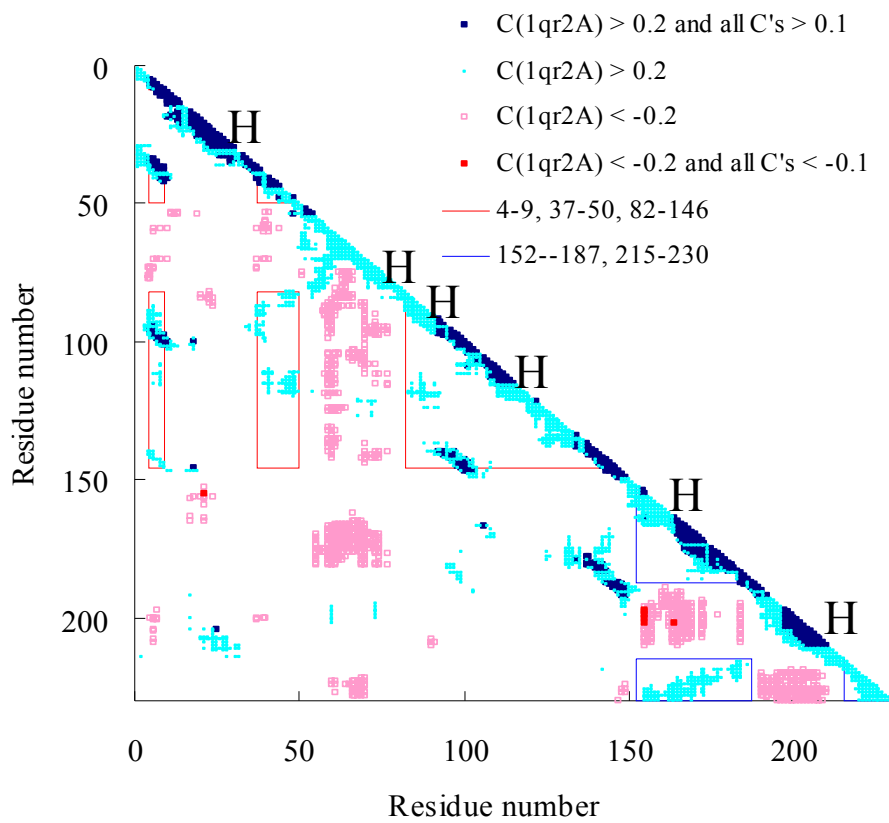


図 4-11 Flavoproteins superfamily の動的ドメイン

$C\alpha$ のゆらぎの相関から定義した動的ドメインを、色分けして示した。中央左は、すべての構造を重ね合わせたもので、できるだけ同じ色のドメインが重なるように、色を指定した。右下の領域に多様性がみられる。



Domain1	4-9, 37-50, 82-146	(red)
Domain2	1-3, 10-36, 147-147	(cyan)
Domain3	51-81	(yellow)
Domain4	148-151, 188-213	(blue)
Domain5	152-187, 215-230	(magenta)

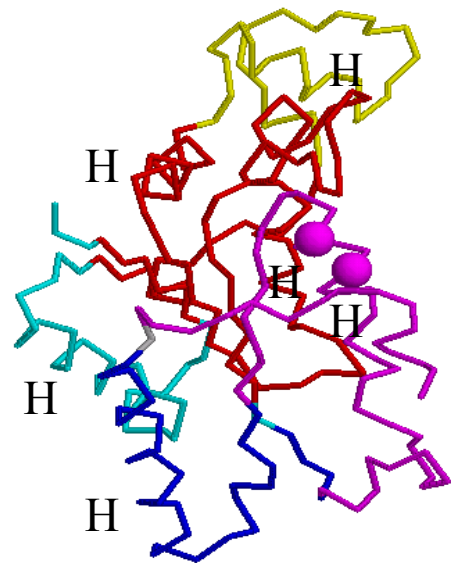


図 4-12 ゆらぎの相関マップと動的ドメイン。

1qr2A を例に示している。(a) 1qr2A の相関係数 C が 0.2 以上の残基ペアを \square で、-0.2 以下の残基ペアを \square で示した。さらに、他のタンパク質の対応する残基ペアもすべて 0.1 以上、あるいは -0.1 以下の場合、それぞれ \blacksquare あるいは \blacksquare で示している。また、定義された 5 つのドメインのうち 2 つについて、それらに含まれる残基ペアをそれぞれ細線で囲んで示した。(b) (a) の三角マップによって定義されたドメインを立体構造上に色分けして示した。

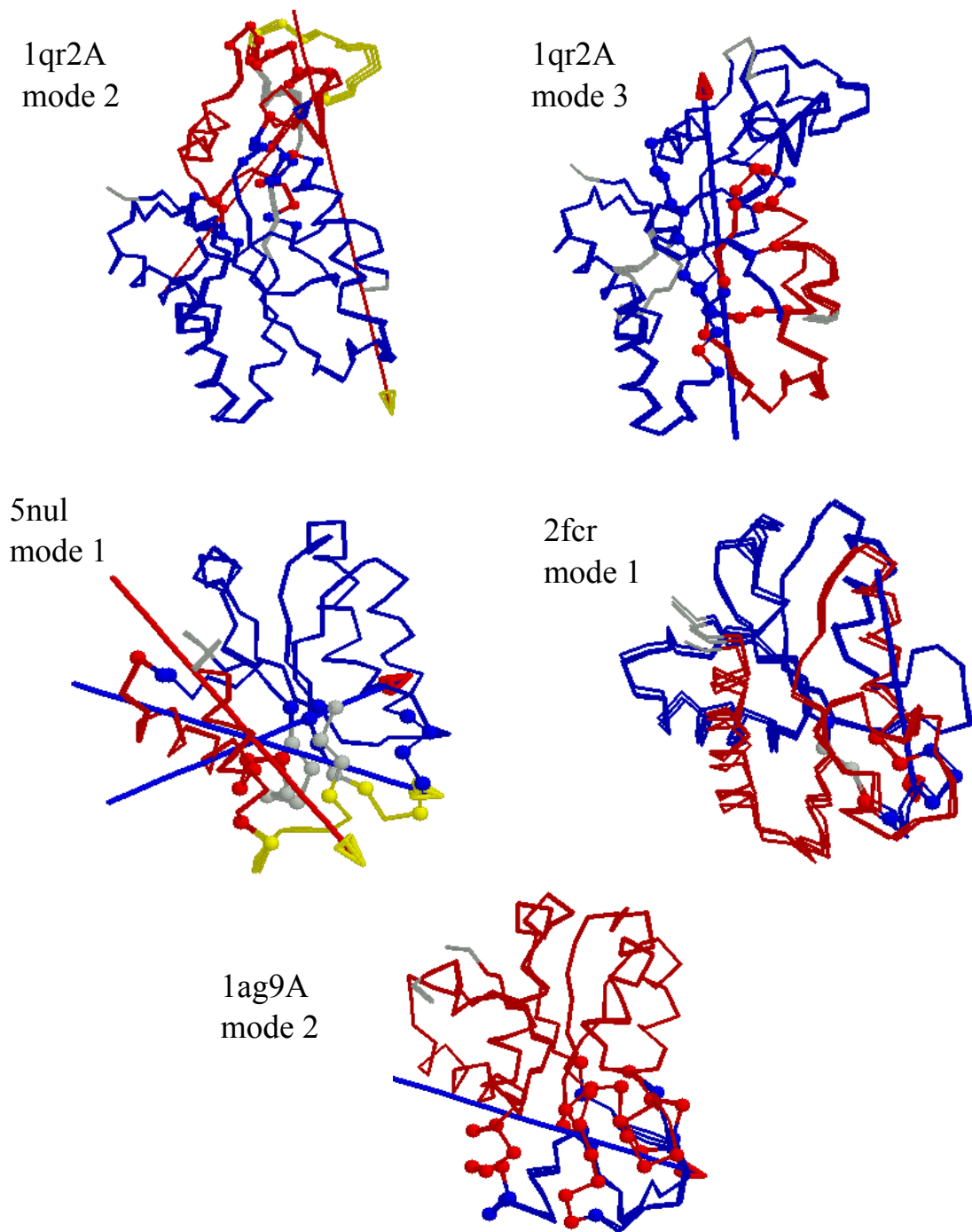


図 4-13 特定の基準振動モードに対して定義された動的ドメインの例

DynDom によって定義された動的ドメインを色分けして示した。ただし、灰色はドメインとして定義されなかった領域を意味している。矢印はそれらドメインの相対運動をらせん運動とみなしたときのらせん軸で、矢尻と軸との色が2つのドメインの色を示している。また、球は、蝶番領域のアミノ酸残基を表している。

4.3 Superfamily: Ribulose-phosphate binding barrel

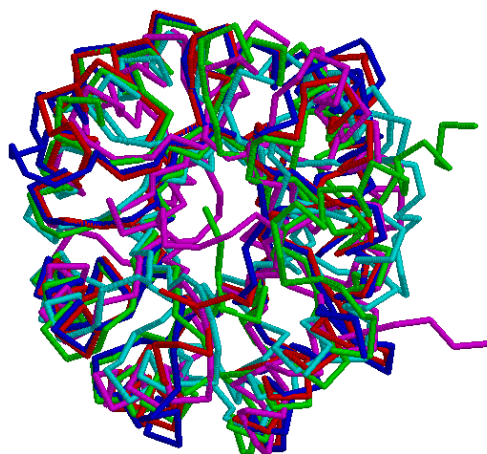


図 4-14 Ribulose-phosphate binding barrel superfamily の立体構造の重ね合わせ。

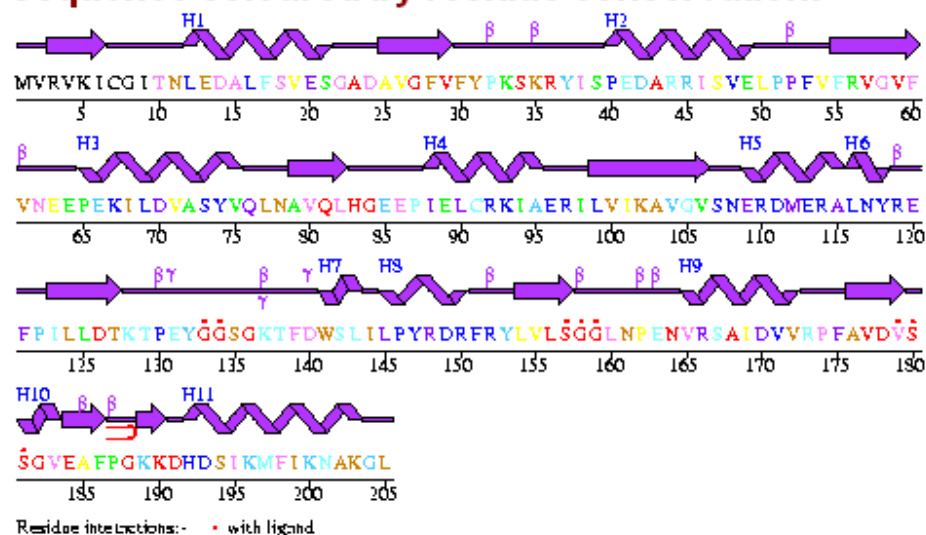
PDB ID

1nsj	Phosphoribosyl anthranilate isomerase
1rpxA	Ribulose-phosphate 3-epimerase
1dvjA	Orotidine 5'-phosphate decarboxylase (Methanobacterium thermoautotrophicum)
1dbtA	Orotidine 5'-phosphate decarboxylase (Bacillus subtilis)
1eixC	Orotidine 5'-monophosphate decarboxylase (Escherichia coli)

SCOP classification

Class: Alpha and beta proteins (a/b)	Mainly parallel beta sheets (beta-alpha-beta units)
Fold: TIM beta/alpha-barrel	contains parallel beta-sheet barrel, closed; n=8, S=8

Sequence coloured by residue conservation:



Conservation colouring: Low 1 2 3 4 5 6 7 8 9 High

図 4-15 アミノ酸配列、二次構造、残基の保存性。

1nsj について PDBsum からコピーした。

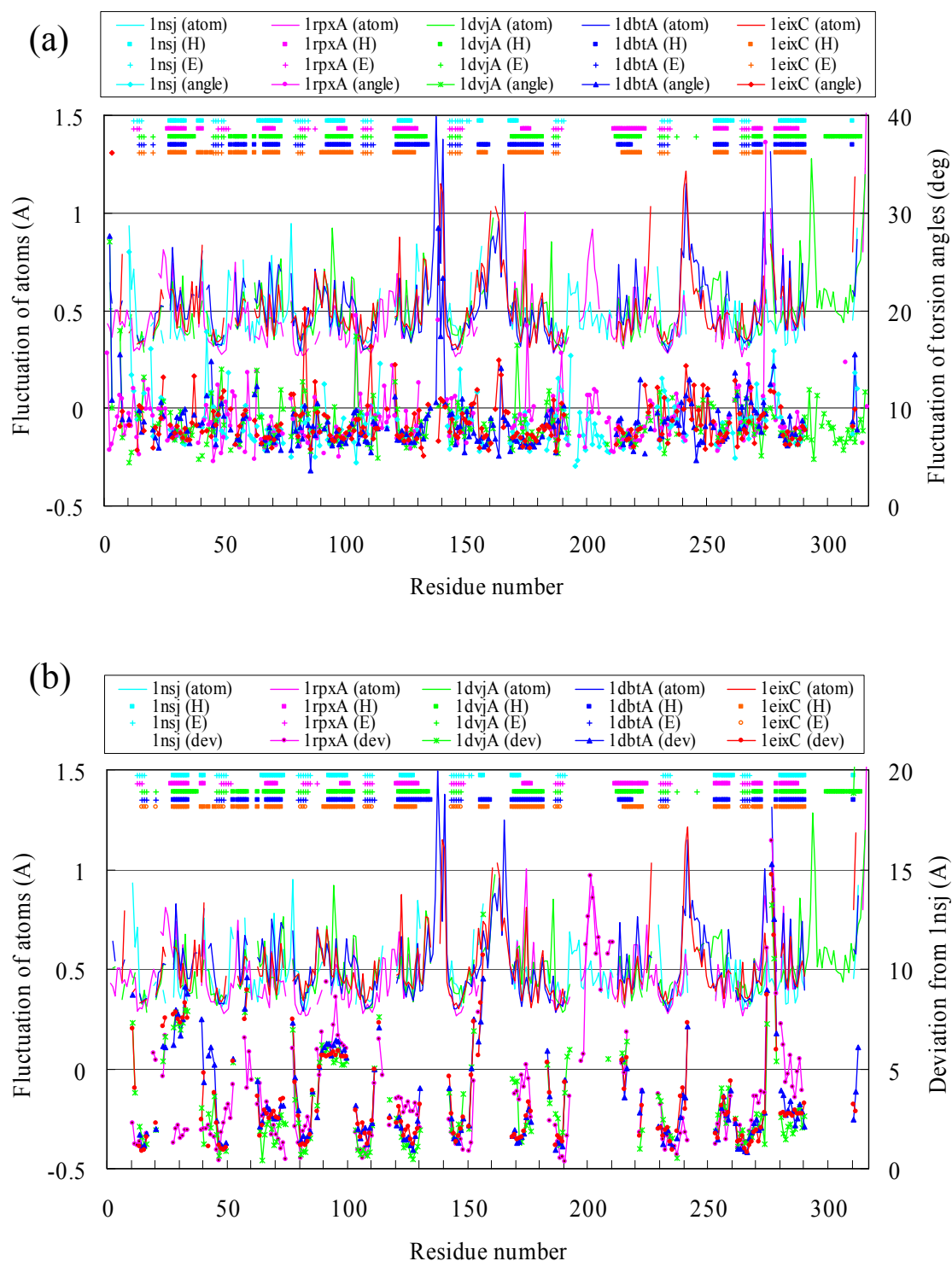


図 4-16 Ribulose-phosphate binding barrel superfamily の原子の位置のゆらぎ、二面角のゆらぎの比較。(a) 上のグラフは原子の位置のゆらぎを、下のグラフは二面角 ϕ のゆらぎを表している。(b) 上のグラフは(a)と同じく原子の位置のゆらぎであり、下のグラフは、立体構造を重ね合わせたときの 1nsj からのずれを表している。それぞれのグラフの上部にある ■ と + はそれぞれ α ヘリックスと β ストランドの位置を示している。また、色は、凡例に示すように、それぞれ特定のタンパク質に対応して用いられている。

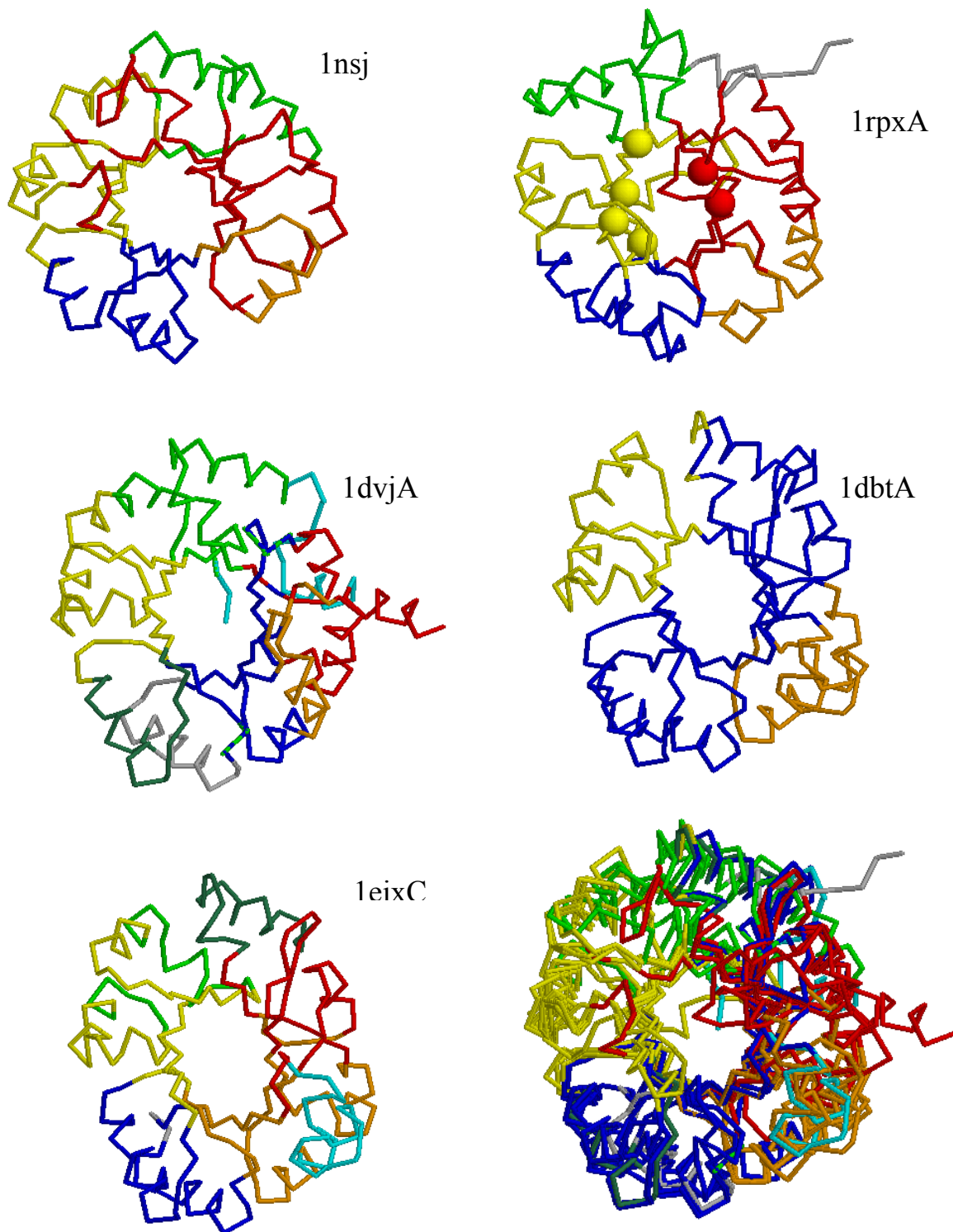
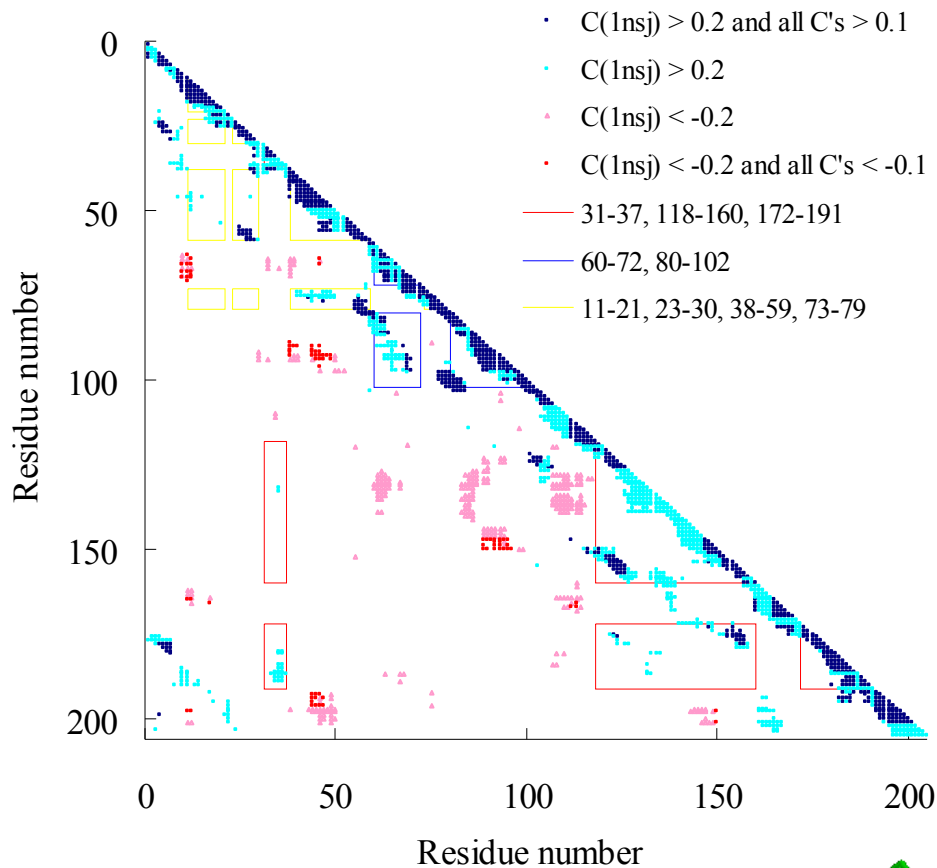


図 4-17 Ribulose-phosphate binding barrel superfamily の動的ドメイン
 $C\alpha$ のゆらぎの相関から定義した動的ドメインを、色分けして示した。右下の図は、すべての構造を重ね合わせたもので、できるだけ同じ色のドメインが重なるように、色を指定した。



Domain1	1-5, 22-22, 161-171, 192-205	(green)
Domain2	60-72, 80-102	(blue)
Domain3	6-8, 10-10, 31-37, 118-160, 172-191	(red)
Domain4	9-9, 11-21, 23-30, 38-59, 73-79	(yellow)
Domain5	103-117	(orange)

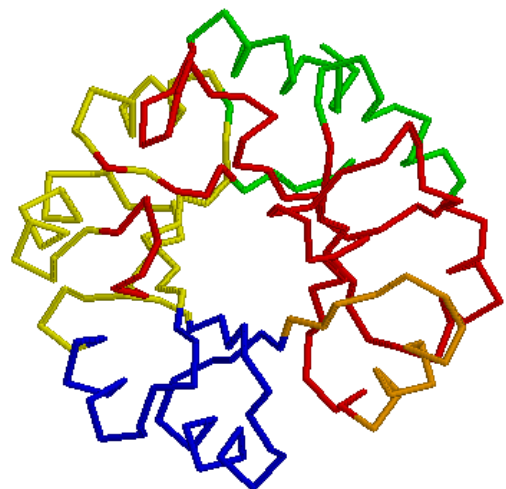


図 4-18 ゆらぎの相関マップと動的ドメイン

1nsj を例に示している。(a) 1nsj の相関係数 C が 0.2 以上の残基ペアを \square で、 -0.2 以下の残基ペアを \square で示した。さらに、他のタンパク質の対応する残基ペアもすべて 0.1 以上、あるいは -0.1 以下の場合、それぞれ \blacksquare あるいは \blacksquare で示している。また、定義された 5 つのドメインのうち 2 つについて、それらに含まれる残基ペアをそれぞれ細線で囲んで示した。(b) (a) の三角マップによって定義されたドメインを立体構造上に色分けして示した。

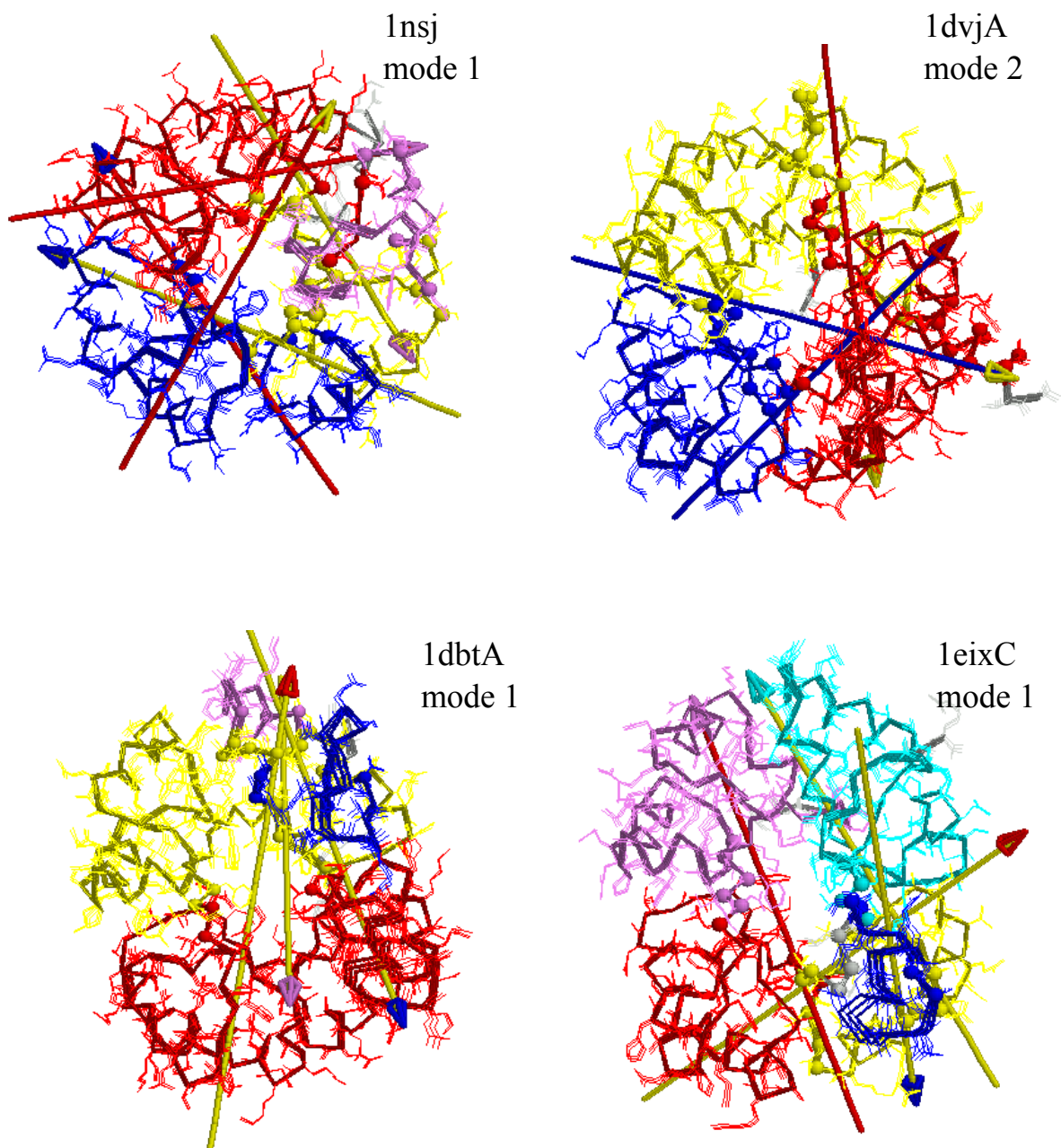


図 4-19 特定の基準振動モードに対して定義された動的ドメインの例

DynDom によって定義された動的ドメインを色分けして示した。ただし、灰色はドメインとして定義されなかった領域を意味している。矢印はそれらドメインの相対運動をらせん運動とみなしたときのらせん軸で、矢尻と軸との色が2つのドメインの色を示している。また、球は、蝶番領域のアミノ酸残基を表している。

4.4 Superfamily: SH2 domain

図 4-20 SH2-domain superfamily の立体構造の重ね合わせ。



PDB ID

2pna	Phosphatidylinositol 3-kinase
2pldA	Phospholipase C-gamma-1
1csyA	Syk tyrosine kinase
1mil	Shc adaptor protein
1ayaA	Tyrosine phosphatase syp
1d4tA	T cell signal transduction molecule sap

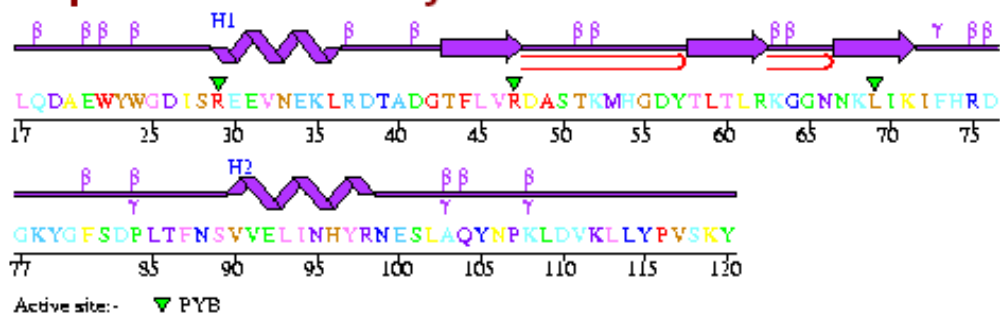
SCOP classification

Class: Alpha and beta proteins (a+b)
 Mainly antiparallel beta sheets
 (segregated alpha and beta regions)

Fold: SH2-like

3 layers: a/b/a; antiparallel beta-sheet of 5 strands is flanked by two helices

Sequence coloured by residue conservation:



Conservation colouring: Low 1 2 3 4 5 6 7 8 9 High

図 4-21 アミノ酸配列、二次構造、残基の保存性。

2pna について PDBsum からコピーした。

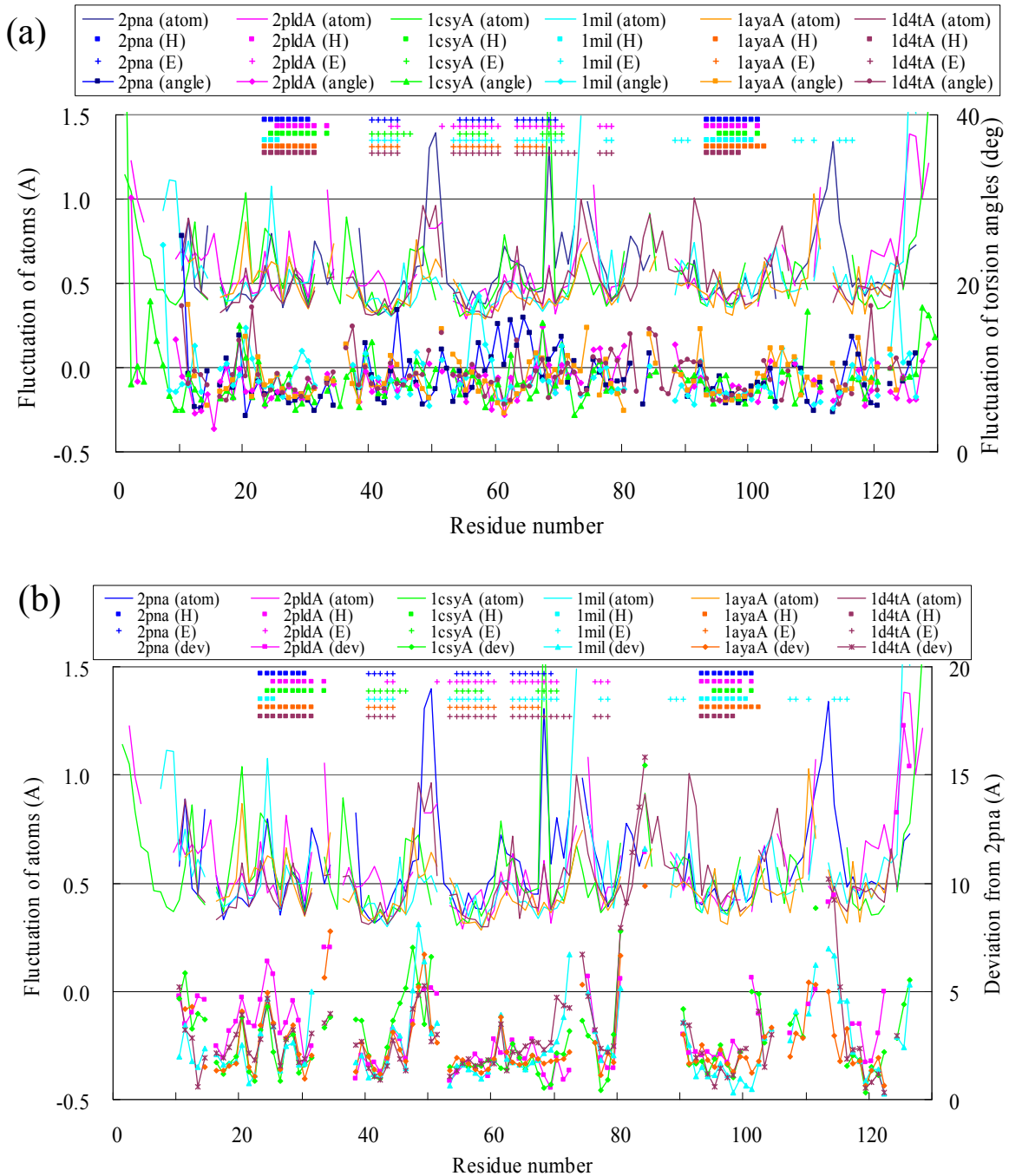


図 4-22 SH2-domain superfamily の原子の位置のゆらぎ、二面角のゆらぎの比較

(a) 上のグラフは原子の位置のゆらぎを、下のグラフは二面角 ϕ のゆらぎを表している。(b) 上のグラフは(a)と同じく原子の位置のゆらぎであり、下のグラフは、立体構造を重ね合わせたときの 2pna からのずれを表している。それぞれのグラフの上部にある ■ と + はそれぞれ α ヘリックスと β ストランドの位置を示している。また、色は、凡例に示すように、それぞれ特定のタンパク質に対応して用いられている。

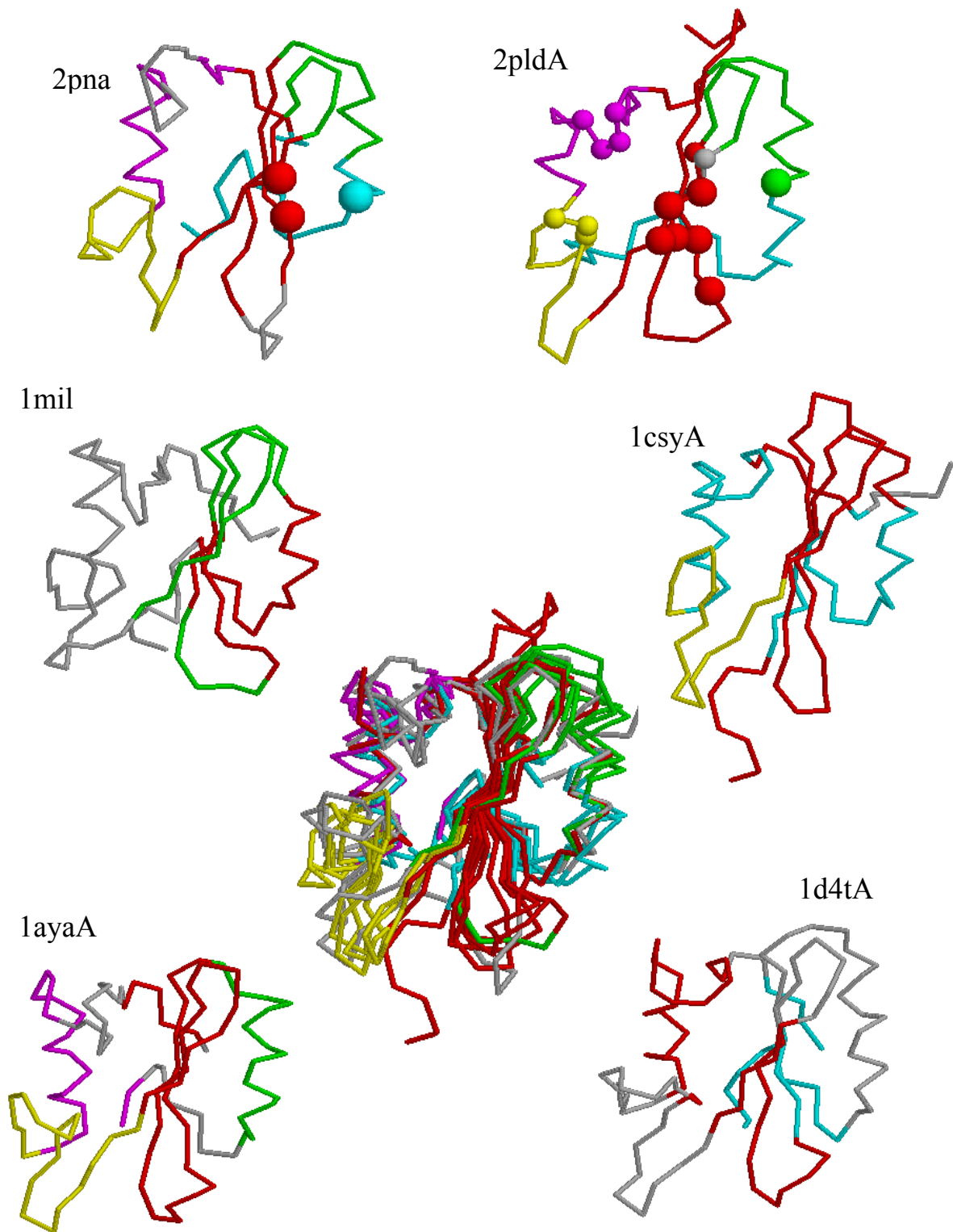
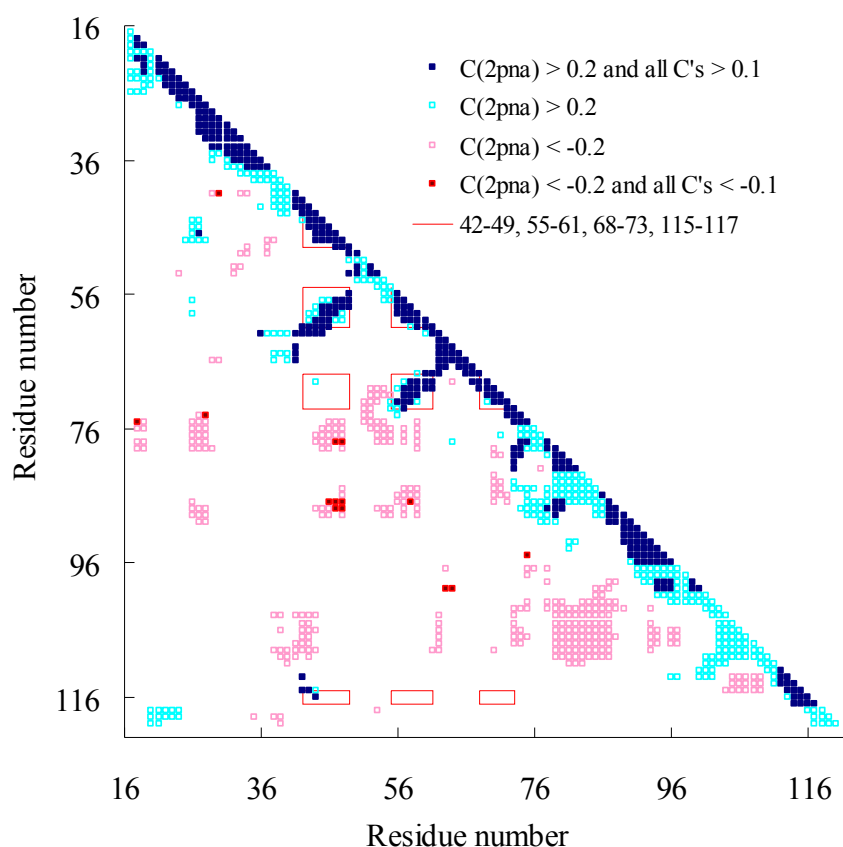


図 4-23 SH2-domain superfamily の動的ドメイン

$C\alpha$ のゆらぎの相関から定義した動的ドメインを、色分けして示した。中央の図は、すべての構造を重ね合わせたもので、できるだけ同じ色のドメインが重なるように、色を指定した。また、2pna、2pldA では、活性部位のアミノ酸残基を球で表した。



Domain1	42-49, 55-61, 68-73, 115-117	(red)
Domain2	90-102, 111-114	(magenta)
Domain3	17-32, 118-120	(cyan)
Domain4	33-41, 62-67	(green)
Domain5	74-89	(yellow)

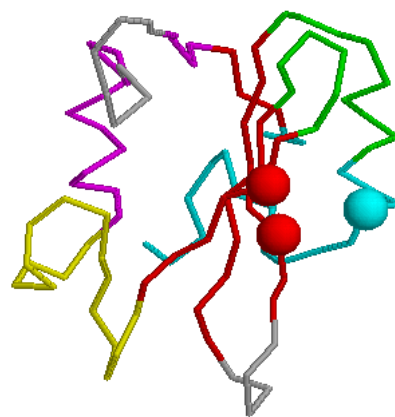


図 4-24 ゆらぎの相関マップと動的ドメイン

2pna を例に示している。(a) 2pna の相関係数 C が 0.2 以上の残基ペアを \square で、 -0.2 以下の残基ペアを \square で示した。さらに、他のタンパク質の対応する残基ペアもすべて 0.1 以上、あるいは -0.1 以下の場合、それぞれ \blacksquare あるいは \blacksquare で示している。また、定義された 5 つのドメインのうち 1 つについて、それに含まれる残基ペアを細線で囲んで示した。(b) (a) の三角マップによって定義されたドメインを立体構造上に色分けして示した。

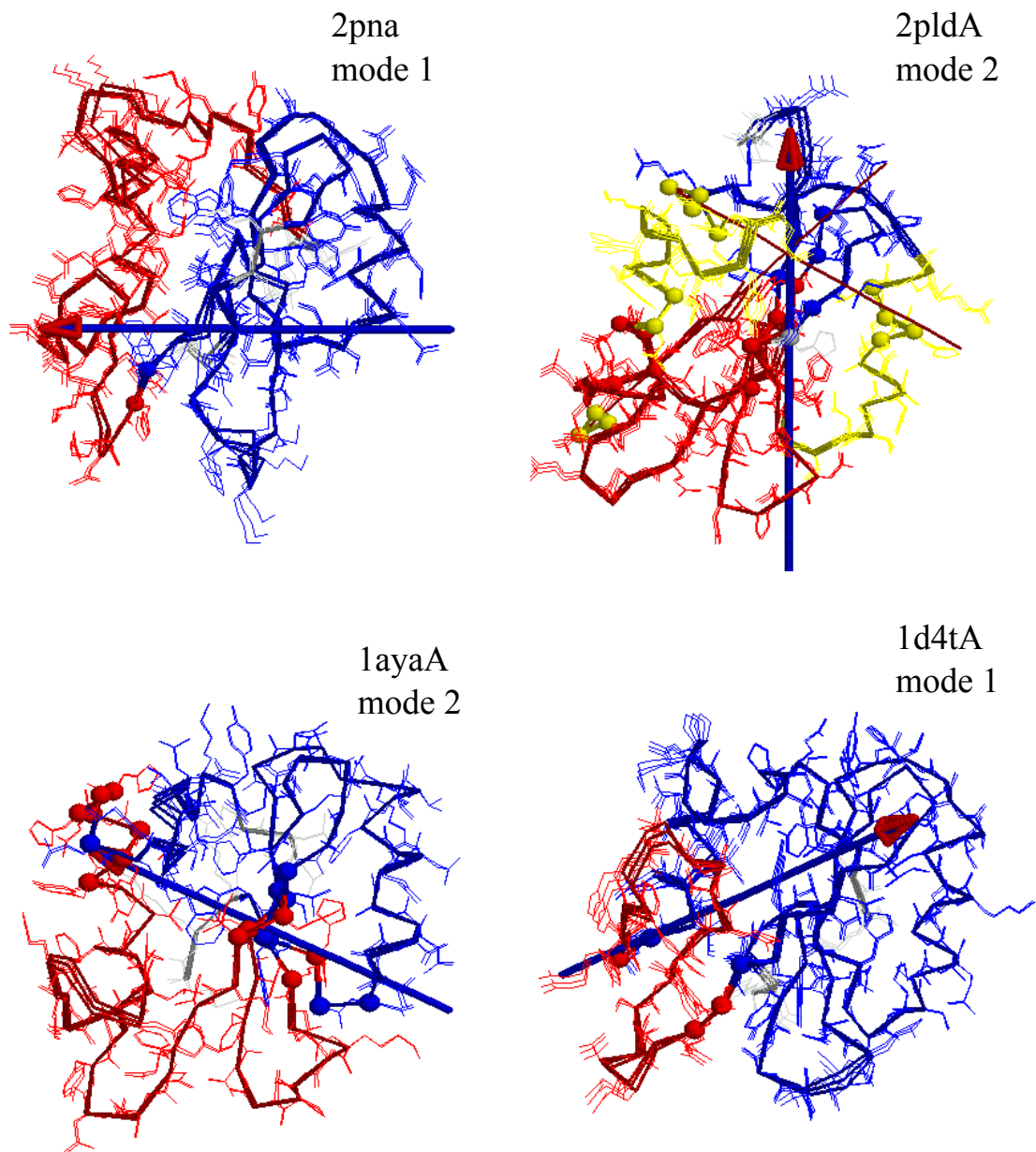


図 4-25 特定の基準振動モードに対して定義された動的ドメインの例

DynDom によって定義された動的ドメインを色分けして示した。ただし、灰色はドメインとして定義されなかった領域を意味している。矢印はそれらドメインの相対運動をらせん運動とみなしたときのらせん軸で、矢尻と軸との色が2つのドメインの色を示している。また、球は、蝶番領域のアミノ酸残基を表している。

4.5 Superfamily: Metalloproteases ("zincins") catalytic domain

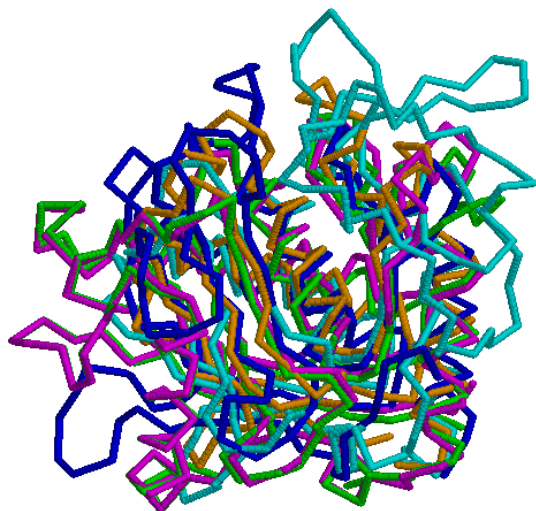


図 4-26 Metalloproteases ("zincins") catalytic domain superfamily の立体構造の重ね合わせ

PDB ID

1bqqM Membrane-type matrix metalloproteinase

1at1A Atrolysin c.

1budA Acutolysin a

1iab Astacin

1kuh Zinc protease

SCOP classification

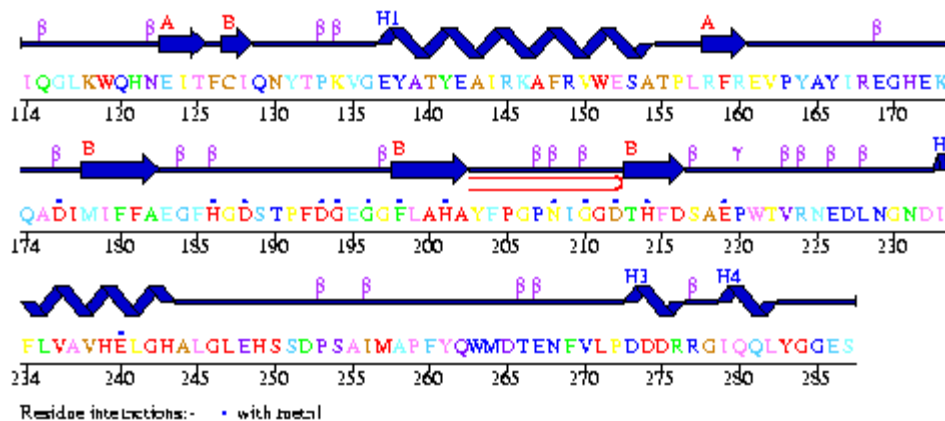
Class: Alpha and beta proteins (a+b)

Mainly antiparallel beta sheets (segregated alpha and beta regions)

Fold: Zincin-like

contains mixed beta sheet with connection over free side of the sheet

Sequence coloured by residue conservation:



Conservation colouring: Low 1 2 3 4 5 6 7 8 9 High

PROSITE: PS00142 - ZINC_PROTEASE

Val236 to Leu245: V A V H E L G H A L

図 4-27 アミノ酸配列、二次構造、残基の保存性、PROSITE パターン

1bqqM について PDBsum からコピーした。

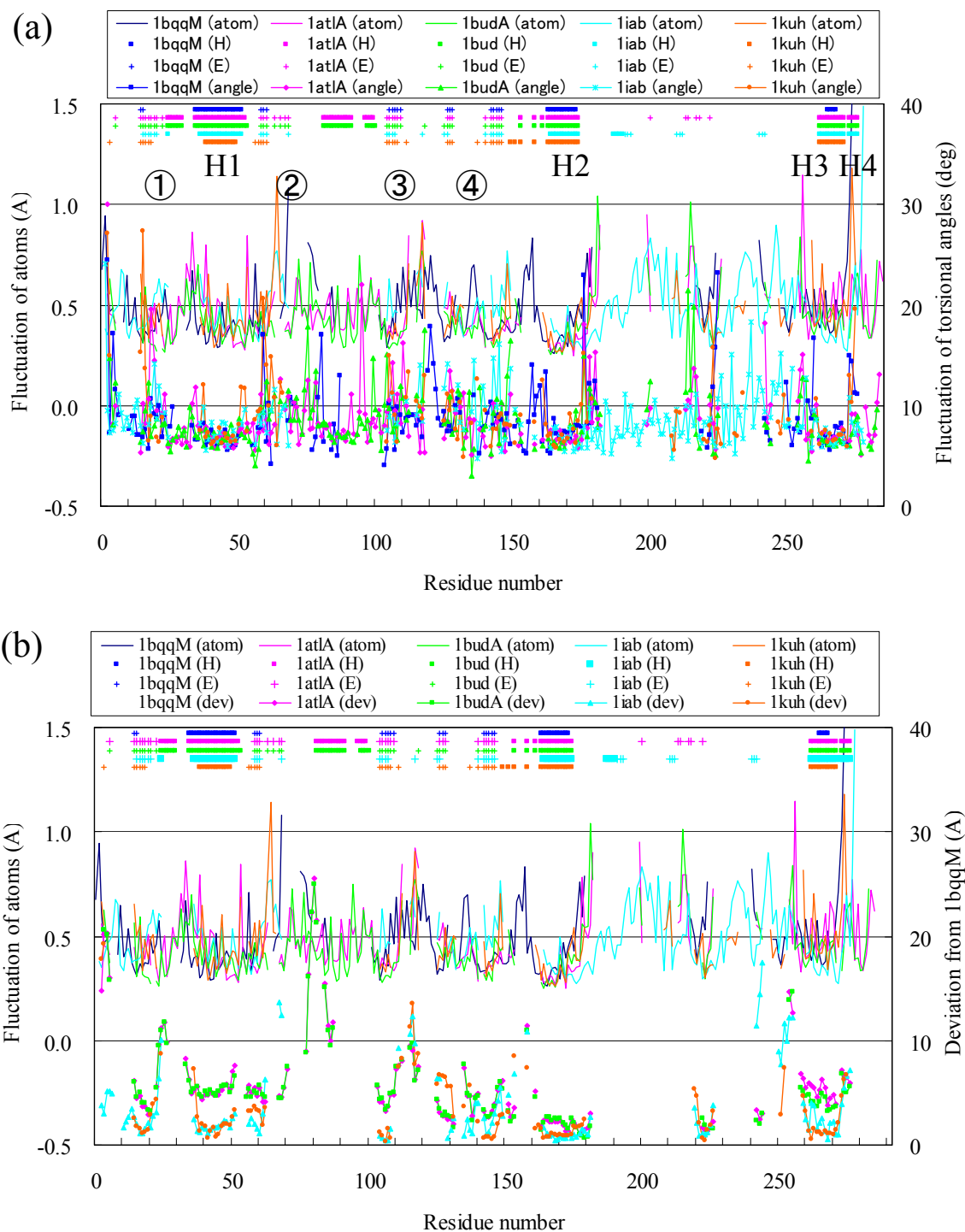


図 4-28 Metalloproteases ("zincins") catalytic domain superfamily の原子の位置のゆらぎ、二面角のゆらぎの比較。(a) 上のグラフは原子の位置のゆらぎを、下のグラフは二面角 ϕ のゆらぎを表している。(b) 上のグラフは(a)と同じく原子の位置のゆらぎであり、下のグラフは、立体構造を重ね合わせたときの 1bqqM からのずれを表している。それぞれのグラフの上部にある ■ と + はそれぞれ α ヘリックスと β ストランドの位置を示している。また、色は、凡例に示すように、それぞれ特定のタンパク質に対応して用いられている。1iab には、C 端側に大きな挿入があることに注意する。

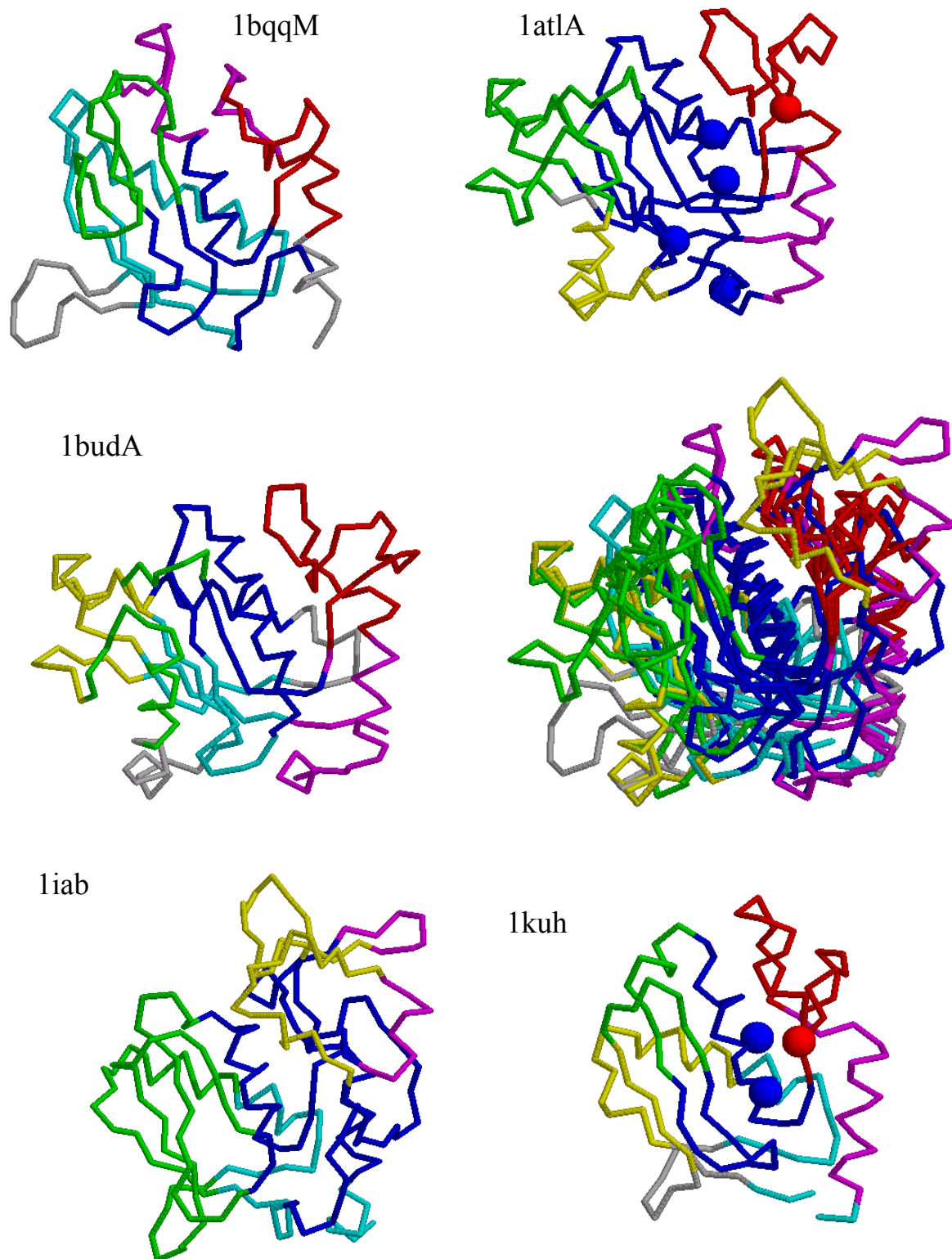
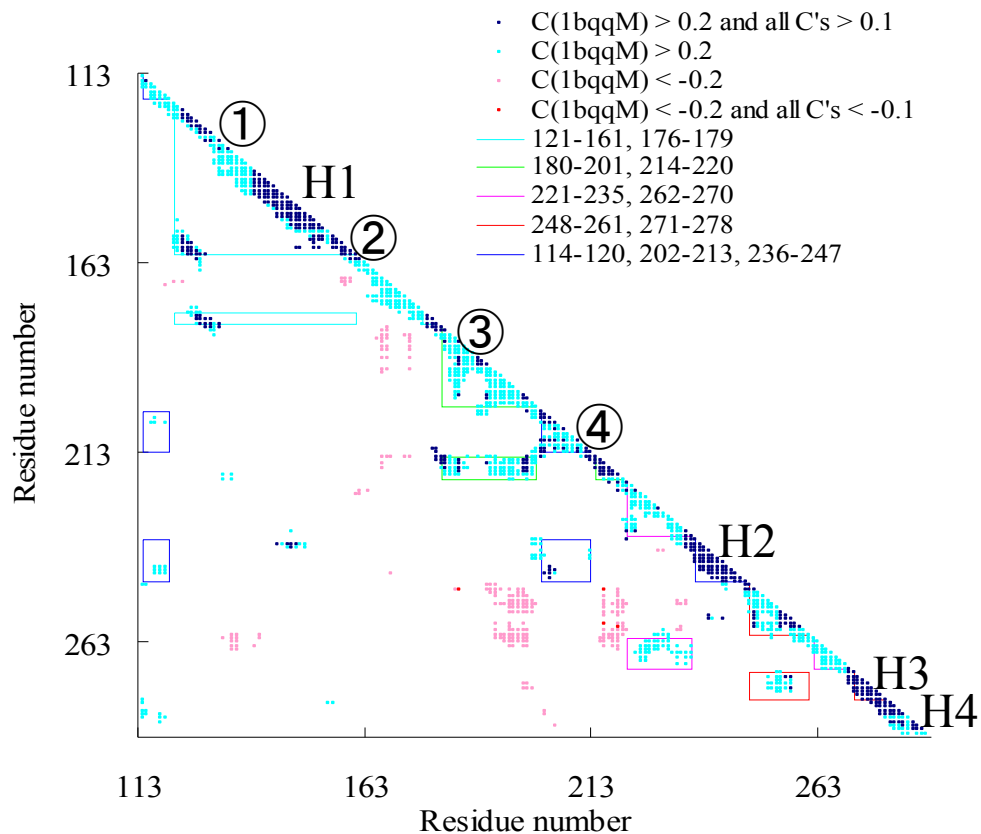


図 4-29 Metalloproteases ("zincins") catalytic domain superfamily の動的ドメイン
 $C\alpha$ のゆらぎの相関から定義した動的ドメインを、色分けして示した。中央右は、すべての構造を重ね合わせたもので、できるだけ同じ色のドメインが重なるように、色を指定した。



Domain1	121-161, 176-179	(cyan)
Domain2	180-201, 214-220	(green)
Domain3	221-235, 262-270	(magenta)
Domain4	248-261, 271-278	(red)
Domain5	114-120, 202-213, 236-247	(blue)

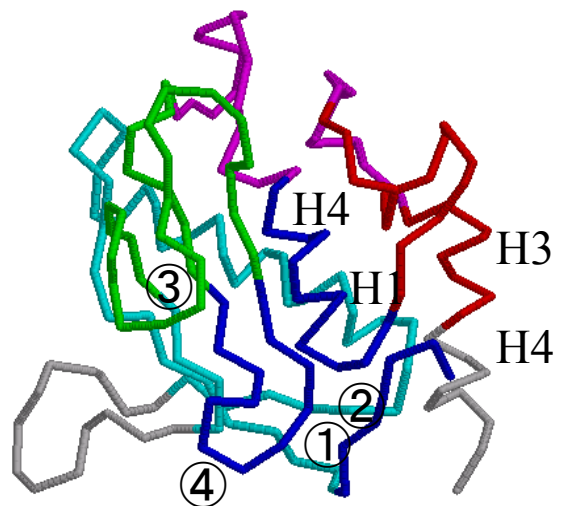


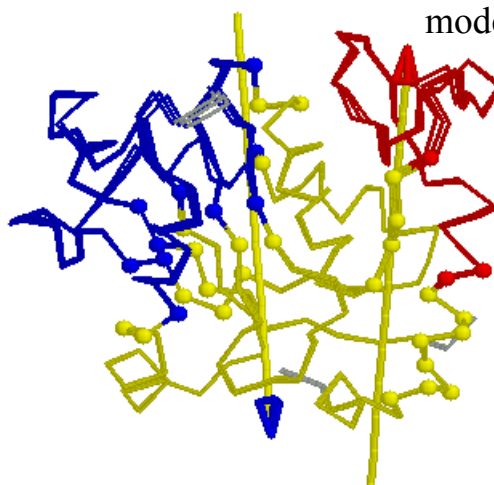
図 4-30 ゆらぎの相関マップと動的ドメイン

1bqqM を例に示している。(a) 1bqqM の相関係数 C が 0.2 以上の残基ペアを□で、-0.2 以下の残基ペアを□で示した。さらに、他のタンパク質の対応する残基ペアもすべて 0.1 以上、あるいは-0.1 以下の場合、それぞれ■あるいは■で示している。また、定義された 5 つのドメインについて、それらに含まれる残基ペアをそれぞれ細線で囲んで示した。(b) (a) の三角マップによって定義されたドメインを立体構造上に色分けして示した。

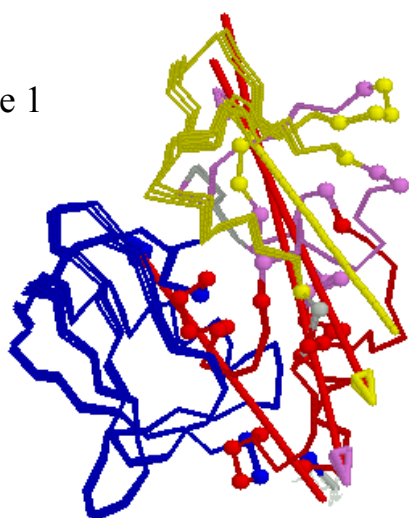
1bqqM
mode 2



1at1A
mode 1



1iab
mode 1



1kuh
mode 2

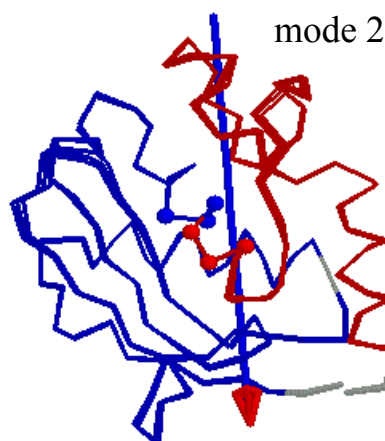


図 4-31 特定の基準振動モードに対して定義された動的ドメインの例

DynDom によって定義された動的ドメインを色分けして示した。ただし、灰色はドメインとして定義されなかった領域を意味している。矢印はそれらドメインの相対運動をらせん運動とみなしたときのらせん軸で、矢尻と軸との色が2つのドメインの色を示している。また、球は、蝶番領域のアミノ酸残基を表している。

付録 ドメインの定義 — クラスタリング法

タンパク質の残基数を N とする ($C\alpha$ 原子のみを考える)。

- (1) クラスタ数: $k=1$ とする
 - (2) $k=k+1$ 。もし、 $k > k_{\max}$ ならば (11) へ。
 - (3) すべての残基を k 個のクラスタのどれかにランダムに割り振る。
 - (4) 残基番号: $j=0$ とおく。
 - (5) $j=j+1$ 。 $j > N$ ならば (10) へ。
 - (6) クラスタ番号: $i=0$ とおく。
 - (7) $i=i+1$ 。 $i > k$ ならば (9) へ。
 - (8) クラスタ i に属するすべての残基 l のゆらぎの単位ベクトルについて、(5) で指定した残基 j のゆらぎの単位ベクトルとの内積を計算し、それらの平均を求める (注1)。 (7) へ。
 - (9) 最も平均値の高かったクラスタが、残基 j が現在属するクラスタと異なる場合、最も平均値の高かったクラスタへ残基 j を移す。 (5) へ。
 - (10) (9) で一度でも変更があった場合は (4) に戻り、(5) ~ (9) を再度実行する。
変更がない場合は、(2) へ。
- (11) 最適の k を選択する (注2)。
各クラスタに属する残基を一つのドメインと定義する。

(注1)
$$\frac{1}{m} \sum_{l=1}^m \vec{v}_l \cdot \vec{v}_j / |\vec{v}_l| |\vec{v}_j|$$

- (注2) 同じクラスタ内であって、ゆらぎベクトルの内積が与えられた閾値よりも大きい残基ペアの数と、互いに異なるクラスタ内であって、ゆらぎベクトルの内積がその閾値よりも小さい残基ペアの数の和が最大となるクラスタ数を最適のクラスタリングとする。
(閾値は 0.3 に設定)

- N. Go, T. Noguti and T. Nishikawa, *Dynamics of a small globular protein in terms of low-frequency vibrational modes*, Proc Natl Acad Sci USA **80** (1983), pp. 3696–3700.
- B. Brooks and M. Karplus, *Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor*, Proc Natl Acad Sci USA **80** (1983), pp. 6571–6575.
- H. Wako, *Dynamic structures of globular proteins with respect to correlative movements of residues calculated in the normal mode analysis*, J Protein Chem **8** (1989) 589-607.
- A. Kidera and N. Go, *Refinement of protein dynamic structure – normal mode refinement*, Proc Natl Acad Sci USA **87** (1990) 3718-3722.
- Y. Seno and N. Go, *Deoxymyoglobin studied by the conformational normal mode analysis. 1. Dynamics of globin and the heme-globin interaction*, J Mol Biol **216** (1990) 95-109.
- Y. Seno and N. Go, *Deoxymyoglobin studied by the conformational normal mode analysis. 2. The conformational change upon oxygenation*, J Mol Biol **216** (1990) 111-126.
- A. Kidera, K. Inaka, K. Matsushima and N. Go, *Normal mode refinement – crystallographic refinement of protein dynamic structure applied to human lysozyme*, **32** (1992) 315-319.
- D.A. Case, *Normal mode analysis of protein dynamics*, Curr Opin Struct Biol **4** (1994), pp. 285–290.
- A. Kitao, S. Hayward and N. Go, *Comparison of normal mode analyses on a small globular protein in dihedral angle space and Cartesian coordinate space*, Biophys Chem **52** (1994) 107-114.
- S. Nakamura and J. Doi, *Dynamics of transfer-RNAs analyzed by normal mode calculation*, Nucl Acid Res **22** (1994) 514-521.
- S. Hayward, A. Kitao and N. Go, *Harmonicity and anharmonicity in protein dynamics – A normal mode analysis and principal component analysis*, Proteins **23** (1995) 177-186.
- M.M. Tirion, *Large amplitude elastic motions in proteins from a single-parameter, atomic analysis*, Phys Rev Lett **77** (1996), pp. 1905–1908.
- H. Wako, M. Tachikawa and A. Ogawa, *A comparative study of dynamic structures between phage 434 Cro and repressor proteins by normal mode analysis*, Proteins **26** (1996) 72-80.
- I. Bahar, A.R. Atilgan and B. Erman, *Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential*, Fold Des **2** (1997), pp. 173–181.
- T. Haliloglu, I. Bahar and B. Erman, *Gaussian dynamics of folded proteins*, Phys Rev Lett **79** (1997), pp. 3090–3093.
- N. Kobayashi, T. Yamato and N. Go, *Mechanical property of a TIM-barrel protein*, Proteins **28** (1997) 109-116.
- I. Bahar, A.R. Atilgan, M.C. Demirel and B. Erman, *Vibrational dynamics of folded proteins: Significance of slow and fast motions in relation to function and stability*, Phys Rev Lett **80** (1998) 2733-2736.
- I. Bahar and R.L. Jernigan, *Vibrational dynamics of transfer RNAs: comparison of the free and synthetase-bound forms*, J Mol Biol **281** (1998) 871-884.
- I. Bahar, A. Wallqvist, D.G. Covell and R.L. Jernigan, *Correlation between native state hydrogen exchange and cooperative residue fluctuations from a simple model*, Biochemistry **37** (1998), pp. 1067–1075.
- K. Hinsen, *Analysis of domain motions by approximate normal mode calculations*, Proteins **33** (1998), pp. 417–429.
- H. Ishida, Y. Jochi and A. Kidera, *Dynamic structure of subtilisin-eglin c complex studied by normal mode analysis*, Proteins **32** (1998) 324-333.
- H. Sumikawa, E. Suzuki, K. Fukuhara, Y. Nakajima, K. Kamiya and H. Umeyama, *Dynamic structures of granulocyte colony-stimulating factor proteins studied by normal mode analysis: Two domain-type motions in low frequency modes*, Chem Pharm Bull **46** (1998) 1069-1077.
- D. Swigon, B.D. Coleman and I. Tobias, *The elastic rod model for DNA and its application to the tertiary structure of DNA minicircles in mononucleosomes*, Biophys J **74** (1998) 2515-2530.
- K. Hinsen, A. Thomas and M.J. Field, *Analysis of domain motions in large proteins*, Proteins **34** (1999)

- 369-382.
- A. Kitao and N. Go, *Investigating protein dynamics in collective coordinate space*, *Curr Opin Struct Biol* **9** (1999), pp. 164–169.
- A. Matsumoto and N. Go, *Dynamic properties of double-stranded DNA by normal mode analysis*, *J. Chem Phys* **110** (1999): 11070-11075.
- A. Matsumoto, M. Tomimoto and N. Go, *Dynamical structure of transfer RNA studied by normal mode analysis*, *Eur Biophys J* **28** (1999): 369-379.
- K. Soejima, Y. Kurihara, K. Kamiya and H. Umeyama, *Dynamic character of the complex of human blood coagulation factor VIIa with the extracellular domain of human tissue factor: a normal mode analysis*, *FEBS Lett* **463** (1999) 19-23.
- A. Hinsen, *The molecular modeling toolkit: A new approach to molecular simulations*, *J Comp Chem* **21** (2000) 79-85.
- O. Keskin, R.L. Jernigan and I. Bahar, *Proteins with similar architecture exhibit similar large-scale dynamics behavior*, *Biophys J* **78** (2000) 2093-2106.
- F. Tama, F.X. Gadea, O. Marques and Y.H. Sanejouand, *Building-block approach for determining low-frequency normal modes of macromolecules*, *Proteins* **41** (2000), pp. 1–7.
- A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin and I. Bahar, *Anisotropy of fluctuation dynamics of proteins with an elastic network model*, *Biophys J* **80** (2001), pp. 505–515.
- F. Tama and Y.H. Sanejouand, *Conformational change of proteins arising from normal mode calculations*, *Protein Eng* **14** (2001), pp. 1–6.
- M. Delarue and Y.-H. Sanejouand, *Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model*, *J Mol Biol* **320** (2002), pp. 1011–1024.
- P. Doruker, R.L. Jernigan and I. Bahar, *Dynamics of large proteins through hierarchical levels of coarse-grained structures*, *J Comput Chem* **23** (2002), pp. 119–127.
- W.G. Krebs, V. Alexandrov, C.A. Wilson, N. Echols, H. Yu and M. Gerstein, *Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic*, *Proteins* **48** (2002), pp. 682–695.
- S. Kundu, J.S. Melton, D.C. Sorensen and G.N. Phillips Jr., *Dynamics of proteins in crystals: comparison of experiment with simple models*, *Biophys J* **83** (2002), pp. 723–732.
- G. Li and Q. Cui, *A coarse-grained normal mode approach for macromolecules: an efficient implementation and application to Ca²⁺-ATPase*, *Biophys J* **83** (2002), pp. 2457–2474.
- D. Ming, Y.F. Kong, M.A. Lambert, Z. Huang and J.P. Ma, *How to describe protein motion without amino acid sequence and atomic coordinates*, *Proc Natl Acad Sci USA* **99** (2002), pp. 8620–8625.
- C. Micheletti, G. Lattanzi and A. Maritan, *Elastic properties of proteins: insight on the folding process and evolutionary selection of native structures*, *J Mol Biol* **321** (2002), pp. 909–921.
- H. Nojima, M. Takeda-Shitaka, Y. Kurihara, M. Adachi, S. Yoneda, K. Kamiya and H. Umeyama, *Dynamic characteristics of a peptide-binding groove of human HLA-A2 class I MHC molecules: Normal mode analysis of the antigen peptide-class I MHC complex*, *Chem Pharm Bull* **50** (2002): 1209-1214.
- F. Tama, W. Wriggers and C.L. Brooks III, *Exploring global distortions of biological macromolecules and assemblies from low-resolution structural information and elastic network theory*, *J Mol Biol* **321** (2002), pp. 297–305.
- M. Adachi, Y. Kurihara, H. Nojima, M. Takeda-Shitaka, K. Kamiya and H. Umeyama, *Interaction between the antigen and antibody is controlled by the constant domains: Normal mode dynamics of the HEL-HyHEL-10 complex*, *Protein Sci* **12** (2003): 2125-2131.
- N. Echols, D. Milburn and M. Gerstein, *MolMovDB: analysis and visualization of conformational change and structural flexibility*, *Nucleic Acids Res* **31** (2003), pp. 478–482.
- O. Miyashita, J.N. Onuchic and P.G. Wolynes, *Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins*, *Proc Natl Acad Sci USA* **100** (2003), pp. 12570–12575.
- S. Nakamura, M. Ikeguchi and K. Shimizu, *Dynamical analysis of tRNA(Gln)-GlnRS complex using*

- normal mode calculation*, Chem Phys Lett **372** (2003) 423-431.
- H. Nojima, M. Takeda-Shitaka, Y. Kurihara, K. Kamiya and H. Umeyama, *Dynamic flexibility of a peptide-binding groove of human HLA-DR1 class II MHC molecules: Normal mode analysis of the antigen peptide-class II MHC complex*, Chem Pharm Bull **51** (2003) 923-928.
- F. Tama, M. Valle, J. Frank and C.L. Brooks III, *Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy*, Proc Natl Acad Sci USA **100** (2003), pp. 9319–9323.
- W. Zheng and S. Doniach, *A comparative study of motor-protein motions by using a simple elastic-network model*, Proc Natl Acad Sci USA **100** (2003) 13253-13258.
- Z.W. Cao, Y. Xue, L.Y. Han, B. Xie, H. Zhou, C.J. Zheng, H.H. Lin and Y.Z. Chen, *MoViES: molecular vibrations evaluation server for analysis of fluctuational dynamics of proteins and nucleic acids*, Nucl Acid Res **32** (2004) W679-685.
- S.C. Chen and I. Bahar, *Mining frequent patterns in protein structures: a study of protease families*, Bioinformatics **20** (2004), pp. i77–i85.
- M. Delarue and P. Dumas, *On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models*, Proc Natl Acad Sci USA **101** (2004), pp. 6957–6962.
- M. Kaledin, A. Brown, A.L. Kaledin and J.M. Bowman, *Normal mode analysis using the driven molecular dynamics method. II. An application to biological macromolecules*, J Chem Phys **121** (2004), pp. 5646–5653.
- J.A. Kovacs, P. Chacon and R. Abagyan, *Predictions of protein flexibility: first-order measures*, Proteins **56** (2004), pp. 661–668.
- S. Kundu, D.C. Sorensen and G.N. Phillips Jr., *Automatic domain decomposition of proteins by a Gaussian network model*, Proteins **57** (2004), pp. 725–733.
- S. Kundu and R.L. Jernigan, *Molecular mechanism of domain swapping in proteins: an analysis of slower motions*, Biophys J **86** (2004), pp. 3846–3854.
- O. Kurkcuoglu, R.L. Jernigan and P. Doruker, *Mixed levels of coarse-graining of large proteins using elastic network model succeeds in extracting the slowest motions*, Polymers **45** (2004), pp. 649–657.
- I. Navizet, R. Lavery and R.L. Jernigan, *Myosin flexibility: structural domains and collective vibrations*, Proteins **54** (2004), pp. 384–393.
- A.J. Rader and I. Bahar, *Folding core predictions from network models of proteins*, Polymers **45** (2004), pp. 659–668.
- A.J. Rader, G. Anderson, B. Isin, H.G. Khorana, I. Bahar and J. Klein-Seetharaman, *Identification of core amino acids stabilizing rhodopsin*, Proc Natl Acad Sci USA **101** (2004), pp. 7246–7251.
- K. Suhre and Y.H. Sanejouand, *ELNémo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement*, Nucleic Acids Res **32** (2004), pp. W610–W614.
- F. Tama, O. Miyashita and C.L. Brooks III, *Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM*, J Struct Biol **147** (2004), pp. 315–326.
- F. Tama, O. Miyashita and C.L. Brooks III, *Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis*, J Mol Biol **337** (2004), pp. 985–999.
- R. Tatsumi, Y. Fukunishi and H. Nakamura, *A hybrid method of molecular dynamics and harmonic dynamics for docking of flexible ligand to flexible receptor*, J Comput Chem **25** (2004), pp. 1995–2005.
- N.A. Temiz and I. Bahar, *E. coli adenylate kinase dynamics: comparison of elastic network model modes with mode-coupling 15N-NMR relaxation data*, Proteins **57** (2004), pp. 468–480.
- H. Wako, M. Kato and S. Endo, *ProMode: a database of normal mode analyses on protein molecules with a full-atom model*, Bioinformatics **20** (2004), pp. 2035–2043.
- Y. Wang, A.J. Rader, I. Bahar and R.L. Jernigan, *Global ribosome motions revealed with elastic network model*, J Struct Biol **147** (2004), pp. 302–314.
- A. Van Wynsberghe, G. Li and Q. Cui, *Normal-mode analysis suggests protein flexibility modulation*

- throughout RNA polymerase's functional cycle*, *Biochemistry* **43** (2004), pp. 13083–13096.
- V. Alexandrov, U. Lehnert, N. Echols, D. Milburn, D. Engelman and M. Gerstein, *Normal modes for predicting protein motions: A comprehensive database assessment and associated Web tool*, *Protein Sci* **14** (2005) 633-643.
- I. Bahar and A.J. Raser, *Coarse-grained normal mode analysis in structural biology*, *Curr Opin Struc Biol* **15** (2005), pp. 1-7.
- C. Chennubhotla, A.J. Rader, L.W. Yang and I. Bahar, *Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies*, *Phys Biol* **2** (2005) S173-180.
- T. Halilou, O. Keskin, B. Ma and R. Nussinov, *How similar are protein folding and protein binding nuclei? Examination of vibrational motions of energy hot spots and conserved residues*, *Biophys J* **88** (2005), pp. 1552–1559.
- K. Hinsén, N. Reuter, J. Navaza, D.L. Stokes and J.J. Lacapère, *Normal mode-based fitting of atomic structure into electron density maps: application to sarcoplasmic reticulum Ca-ATPase*, *Biophys J* **88** (2005), pp. 818–827.
- S.M. Hollup, G. Salensminde and N. Reuter, *WEBnm@: a web application for normal mode analyses of proteins*, *BMC Bioinformatics* **6** (2005), pp. 1–8.
- A. Leo-Macias, P. Lopez-Romero, D. Lupyán, D. Zerbino and A.R. Ortiz, *An analysis of core deformations in protein superfamilies*, *Biophys J* **88** (2005), pp. 1291–1299.
- J. Ma, *Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes*, *Structure* **13** (2005), pp. 373–380.
- S. Maguid, S. Fernandez Alberti, L. Ferrelli and J. Echave, *Exploring the common dynamics of homologous proteins. Application to the globin family*, *Biophys J* **89** (2005), pp. 3–13.
- A. Taly, M. DeLauro, T. Grutter, M. Nilges, N. Le Novère, P.J. Corringer and J.P. Changeux, *Normal mode analysis suggests a quaternary twist model for the nicotinic receptor gating mechanism*, *Biophys J* **88** (2005), pp. 3954–3965.
- F. Tama and C.L. Brooks III, *Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis*, *J Mol Biol* **345** (2005), pp. 299–314.
- F. Tama, M. Feig, J. Liu, C.L. Brooks III and K.A. Taylor, *The requirement for mechanical coupling between head and S2 domains in smooth muscle myosin ATPase regulation and its implications for dimeric motor function*, *J Mol Biol* **345** (2005), pp. 837–854.
- M.M. Wang, R.T. Borchardt, R.L. Schowen and K. Kuczera, *Domain motions and the open-to-closed conformational transition of an enzyme: a normal mode analysis of S-adenosyl-L-homocysteine hydrolase*, *Biochemistry* **44** (2005), pp. 7228–7239.
- A.W. Van Wynsbeghe and Q. Cui, *Comparison of mode analyses at different resolutions applied to nucleic acid systems*, *Biophys J* **89** (2005) 2939-2949.
- L.W. Yang and I. Bahar, *Coupling between catalytic site and collective dynamics: A requirement for mechanochemical activity of enzymes*, *Structure* **13** (2005), pp. 893–904.
- L.W. Yang, X. Liu, C.J. Jursa, M. Holliman, A.J. Rader, H.A. Karimi and I. Bahar, *iGNM: a database of protein functional motions based on Gaussian network model*, *Bioinformatics* **21** (2005), pp. 2978–2987.
- W. Zheng and B.R. Brooks, *Normal-modes-based prediction of protein conformational changes guided by distance constraints*, *Biophys J* **88** (2005), pp. 3109–3117.
- W. Zheng and B.R. Brooks, *Identification of dynamical correlations within the myosin motor domain by the normal mode analysis of an elastic network model*, *J Mol Biol* **346** (2005), pp. 745–759.
- L.W. Yang, A.J. Rader, X. Liu, C.J. Jursa, S.C. Chen, H. Karimi and I. Bahar, *oGNM: Online computation of structural dynamics using the Gaussian network model*, *Nucl Acid Res* (2006) in press.
- C. Chennubhotla and I. Bahar, *Markov methods for hierarchical coarse-graining of large structures*, *J Comp. Biol* (2006) in press.

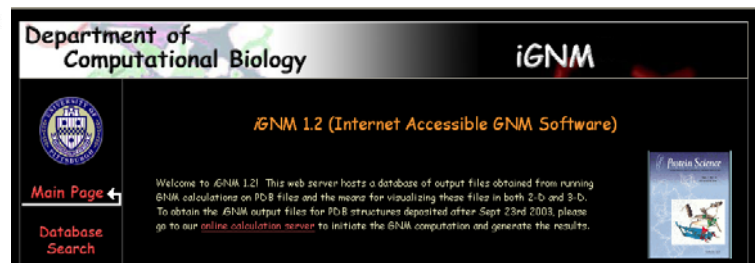
公開されている基準振動解析関連の Database あるいは Web サイト

elNemo	http://igs-server.cnrs-mrs.fr/elnemo/
iGNM	http://ignm.ccbb.pitt.edu/
MolMovDB	http://molmovdb.mbb.yale.edu/molmovdb/
MoViEs	http://ang.cz3.nus.edu.sg/cgi-bin/prog/norm.pl
oGNM	http://ignm.ccbb.pitt.edu/GNM_Online_Calculation.htm
ProMode	http://promode.socs.waseda.ac.jp/
WEBnm@	http://www.bioinfo.no/tools/normalmodes



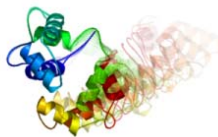
Welcome to elNémo !

elNémo is the Web-interface to *The Elastic Network Model*, a fast and simple tool to compute the low-frequency normal modes of a protein by Yves-Henri Sanejouand and co-workers (Tirion, 1996; Tama et al., 2000; Delarue and Sanejouand, 2002).



Database of Macromolecular Movements

with Associated Tools for Flexibility and Geometric Analysis



This describes the motions that occur in proteins and other macromolecules, particularly using movies. Associated with it are a variety of free software tools and servers for structural analysis.

[[Citation info](#) | [Old front page](#)]

Bioinformatics & Drug Design group [BIDD]

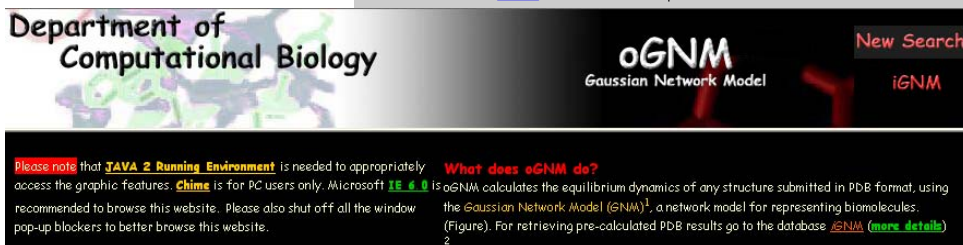
MoViES: Molecular Vibrations Evaluation Server

for Analysis of Vibrational Dynamics of Proteins and Nucleic Acids

This software only accepts protein or nucleic acid 3D structural entries. The 3D structure **MUST** be provided in PDB format. Five output files (out_hbnds, out_therm, out_modes, out_eigen and out_dist) will be sent to user via email.

Please click [here](#) for the explanation of INPUT and OUTPUT files.

Please click [here](#) for ADVANCED options.



Please note that **JAVA 2 Running Environment** is needed to appropriately access the graphic features. **Chime** is for PC users only. Microsoft **IE 6.0** is recommended to browse this website. Please also shut off all the window pop-up blockers to better browse this website.

What does oGNM do?
oGNM calculates the equilibrium dynamics of any structure submitted in PDB format, using the Gaussian Network Model (GNM)¹, a network model for representing biomolecules. (Figure). For retrieving pre-calculated PDB results go to the database [iGNM \(more details\)](#)².

Platform partners:

- LIB
- LIQ
- NTNU

Research:

WEBnm@
Web-server for Normal Mode Analysis of proteins - v2.0

Please submit a structure file in the pdb format and our server will then calculate the lowest frequency normal modes of your molecule. You will then be offered different types of calculations to analyse the modes that were calculated. Currently, the following analyses are available:

About normal modes calculations and analysis at bioinfo.no

The server is meant to provide users with simple and automated computation and analysis of low-frequency normal modes for proteins.

Improvements in *ProMode* (a database of normal mode analyses of proteins)

Hiroshi Wako¹
wako@waseda.jp

Masaki Kato²
kato@tsurumi.yokohama-cu.ac.jp

Shigeru Endo³
endo@sci.kitasato-u.ac.jp

¹ School of Social Sciences, Waseda University, 1-6-1 Nishi-Waseda, Shinjuku-ku, Tokyo 169-8050, Japan.

² Department of Science of Biological Supramolecular Systems, Graduate School of Integrated Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan.

³ Department of Physics, School of Science, Kitasato University, 1-15-1 Kitasato, Sagami-hara 228-8555, Japan.

Keywords: normal mode analysis, dynamic domain, database, computer simulation, molecular graphics

1 Introduction

ProMode (<http://promode.socs.waseda.ac.jp/>) is a database collecting the results from normal mode analyses (NMA) of various protein molecules [3]. Although NMA is based on the harmonic approximation, it has been shown by many studies that the results from NMA are not only reasonable qualitatively in most cases, but also can provide a proper description of the functionally important motions of the protein. The collection of the NMA results for various proteins is useful for comparative and statistical studies of the dynamic structures of proteins.

For further details of the method of NMA and the contents of *ProMode*, see the reference 3.

2 Improvements in *ProMode*

(1) In *ProMode*, we have collected more than 1,000 proteins, which have relatively low sequence similarity (< 50%) to each other and the sizes of which are less than 200 amino acid residues. We will collect data as many as possible in the future. Since a comparative study among homologous proteins is interesting, we are planning to perform NMA for proteins belonging to the same family. The calculations for larger-sized proteins are also planned.

(2) One of the distinctive features of *ProMode* is the animations of vibrating protein molecules which can be seen with a free plug-in, Chime (MDL Information Systems, Inc.). Figure 1 shows a screenshot of *ProMode*. A protein is displayed in the two windows: 3D animation in the left and 3D static images (the energy-minimum conformation and the two fluctuated ones superimposed on it) in the right. The mouse events for translation, rotation, and zoom on the one window affect the other simultaneously. The dynamic domains defined by DynDom are distinguished by different colors in both windows. The axes of screw motions between the domains also defined by DynDom are shown in the right window.

Since the input data file for the animation has no information about amino acid residues in Chime (which contains only atom types and their coordinates), while the static image contains such information, a user cannot manipulate the molecule displayed in the animation window. To overcome this inconvenience, we developed the manipulation dialog (superimposed in the upper left in Fig. 1) which can be used for selecting residues and/or atoms and then changing color and atomic representation to highlight them in the animation window by referring the information contained in the static-image window.

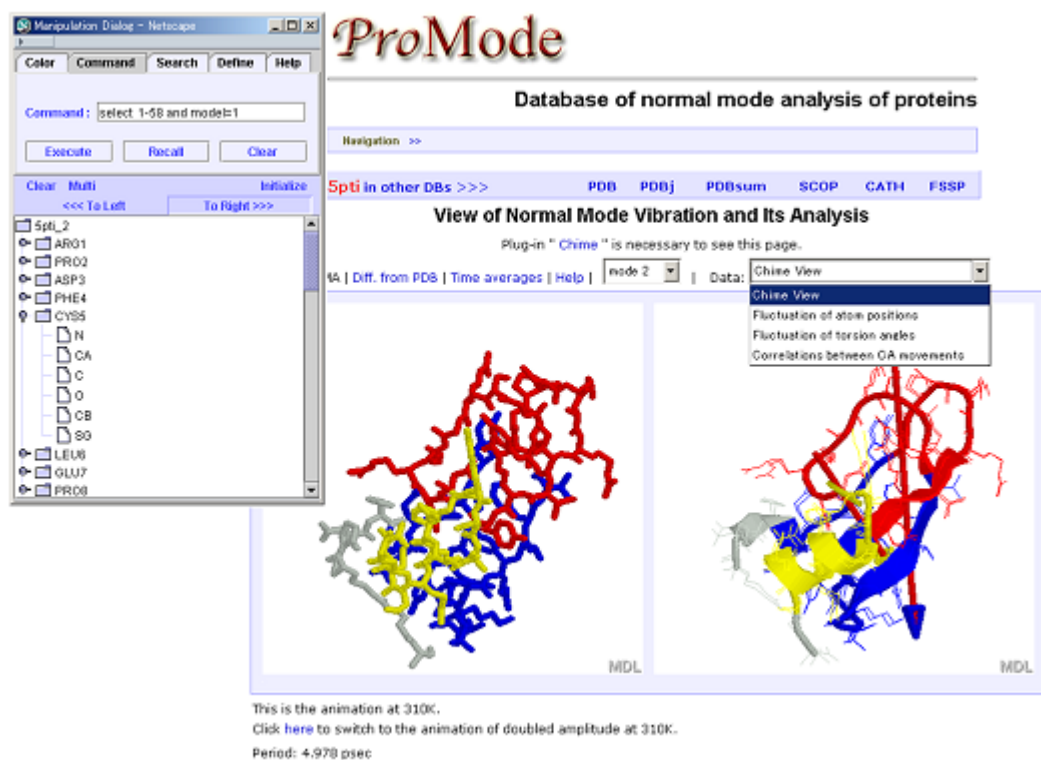


Figure 1: Screenshot of *ProMode*

3 Discussions

Recently, a couple of databases for NMA of proteins were developed [1,2]. They used a simplified molecular model and a simplified energy potential function (e.g., an elastic proteins model in which atoms representative of residues are connected with harmonic springs). In *ProMode*, however, a full-atom molecular model is used together with the program FEDER/2 [4]. This makes it possible to provide realistic molecular motions in full-atom details. We hope that researchers, especially who are not familiar to the simulation of protein dynamics and have seen only static images of protein structures, will get access to *ProMode*.

This work was financially supported by BIRD-JST (Institute for Bioinformatics Research and Development of Japan Science and Technology Corporation), and is running as a subgroup of one of the BIRD project PDBj (supervised by H. Nakamura, Osaka University; URL is <http://www.pdbj.org/>).

References

- [1] Chacon, P., Tama, F. and Wrighers, W., Mega-Dalton biomolecular captured from electron microscopy reconstructions, *J. Mol. Biol.*, 326:485-492, 2003.
- [2] Echols, N., Milburn, D. and Gerstein, M., MolNovDB: analysis and visualization of conformational change and structural flexibility, *Nucleic Acid Res.*, 31:478-482, 2003.
- [3] Wako, H. and Endo, S., *ProMode*: a database of normal mode analysis of proteins, *Genome Informatics*, 13:519-520, 2003.
- [4] Wako, H., Endo, S., Nagayama, K. and Go, N., FEDER/2: program for static and dynamic conformational energy analysis of macro-molecules in dihedral angle space, *Comp. Phys. Comm.*, 91:233-251, 1995.

Reflection of Knowledge Information in *ProMode* (A Database of Normal Mode Analyses on Proteins)

Hiroshi Wako,¹
wako@waseda.jp

Motohiro Otsuka,²
otsuka@mxs.nes.nec.co.jp

Yuhki Tomizawa,³
tomizawa@sciencehouse.co.jp

Masaki Kato,⁴
kato@tsurumi.yokohama-cu.ac.jp

Shigeru Endo⁵
endo@sci.kitasato-u.ac.jp

¹ School of Social Sciences, Waseda University, 1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050, Japan

² NEC Soft, Ltd., 1-18-6, Shinkiba, Koto-ku, Tokyo 136-8608, Japan

³ Sciencehouse, Inc., 1-4-17, Higashi-Tabata, Kita-ku, Tokyo 114-0013, Japan

⁴ Department of Science of Biological Supramolecular Systems, Graduate School of Integrated Science, Yokohama City University, 1-7-29, Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

⁵ Department of Physics, School of Science, Kitasato University, 1-15-1, Kitasato, Sagami-hara 228-8555, Japan

Keywords: normal mode analysis, protein conformation, SCOP, eF-site, Catalytic Residue Dataset, PROSITE,

1 Introduction

ProMode (<http://promode.socs.waseda.ac.jp>) is a database collecting results from normal mode analyses (NMA) of various protein molecules [5]. *ProMode* is designed so as to play an animation of a vibrating protein molecule with a free plug-in soft, Chime (MDL Information Systems, Inc.), and to graph out the results from NMA, such as root-mean-square fluctuations of atoms and dihedral angles. A user can learn the dynamic structure of the protein through the animation and the graphs.

In this paper we will report the recent improvements in *ProMode*.

2 Use of the SCOP classification in a protein search

The collection of the NMA results for various proteins in *ProMode* may be useful for comparative and statistical studies of dynamic structures of proteins. Although *ProMode* has collected the NMA data of more than 1,500 protein chains, it is still hard to carry out NMA for the entire PDB data, because of computation time required. In such a situation, for proteins not contained in *ProMode*, it is useful to provide some information about the relative proteins in *ProMode*. For this purpose we utilize the SCOP classification of protein conformations [4]. If a user searches *ProMode* for a protein with some PDB ID and no data is found there, the search tries to find an alternative protein in the same superfamily or family in *ProMode*. The user may obtain some information about the protein of interest through the alternative one.

Additionally, a user can also search the *ProMode* database along the SCOP tree structure like the original SCOP web page. At any level (class, fold, superfamily and family), the user is informed whether or not a given category, e.g. superfamily, has an entry data. With this information we can consider which proteins should be a target in the next NMA calculation. It may be also useful for a user to compare the NMA results of the proteins in a given superfamily or family.

3 Reflection of knowledge information in an animation

It has been shown by many studies that the results from NMA can provide a proper description of the functionally important motions of the proteins. Therefore, in *ProMode*, an animation can be manipulated by a control panel we developed to highlight residues in which a user is interested. As a result, we can observe

the relationship between the vibrational motion and the functionally important residues in the animation. However, information about, for example, which residues are functionally important, had to be provided by the user. It is inconvenient, especially for a user not familiar to the protein displayed. In the new version of *ProMode* described here, some knowledge information is prepared and can be reflected easily in the animation just by clicking on a button. The knowledge information from Catalytic Residue Dataset [1], PROSITE [2], and eF-site [3] is available at present. An example is shown in Figure 1. A user may be able to relate functionally important residues to the dynamic structure of the protein through the observation of the animation.

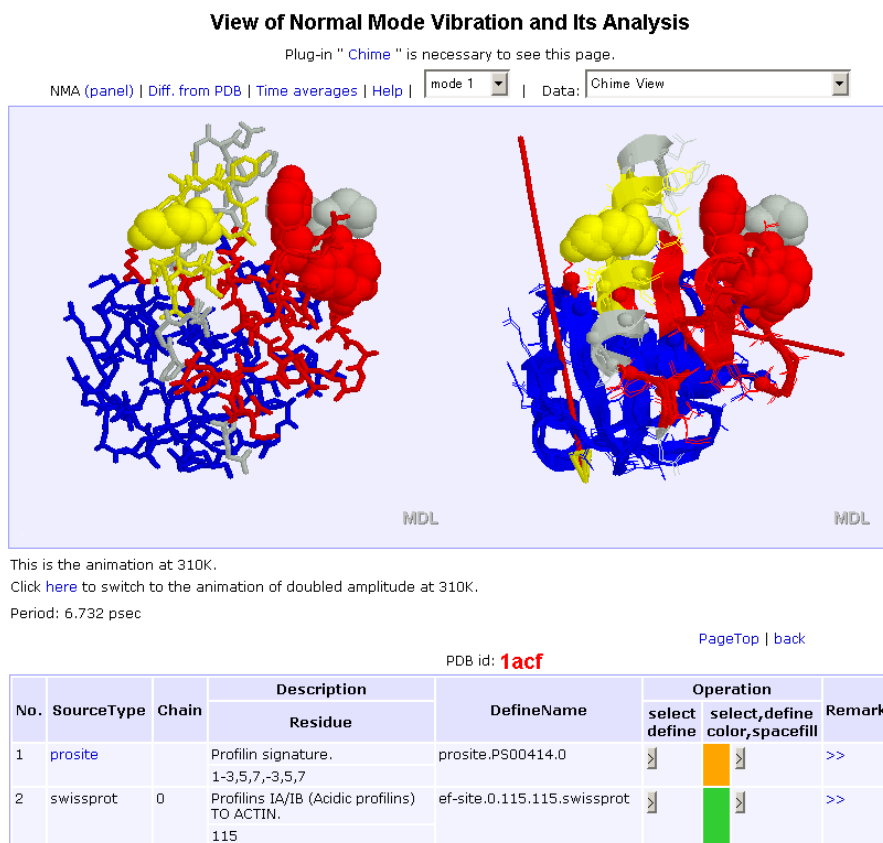


Figure 1. Example for the reflection of knowledge information in the Chime views (1acf: profilin I). The left Chime view is an animation and the right one is a static image displayed in cartoon. The residues are colored depending on dynamic domains they belong to. The arrows in the right view indicate a screw axis related to the screw motion between the dynamic domains. The knowledge information is given in the table below. The specified residues in the table (the PROSITE motif and the binding site to actin in this case) are displayed in a space-filling model in both Chime views.

References

- [1] Bartlett, G. J., Porter, C. T., Borkakoti, N. and Thornton, J.M., Analysis of catalytic residues in enzyme active sites, *J. Mol. Biol.*, **324**:105-121, 2002.
- [2] Hulo, N., Sigrist, C. J. et al., Recent improvements to the PROSITE database, *Nucleic Acids Res.*, **32 Database issue**:D134-137, 2004.
- [3] Kinoshita, K. and Nakamura, H., eF-site and PDBjViewer: database and viewer for protein functional sites, *Bioinformatics*, **20**:1329-1330, 2004.
- [4] Murzin A. G., Brenner S. E., Hubbard T. and Chothia C., SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.*, **247**:536-540, 1995.
- [5] Wako, H., Kato, M. and Endo, S., ProMode: a database of normal mode analyses on protein molecules with a full-atom model, *Bioinformatics*, **20**:2035-2043, 2004.

Comparisons between Dynamic Properties of Homologous Protein Structures in *ProMode* (Database of Normal Mode Analyses on Proteins)

Hiroshi Wako,¹
wako@waseda.jp

Motohiro Otsuka,²
otsuka@mxs.nes.nec.co.jp

Yuhki Tomizawa,³
tomizawa@sciencehouse.co.jp

Masaki Kato,⁴
kato@tsurumi.yokohama-cu.ac.jp

Shigeru Endo⁵
endo@sci.kitasato-u.ac.jp

¹ School of Social Sciences, Waseda University, 1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050, Japan

² NEC Soft, Ltd., 1-18-7, Shinkiba, Koto-ku, Tokyo 136-8627, Japan

³ Sciencehouse, Inc., 1-4-17, Higashi-Tabata, Kita-ku, Tokyo 114-0013, Japan

⁴ Department of Science of Biological Supramolecular Systems, Graduate School of Integrated Science, Yokohama City University, 1-7-29, Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

⁵ Department of Physics, School of Science, Kitasato University, 1-15-1, Kitasato, Sagami-hara 228-8555, Japan

Keywords: normal mode analysis, protein conformation, dynamic structures of homologous proteins, structural alignment

1 Introduction

ProMode (<http://promode.socs.waseda.ac.jp>) is a database collecting results from normal mode analyses (NMA) of various protein molecules [3]. *ProMode* is designed so as to play an animation of a vibrating protein molecule with a free plug-in soft, Chime (MDL Information Systems, Inc.) and PDBjViewer [1], and to graph out the results from the NMA, such as root-mean-square fluctuations of atoms and dihedral angles. Furthermore, a user can learn the dynamic structure of the protein not only through the animation and the graphs, but also by manipulating the animation to reflect some knowledge information (e.g., information from Prosite, eF-site, and Catalytic Residue Dataset). A search for a target protein along the tree of the SCOP classification is also useful [4]. In this paper we will report the recent improvements in *ProMode*.

2 Comparisons between homologous proteins

The collection of the NMA results for various proteins in *ProMode* may be useful for comparative and statistical studies of dynamic structures of proteins. For example, a comparative study of dynamic structures by the NMA was carried out for the two DNA binding proteins, Cro and the amino-terminal domain of the repressor of bacteriophage 434 that adopt very similar three-dimensional structures to each other [5]. For bovine pancreatic trypsin inhibitor (BPTI), the NMA results for the mutants with one to three amino acid substitutions were compared with those for the wild-type BPTI [3].

Although a structural comparison of proteins has been performed mainly from a static or geometrical point of view, the above studies demonstrated that the comparison from a dynamic point of view, using the NMA, is useful and convenient to explore a difference that is difficult to find only from a geometrical point of view, especially for proteins very similar in structure, because the NMA is sensitive to the changes in the interactions between atoms and between amino acid residues.

Such a kind of information is now available in *ProMode*. For the superfamily in which NMA data for more than one protein are available in *ProMode*, firstly, the member proteins are structurally aligned by the program ASH [2], and then, the fluctuations of atom positions or torsion angles for all the member proteins are plotted against the aligned sequences in the same graph (see Figure 1). In some cases, although there

exists a large insertion/deletion region, it is found that their fluctuation for the common regions are very similar to each other. On the other hand, the differences between the member proteins provide information about differences in the interactions arising from the differences in the amino acid sequences. The analyses of such a kind of information are useful for understanding the three-dimensional structures of the proteins.

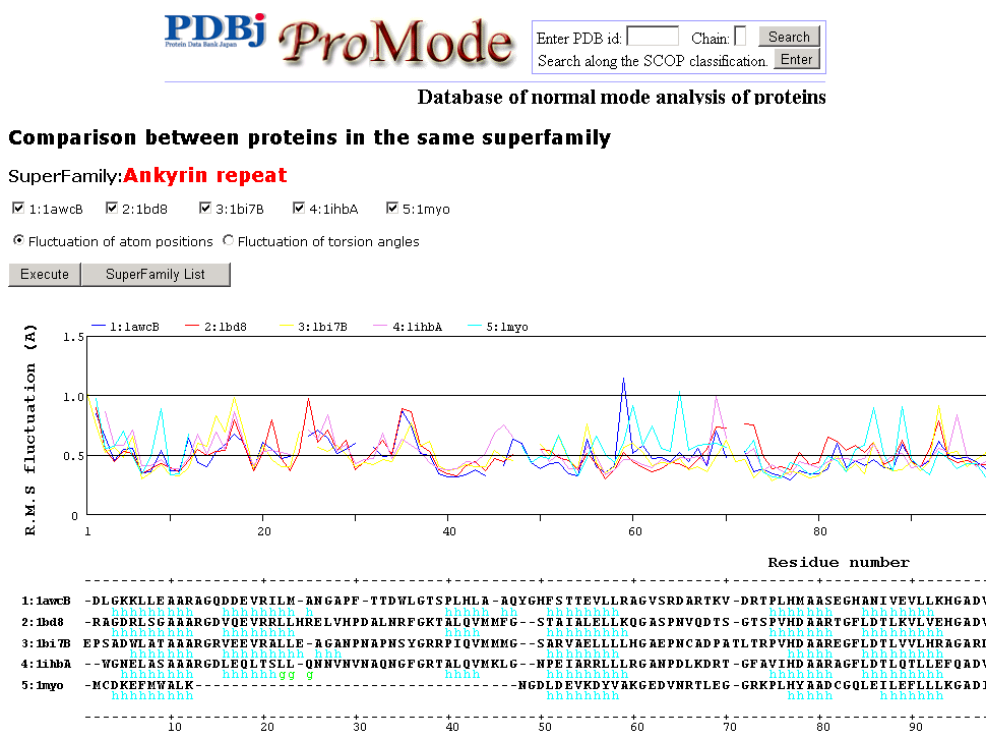


Figure 1. “Comparison between proteins in the same superfamily” page in *ProMode*. Fluctuations of atom positions of five proteins in the same superfamily are plotted against residue numbers. The structurally aligned amino acid sequences of the five proteins by the program ASH [2] are given below.

Acknowledgement

This work was supported by BIRD-JST (Institute for Bioinformatics Research and Development of Japan Science and Technology Corporation).

References

- [1] Kinoshita, K., and Nakamura, H., eF-site and PDBjViewer: database and viewer for protein functional site, *Bioinformatics* 20:1329-1330, 2004.
- [2] Toh, H., Introduction of a distance cut-off approximation into structural alignment with the double dynamic programming algorithm. *CABIOS* 13:387-396, 1997.
- [3] Wako, H., Kato, M. and Endo, S., *ProMode*: a database of normal mode analyses on protein molecules with a full-atom model, *Bioinformatics*, **20**:2035-2043, 2004.
- [4] Wako, H. Otsuka, M., Tomizawa, Y., Kato, M. and Endo, S., Reflection of knowledge information in *ProMode* (a database of normal mode analyses on proteins), *Genome Informatics 2004, Poster and Software Demonstrations*, S03, 2004.
- [5] Wako, H., Tachikawa, M. and Ogawa, A., A comparative study of dynamic structures between phage 434 Cro and repressor proteins by normal mode analysis. *Proteins* 26:72-80, 1996.



ProMode: a database of normal mode analyses on protein molecules with a full-atom model

Hiroshi Wako^{1,*}, Masaki Kato² and Shigeru Endo³

¹School of Social Sciences, Waseda University, Tokyo 169-8050, Japan, ²Department of Science of Biological Supramolecular Systems, Graduate School of Integrated Science, Yokohama City University, Yokohama 230-0045, Japan and ³Department of Physics, School of Science, Kitasato University, Sagami-hara 228-8555, Japan

Received on September 19, 2003; revised on February 10, 2004; accepted on February 27, 2004
Advance Access publication April 1, 2004

ABSTRACT

Motivation: Although information from protein dynamics simulation is important to understand principles of architecture of a protein structure and its function, simulations such as molecular dynamics and Monte Carlo are very CPU-intensive. Although the ability of normal mode analysis (NMA) is limited because of the need for a harmonic approximation on which NMA is based, NMA is adequate to carry out routine analyses on many proteins to compute aspects of the collective motions essential to protein dynamics and function. Furthermore, it is hoped that realistic animations of the protein dynamics can be observed easily without expensive software and hardware, and that the dynamic properties for various proteins can be compared with each other.

Results: *ProMode*, a database collecting NMA results on protein molecules, was constructed. The NMA calculations are performed with a full-atom model, by using dihedral angles as independent variables, faster and more efficiently than the calculations using Cartesian coordinates. In *ProMode*, an animation of the normal mode vibration is played with a free plug-in, Chime (MDL Information Systems, Inc.). With the full-atom model, the realistic three-dimensional motions at an atomic level are displayed with Chime. The dynamic domains and their mutual screw motions defined from the NMA results are also displayed. Properties for each normal mode vibration and their time averages, e.g. fluctuations of atom positions, fluctuations of dihedral angles and correlations between the atomic motions, are also presented graphically for characterizing the collective motions in more detail.

Availability: <http://promode.socs.waseda.ac.jp>

Contact: wako@waseda.jp

INTRODUCTION

Protein dynamics simulation is important to understand principles of architecture of a protein structure and its function. We have several conventional methods to simulate the protein

dynamics by computer, e.g. normal mode analysis (NMA), Monte Carlo simulation (MC), and molecular dynamics simulation (MD) (Brooks *et al.*, 1988). MC and MD are more reliable than NMA, because NMA is based on the harmonic approximation (i.e. MC and MD can cover much wider area of the conformational space of the protein than NMA). However, it has been shown by many studies that the results from NMA are not only qualitatively reasonable in most cases, compared to those from MC and MD, but also can provide a proper description of the functionally important motions of the protein (Go *et al.*, 1983; Brooks and Karplus, 1983; Levitt *et al.*, 1985; Gibrat and Go, 1990; Horiuchi and Go, 1991; Marques and Sanejouand, 1995; Hayward and Go, 1995; Wako *et al.*, 1996; Thomas *et al.*, 1996a,b; Hinsen, 1998; Kitao and Go, 1999; Miller and Agard, 1999; Kikuchi *et al.*, 2000; Berendsen and Hayward, 2000; Chacon *et al.*, 2003). Fortunately, NMA is much less time consuming in computation and can be performed in more systematic ways than MC and MD. Consequently, NMA is more adequate to carry out dynamics calculations routinely for many proteins and then, to describe the protein dynamics concisely.

ProMode is a database collecting the results from NMA carried out for many proteins. Although we should be careful in interpreting the results from NMA, owing to the approximation described above, the data presented in *ProMode* provide fruitful information about protein dynamics, especially for large-scale collective motions. The collection of the NMA results for various proteins is useful for comparative and statistical studies of the dynamic structures of proteins.

Another distinctive feature of *ProMode* is an animation of a vibrating protein played with a free plug-in, Chime (MDL Information Systems, Inc. <http://www.mdli.com/>). It is possible for a user with Chime not only to observe a concrete three-dimensional (3D) image in atomic detail about protein dynamics, but also to manipulate the animated protein molecule. The several kinds of time-averaged properties to characterize the motions are also provided graphically.

*To whom correspondence should be addressed.

Recently, a couple of databases for NMA of proteins were developed (Krebs *et al.*, 2002; Chacon *et al.*, 2003; Echols *et al.*, 2003). They used a simplified molecular model and an energy potential function [e.g. an elastic protein model in which atoms representative of residues are connected with harmonic springs (Hinsen, 2000)]. In *ProMode*, however, a full-atom model is used. We used dihedral angles rather than Cartesian coordinates of atoms as independent variables. Since the number of variables can be reduced to one-eighth of the number of Cartesian coordinates, it is possible to perform NMA with the full-atom model even for relatively larger proteins. Consequently, *ProMode* can provide realistic molecular motions in full-atomic detail.

METHODS

Regularization

The most time-consuming part of NMA is regularization of the Protein Data Bank (PDB; Berman *et al.*, 2000) data. We used a program, FEDER, developed by Go's group for this purpose (Noguti and Go, 1983a,b; Wako and Go, 1987; Wako *et al.*, 1995).

FEDER carries out a conformational energy minimization of a protein rapidly and efficiently with dihedral angles as independent variables (i.e. bond lengths and angles are fixed to their standard values). The number of the dihedral angles is about one-eighth of the number of atomic coordinates, and about $6N$ for a protein with N residues (compare to a simplified model, in which each residue is represented by one atom; in such a model, the number of variables is $3N$). Using the algorithm for fast calculation of the first and second derivatives of a conformational energy function with respect to the dihedral angles, FEDER can precisely obtain a minimum point in the conformational energy space. It is strictly required in NMA that the second derivative matrix of the conformational energy function at the minimum should be positive definite. Since our calculations always satisfied this requirement, all normal modes could be obtained for any proteins.

FEDER generates various conformations with dihedral angles as independent variables. It is necessary to obtain a proper set of dihedral angles that generates a conformation not only at an energy minimum, but also very close to the PDB data whose bond lengths and bond angles are slightly deviated, in general, from their standard values used in FEDER, before NMA is performed. This procedure is referred to as regularization. For this requirement, we carried out the restrained energy-minimization with the following objective function.

$$F = w \sum_{i,j} (r_{ij} - r_{ij}^X)^2 + pE, \quad (1)$$

where r_{ij} and r_{ij}^X are the distances between atoms i and j in a calculated conformation and in the PDB conformation, respectively. E is a conformational energy function containing

van der Waals, electrostatic, hydrogen bonding and torsional potentials (Wako *et al.*, 1995).

If the objective function F is minimized, the first term works as a geometrical constraint to hold the molecular conformation close to the PDB data and the second term to lower the conformational energy. w and p are weight factors for the geometrical constraint and conformational energy terms, respectively. The minimization of the objective function is carried out, as p is set to zero at first. After the minimization converges, p and w are set to one and a large value, respectively, and then the minimization is resumed. The minimization is repeated, as w is decreased, step by step. Finally, by setting $w = 0$, we can obtain the conformation located at a local minimum and very close to the PDB conformation. NMA is applied to such an energy-minimum conformation.

Normal mode analysis

Normal mode analysis is based on the assumption that a conformational energy surface can be characterized by a parabolic approximation at a minimum. One of the merits of NMA is that the derivation of an analytical solution to the equation of motion subject to this assumption is well established (Levitt *et al.*, 1985). NMA divides the atomic motion into vibrational components called normal modes. A variety of atomic motions of a hundred normal modes (the number of normal modes is around six times the number of residues) in a protein, characterizes the dynamic structure of the protein. A time-averaged property of the motion, which is a superposition of all the normal modes, also provides important information on the protein dynamics (Go *et al.*, 1983; Brooks and Karplus, 1983). The time-averaged properties of the motion depend only on the amplitudes of the normal mode vibrations, and the amplitude depends on the temperature. Calculations are carried out at 310 K (37°C) in *ProMode*.

The properties provided in *ProMode* are described below.

RESULTS

Proteins

We have collected NMA data mainly for proteins with less than 200 residues and with relatively low sequence similarity (<50%) to each other at present. The data for more than 1000 proteins have been collected already. The most time-consuming part of the calculations was the regularization. In Table 1, an example of the regularization schedule for human lysozyme (the number of residues is 130) is shown. The CPU time depends not only on protein size, but also on protein conformation. Especially, when the regularized conformation was largely deviated from the PDB one (see below for the reason why such a deviation occurred), the CPU time was much longer than the average one. However, once the regularized conformation was obtained for a given protein, the NMA calculation could be performed in short time [remember

Table 1. An example of regularization schedule^a

Range ^b	Weighting factors ^c (kcal/mol/Å ² for w)	Convergence ^d (kcal/mol)	CPU ^e (s)	No. of iterations ^f
1	$p = 0$	10^{-3}	12.4 ± 2.9	25 ± 0.7
2	$p = 0$	10^{-3}	18.0 ± 4.5	29 ± 2.4
3	$p = 0$	10^{-3}	16.1 ± 1.4	21 ± 1.0
4	$p = 0$	10^{-3}	11.1 ± 2.6	20 ± 1.1
5	$p = 0$	10^{-3}	9.4 ± 0.5	18 ± 0.9
6	$p = 0$	10^{-3}	9.0 ± 0.9	18 ± 1.6
10	$p = 0$	10^{-3}	9.1 ± 0.3	17 ± 0.6
14	$p = 0$	10^{-3}	9.6 ± 1.2	17 ± 0.6
18	$p = 0$	10^{-3}	8.8 ± 0.8	15 ± 1.3
30	$p = 0$	10^{-3}	9.8 ± 0.8	15 ± 1.3
60	$p = 0$	10^{-3}	12.1 ± 0.9	15 ± 0.1
90	$p = 0$	10^{-3}	13.6 ± 1.2	14 ± 1.1
120	$p = 0$	10^{-3}	16.5 ± 6.5	16 ± 4.6
150	$p = 0$	10^{-3}	17.5 ± 7.4	14 ± 4.7
200	$p = 0$	10^{-3}	18.7 ± 4.5	13 ± 0.7
300	$p = 0$	10^{-3}	27.0 ± 8.5	17 ± 2.8
All	$p = 0$	10^{-3}	21.5 ± 5.7	13 ± 1.5
All	$p = 1, w = 100$	10^{-10}	482.2 ± 181.2	119 ± 40.7
All	$p = 1, w = 10$	10^{-10}	189.2 ± 89.8	48 ± 19.4
All	$p = 1, w = 1$	10^{-10}	254.3 ± 91.3	66 ± 23.0
All	$p = 1, w = 0.1$	10^{-10}	419.4 ± 115.2	109 ± 15.3
All	$p = 1, w = 0.01$	10^{-10}	439.9 ± 185.9	114 ± 38.2
All	$p = 1, w = 0$	10^{-12}	141.1 ± 84.0	57 ± 27.5

Mean total CPU time: 2166 s

^aThe data of the regularization for wild type and 19 mutants of human lysozyme are averaged.

^bThe range within which two atoms are possible to interact with each other in the calculation is shown. The distance between the two dihedral angles along the chain [defined in the paper by Wako *et al.* (1995)] is used to specify the range. The range is gradually expanded to avoid being trapped at a local minimum.

^cSee the objective function (1).

^dThe convergence criterion is given with an energy difference between the two conformations subsequently generated in the regularization.

^eThe calculation was performed on the Pentium 4 processor (2.13 GHz) with 512 MB memory. The mean and SD over 20 human lysozymes are given.

^fThe mean and SD of iteration times in the optimization of the objective function (1) over 20 human lysozymes are given.

that the number of variables (dihedral angles) is about $6N$ for a protein with N residues].

Owing to the computational difficulty in FEDER, any ligands, DNA and water molecules are omitted in our computations. For the proteins comprising two or more chains, each chain is treated independently. The calculation for a multimer system will be a future work.

The conformations of some chains were largely deviated from the PDB data after the regularization, mainly because the hetero molecules or other accompanying chains were ignored. The root-mean-square (RMS) deviations of about 60, 25 and 15% of the calculated proteins from the PDB conformations are less than 2, 2–4 Å and greater than 4 Å, respectively. In *ProMode*, the data of the proteins with the RMS deviation >2.0 Å are given with a warning message. It is interesting, however, to study which parts of such proteins were changed by the regularization. We are planning to study this problem in the near future. (For the problems in the regularization, see also Brunger 1993).

We will collect data on as many proteins as possible, including larger sized ones. Since a comparative study among

homologous proteins is interesting (Wako *et al.*, 1996), we will also add NMA results for such proteins to *ProMode*.

Properties presented

ProMode provides several properties characterizing the dynamics of the proteins. The properties that can be calculated in NMA, however, have a variety, and the total sizes of the resulting files are too large to archive in general. Consequently, the data provided in *ProMode* have to be limited. What follows are the properties provided in *ProMode* at present.

- (i) Time-averaged properties.
 - (a) *Positional fluctuations of atoms*: The absolute values of the normal mode displacement vectors are averaged over all atoms or over backbone atoms in each residue, respectively, and plotted against residue numbers (Fig. 1a).
 - (b) *Fluctuations of dihedral angles*: The data on the dihedral angles, ϕ , φ and χ_1 , are plotted against residue numbers (Fig. 1b).

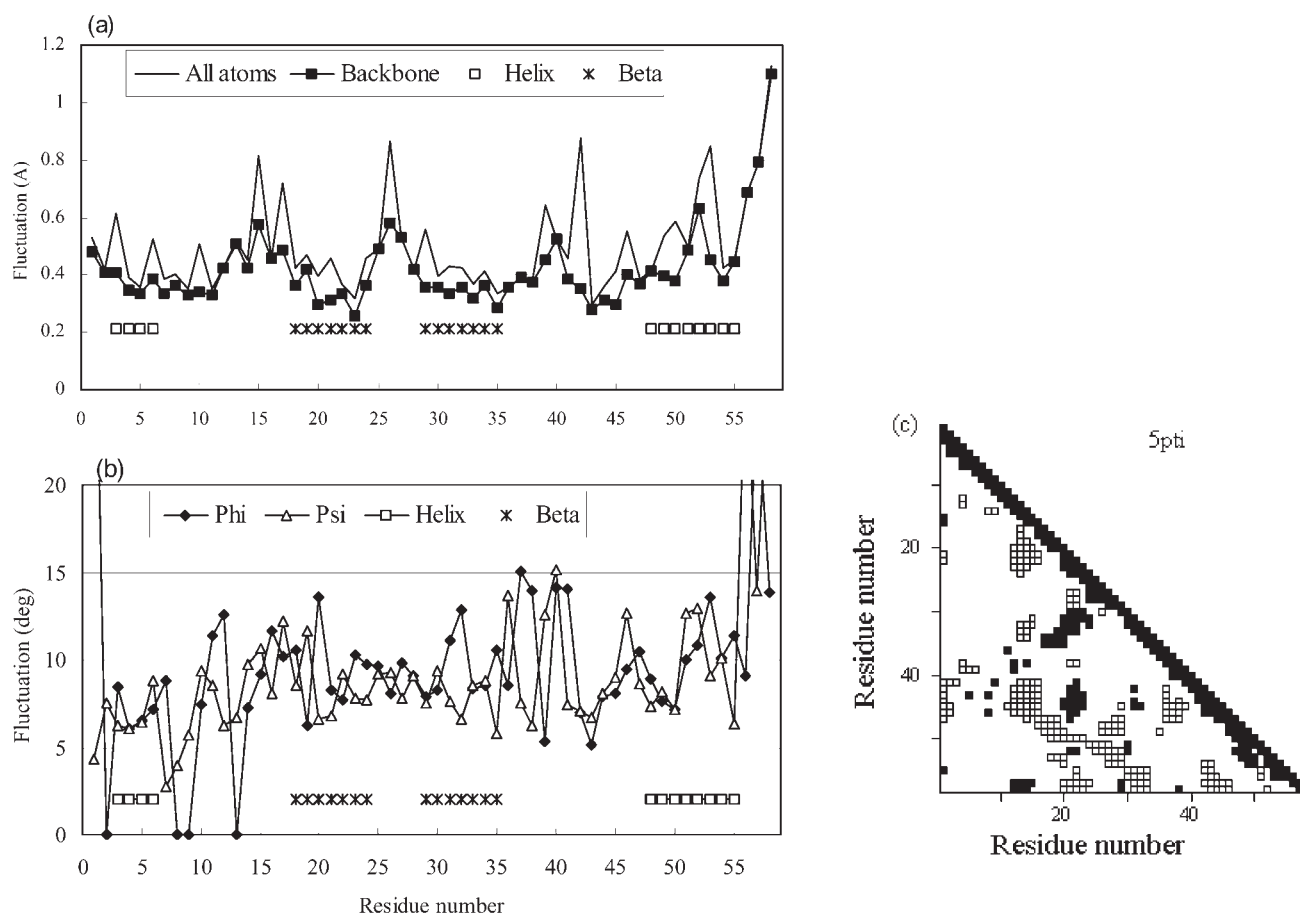


Fig. 1. Examples of time-averaged properties represented graphically in *ProMode* (these figures are not a screenshot of *ProMode*, but reproduced with some modifications). (a) Positional fluctuations of atoms, which are absolute values of normal mode displacement vectors averaged over all atoms and averaged over backbone atoms in each residue, respectively. (b) Fluctuations of dihedral angles, ϕ and ψ . (c) Correlations between the positional fluctuations of atoms, which are defined as normalized dot products of the normal mode displacement vectors of $C\alpha$ atoms. The closed and open squares in the triangle map indicate that the correlations of these pair of residues are greater than 0.2 and less than -0.2 , respectively. The protein shown here is BPTI (PDB ID=5pti).

- (c) *Correlations between the positional fluctuations of atoms:* The normalized dot products of the normal mode displacement vectors of $C\alpha$ atoms are presented in a triangle map (Fig. 1c).
- (ii) Properties of each normal mode vibration (only for the 20 lowest-frequency modes).
- (a) *Positional fluctuations of atoms:* The absolute values of the normal mode displacement vectors are averaged over all atoms or over backbone atoms in each residue, respectively, and plotted against residue numbers.
- (b) *Fluctuation of dihedral angles:* The data on the dihedral angles, ϕ , ψ and χ_1 , are plotted against residue numbers.
- (c) *Correlations between the atomic motions:* The normalized dot products of the normal mode displacement vectors of $C\alpha$ atoms are presented in a triangle map.
- (d) Results from DynDom.
- (e) Animation of normal mode vibration at 310 K. (The amplitude is too small to observe it in most cases. To exaggerate the atomic movements, the animation in which the amplitude at 310 K is doubled can be also played.)

In *ProMode*, the provided data are limited to the 20 lowest-frequency normal modes, because it is well known that atomic fluctuations within the subspace spanned by a small number of the lowest-frequency normal modes dominantly contribute to total fluctuation (Go *et al.*, 1983; Levy *et al.*, 1984; Thomas *et al.*, 1996a,b; Kidera and Go, 1992; Krebs *et al.*, 2002). Although some particular higher frequency modes may contribute to the local fluctuations significant for a protein

The screenshot shows the ProMode web interface. At the top, the title is "ProMode Database of normal mode analysis of proteins". Below the title, there are navigation links for "Spti in other DBs" (PDB, PDBj, PDBsum, SCOP, CATH, FSSP). The main heading is "View of Normal Mode Vibration and Its Analysis". A note states "Plug-in 'Chime' is necessary to see this page." Below this, there are controls for "mode 2" and a "Data:" dropdown menu. The dropdown menu is open, showing options: "Chime View", "Fluctuation of atom positions", "Fluctuation of torsion angles", and "Correlations between CA movements". The main content area contains two Chime windows, both labeled "MDL". The left window shows a 3D animation of a protein structure, and the right window shows a static 3D image of the same protein structure. The protein is colored by dynamic domains (red, blue, yellow, grey). Below the windows, text indicates "This is the animation at 310K. Click here to switch to the animation of doubled amplitude at 310K. Period: 4.978 psec". On the left side, there is a "Manipulation Dialog" window with tabs for "Color", "Command", "Search", "Define", and "Help". The "Command" tab is active, showing the command "select 1-58 and model=1" and buttons for "Execute", "Recall", and "Clear". Below the command field, there are buttons for "Clear", "Multi", "Initialize", and "To Right >>>". A tree view shows the protein structure with residues listed: 5pti_2, ARG1, PRO2, ASP3, PHE4, CYS5, N, CA, C, O, CB, SG, LEU6, GLU7, and PRO8.

Fig. 2. Screenshot of *ProMode*. A protein is displayed in the two Chime windows: (left) 3D animation; (right) static 3D image; the energy-minimum conformation and the two most fluctuated ones superimposed onto it. The mouse event for translation, rotation and zoom on one window affects the other window simultaneously. The dynamic domains defined by DynDom are distinguished by different colors. The axes of screw motions between the domains are shown in the right window. The manipulation dialog (superimposed in the upper left) can be used for selecting residues and/or atoms and executing commands not available in the pop-up menu of Chime.

function, they are not presented in *ProMode*. If we could detect such modes from a hundred modes automatically, we would present the data for such modes in addition to the lowest ones. We are also thinking about how to provide the data for all the modes.

The program DynDom, developed by Hayward *et al.* (1997), is used in *ProMode*. It is a program to determine dynamic domains, hinge axes and amino acid residues involved in the hinge bending motion, when the two different conformations for a protein are given. In *ProMode*, the energy minimum conformation and the most fluctuated one in each normal mode motion are used as input data for DynDom. The results characterize the normal mode vibrations well.

To see the animation, DynDom results, and the energy-minimum conformation superimposed onto the PDB one, a plug-in, Chime, is necessary to be installed. Chime is free, and can play an animation. Since the input data file for the

animation has no information about amino acid residues in Chime (which requires only atomic types and their Cartesian coordinates), it is difficult for a user to manipulate the protein displayed in the Chime animation beyond the original pop-up menu. For this problem, we created a manipulation dialog to execute scripts and commands with which a user can specify particular atoms and/or residues (e.g. an active site, a hydrophobic core and a sequence/structural motif) to highlight by changing color and atomic representation (Fig. 2).

Example for a comparative study

With the collection of the results from the NMA for various proteins in *ProMode*, e.g. it is possible to make a comparative study among a wild type protein and its mutants, or among homologous proteins (Wako *et al.*, 1996; Kikuchi *et al.*, 2000). Here, we will illustrate an example for such a study with bovine pancreatic trypsin inhibitor (BPTI). The 3D structures

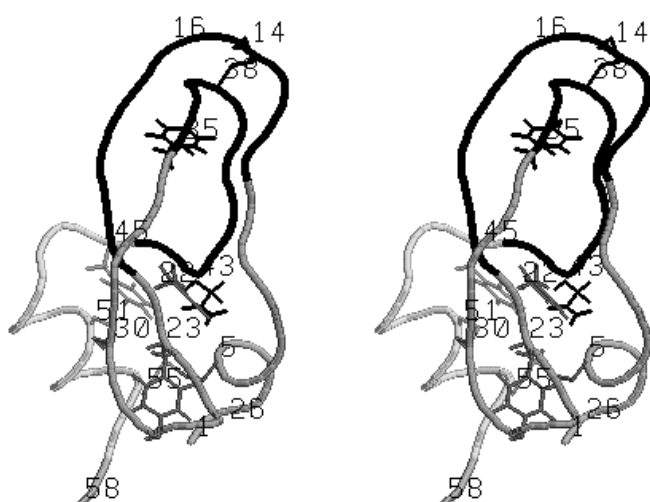


Fig. 3. Stereoscopic view of the 3D structure of BPTI (5pti) drawn by RASMOL (Sayle, 1995). The chain is shaded according to the three dynamic domains defined for the second lowest frequency normal mode by DynDom (Fig. 4). The residue numbers and side chains are shown for the replaced residues and the residues forming a disulfide bond.

of a wild type and several mutant BPTIs have been determined. The following data are used in this paper: the wild type (PDB ID=5pti; Wlodawer *et al.* 1984; Fig. 3), Y23A mutant (1bpt; Danishefsky *et al.* 1993), F22A mutant (1bti; Danishefsky *et al.* 1993), F45A mutant (1fan; Danishefsky *et al.* 1993), N43G mutant (1nag; Danishefsky *et al.* 1993), C30A and C51A mutant (7pti; Eigenbrot *et al.* 1990) and Y35G mutant (8pti; Housset *et al.* 1991). The number of residues of BPTI is 58. In the data of 1bpt and 1nag, however, the two C-terminal residues, Gly-57 and Ala-58, are not visible.

Figure 1 shows the positional fluctuations of atoms and dihedral angles, ϕ and φ for the wild type BPTI. The positional fluctuations in the secondary structures (the two α -helices, 3–6 and 48–55 and the two β -strands 18–24 and 29–35) are relatively small, whereas the larger fluctuations are observed in the loop regions, 15–17, 25–27 and 39–42. Exceptionally, the residues 50–53 in the α -helix fluctuate largely. According to the animations in *ProMode*, it seems that the large fluctuations of the loop region 25–27 affect the C-terminal α -helix mainly through the disulfide bond 30–51. The fluctuations of dihedral angles are larger not only in the loop regions, but also in some residues in the secondary structures. It is notable that the fluctuations of dihedral angles of the residues 25–27 are not very large, whereas their positional fluctuations are large. This region is a hairpin loop located between the two β -strands forming a β -sheet. It seems that the larger fluctuations of dihedral angles in the two β -strands are responsible for the larger positional fluctuations of the loop.

Figure 4 shows the dynamic domains defined by DynDom for the four lowest frequency normal modes. For most of the

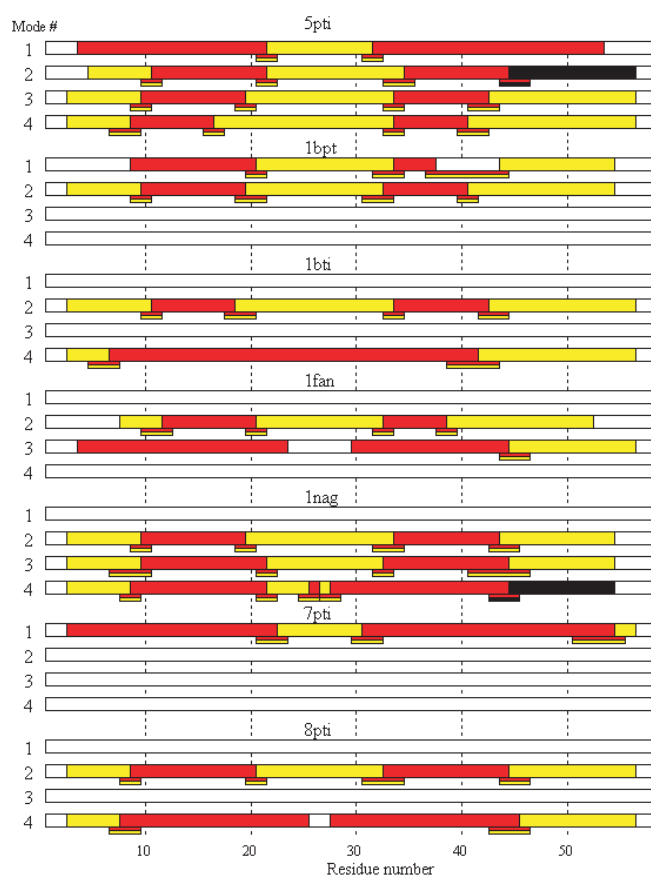


Fig. 4. Dynamic domains defined by DynDom for the four lowest frequency normal modes of the seven BPTIs, 5pti, 1bpt, 1bti, 1fan, 1nag, 7pti and 8pti. In each mode, the region with the same color indicates that the residues in the region belong to the same dynamic domain. The white regions are not defined as dynamic domains. For some modes, no dynamic domains are defined. The short double bars indicate hinge regions.

modes, two dynamic domains (three dynamic domains for two of the modes) are defined. For some modes, however, no dynamic domains are defined. One dynamic domain consists of the half of the β -sheet containing the hairpin loop (around 21–32) and the N- and C-terminals connected to each other with the disulfide bond 5–55. In two of the modes, the C-terminal region is distinguished as a different domain. The other dynamic domain consists of the other half of the β -sheet and the two long loops connected to each other with the disulfide bond 14–38. Each domain consists of more than one region of consecutive residues. In Figure 3, we can see that these regions, which are separated along the chain, are put together into the domains reasonably in the 3D structure of BPTI (more detail views related to Figs 3 and 4 are available in *ProMode*).

Figure 5 shows the differences in positional fluctuations of atoms and dihedral angle φ of the six mutant BPTIs from the wild type. We notice at first that the changes in the fluctuations

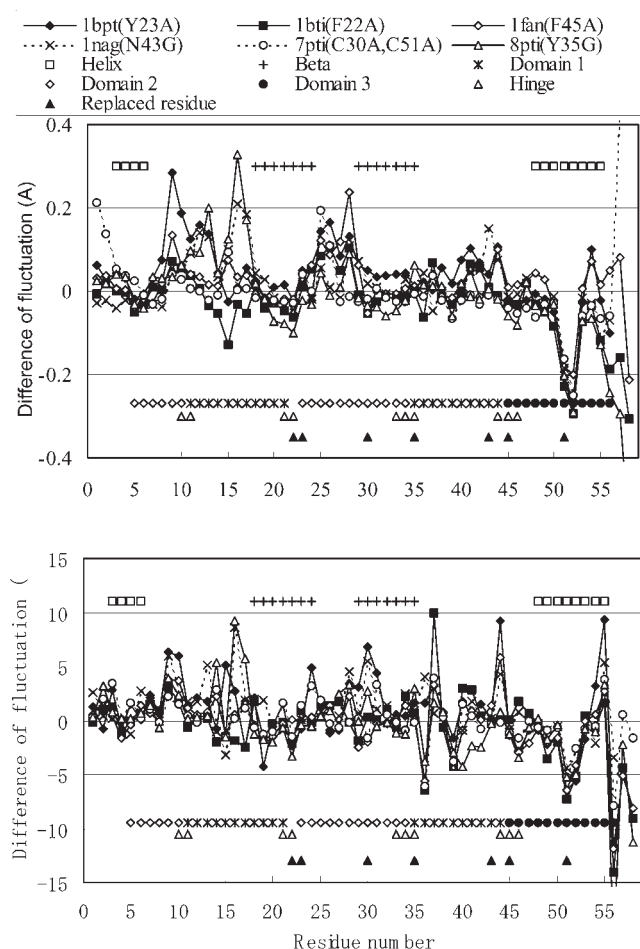


Fig. 5. Differences in positional fluctuations of atoms (upper panel) and dihedral angle φ (lower panel). The differences of the six mutant BPTIs (1bpt, 1bti, 1fan, 1nag, 7pti and 8pti) from the wild type (5pti) are plotted. The replaced residues are indicated in the legend above the upper panel. The secondary structures (the two α -helices and the two β -sheets), and the dynamic domains and hinge residues defined for the second-lowest normal mode of 5pti by DynDom (Fig. 4), are also indicated.

do not necessarily occur around the replaced residues. It is expected that the substituted residues can fluctuate much more, because the amino acids with larger sidechains in the wild type were replaced by those with smaller sidechains in any mutations. In addition, comparing the fluctuations of atoms and dihedral angles for each residue in Figure 5, we found that the residues with larger fluctuations of the dihedral angles are not necessarily the residues with larger fluctuations of the atoms. Incidentally, the replaced residues 22, 23, 35, 43 and 45 are located near hinge regions (Fig. 4), and most of them are hydrophobic residues forming a hydrophobic core. As a result, the changes in these residues seem to affect the movements of the dynamic domains such as regions 9–17, 25–28 and 40–44.

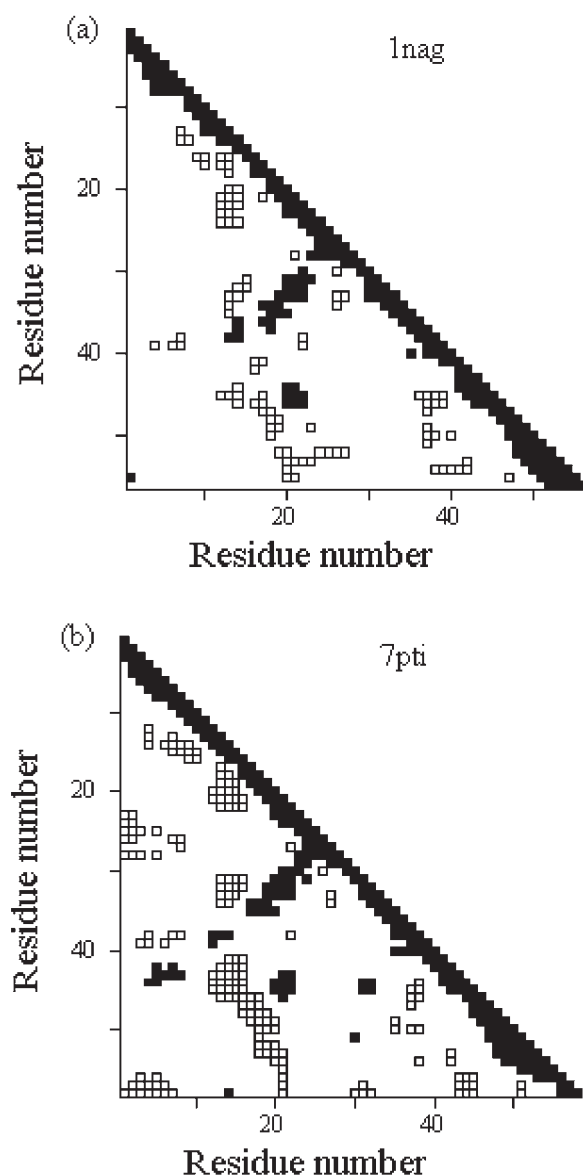


Fig. 6. Correlations between the positional fluctuations of atoms, which are defined as normalized dot products of the normal mode displacement vectors of $C\alpha$ atoms for the mutant BPTIs, 1nag and 7pti. See also Figure 1c.

Finally, the suppression of the fluctuations in both atomic positions and dihedral angles at residues 51–52 is remarkable in any mutants (Fig. 5). In other words, these fluctuations are larger only in the wild type BPTI. According to Figure 1c, the two regions, 26–34 and 52–58, move in a negatively correlated manner on average. Such correlative motions of these residues, however, become weaker in the mutant BPTIs (Fig. 6). In general, the number of pairs of residues with relatively larger positive or negative correlations decreases in the mutant BPTIs, compared with the wild type BPTI. Several

factors, such as the changes in the hinge region, the hydrophobic core and the disulfide bond, are considered responsible for the changes in the correlative motions. Further investigation is necessary to reveal to what extent and how these factors affect the changes.

In this kind of study, we observe the responses to the perturbations of the amino acid substitutions. By collecting several cases, not only the characteristic responses in the individual cases, but also common features of the dynamic properties over the mutant proteins are revealed. Since our aim here is to show the analyses of BPTIs and to give a demonstration on how to use the collection of the NMA data in *ProMode*, we will leave discussion on the detail of the individual mutant BPTIs for another occasion.

DISCUSSION

ProMode provides dynamic properties of various proteins obtained from NMA. By using the dihedral angles as independent variables, we can perform NMA with a full-atom model. Consequently, the animation of the vibrational motions of the proteins can be displayed with a full-atom model. However, one of the problems in *ProMode* is computational expense. Especially, the regularization performed before the NMA calculation requires much CPU time. On the other hand, in the simplified models (Krebs *et al.*, 2002; Chacon *et al.*, 2003; Echols *et al.*, 2003), the regularization is not necessary, because a PDB conformation is already at an energy minimum by their definition. As a result, it is possible to execute the NMA calculations for the entire PDB and to immediately provide a result on demand by the simplified models, while it is hard in *ProMode*. Alternatively, information on NMA with the full-atom model is obtained in *ProMode*. The simplified models and *ProMode* are complementary to each other in the NMA databases.

A comparison with the MD calculations is important. As for the CPU time, since the PDB conformation is not at a minimum of a potential energy function one uses usually (except for the simplified models), the energy minimization must be performed even in MD. In addition, equilibration is required in advance of the simulation. These procedures are comparable to the regularization in *ProMode*. The NMA calculation after the regularization can be performed very fast, whereas the MD simulation for statistical analysis after the equilibration requires much more CPU time. As for the memory requirement, it is proportional only to the number of variables, N , in MD, but to N^2 in NMA. Even though N^2 in *ProMode* is reduced to about 1/64 of the NMA using Cartesian coordinates as independent variables, it is still much larger than N in MD. Since MD can cover much wider area of the conformational energy space, MD may replace NMA in the databases of protein dynamics in future, when the CPU time required for MD will become much less than at present.

Visual expression of the NMA results is another characteristic point of *ProMode*. Not only the global picture of the protein dynamics such as domain motions, but also the fluctuation of the local structures and side chain motions, can be observed realistically through the animation of the full-atom model. By highlighting specific atoms and/or residues (e.g. an active site, a hydrophobic core and sequence/structural motif, with the manipulation dialog we made), it is possible to obtain more detailed information about the dynamics of the protein. The atomic motions and dihedral angle fluctuations plotted against residue numbers and the triangle map for the correlations between the positional fluctuations of atoms are useful to characterize the dynamics of the protein at residue and atom levels. The dynamic domains defined by DynDom and their mutual motions are also helpful to understand the dynamics of the proteins. With the collection of the NMA results for the various proteins, we can compare those results with each other and perform statistical analyses. We expect that such a study will provide new aspects of dynamics of protein structures.

ACKNOWLEDGEMENTS

We gratefully acknowledge Professor Haruki Nakamura, Osaka University, who is a supervisor of one of the BIRD projects, PDBj (Protein Data Bank Japan: <http://www.pdbj.org>). We also gratefully acknowledge Professor Nobuhiro Go, Japan Atomic Energy Research Institute, who was a supervisor in creating the program FEDER and encouraged us to construct *ProMode*.

This work was financially supported by BIRD-JST (Institute for Bioinformatics Research and Development of Japan Science and Technology Corporation). This work is running as a subgroup of the PDBj project. *ProMode* is designed and maintained in collaboration with NEC Soft, Ltd. and Science House, Inc.

REFERENCES

- Berendsen, H.J.C. and Hayward, S. (2000) Collective protein dynamics in relation to function. *Curr. Opin. Struct. Biol.*, **10**, 165–169.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
- Brooks, B.R. and Karplus, M. (1983) Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl Acad. Sci., USA*, **80**, 6571–6575.
- Brooks III, C.L., Karplus, M. and Pettitt, B.M. (1988) *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, John Wiley & Sons, New York.
- Brunger, A.T. (1993) *X-Plor Version 3.1: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven and London.
- Chacon, P., Tama, F. and Wriggers, W. (2003) Mega-Dalton biomolecular motion captured from electron microscopy reconstructions. *J. Mol. Biol.*, **326**, 485–492.
- Danishefsky, A.T., Housset, D., Kim, K.S., Tao, F., Fuchs, J., Woodward, C., and Wlodawer, A. (1993) Crevice-forming mutants

- in the rigid core of bovine pancreatic trypsin inhibitor: crystal structures of F22A, Y23A, N43G, and F45A. *Protein Sci.*, **2**, 577–587.
- Echols,N., Milburn,D. and Gerstein,M. (2003) MolNovDB: analysis and visualization of conformational change and structural flexibility. *Nucleic Acid Res.*, **31**, 478–482.
- Eigenbrot,C., Randal,M., and Kossiakoff,A.A. (1990) Structural effects induced by removal of a disulfide-bridge: the X-ray structure of the C30A/C51A mutant of basic pancreatic trypsin inhibitor at 1.6 Å. *Protein Eng.*, **3**, 591–598.
- Gibrat,J.F. and Go,N. (1990) Normal mode analysis of human lysozyme: study of the relative motion of the two domains and characterization of the harmonic motion. *Proteins*, **8**, 258–279.
- Go,N., Noguti,T. and Nishikawa,T. (1983) Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc. Natl Acad. Sci., USA*, **80**, 3696–3700.
- Hayward,S. and Go,N. (1995) Collective variable description of native protein dynamics. *Annu. Rev. Phys. Chem.*, **46**, 223–250.
- Hayward,S., Kitao,A. and Berendsen,H.J. (1997) Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins*, **27**, 425–437.
- Hinsen,K. (1998) Analysis of domain motions by approximate normal mode calculations. *Proteins*, **33**, 417–429.
- Hinsen,K. (2000) The molecular modeling toolkit: A new approach to molecular simulations. *J. Comp. Chem.*, **21**, 79–85.
- Horiuchi,T. and Go,N. (1991) Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes: human lysozyme. *Proteins*, **10**, 106–116.
- Houssert,D., Kim,K.S., Fuchs,J., Woodward,C. and Wlodawer,A. (1991) Crystal structure of a Y35G mutant of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.*, **220**, 757–770.
- Kidera,A. and Go,N. (1992) Normal mode refinement: crystallographic refinement of protein dynamics structure. I. Theory and test by simulated diffraction data. *J. Mol. Biol.*, **225**, 457–475.
- Kikuchi,H., Wako,H., Yura,K., Go,M. and Mimuro,M. (2000) Significance of a two-domain structure in subunits of phycobiliproteins revealed by the normal mode analysis. *Biophys. J.*, **79**, 1587–1600.
- Kitao,A. and Go,N. (1999) Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol.*, **9**, 164–169.
- Krebs,W.G., Alexandrov,V., Wilson,C.A., Echols,N. Yu,H. and Gerstein,M. (2002) Normal mode analysis of macromolecular motions in a database framework: Developing mode concentration as a useful classifying statistic. *Proteins*, **48**, 682–695.
- Levitt,M., Sander,C. and Stern,P.S. (1985) Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.*, **181**, 423–447.
- Levy,R.M. Srinivasan,R. and Olson,W.K. (1984) Quasi-harmonic method for studying very low frequency modes in proteins. *Biopolymers*, **23**, 1099–1112.
- Marques,O. and Sanejouand,Y.H. (1995) Hinge-bending motion in citrate synthase arising from normal mode calculations. *Proteins*, **23**, 557–560.
- Miller,D.W. and Agard,D.A. (1999) Enzyme specificity under dynamic control: a normal mode analysis of alpha-lytic protease. *J. Mol. Biol.*, **286**, 267–278.
- Noguti,T. and Go,N. (1983a) Dynamics of native globular proteins in terms of dihedral angles. *J. Phys. Soc. Jpn.*, **52**, 3283–3288.
- Noguti,T. and Go,N. (1983b) A method of rapid calculation of a second derivative matrix of conformational energy for large molecules. *J. Phys. Soc. Jpn.*, **52**, 3685–3690.
- Sayle,R.A. (1995) RASMOL: biomolecular graphics for all. *TIBS*, **20**, 374–376.
- Thomas,A., Field,M.J., Mouawad,L. and Perahia,D. (1996a) Analysis of the low-frequency normal modes of the T-state of aspartate transcarbamylase. *J. Mol. Biol.*, **257**, 1070–1087.
- Thomas,A., Field,M.J. and Perahia,D. (1996b) Analysis of the low-frequency normal modes of the R state of aspartate transcarbamylase and a comparison with the T state modes. *J. Mol. Biol.*, **261**, 490–506.
- Wako,H., Endo,S., Nagayama,K. and Go,N. (1995) FEDER/2: program for static and dynamic conformational energy analysis of macro-molecules in dihedral angle space. *Comp. Phys. Comm.*, **91**, 233–251.
- Wako,H. and Go,N. (1987) Algorithm for rapid calculation of hessian of conformational energy function of protein by supercomputer. *J. Comp. Chem.*, **8**, 625–635.
- Wako,H., Tachikawa,M. and Ogawa,A. (1996) A comparative study of dynamic structures between phage 434 Cro and repressor proteins by normal mode analysis. *Proteins*, **26**, 72–80.
- Wlodawer,A., Walter,J., Huber,R. and Sjolín,L. (1984) Structure of bovine pancreatic trypsin inhibitor. Results of joint neutron and X-ray refinement of crystal form II. *J. Mol. Biol.*, **180**, 301–329.