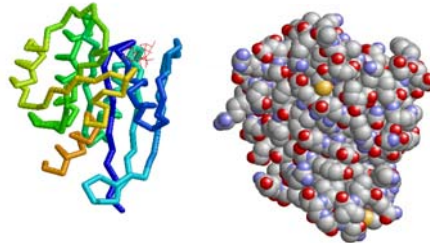


近畿大学・農学部・生命情報学

生体高分子の立体構造とその予測

Ras

```
MTEYKLVVVGAGGVGKSAL  
TIQLIQNHVFVDEYDPTIED  
SYRKQVVIDGETCLLDILD  
TAGQEEYSAMRDQYMRTE  
GFLCVFAINNTKSFEDIHQ  
YREQIKRVKDSDDVPMVLV  
GNKCDLAARTVESRQAQDL  
ARSYGIPYIETSAKTRQGV  
EDAFYTLVREIRQH
```



2009年5月26日(火)

奈良先端大・情報・蛋白質機能予測学講座

川端 猛

takawaba@is.naist.jp

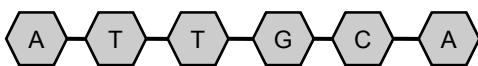
<http://isw3.naist.jp/IS/Kawabata-lab/home-ja.html>


6/2(火)3限13:00～に
311の部屋で
テスト(川端担当分)をします

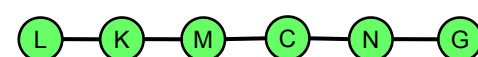
6/2(火)1限の講義は休講となります。

三つの重要な高分子: DNA, RNA, 蛋白質

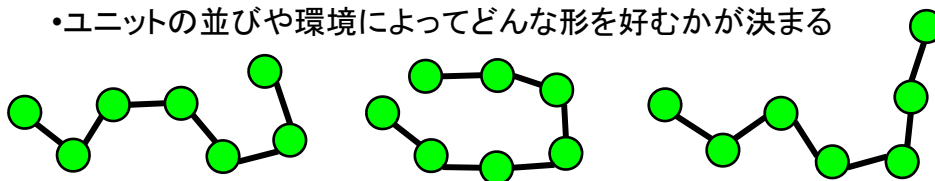
これら三つはいずれも重合体(polymer)、つまりある単位となる分子(monomer)が一行に並んだ形

DNA  4種類

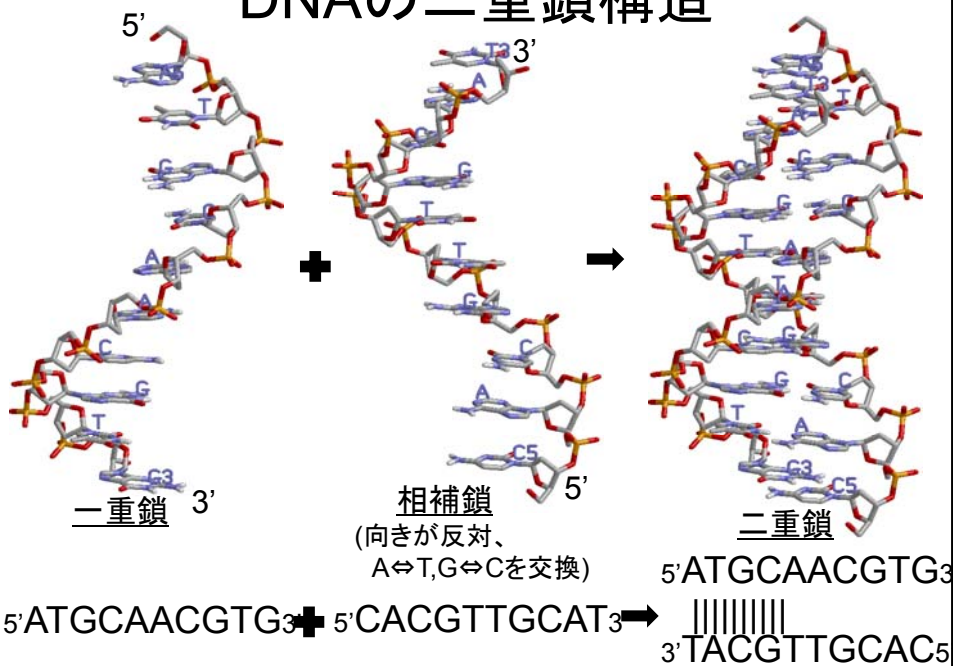
RNA  4種類

蛋白質  20種類

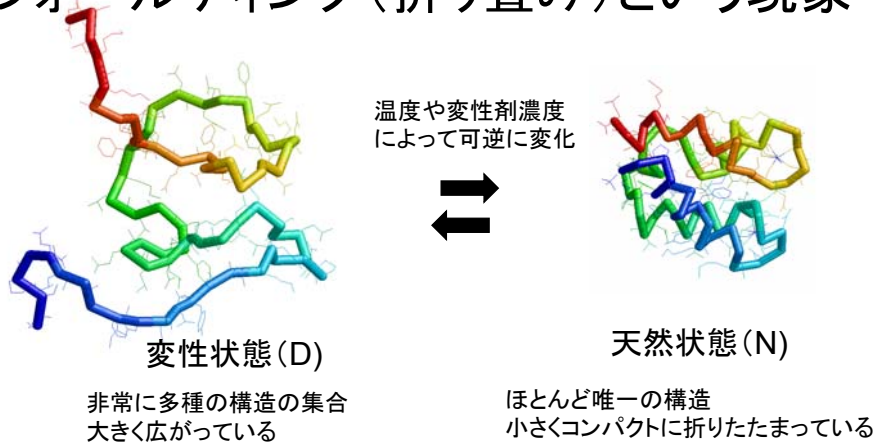
- ひもなので、とにかくフレキシブル
- 原理的にどんな形もとれる。
- ユニットの並びや環境によってどんな形を好むかが決まる



DNAの二重鎖構造



フォールディング(折り畳み)という現象

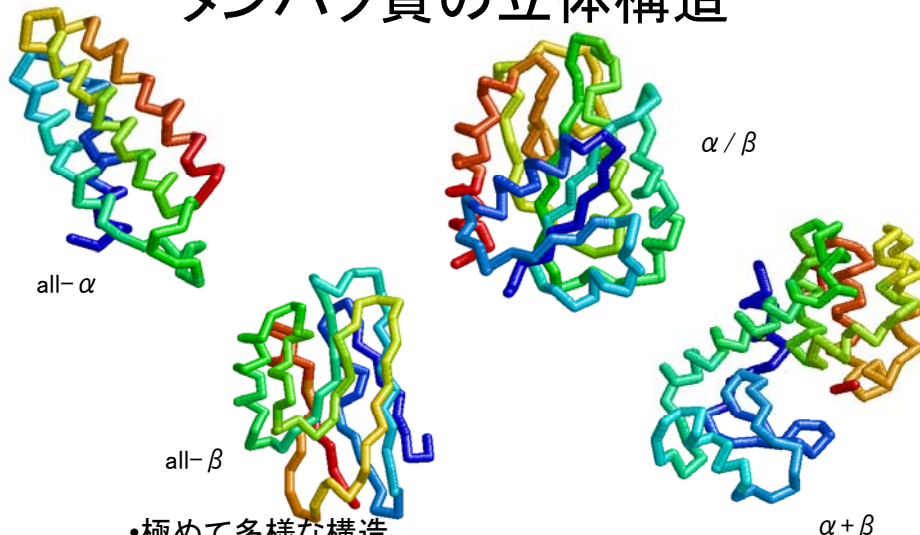


・折り畳みは、原則としてその蛋白質以外の分子の介助を必要としない

アミノ酸配列の情報だけで、天然状態の立体構造が決定される

- ・どうやってコンパクトになるか? → (1)主鎖の水素結合(2)側鎖間の疎水性相互作用
- ・あるアミノ酸配列がどうやって一つの構造を決めるのか???

タンパク質の立体構造



- ・極めて多様な構造
- ・構造はそのアミノ酸配列によって決定される
- ・立体構造の多様性はその分子機能の多様性と関係

立体構造データからわかること

1. 機能(主に結合)に関する理解が深まる

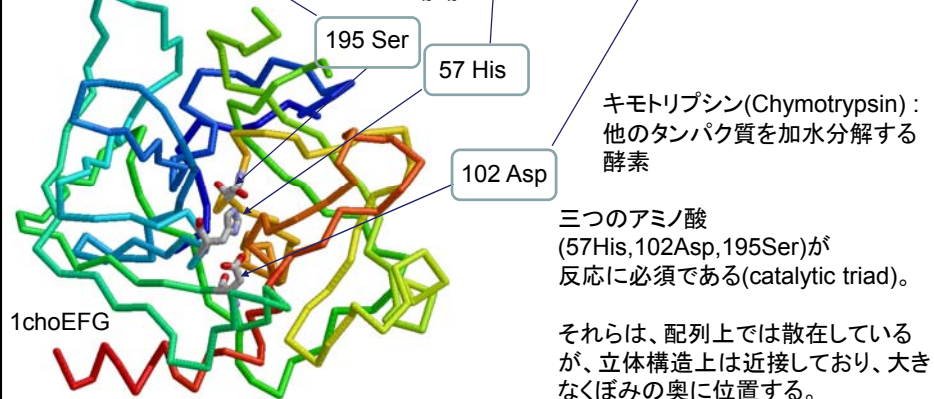
- 原子レベルの結合・反応メカニズムの理解
- 機能に重要なアミノ酸の配置 (配列上散らばっていても、空間的には近接している)
- 立体構造データをもとにした、結合分子の推定・設計

2. 立体構造の類似から、遠い相同性の認識・大きなファミリーの構成が可能

- 立体構造は配列に比べて進化的に保存しやすい

キモトリプシンの活性部位

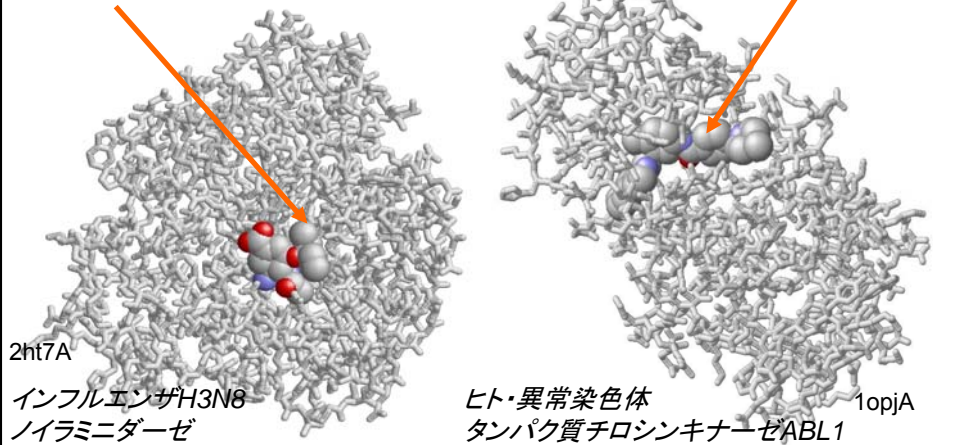
```
>CTRA_BOVIN "Chymotrypsinogen A (EC 3.4.21.1)
C G V P A I Q P V L S G L S R I V N G E E A V P G S W P W Q V S L Q D K T
G F H F C G G S L I N E N W V V T A A (H) C G V T T S D V V V A G E F D Q G
S S S E K I Q K L K I A K V F K N S K Y N S L T I N N (D) I T L L K L S T A
A S F S Q T V S A V C L P S A S D D F A A G T T C V T T G W G L T R Y T N
A N T P D R L Q Q A S L P L L S N T N C K K Y W G T K I K D A M I C A G A
S G V S S C M G D (S) G G P L V C K K N G A W T L V G I V S W G S S T C S T
S T P G V Y A R V T A L V N W V Q Q T L A A N
```



タンパク質に選択的に結合し機能を阻害する分子

タミフル
(インフルエンザ治療薬)

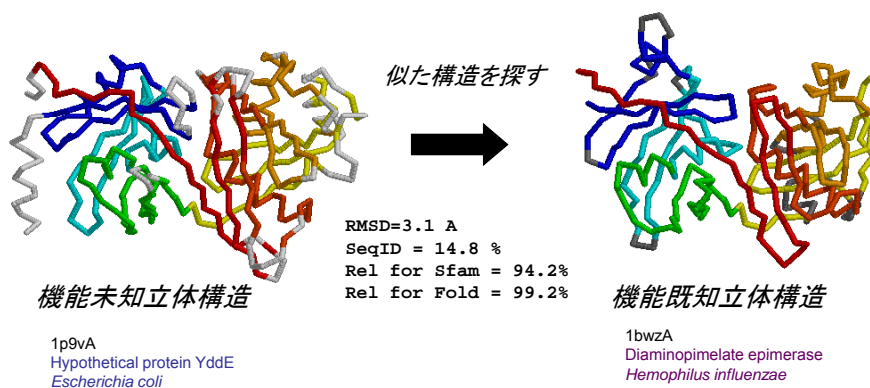
グリベック
(白血病治療薬)



標的タンパク質の結合部位の立体構造をもとに、そこに選択的に結合する分子を設計(ドラッグデザイン)

「形」の比較による弱い相同性認識

機能未知の立体構造データから機能を予測するには？



論理の流れ

1. 立体構造が似ているなら相同(ホモガス)
2. 相同(ホモガス)なら分子機能も似ているはず
3. 構造類似性は機能の類似性を意味する

立体構造データの計測と データフォーマット

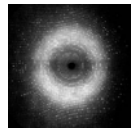
立体構造の決定法

X線結晶解析

大量発現
精製
結晶化

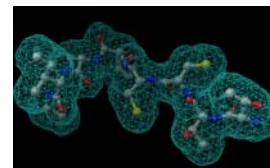


X線回折強度の測定



3次元
電子密度マップ

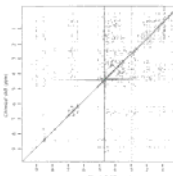
原子モデルの構築
原子モデルの精密化



NMR(核磁気共鳴法)

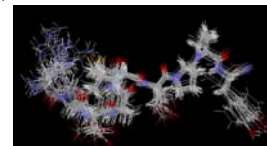
大量発現
精製

核磁気共鳴解析
(NOE解析, 帰属と距離拘束の抽出)



原子間
距離拘束

原子モデルの構築
原子モデルの精密化



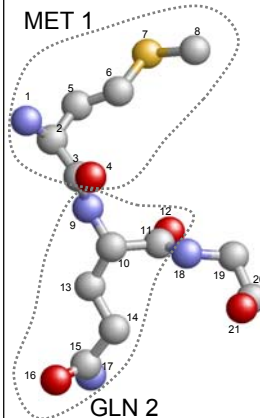
- (1) 多数分子の平均構造の観察(精製が重要。大きな分子、複合体はより難しくなる)
- (2) 発現・精製・結晶化のステップで、各タンパク質によって実験条件の調整が必須
- (3) 実験データの収集・原子モデル構築の段階で計算機の支援が不可欠

PDBフォーマット

RCSB PDB <http://www.rcsb.org/pdb/>
 MSD EBI <http://www.ebi.ac.uk/msd>
 PDBj <http://www.pdbj.org>

wwPDB(World Wide Protein Data Bank) : 立体構造データの収集するデータバンク

| HEADER | | UBIQUITIN | | 17-APR-92 | | 1AAR | | | |
|--------|-----|--|-------|-----------|--------|--------|--------|------|-------|
| COMPND | | DI-UBIQUITIN | | | | | | | |
| SOURCE | | BOVINE (BOS TAURUS) ERYTHROCYTES | | | | | | | |
| AUTHOR | | W. J. COOK, L. C. JEFFREY, M. CARSON, Z. CHEN, C. M. PICKART | | | | | | | |
| 原子番号 | 残基名 | 鎖識別子 | 残基番号 | X座標 | Y座標 | Z座標 | 占有率 | 温度因子 | |
| ATOM | 1 | N | MET A | 1 | 15.493 | 30.088 | 14.694 | 1.00 | 8.36 |
| ATOM | 2 | CA | MET A | 1 | 14.600 | 29.031 | 15.110 | 1.00 | 8.15 |
| ATOM | 3 | C | MET A | 1 | 15.476 | 27.793 | 15.419 | 1.00 | 9.30 |
| ATOM | 4 | O | MET A | 1 | 16.571 | 27.561 | 14.871 | 1.00 | 8.96 |
| ATOM | 5 | CB | MET A | 1 | 13.500 | 28.837 | 14.105 | 1.00 | 9.89 |
| ATOM | 6 | CG | MET A | 1 | 13.823 | 27.997 | 12.931 | 1.00 | 10.21 |
| ATOM | 7 | SD | MET A | 1 | 12.312 | 27.711 | 11.891 | 1.00 | 10.33 |
| ATOM | 8 | CE | MET A | 1 | 13.174 | 26.595 | 10.726 | 1.00 | 7.30 |
| ATOM | 9 | N | GLN A | 2 | 14.968 | 27.014 | 16.326 | 1.00 | 9.75 |
| ATOM | 10 | CA | GLN A | 2 | 15.552 | 25.806 | 16.852 | 1.00 | 11.92 |
| ATOM | 11 | C | GLN A | 2 | 15.000 | 24.553 | 16.168 | 1.00 | 11.71 |
| ATOM | 12 | O | GLN A | 2 | 13.787 | 24.387 | 16.086 | 1.00 | 10.00 |
| ATOM | 13 | CB | GLN A | 2 | 15.368 | 25.715 | 18.386 | 1.00 | 12.39 |
| ATOM | 14 | CG | GLN A | 2 | 15.858 | 24.413 | 19.051 | 1.00 | 14.00 |
| ATOM | 15 | CD | GLN A | 2 | 15.676 | 24.587 | 20.561 | 1.00 | 15.59 |
| ATOM | 16 | OE1 | GLN A | 2 | 16.525 | 25.209 | 21.205 | 1.00 | 18.67 |
| ATOM | 17 | NE2 | GLN A | 2 | 14.564 | 24.203 | 21.152 | 1.00 | 15.39 |
| ATOM | 18 | N | ILE A | 3 | 15.960 | 23.745 | 15.687 | 1.00 | 11.79 |
| ATOM | 19 | CA | ILE A | 3 | 15.593 | 22.437 | 15.130 | 1.00 | 12.14 |
| ATOM | 20 | C | ILE A | 3 | 16.491 | 21.342 | 15.720 | 1.00 | 11.99 |
| ATOM | 21 | O | ILE A | 3 | 17.701 | 21.540 | 15.767 | 1.00 | 12.91 |

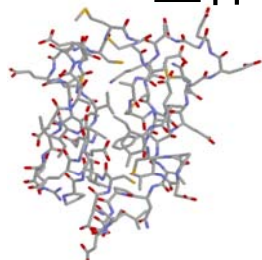


PDBは1971年に設立

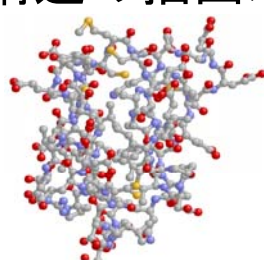
立体構造描画のフリーウェア・シェアウェア

| ソフト名 | WebSite | 対応機種 | 特徴 |
|-------------------------|--|---------------------------|--|
| RasMol | www.openrasmol.org | Win Mac Linux (ソース公開) | 独自の高速描画アルゴリズムを用い、非力なハードでも同様に高速描画が可能。コマンド言語は強力が多機能。基本的な描画法をカバーするが、分子表面の描画はできない。描画の質はやや低い。 |
| Chime | www.mdchime.com | Win Mac (ソース非公開) | RasMolをベースにしたWEBブラウザ内で動作するプラグイン。分子表面表示など描画機能はRasMolより拡張されている。WEBページに分子描画機能を加えることが可能。 |
| Jmol | jmol.sourceforge.net | Win Mac Linux (ソース公開) | Java Applet によりWEBブラウザ内で動作。Chimeと同様に、WEBページに分子描画機能を加えることが可能。Javaを実行できる環境(JRE)があれば、特別なソフトウェアのインストールを必要としない。 |
| PyMOL | pymol.sourceforge.net | Win Mac Linux (ソース公開) | OpenGLを用いた質の高い描画。Python言語で開発されており、ソースも公開。Python言語を用いた機能拡張も可能。 |
| CHIMERA | www.cgl.ucsf.edu/chimera/ | Win Mac Linux (ソース非公開) | OpenGLを用いた質の高い描画。電顕画像などの多様なデータの取り込み、ドッキング、静電計算などのインタフェース有り。Python言語を用いた機能拡張。 |
| VMD | www.ks.uiuc.edu/Research/vmd/ | Win Mac Linux (ソース非公開) | OpenGLを用いた質の高い描画。Tcl言語を用いた機能拡張が可能。分子力学の可視化を主要目的として幅広い機能。 |
| Swiss PDB Viewer | au.expasy.org/spdbv | Win Mac Linux (ソース非公開) | OpenGLを用いた質の高い描画。幅広い機能をサポートしているが、現在あまり活発にアップデートが進んでいない。 |

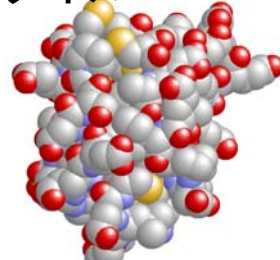
立体構造の描画スタイル



ワイアフレーム



ボール&スティック



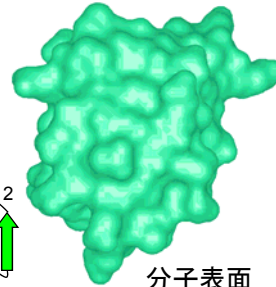
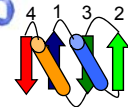
空間充填



バックボーン



リボン



分子表面

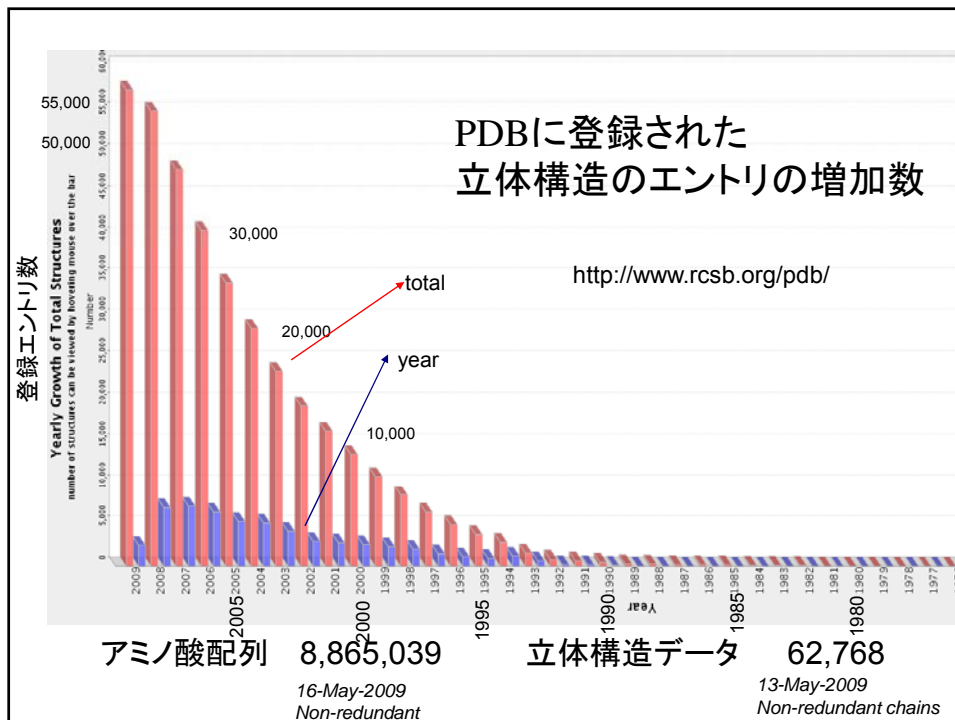
PDBコード: 1fxd(ferredoxin II), Rasmol / Chimeプラグインで描画

立体構造の進化と 立体構造の比較分類

タンパク質立体構造の分類と比較

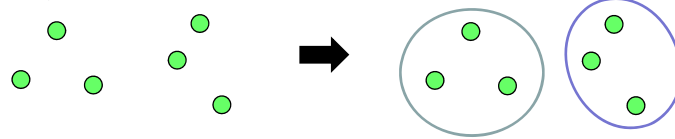
立体構造比較の重要性

1. 「データ整理」: 立体構造データ数は近年急増(約6万エントリ)
2. 「進化」: 立体構造は配列より進化的に保存しやすい
 - (1) 配列では見つからない遠いホモログの発見
 - → 機能推定につながる
 - → 生物の初期進化の解明
 - (2) 精確なアライメント
3. 「物理化学」: 構造データベースの統計から、タンパク質の物理化学的な性質が明らかになる可能性
 - (1) アミノ酸の2次構造傾向
 - (2) 安定な2次構造のパッキング
 - (3) リガンド結合のパターン



配列比較によるタンパク質分類

1970年代から、配列の一致度(sequence identity)に基づいて、グループを作る作業が行われてきた。



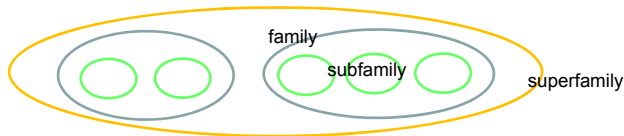
どのくらいの類似性で一つにまとめるかは様々な考えがあるが、少なくとも「**相同なタンパク質群**」を一つのグループにまとめるべき

ファミリー (family, 族):

: 進化的に関係のある (相同な) タンパク質グループのこと

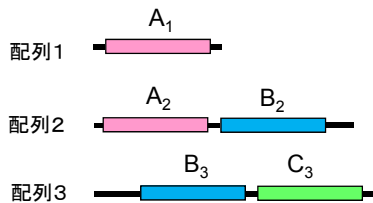
複数のファミリーがあつまった大きなグループ: **スーパーファミリー (superfamily, clan)**

一つのファミリーの中の小さなグループ: **サブファミリー (subfamily)**

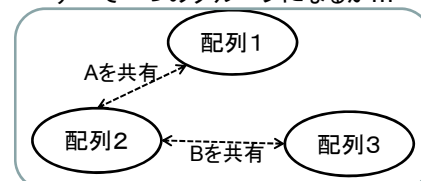


ドメイン単位による分類

配列の部分どうしだけが類似していた場合

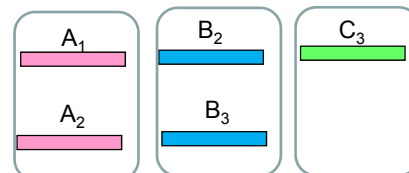
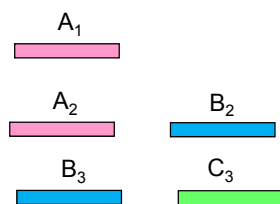


部分一致を重視すると、
すべて一つのグループになるが...



1と2が類似、2と3が類似していても
1と3は類似していない!

ドメインに分割すると



きれいに分類することができる!!

Pfam : 蛋白質ファミリーのデータベース

<http://pfam.sanger.ac.uk>

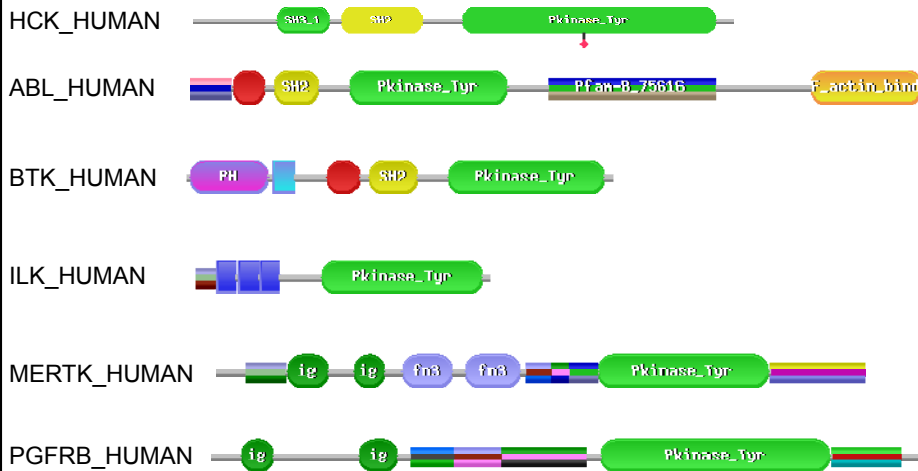
各蛋白質ファミリーのマルチプル
アライメント、HMMなどを
集めたデータベース

The screenshot shows the Pfam website interface. On the left, there is a table titled "Browse Pfam families" with columns for ID, Accession, Type, Number of sequences, Average length, Average %id, Average coverage, Hits, Change status, and Description. The table lists various families such as GP120, BVT_1, ABC_tran, Ank, COX1, BVT, LMR_1, zf-C2H2, Cytochrom_B_H, WQ90, TIR_1, and Oudored_41.

On the right, there is a detailed view of the "Family: zf-C2H2 (PF00096)". This view includes a summary, domain organization, alignments, HMM logo, trees, curation & models, species, interactions, structures, and literature references. A 3D ribbon diagram of the protein structure is also shown.

Pkinase_Tyrドメインをもつタンパク質の例

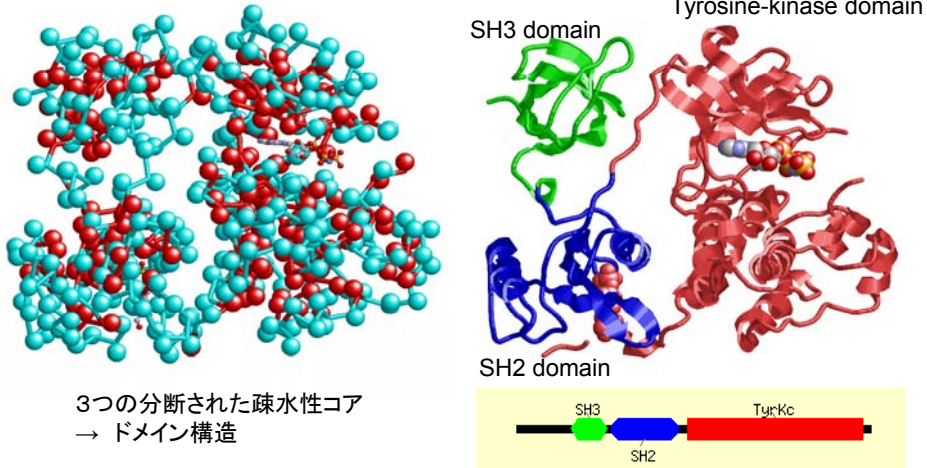
Family : Pkinase_Tyr (PF07714) : Protein tyrosine kinase



Pfamデータベース (<http://pfam.sanger.ac.uk/Software/Pfam/>) からの引用

ドメイン構造と疎水性の関係

Tyrosine protein kinase HCK (1ad5A, HCK_HUMAN)



大きなタンパク質は、いくつかの構造上のまとまり(ドメイン)を持つことが多い

立体構造によるタンパク質の分類

- 配列による分類の延長。ドメイン単位の分類、進化的に関係のある(相同な)タンパク質を群にする(ファミリー)という考え方は同じ。
- 立体構造の類似性から、より遠縁の相同性を見つけ、より大きな群を作る(スーパーファミリー)
- スーパーファミリーをまとめたさらに大きなグループ(フォールド、クラス)
- 主にアミノ酸(C α レベル)のラフな構造の類似性に着目
- 分類対象は主に球状タンパク質

立体構造分類データベースSCOP

<http://scop.mrc-lmb.cam.ac.uk/scop/>

・4階層で分類

Class (クラス) <アナロジー>

Fold (フォールド)

Superfamily (スーパーファミリー)

Family (ファミリー) <ホモロジー>

どの階層に分類するかは、
配列の類似性
立体構造の類似性
分子機能の類似性
機能部位の類似性
を専門家が総合的に判断して決める

Protein: Flavodoxin from *Desulfovibrio vulgaris*

Lineage:

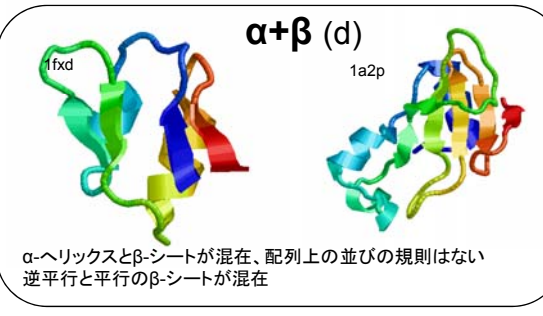
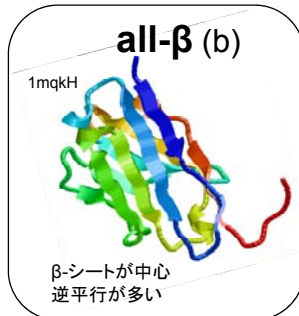
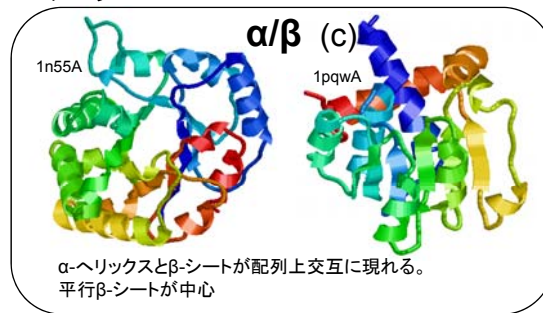
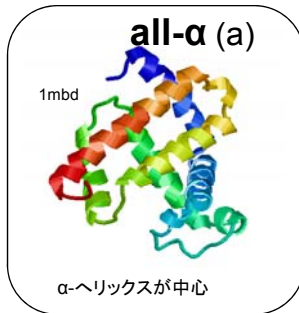
- Root: [scop](#)
- Class: [Alpha and beta proteins \(a/b\)](#) [51349]
Mainly parallel beta sheets (beta-alpha-beta units)
- Fold: [Flavodoxin-like](#) [52171]
3 layers, a/b/a; parallel beta-sheet of 5 strand, order 21345
- Superfamily: [Flavoproteins](#) [52218]
- Family: [Flavodoxin-related](#) [52219]
binds FMN
- Protein: [Flavodoxin](#) [52220]
- Species: [Desulfovibrio vulgaris](#) [52222]

PDB Entry Domains:

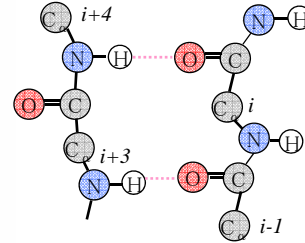
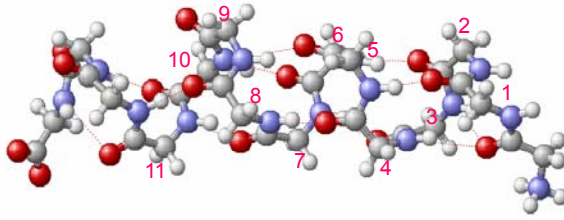
- [1f4p](#)
complexed with fmn; mutant
1. [chun a](#) [59651]
- [1j8q](#)
complexed with fmn
1. [chun a](#) [62744]
- [1j9e](#)
complexed with fmn; mutant
1. [chun a](#) [62756]
- [1akl](#) [31152]
complexed with fmn; mutant

1akr

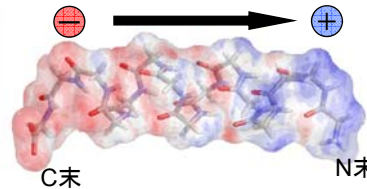
クラス



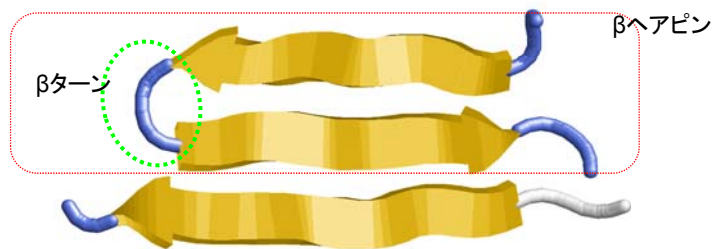
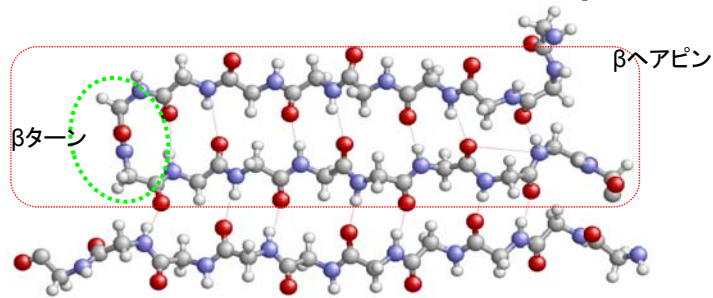
主鎖の水素結合： α ヘリックス



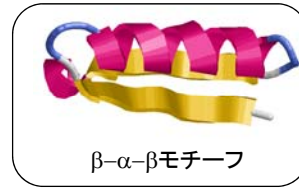
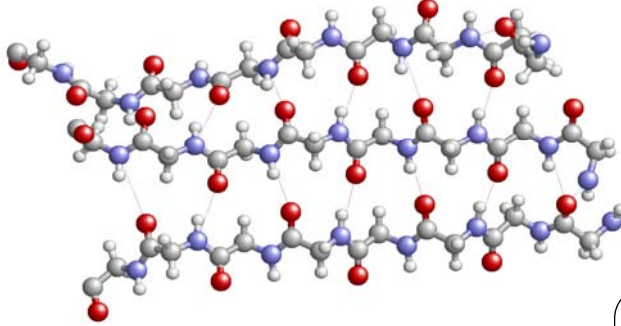
- i 番目のCOが $i+4$ 番目のNHと水素結合を形成
- 3.6残基が1周期の右巻きらせん
- $\text{NH}\cdots\text{O}=\text{C}$ が同じ向きに並ぶため、ヘリックス全体に電気双極子が発生



主鎖の水素結合：逆平行 β シート



主鎖の水素結合：平行βシート



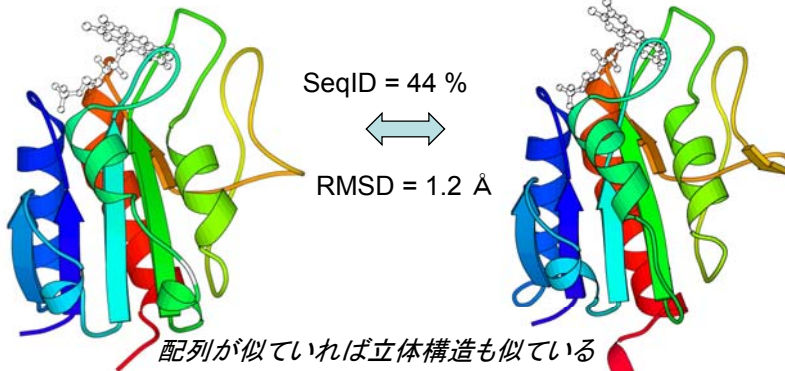
平行βシートがαヘリックスで
接続された構造モチーフ

アミノ酸配列と立体構造の関係

ピロリ菌と大腸菌のフラボドキシシン (同一残基率 SeqID 44%)

ピロリ菌(1fueA)

大腸菌(1ag9A)

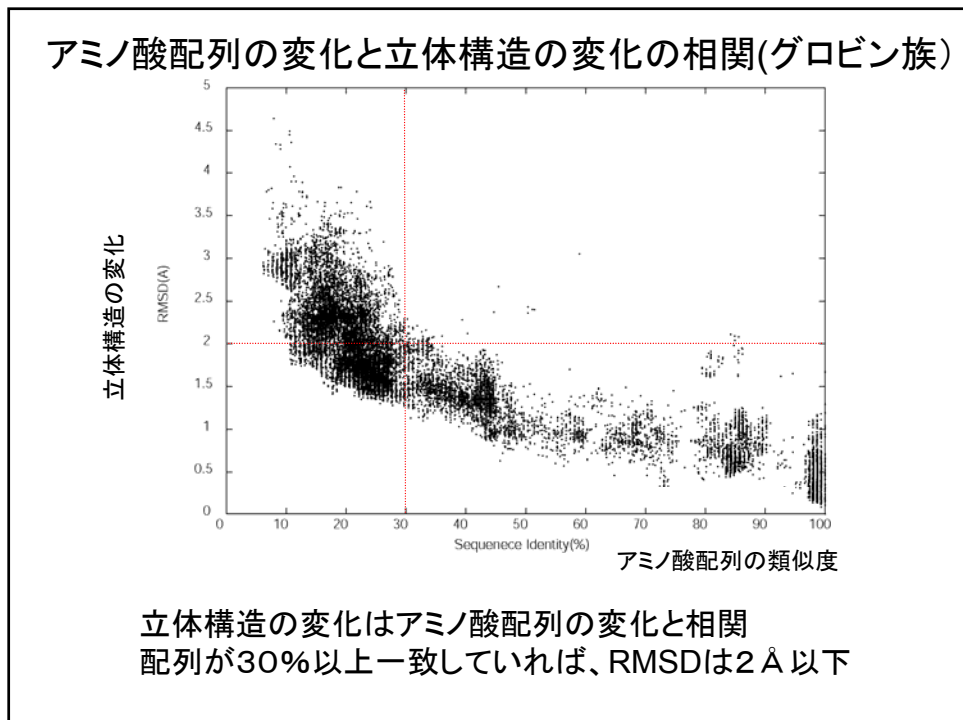
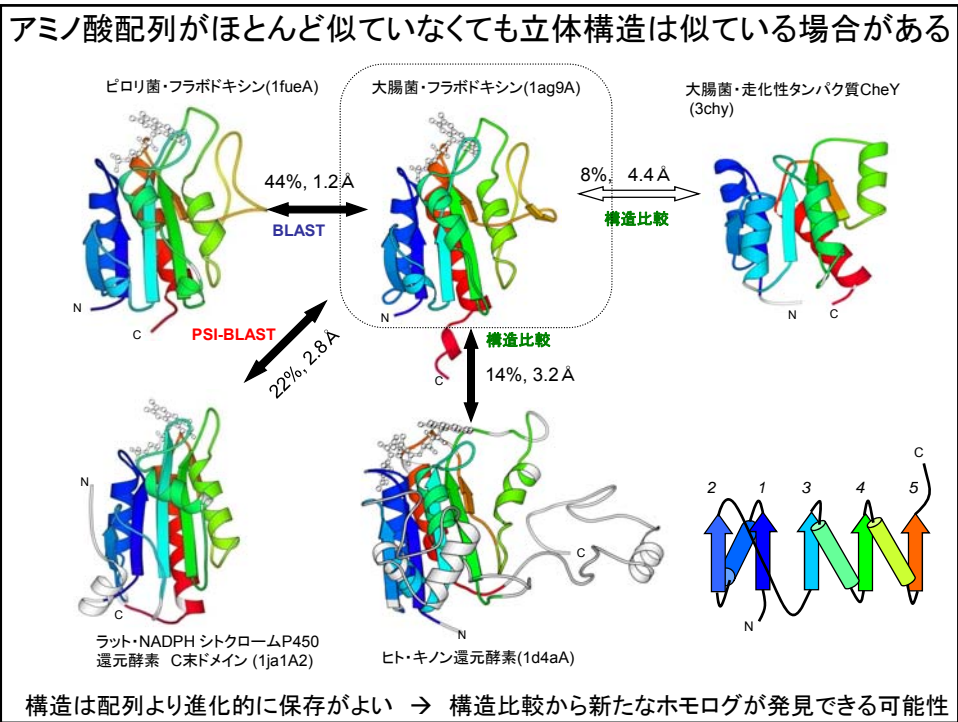


1fueA 2:GKIGIFFGTDSGNAEAI AEKISKAI G--NAEVVDVAKASKEQFNGFTK VILVAPTAGAGD:59

1ag9A 2:AITGIFFGSDTGNTENIAKMIQKQLGKDVADVHDI AKSSKEDLEAYDILL LLIPTWYYGE:61

1fueA 60:LQTDWEDFLGTLEASD-FANKTIGLVGLGQDQTYSETFAEGIFHIYEKAK--AGKVVGQT:116

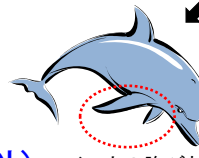
1ag9A 62:AQCDWDDFFPTLE-EIDFNGKLVALFGCGDQEDYAEYFCDALGTIRDIIEPRGATIVGHW:120



2つの構造の類似性

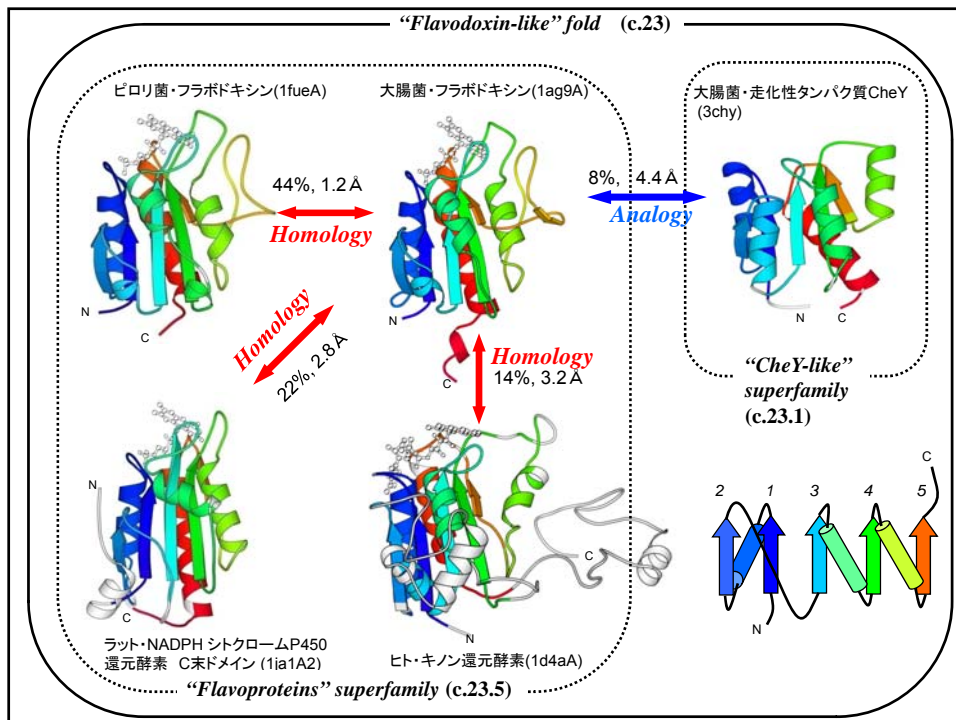
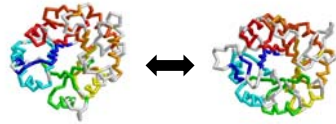
ホモロジー (homology, 相同):

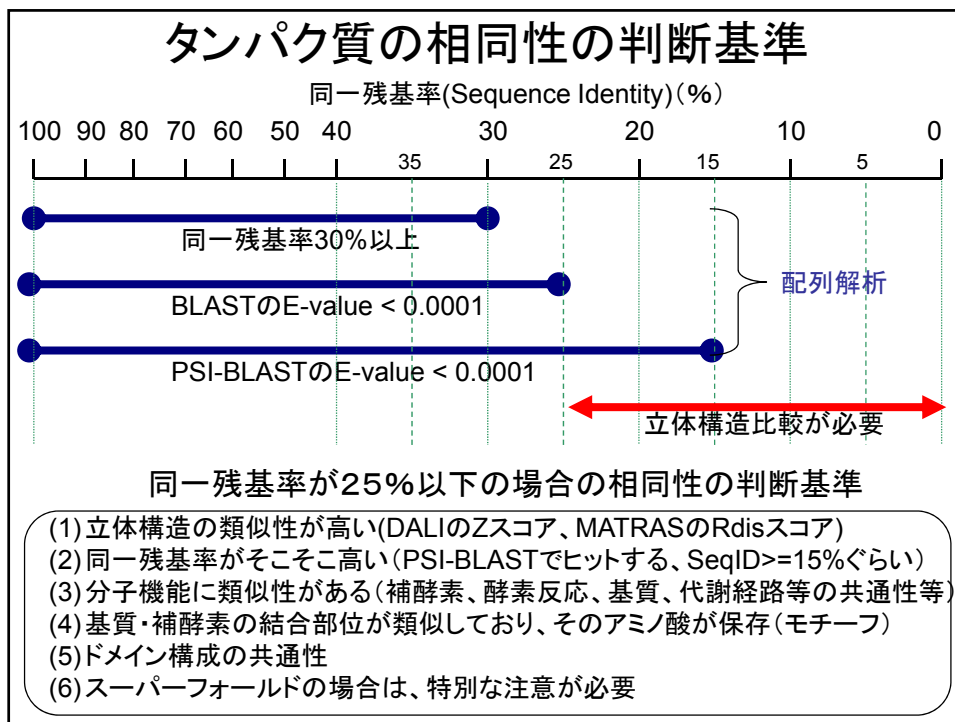
進化的起源を共有することによる類似
多くの場合、分子機能なども類似している



アナロジー (analogy, 相似)

進化的起源とは無関係な類似
多くの場合、分子機能など他の属性は似ていない。
物理化学的な構造の偏好が原因とされる。

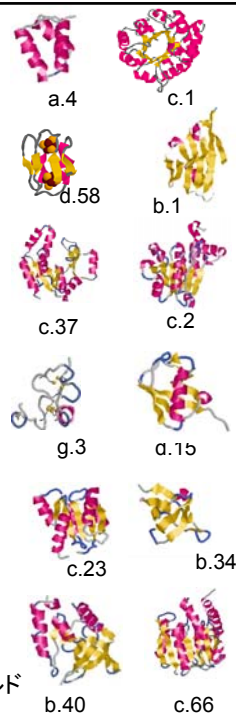




タンパク質の代表的な フォールドの紹介

多く現れるフォールド(出現ドメイン数順)

| | フォールド名 (SCOP1.73による。クラスfhijklは除く) | ドメインの数 | スーパーファミリーの数 |
|-------------|--|------------|-------------|
| a.4 | DNA/RNA-binding 3-helical bundle | 292 | 14 |
| c.1 | TIM beta/alpha-barrel | 269 | 33 |
| d.58 | Ferredoxin-like | 248 | 55 |
| b.1 | Immunoglobulin-like beta sandwich | 246 | 27 |
| c.37 | P-loop nucleoside triphosphate hydrolase | 235 | 1 |
| c.2 | NAD(P)-binding Rossmann-fold domains | 168 | 1 |
| g.3 | Knottins | 154 | 19 |
| d.15 | Beta-Grasp(ubiquitin-like) | 131 | 13 |
| c.23 | Flavodoxin-like | 118 | 15 |
| b.34 | SH3-like barrel | 117 | 18 |
| b.40 | OB-fold | 100 | 12 |
| c.66 | SAM-dependent methyltransferases | 95 | 1 |

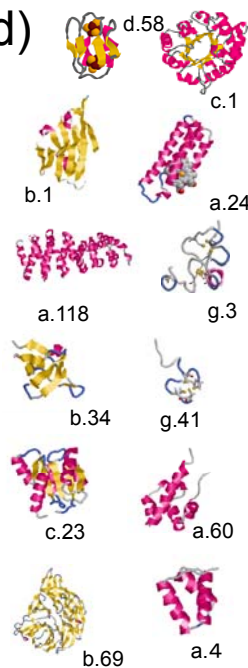


太字はスーパーフォールド: 多くのスーパーファミリーを含むフォールド

スーパー・フォールド(Superfold)

スーパーフォールド: 多くのスーパーファミリーを含むフォールド

| | フォールド名 (SCOP1.73による。クラスfhijklは除く) | その下のスーパーファミリーの数 |
|-------|--------------------------------------|-----------------|
| d.58 | Ferredoxin-like | 55 |
| c.1 | TIM beta/alpha-barrel | 33 |
| b.1 | Immunoglobulin-like beta sandwich | 27 |
| a.24 | Four-helical up-and-down bundle | 27 |
| a.118 | Alpha-alpha superhelix | 23 |
| g.3 | Knottins | 19 |
| b.34 | SH3-like barrel | 18 |
| g.41 | Rubredoxin-like | 16 |
| c.23 | Flavodoxin-like | 15 |
| a.60 | SAM(sterile alpha motif) domain-like | 15 |
| b.69 | 7-bladed beta-propeller | 14 |
| a.4 | DNA/RNA-binding 3-helical bundle | 14 |



スーパーフォールドは「構造の類似と機能の類似が対応しにくいフォールド」ともいえる

フォールド・ファミリーの数は有限

Chothiaは、生物界には約 1000 のファミリーが存在すると推定
(Nature. 1992年, 357,543-544)

SCOP 1.73 (Nov 2007)

(Class fhijkを除く)

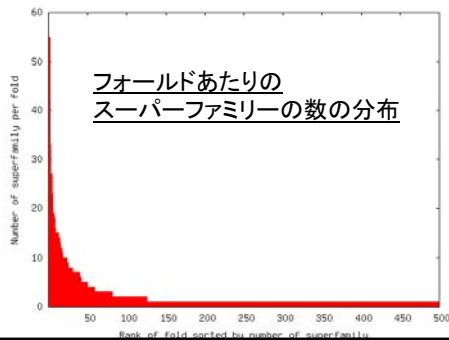
フォールドの数 : 1036

スーパーファミリーの数 : 1685

ファミリーの数 : 3360

1000個程度の立体構造を解けば、
生物界のすべての基本構造
がわかる？

(「構造ゲノム科学」の発想の原点)



ごく少数のフォールドが多数のスーパー
ファミリーをかかえる

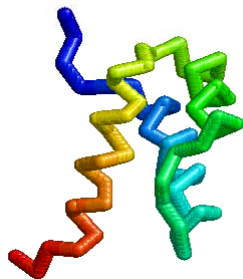
一つしかスーパーファミリーを含まない
フォールドが圧倒的に多い

大多数の蛋白質はメジャーな100程
度のフォールドに含まれるが、残りの
蛋白質のフォールドは多様

三本ヘリックスバンドル

DNA/RNA-binding 3-helical bundle(a.4)

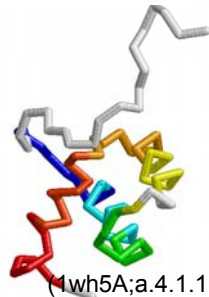
三本のヘリックスが束になった構造
DNAやRNAに結合する蛋白質が多い



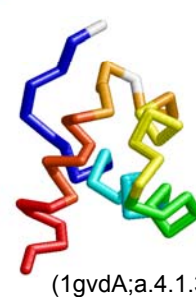
Mating type protein A2,
yeast(1k61A;a.4.1.1)



DNAと
結合した
構造(1k61)



(1wh5A;a.4.1.1)



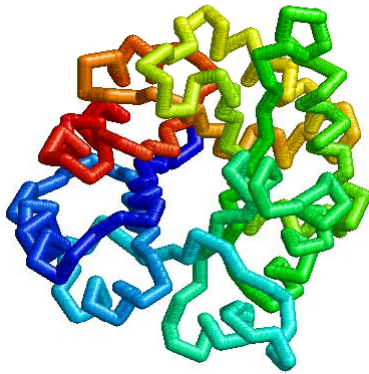
(1gvdA;a.4.1.3)

TIMバレル

TIM beta/alpha barrel (c.1)

八本のβストランドとαヘリックスが交互に組み合わさり、対称的な樽(barrel)構造を形成する。中央の樽構造は平行のβシートからなる。

全般に解糖系の酵素が多い。基質、酵素反応は極めて多彩。



Triosephosphate isomerase
1n55A(c.1.1.1)

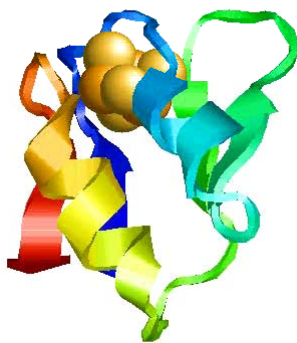


Imidazole glycerol
phosphate synthase
subunit hisF
1thfD(c.1.2.1)

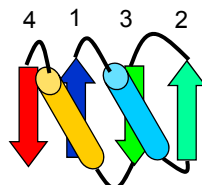


KHG/KDPG aldolase
1euaA(c.1.10.1)

フェレドキシン様 Ferredoxin-like (d.58)



4Fe-4S Ferredoxin(1fxd,d.58.1)

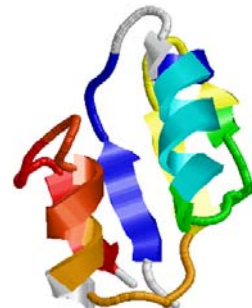


4本の逆平行のβストランドと
2本のαヘリックスからなる。

分子機能は極めて多彩。



U1A RNA binding domain
(1nu4A,d.58.7.1)



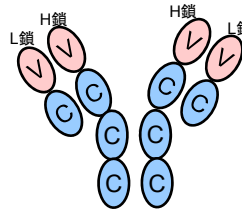
CheY binding domain of CheA
(1eayC,d.58.24)

免疫グロブリン様

Immunoglobulin-like beta sandwich(b.1)



Immunoglobulin
Heavy chain variable
domain. 1mjuH1(b.1.1.1)



免疫グロブリン分子はH鎖とL鎖からなり、それぞれ、免疫グロブリン様ドメインからできている。

細胞表面あるいは分泌蛋白質が多い。免疫や分子認識に関する機能を担うことが多い。



Immunoglobulin
Heavy chain constant
domain 1mjuH2(b.1.1.2)

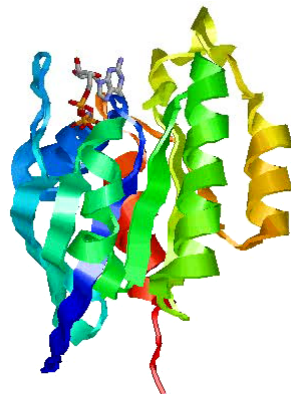


Macromomycin
1noaA (b.1.7.1)

4本のβストランドからなるβシートが二枚、サンドイッチ状に重なった構造をとる。逆平行のβシートが主。

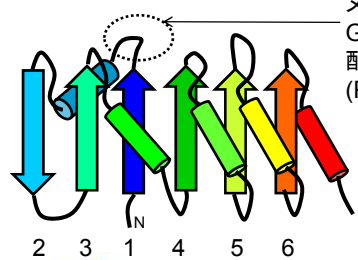
P-loop スーパーファミリー

P-loop nucleoside triphosphate hydrolase(c.37)



Ras-related protein RAB-5A
1r2qA(c.37.1.8)

シグナル伝達、DNA結合、モーター機能など、多彩な分子機能を担うが、ATP、GTPなどのヌクレオシドの加水分解を行う点は共通。



ヌクレオシド結合部。
GxxxxGKTの配列モチーフ(P-loop)を持つ。

平行のβシートが主。トポロジーには様々な変異あり。



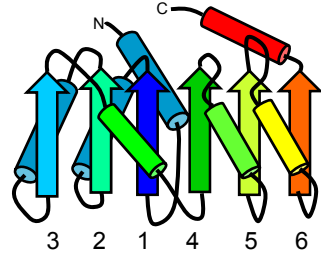
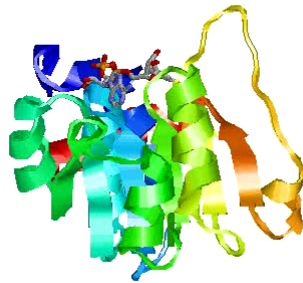
Guanylate kinase
1gkyA(c.37.1.1)



RecA protein.Domain1
1xmsA1(c.37.1.11)

ロスマン フォールド

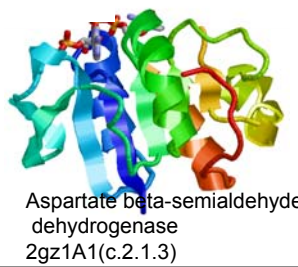
NAD(P)-binding Rossmann-fold domains (c.37)



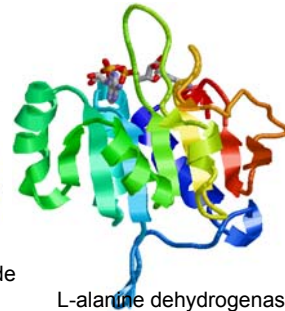
平行の
βシートが主。

Alcohol dehydrogenase gamma chain domain2, 1u3wA(c.37.1.8)

補酵素を利用したNAD(P)を脱水素酵素を持つ。基質を認識するドメインは別に存在する。



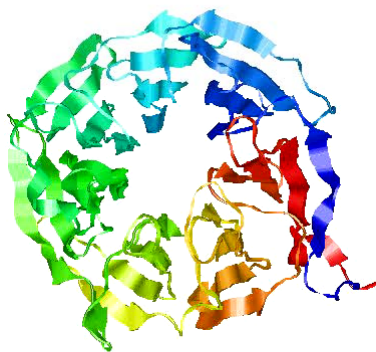
Aspartate beta-semialdehyde dehydrogenase 2gz1A1(c.2.1.3)



L-alanine dehydrogenase 1pjcA1(c.2.1.4)

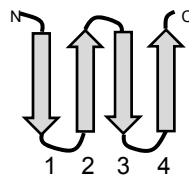
7枚羽 βプロペラ

7-bladed beta-propeller (b.69)

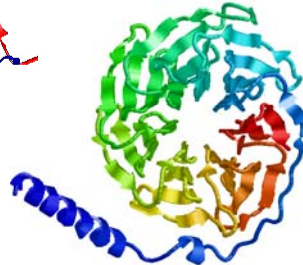


Garactose oxydase domain3, 1k3iA3(b.69.1.1)

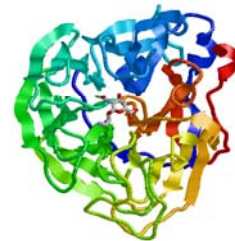
細胞膜付近の、シグナル伝達や糖結合に関わることが多い。



のユニットが7回、円環状に繰り返されたフォールド



Guanine nucleotide-binding protein subunit beta-1 1gotB(b.69.4.1)



Influenza Neuraminidase 1f8dA(b.68.1.1)
6枚羽 βプロペラ

蛋白質の立体構造比較: 学籍番号: _____ 氏名: _____

上の蛋白質群の中から、以下のフォールドに属する蛋白質の番号を記せ。

TIMバレル: _____ 免疫グロブリン様: _____
 βプロペラ: _____ フェルドキシン様: _____

H21 生命情報学
2009.5.26

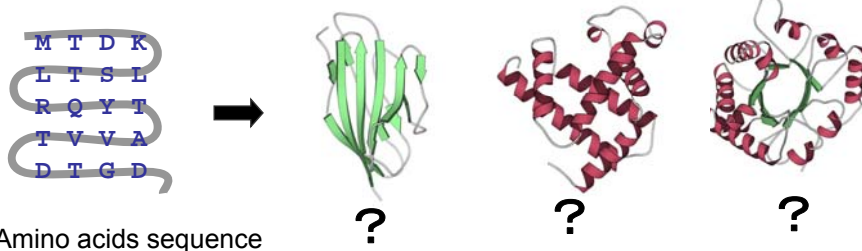
蛋白質の立体構造比較: 学籍番号: _____ 氏名: _____

上の蛋白質群の中から、以下のフォールドに属する蛋白質の番号を記せ。

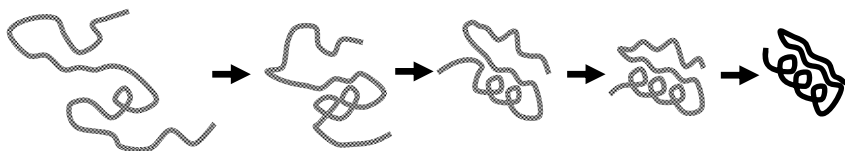
TIMバレル: 1, 10 免疫グロブリン様: 8, 17
 βプロペラ: 6, 7 フェルドキシン様: 12, 15

H21 生命情報学
2009.5.26

タンパク質の立体構造予測



3次構造予測(1): *Ab initio* 予測



原子モデルとポテンシャルエネルギー関数を設定
より低いエネルギー値になるように構造を変形していく

分子動力学法、モンテカルロ法、
エネルギー最小化計算。

大きな計算量が必要。

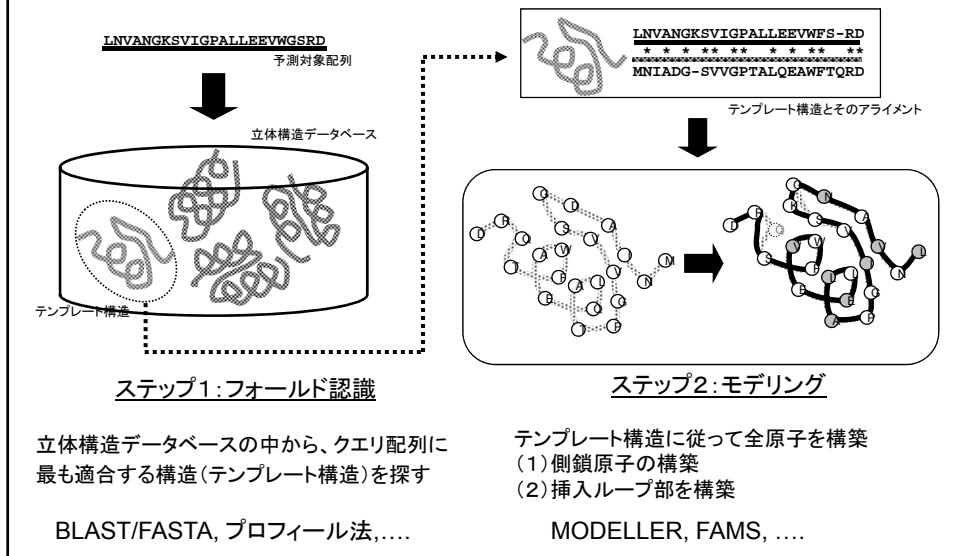
近年フラグメントアセンブリの手法が進展

Baker グループ (U. Washington) <http://depts.washington.edu/bakerpg/newindex.html>

高田グループ (神戸大学) <http://theory.chem.sci.kobe-u.ac.jp/indexj.html>

3次構造予測(2): 比較モデリング (ホモロジー・モデリング)

原理: 立体構造はアミノ酸配列より保存しやすい。

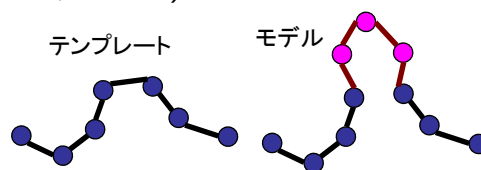


モデリング

テンプレート構造を元にした全原子の構築
(MODELLER, FAMS)

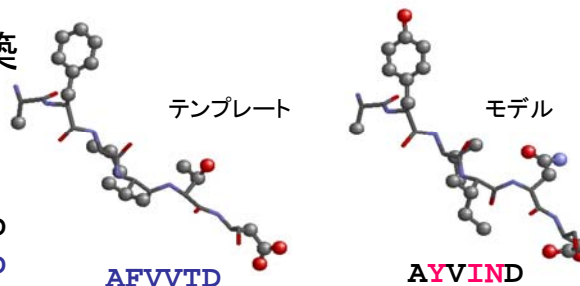
(1) ループの構築

Sequence ALIM**ST**KGFVS
Structure LLLM---GFIT



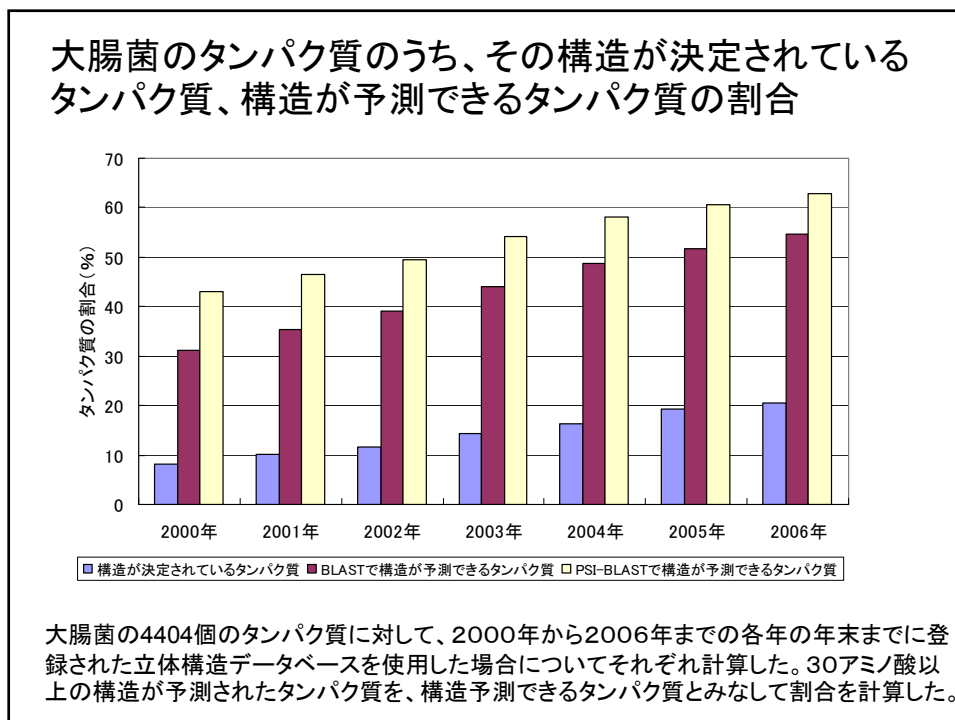
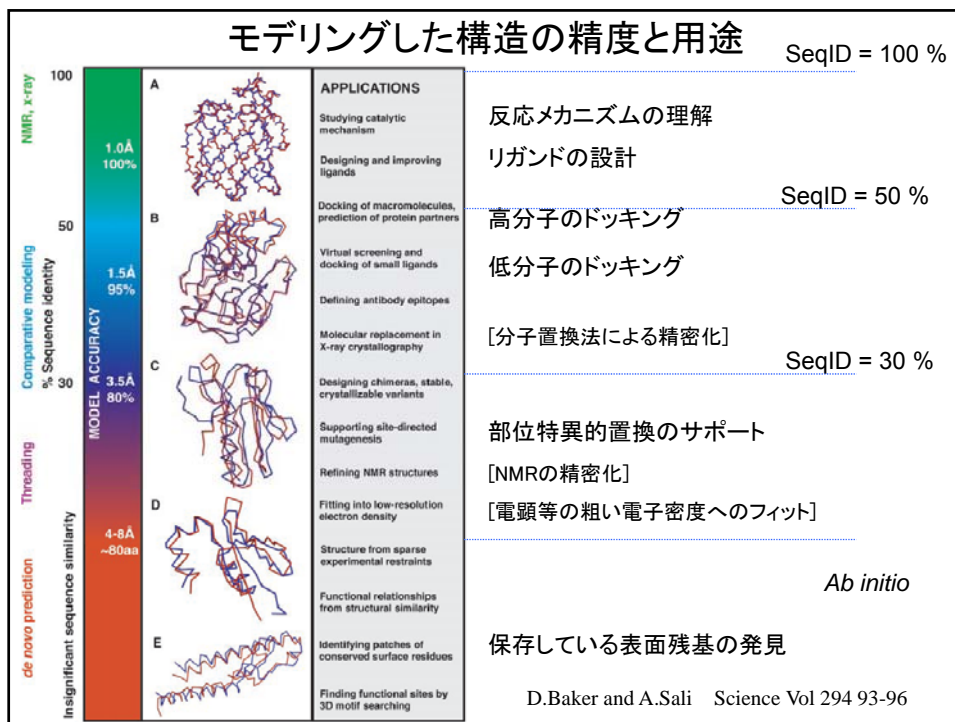
(2) 側鎖原子の構築

Sequence A**Y**VIND
Structure AFV**V**TD



MODELLER : <http://www.salilab.org/modeller/modeller.html>

FAMS <http://www.pharm.kitasato-u.ac.jp/biomoleculardesign/>



構造予測コンテスト CASP

Critical Assessment of techniques for protein Structure Prediction

<http://predictioncenter.llnl.gov/>

(1) ターゲットタンパク質の募集

半年以内に立体構造が解ける予定のタンパク質を広く構造生物学者から募集

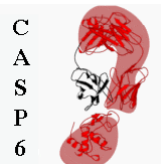
(2) ターゲットタンパク質のアミノ酸配列だけをWEBで公示

(3) 予測者は、期日までに、予測立体構造をサブミット

(4) ターゲット立体構造が公表された後、予測立体構造と比較・優劣を判断

- John Moultらが主催で1994年に開始。2年おきに開催。
- 2004年にCASP6が開催。世界各国から 224チーム、65サーバが参加。
- 予測の手法・難易度によっていくつかのカテゴリに分けて審査

Comparative Modeling
Fold Recognition (Homologous)
Fold Recognition (Analogous)
New Fold



- CASP6ではいくつかの日本人研究者のグループが入賞

CBRC-3D (Kentaro Tomii) :Fold Recognition(Homologous)

CHIMERA (Mayuko Takeda-Shitaka) :Fold Recognition(Homologous)

参考図書

構造生物学一般について

- Bluce Alberis他著、中村桂子、松原謙一監訳 「Essential 細胞生物学 原書第2版」 第2章、第4章、2005年、南江堂
- C.Branden & J.Tooze (勝部幸輝ら訳)「タンパク質の構造入門」(2000)、ニュートンプレス
- 松澤 洋編 「タンパク質工学の基礎」 2004年、東京化学同人
- 後藤祐児、桑島邦博、谷澤克行 「タンパク質科学 – 構造・物性・機能 -」 2005、化学同人

構造バイオインフォマティクス

- 郷通子・高橋健一 編集 「基礎と実習 バイオインフォマティクス」 2004年 共立出版
- 藤博幸 編集「はじめてのバイオインフォマティクス」第2.2章 2006年 講談社
- Arthur M. Lesk (高木淳一 訳)「ポストゲノム時代のタンパク質科学」2007年 化学同人
- Arthur M. Lesk (岡崎康司・坊農秀雄 監訳) 「バイオインフォマティクス基礎講義 一歩進んだ発想をみかくために」 2003年 メディカル・サイエンス・インターナショナル
- 美宅成樹・榊佳之 「バイオインフォマティクス」 2003年 東京化学同人 第6章、第7章
- 中村春木・有坂文雄 編 「シリーズ・ニューバイオフィジクス1 タンパク質のかたちと物性」 1997年、共立出版
- 後藤祐児、桑島邦博、谷澤克行 「タンパク質科学 – 構造・物性・機能 -」 2005年 化学同人 第2、4、5、6章