

HOMCOS: 複合体立体構造の 検索・モデリングサーバ

HOMCOS: a server to search and model complex 3D structure

○川端 猛、中村春木、金城玲 (大阪大・蛋白質研究所)

Takeshi Kawabata, Haruki Nakamura, Akira Kinjo (Inst.Prot.Res., Osaka Univ.)

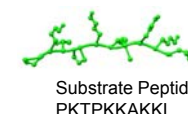
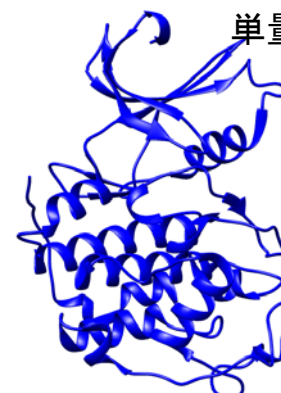


<http://homcos.pdbj.org>



本サーバは創業等支援技術プラットフォーム事業
の支援を受けて開発されています

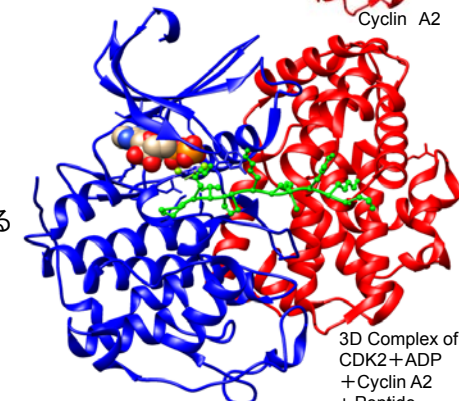
複合体立体構造は
単量体構造より機能情報が豊富



Cyclin-dependent protein kinase (CDK2)

複合体立体構造から以下のことがわかる

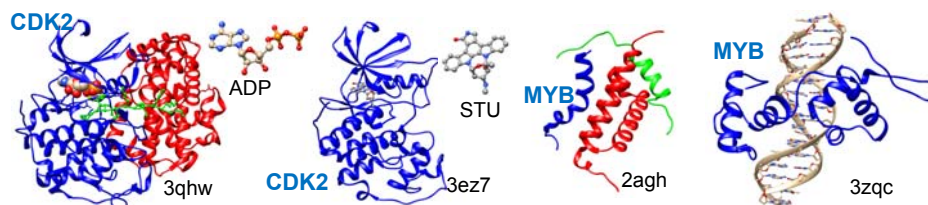
- (1)他の分子との結合部位
→変異体の解釈・設計
→阻害剤の設計・改変
- (2)結合・反応のメカニズムの理解



3D Complex of
CDK2+ADP
+Cyclin A2
+ Peptide
(PDBcode:3qhw)

複合体立体構造データの有効活用の問題点

PDBの複合体3D構造の有効活用は、蛋白質の機能解析に重要。しかし、



(1)PDBに含まれている分子単体を検索するサーバは数多くあるが、複合体の立体構造が検索できるサーバはほとんどない。

ある二つのタンパク質の複合体構造が得られているかどうかの検索は？

あるタンパク質と複合体構造が得られている分子のリストの検索は？

(2)PDBに含まれている分子単体のホモロジー・モデリングを行う予測サーバは数多くあるが、複合体のホモロジー・モデリングができる予測サーバはほとんどない。

相互作用が報告されている分子ペアのうち、複合体立体構造データがあるのはごくわずか

→ホモロジー・モデリング法で、できるだけ予測を試みる必要

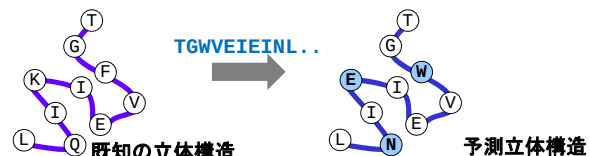
立体構造予測法の二つのアプローチ

名称	ホモロジー・モデリング法 比較モデリング法 鋳型ベース予測法	非経験的方法 Ab initio 予測法 De novo 予測法
手法の概要	鋳型立体構造にできるだけ似た形で、立体構造を予測	鋳型構造を用いずに、物理化学的な原理(分子シミュレーションの技法)に基づいて立体構造を予測
鋳型立体構造	必要	不要
一般性	低い	高い
計算量	少ない	多い
予測精度	似た鋳型があれば高い	高い精度を得るには大きな計算量が必要
単体の立体構造予測	MODELLER, SWISS-MODEL	ROSETTA, EVfold,...
蛋白質複合体予測	MODELLER, HOMCOS	ZDOCK, HADDOCK,...
低分子-タンパク質複合体予測	MODELLER, HOMCOS, fkcombu	DOCK, AutoDock, sievgene, Glide,...

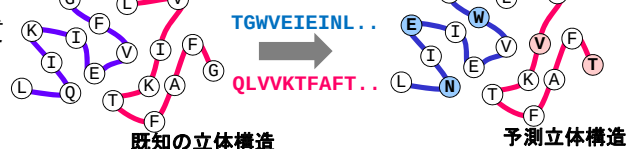
鋳型ベースのモデリング: Template-based Modeling

複合体のホモロジー・モデリング

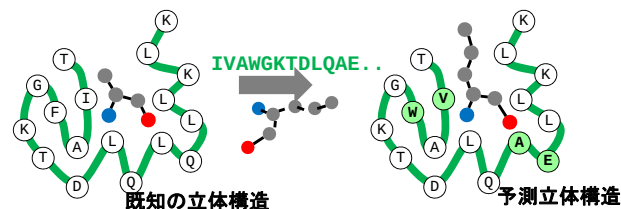
タンパク質単量体



タンパク質-タンパク質複合体

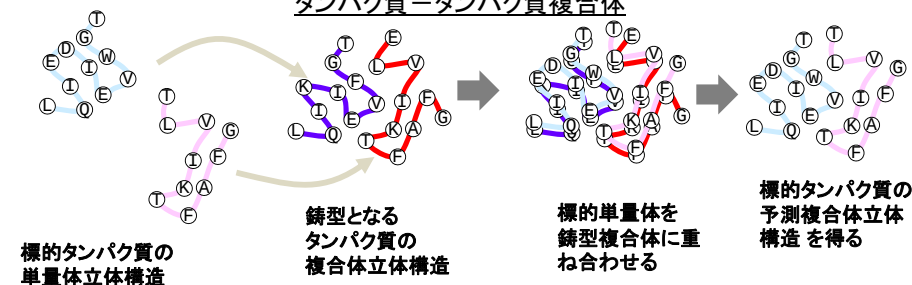


化合物-タンパク質複合体

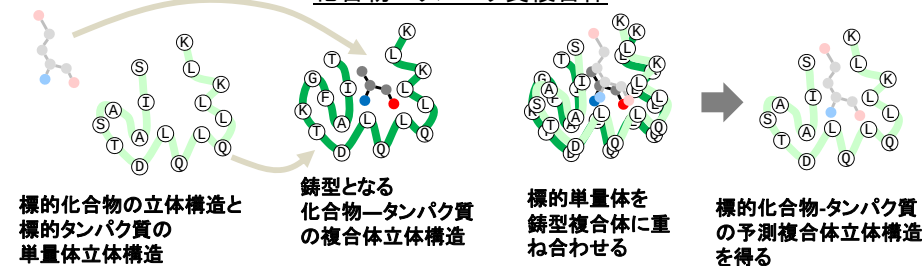


鋳型ベースのドッキング: Template-based Docking

タンパク質-タンパク質複合体



化合物-タンパク質複合体



HOMCOS: 複合体立体構造の検索・ホモロジーモデリングのサーバ

・PDB内の複合体の立体構造データを検索し、それを鋳型にモデリングする

・配列相同性検索はBLAST、化学構造類似性検索はKCOMBUを使用

“HOMCOS”でグーグル検索 ⇒ <http://homcos.pdbj.org>

サービス	HOMCOS	入力1	入力2
PDB内の結合分子の検索	タンパク質に対する結合分子の検索	アミノ酸配列	
	化合物に対する結合分子の検索	化合物構造	
複合体立体構造のホモロジーモデリング	ホモ多量体モデル	アミノ酸配列	
	ヘテロ多量体モデル	アミノ酸配列A	アミノ酸配列B
	化合物-タンパク質複合体のモデル	アミノ酸配列	化合物構造

タンパク質に対する結合分子検索

化合物-タンパク質複合体のモデル

アミノ酸配列と化合物構造を入力、アミノ酸配列はBLASTで、化学構造はKCOMBUでPDBに対して検索

予測構造

鋳型構造

予測構造

鋳型構造

タンパク質に対する結合分子の検索

問い合わせ配列: TGWVEIEINL...

BLAST検索

PDBに登録されたアミノ酸配列のデータベース

相同なタンパク質のリスト

2g9xA
1w98A
1fq1B
:

1vwg_1 A1 B1
2g9x_1 A1 B1
:

コンタクトしている分子の表

問い合わせ配列とコンタクトする分子の予測リスト

化合物に対する結合タンパク質の検索

問い合わせ化合物: C39, SHL, GBC

KCOMBU検索

PDBに登録された化合物のデータベース

類似した化合物のリスト

1vwg_1 A1 B1
2g9x_1 A1 B1
:

SHL
C39
GBC
:

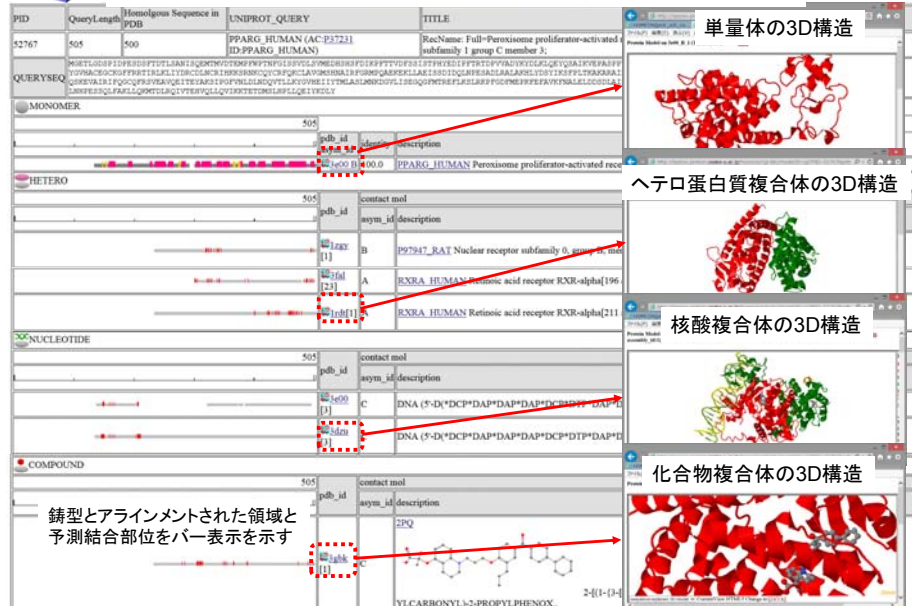
コンタクトしている分子の表

問い合わせ化合物とコンタクトする分子の予測リスト



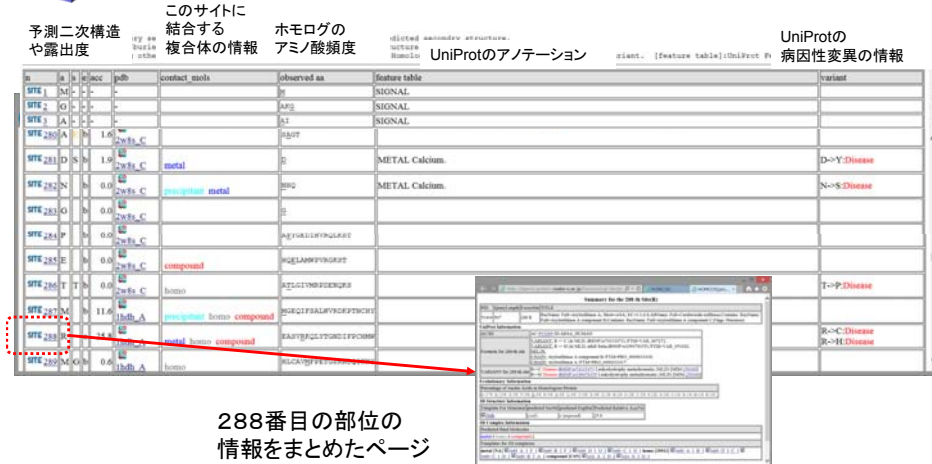
タンパク質に対する結合分子検索の結果

PPARG_HUMAN (Peroxisome proliferator-activated receptor gamma)



部位ごとの解析をまとめたSiteTable(ARSA)

ARSA_HUMAN (Arylsulfatase A) [セレブロシド-3-硫酸→セレブロシド+硫酸]

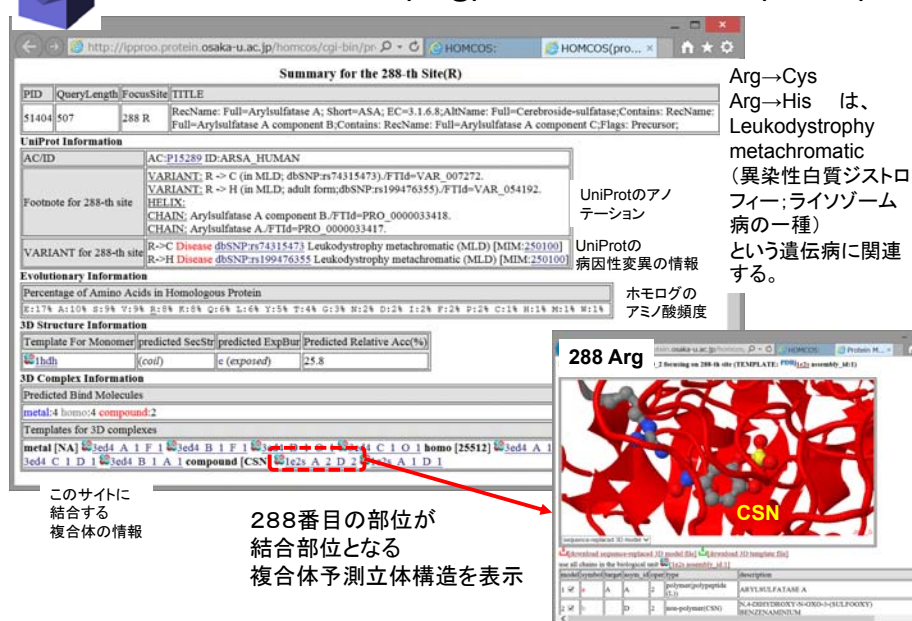


溶媒露出度: 溶媒露出度の小さい埋もれた部位への変異は、病変性であることが多いことが経験的に知られている。タンパク質の安定性を損ない、失活しやすくなるからとされる。

ホモログのアミノ酸頻度: 頻度の低いアミノ酸への変異は、病変性であることが多いことが経験的に知られている(SIFTスコアなどの動作原理)



288番目の部位(Arg)の解析のまとめ(ARSA)



ヘテロ蛋白質複合体のモデリング

配列Aと相同なタンパク質のリスト

問い合わせタンパク質A

配列 TGWVEIEINL...

or
単量体構造

BLAST検索

>1vwg_A
>1vwg_B
>2g9x_A

PDB内のアミノ酸配列のデータベース

結合している分子の表

問い合わせタンパク質B

配列 QLVVKTFAFT...

or
単量体構造

BLAST検索

1vwgB
2g9xB
2fi5I
2eufA

配列Bと相同なタンパク質のリスト

相同な複合体の一つを
鋳型として取り出す

鋳型構造: 1vwg_1 A1 B1

or

単量体の重ね合わせ

予測モデル構造

Template-based Model (Sequence-replaced model)

予測モデル構造

Template-based docking

ヘテロ多量体のモデリング(2本の配列から)

CDK5_HUMAN: Cyclin-dependent protein kinase 5
CCNB1_HUMAN: G2/mitotic-specific cyclin B1

複合体のsequence-replaced 3D model

sequence-replaced model
template 3D structure
の二つを、表示・ダウンロード可
Modeller用のスクリプトもダウンロード可

Sequence-replaced 3D model(簡易ホモロジーモデル構造):
鋳型構造と座標は同じ。残基名と残基番号を標的配列と入れ替えてある。側鎖原子や挿入残基は正しくモデリングされていない。

問い合わせ蛋白質の配列は以下の4通りで入力可
(i) PDB_ID+鎖 (ii) PDBファイルのアップロード
(iii) UniProt ID (iv) アミノ酸配列

配列Aと配列Bについて、対PDBのBLAST検索が実行される

※クエリとして2つの単量体の立体構造を入力して、Template-based Dockingを実行することもできる。

化合物-タンパク質複合体のモデリング

問い合わせタンパク質 BLAST検索

配列 TVAWGKTDLQL...
or
単量体構造

PDB内のアミノ酸配列のデータベース

SHL
C39
GBC

問い合わせ化合物構造 KCOMBU 検索

SHL
C39
GBC

問い合わせ化合物と類似した化合物のリスト

問い合わせ配列と相同なタンパク質のリスト

鋳型となる複合体構造の選択

SHL
C39
GBC

鋳型: 2g9x_A1 B1

配列の置き換え

単量体の重ね合わせ

予測モデル構造
Template-based Model
(Sequence-replaced model)

予測モデル構造
Template-based docking

※問い合わせ化合物はfkcombuを用いフレキシブルに鋳型化合物に重ね合わせる

化合物-タンパク質複合体のモデリング

問い合わせ蛋白質の入力
(i) PDB_ID+鎖
(ii) PDBファイルのアップロード
(iii) UniProt ID
(iv) アミノ酸配列

問い合わせ化合物の入力
(i) PDBの3文字コード
(ii) SMILESの文字列
(iii) SDF/MOL2/PDBファイルのアップロード

CDK3_HUMAN
IRE (iressa)

アミノ酸配列対PDBのBLAST検索、化合物 対PDBのKCOMBU検索が実行される

複合体のモデル構造が表示される

標的化合物(IRE) 鋳型化合物(DTQ)

※クエリとして2つの単量体の立体構造を入力して、Template-based Dockingを実行することもできる。

VaProS

創薬等の情報拠点で作成している統合データベース
<http://pford.info/vapro/>

タンパク質名、遺伝子名、アミノ酸配列を入力可

ゲノム 相互作用 発現
病気 パスウェイ 立体構造

3D Interaction

VaProSの 3D Interaction

HOMCOSの タンパク質に対する検索