

PSILO -アノテーション機能のご紹介-

PSILOはタンパク質立体構造データベースシステムです。公共データ、社内データ、モデリングデータなどのタンパク質立体構造データを統合的に管理します。Webブラウザからアクセスでき、実験研究者、構造解析研究者、モデリング研究者が、それぞれ必要に応じたデータを即座に参照することができます。ここでは、PDBデータには含まれない情報をPSILOが付与するアノテーション機能をご紹介します。

タンパク質立体構造の評価

一般的にPDBデータは安易に信頼される傾向がありますが、以下のプロットを比較、参照することで、X線結晶解析が専門でないユーザでも簡単に3D構造の妥当性を評価することができます。

◆主鎖二面角

座標情報から算出した主鎖(ϕ , ψ)二面角を表示し、不適切な残基に印を付けます。側鎖については、残基ごとのポテンシャルエネルギーとタンパク質の折りたたみに伴う脱溶媒和エネルギーを表示します。

◆密度相関と温度因子

電子密度データがある場合は電子密度図を描画(図1)するとともに、残基ごとの動きやすさの傾向や密度相関(2Fo-FcとFcの相関)を表示します。精度よく求められていない残基は、2Fo-FcとFcの差が大きくなります。Density Indexは、全原子の電子密度に対する各残基の2Fo-Fcの中の相対的な電子密度の大きさを表します。通常、温度因子が大きい原子はDensity Indexが小さいので、温度因子とDensity Indexが共に大きい残基は精密化が適切でないことを意味しています。

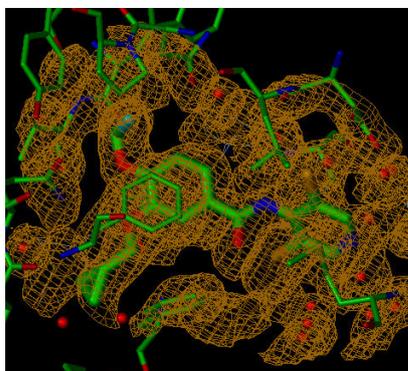


図1 PDB:1XOQのリガンド周辺の2Fo-Fcマップ

類縁タンパク質のアノテーション

PDBファイル内で指定された関連するタンパク質構造データ、BLAST検索による類縁配列、SCOPによるファミリー構造のアノテーションを

計算します。アノテーションはローカルサーバ内で計算するため、データが外部に漏れることはありません。全ての類縁構造についてリンクが作成されるので、1クリックで即座に類縁構造を参照することができます。また、SCOPデータベースへの外部リンクが作られるので、スーパーファミリーに関する情報も手軽に参照可能です。

InterProScanのアノテーション

InterProScanはアミノ酸配列についてGene Ontology、ファミリー、モチーフ、クラスなどの複数のデータベースに対して検索することで、タンパク質の特徴付けを支援します。PSILOはInterProScanアノテーション計算をあらかじめ行い、それぞれの公共データベースへの外部リンクを作成します。アノテーション計算はすべてローカルサーバ内で行われるので、社内データが外部に漏れることはありません。ユーザは面倒な個々のデータベースへの検索を外部に対して行う必要が無く、タンパク質から得られる出来る限りの情報を即座に参照することができます。

リガンドに対するアノテーション

低分子に対するアノテーションとして、PDBファイル内に書かれたシノニム、SMILES、InChiを表示します。InChiはPubChemなどの化合物検索サイトでクエリーなどに利用できます。

化合物の結合次数は、多くのPDBファイル内には書かれていませんが、PDBで公開されているcomponents.cifを参照することにより、正しいリガンド構造で登録しています。

まとめ

PSILOはPDBファイル内に書かれた全ての情報と、独自に収集する関連情報を合わせて保存し、ユーザフレンドリーなインターフェイスで表示します。生物情報学やX線結晶解析に精通してなくても、それらの情報を手軽に参照することが可能です。