

## 研究開発課題別中間評価結果

### 1. 研究開発課題名

蛋白質立体構造データベースの高度化

### 2. 代表研究者名

中村 春木（国立大学法人大阪大学蛋白質研究所 教授）

### 3. 研究開発概要

これまでに決定された蛋白質の立体構造は、蛋白質立体構造データベース（PDB）に整理・蓄積されてきた。PDBは地球規模で利用されているものの、構造生物学研究の成果をまとめあげただけのデータベースに留まってきた。そこで、本研究開発では、蛋白質の立体構造とゲノム情報との結びつきを強める一方、XMLなどの最新情報技術を利用することにより、国際的な連携のもとに世界標準としての新しいデータ記述（PDBML）と解析ツールや二次データベースを開発して付加価値を付け加え、単に構造生物学者だけを対象とする専門的なデータベースから、広く生命科学の研究者、産業界、さらには一般の人にも役立つデータベースに高度化する。

### 4. 中間評価

#### 4-1. 研究開発の進捗状況と今後の見込み

新たに国内外の組織を設立している。国内では、「日本蛋白質構造データバンク（PDBj:Protein Data Bank japan）」を設立している他、複数の大学の研究室をサテライト・サイトとし独自の研究開発体制を構築している。国際的には、米国RCSBや欧州EBIと共同で「国際蛋白質構造データバンク（worldwide Protein Data Bank:wwPDB）」を設立し、国際的な連携を強化させ、研究開発を推進している。

当初計画目標である、（1）新しい蛋白質立体構造に関するデータベースの構築、（2）蛋白質構造を基にした解析システムの開発と二次データベースの構築、（3）PDBデータベース業務の日本の分担作業の実施、（4）生体分子磁気共鳴データバンク（BioMagResBank:BMRB）の日本の分担作業の実施、（5）日本国内での教育用データベースの作成と公開、はいずれの項目も実施され、その結果はWebを通して公開されている。

当初想定されていた以上の進捗であり、今後の展開も非常に期待できる。

#### 4-2. 研究開発成果の現状と今後の見込み

Web公開状況 [PDBj http://www.pdb.j.org/index\\_j.html](http://www.pdb.j.org/index_j.html)

PDBjはアジア・オセアニア地区からのPDB登録データの受付・研修・公開作業を行っているが、PDBjにおける登録処理件数は年々大幅に増加し、平成15年度には世界全体の登録数の2割をしめるほどになり、PDBjの国際的な役割が増大している。また、BMRBのミラーサーバーの運営を平成14年より開始している。それに伴い日本語の登録マニュアルを作成し、日本の研究者が容易にデータ登録を行うことを可能とした。BMRBの有用性を高めるために登録促進活動が必要であると考えられる。

本研究開発では、蛋白質構造データのXML記述をいち早く進め、XML記述の世界標準（PDBML）を構築し、本技術を応用したXPathによる検索のWebサービスを公開・実現している。今後は階層構造化を可能とする新たなXML記述の開発が考えられている。また、PDBMLデータを基にした、native XML-DBによる蛋白質構造データ検索システム（xPSSS:xml-based Protein Structure Search Service）が開発され、世界に先駆けて公開されている。

他のデータベースや原著文献等より抽出した蛋白質構造解析実験及び蛋白質機能に関する情報が、XML記述した上でPDBMLに追加されている。これらの付加された情報は、xPSSS上にて公開されている。

蛋白質立体構造に関する新規二次データベースとしては、蛋白質表面形状と物性に関するデータベース（eF-site）や、3次元蛋白質分子グラフィックビューア（jV）、蛋白質ダイナミクス・データベース（ProMode）、立体構造比較ツール（ASH）、全自動分子置換法解析システム（MrPDB）、教育用蛋白質データベースである蛋白質構造百科（eProtS）等、数多くの開発が進められている。これらは今後さらなる改良が加えられる予定である。なお、本研究開発の成果はweb上での公開やNature Structural Biology等の国際雑誌への掲載、日本蛋白質学会を始

めとする様々な学会での発表等、広範に発表されている。

当初想定されていた以上の成果がでており、今後の成果も期待される。

#### 4-3. 今後の研究開発にむけて

日本・アジアから登録される蛋白質立体構造情報は今後さらなる増加が予測される。そのため、データベースの安全性管理をさらに厳密に行うとともに、アクセスの高速化を進めるなど、国際的な協力の基で高品質の蛋白質立体構造データ登録を継続・推進していくことが考えられている。

BMRBは環境が整ったので、ミラーサイトの運営を推進し、アジア地域におけるデータ登録を進めるとともに、Web上での登録用マニュアルの充実や、登録データの高度利用ツールの開発が計画されている。

階層性というXMLの特徴を活かした記述システムによる新たなデータベースを開発することや、他の研究機関が独自に開発したデータベースとのマッピングを図り互換性の改善に努めること、GUI等を用いてユーザーが利用しやすい環境を構築すること、などさらにデータベースの整備を進めることが計画されている。

本研究開発で開発された新規二次データベース及び解析ツールはより高度化された信頼性の高いシステムを目指している。中でもeProtSについては、対象データ数をさらに増やし、蛋白質の構造と機能についての豊富な情報源として確立し、より教育的な内容を発信することを目指している。

#### 5. 国内外のデータベース高度化・標準化の動向・状況と本課題の位置づけ

PDBのデータ量は、ここ4年間で約2倍になっている。PDBjによるアジア・オセアニア地域からの登録も年々増加し、世界全体の登録数の2割に達しており、PDBjの国際的な役割の重要性が増している。PDBjは、いち早くXMLの利用に着目し、その開発において主導的な役割を果たしてきた。本研究開発で開発されたその他の二次データベース及び解析ツールは、既存のデータベース等の弱点を改善し、いずれも有用なものとなっている。また、教育用蛋白質データベースである蛋白質構造百科(eProtS)は一般の人々を対象として、分かりやすく記述したデータベースとして、国内では唯一のものとなっている。

PDBjはwwPDBを構成する三極の一つとして認知されており、我が国唯一の国際拠点として重要である。日本は蛋白質の構造解析に特に力を入れており、タンパク3000プロジェクトとの連携も含め、今後の継続性が不可欠である。

#### 6. 研究開発成果の社会への貢献

本研究開発で構築された蛋白質立体構造データベースは、医・薬学を含む生物学系研究者に蛋白質の立体構造データを無料かつ利用しやすい形で一元的に提供することで、産学を問わず、社会に大きく貢献している。また、分子グラフィックス・ビューアや教育用データベースの開発・普及は、生体高分子の立体構造に関する一般の人々の啓蒙を目指しており、このことは社会への直接的な貢献と言える。

#### 7. 総合的評価

当初の想定を大きく上回る成果がでてきている。PDBjの設立とメンテナンスは大きな意義が認められ、wwPDBの一員として世界レベルで協力できるようになったのは特に注目すべき成果である。代表研究者のマネジメント努力も評価できる。蛋白質構造は生物機能との関係が深いので、今後生物学、生化学的分野のデータへのリンクを充実させ、引き続きデータベースを発展させることを期待する。

基礎研究および企業における研究開発に対する貢献度は大いに期待でき、また、本事業が確立した蛋白質構造データベースの国際拠点としての役割からみて、国家戦略として大変重要である。さらに体制を充実させて継続して行うべき事業である。