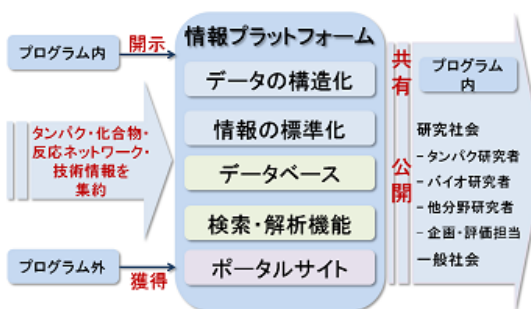


成果報告票(技術開発研究・代表機関のマネジメント)

課 題 名	ターゲットタンパク研究情報プラットフォームの構築運用
機 関 名	大学共同利用機関法人情報・システム研究機構国立遺伝学研究所
代 表 研 究 者 名	菅原 秀明

1. 課題開始時における達成目標

情報プラットフォーム（以下、情報 PF）は、ターゲットタンパク研究プログラム内のターゲットタンパク研究課題と技術開発3領域（生産・解析・制御）の研究拠点ならびに情報 PF 自身の成果情報を集約し、事務局機能と連動して、プロジェクト内での**情報共有**に加えて、プログラム外への適切な**情報発信**を実現する。右図に、情報 PF によるプログラム外の研究情報を含む研究情報の集約・獲得から、情報共有ならびに多様なコミュニティへの情報公開への流れをまとめた。



2. 平成23年5月末時点における課題全体の事業計画に対する研究成果

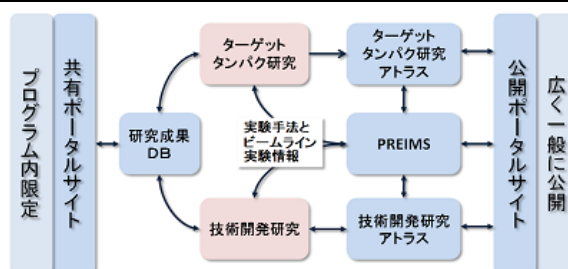
(1) 成果概要

情報共有：非公開情報も含む研究成果 DB、ビームタイム予約システム、ビームライン情報システムおよび共有ポータルサイトの構築運用によって実現した。評価のためのリストや指標を生成する機能も実現した。

情報発信：公開可能な研究成果を構造化・標準化して

集約したアトラス、PREIMS ならびに公開ポータルサイトの構築運用によって実現した。加えて、公開シンポジウム、産学懇談会、高校生ワークショップなどの開催、YouTube の利用、産学官向けのパンフレット、一般向けのリーフレット、ニュースレターなどの配布によって成果の普及に努めた。

タンパク質研究を支援する情報資源を情報 PF 自ら構築（右上図では「技術開発研究」のボックスに包含される）：日米欧のアミノ酸配列データベースを集約したデータベース（CASA db）、アミノ酸配列からタンパク質の構造・機能を推定するオンデマンド解析システム（FUJI db）、タンパク質-化合物相互作用の情報を集約したデータベース（PCI db）、ならびに有用なサイトを分類して短評を付けたリンク集（TP リンク）を構築し、D1 課題の成果であるタンパク質複合体構造推定（ToSY）と共に公開。



(2) 技術開発の進捗及び成果（前項の図参照。データ件数などの詳細は各機関の成果報告票参照）

1. 研究成果の集約

1) 研究の進捗を公開ポータルサイト (<http://www.tanpaku.org/>) に集約（H19年10月に運用開始）

- ・ 継続的に改良を加えてきたが現行サイトは、階層を浅くして多様なコンテンツの一覧性を高めた構造をとっている。3カラム構成の左カラムにターゲットタンパク研究課題紹介ページへの直行メニューを配置し、右カラムにサイト内検索、問い合わせ窓口、論文数/PDB 登録・公開数/特許申請数一覧表へのメニュー、技術開発課題の紹介ページへの直行メニュー、情報検索・解析メニューを配置し、幅の広い中央カラムにてプログラム内外の動向を紹介している。

- ・ 中央カラムの項目構成：プログラムの最新動向を紹介する「お知らせ・トピックス」；毎週水曜日に PDB を精査してプログラム由来の新規構造を集約している「TP 構造ギャラリー」；タンパク 3000 の成果を追跡調査した結果である「タンパク 3000 構造ギャラリー」；事務局への通知と新聞報道や拠点機関の Web サイトのチェック結果を併せて集約した「TP プレスリリース」；各ジャーナルの電子版などを日々確認して集約した各課題の発表論文を速報する「TP 論文オンライン」；構造生物学に関連する論文・記事を日々幅広く集約し選択して速報する「タンパクニュースウォッチ」
- 2) 非公開情報は共有ポータルサイト（公開ポータルサイトからログイン）に集約
- ・ 研究成果 DB：公開と非公開を問わずに各課題の成果を登録・蓄積・検索・閲覧・集計可能とするシステム群の総称。利用者権限に応じて研究実施者による登録が可能であるが、情報 PF による代行登録も行った。H22 年度の内部評価の際には、情報 PF が研究成果 DB から論文リスト、特許出願リストおよび PDB エントリーリストを用意して各課題にチェックを依頼した。
 - ・ ビームタイム管理システム：SPring-8 ならびに Photon Factory と共同で、両施設のビームラインの一括申請・審査・承認システムを構築し平成 21 年 4 月以来安定に運用してビームタイム割り振りの効率化に貢献した。
 - ・ ビームライン情報システム：SPring-8 ならびに Photon Factory と共同で両施設のビームラインでの実験情報を対象とする LIMS を実現した。実験情報の記述を構造化・標準化して研究成果 DB にファイルとして登録し、次項の PREIMS にも送信して実験情報の再利用を可能にした。
 - ・ ビームタイム管理システムとビームライン情報システムはこれまでプログラム内公開であったが、今後のプログラム外への展開を視野に入れて、PREIMS と研究成果 DB（進捗 MS）との間における情報の流れと各システムの操作の理解を助ける動画マニュアルを作成し YouTube の TargetTanpaku チャンネル (<http://www.youtube.com/user/TargetTanpaku?gl=JP>) で公開した。
- 3) 研究成果の詳細を DB に集約
- ・ 蛋白質実験情報マネジメント・システム (PREIMS: PRotein Experimental Information Management System)：技術開発研究課題と連携して、クローニングから構造と機能の解析同定までのフローの各フェーズにおける実験プロトコルを集約した。オントロジーを構築して標準の形式でノウハウまで含む詳細情報を集約した PREIMS によって、技術開発研究の成果である先端の実験手法の普及を加速することを期待できる。また、実験プロトコルに加えて、前項のビームライン情報システムからネットワーク経由で取得した実験情報を DB 化し、実験プロトコルと実験情報の相互参照を可能とした。PREIMS は、H21/3～H23/5 の間に 68,669 ページ閲覧された。
 - ・ eSOL (Solubility database of all *E.coli* proteins)：再構築型の試験管内タンパク質合成系を用いて大腸菌の全てのタンパク質を発現させた際の凝集の度合い（可溶率）と合成量を集約した（生産領域 PPC1 課題上田研究室と共同で開発）
 - ・ オートファジーDB：オートファジーのタンパク質として実験で検証済みのタンパク質とオーソログならびにホモログを収録し、関連論文を日々追加更新している（基本的生命 FBA3 課題ならびにプログラム外のオートファジー研究グループの協力を得て開発）
 - ・ トリパノソーマ DB: トリパノソーマに関するタンパク質ならびに阻害剤を集約し、関連論文を日々更新している（医学薬学 MPA5 課題と共同で構築）
 - ・ ネットワーク DB：ターゲットタンパク研究 35 課題の協力を得て構築した DB。各研究対象の反応経路を Cell Illustrator で描いた graphical summary、研究課題の背景、特徴、実施者などをテキストで記述した general summary ならびに経路に登場するタンパク質と化合物や、関連論文の情報を統

合した tabular summary で構成した。研究の進展に応じて随時改訂してきたが、DB の内容をより良く表現するために、この DB を平成 22 年 10 月以来ターゲットタンパク研究アトラス (TP Atlas) と呼ぶことにし、その動画マニュアルを作成して YouTube の TargetTanpaku チャンネルから公開した。その後、同じ枠組みを技術開発研究課題に適用した技術開発アトラス (Advanced Technologies Atlas、以下 AT Atlas) の構築も進めている (年内公開予定)。

4) 統合検索システムの構築

平成 22 年度までに構築した DB 群を平成 23 年度に統合検索可能とした。統合検索機能に関連して、前述第 2) 項の成果にならって、情報 PF 全体における情報の流れと各サブシステムの利用法の理解を助ける動画マニュアル作成を進めている (年内公開予定)。

2. 研究成果の発信

公開ポータルサイトの利用は年々増加している。平成 22 年度の年間訪問者数 (ユニーク IP 数) / 訪問数が 23,166 人/81,828 回に達し、平成 23 年 5 月の月間訪問者数/訪問件数は、計画停電期間の落ち込みから回復して過去最高となり 3,500 人/8,000 件を超えた。プログラムの研究実施者数がおよそ 700 名であることから、プログラム外からのアクセスが多数を占めていると推定できる。また、IP アドレスから判断すると、その約 60% が産業界など、40% 程度がアカデミアや政府機関である。

平成 19 年度から毎年行っている公開シンポジウムの参加者数も回を追って増加し、平成 21 年度は会場の限界に達した (564 名)。参加者のほぼ 50% が企業や一般の参加者である。なお、平成 22 年度は東日本大震災のためシンポジウム途中で解散した。

3. 研究支援情報資源の整備

- 配列 DB (CASA: Comprehensive Amino Acids Sequences Annotated database) : UniRef、NCBI NR、DDBJ DAD などから網羅的に取得したアミノ酸配列 2,000 万件を超える和集合に対して、種々のデータベースを参照してアノテーションを付与した。次項 FUJI DB と統合して、既知アミノ酸配列に基づく迅速アノテーションとアミノ酸配列解析による新奇配列へのアノテーション付与を実現した (<http://sqpr-web.genes.nig.ac.jp/casafuji/casafuji.html>)。
- 構造・機能 DB : 利用者が与える任意のタンパク質のアミノ酸配列に対し、配列の類似性のみならず立体構造情報も考慮して機能予測を行い、結果を 2 次元ならびに 3 次元表示する。配列の特徴分析に加え立体構造や天然変性 (intrinsically disordered) 領域も高速で予測する。オンデマンド解析 (FUJI: Functionally Annotated Japanese Protein Structural Information database) として公開している。
- TP リンク : タンパク質の配列、構造、解析技術等に関するサイトと、タンパク質の構造や相互作用等を解析するツールのサイトに短評を付け、サイトが提供する機能の観点で分類整理して提供。また、随時サイトの追加更新をしている。
- 化合物 DB : 研究対象となっているタンパク質と相互作用する化合物の発見を支援するタンパク質 - 化合物相互作用 DB (PCI DB) の観点で構築した。PubChem (Bioassay のデータ)、DrugBank、Comparative Toxicogenomics DB および ChEMBL を集約し、タンパク質と化合物の双方向検索を実現した。タンパク質ファミリーや化合物の構造情報からの検索の年内実現を予定している。
- 複合体構造推定 : アミノ酸配列からタンパク質単体の構造を予測し、タンパク質単体の立体構造から複合体の構造を推定する (情報課題 D1 の成果を平成 22 年 9 月公開ポータルサイトから公開)。

3. タンパク質の構造解析件数

PDB 登録数	該当なし
---------	------

情報C 1 代表機関のマネジメント

PDB 登録の有無にかかわらず構造解析したタンパク質数	CASP9 (9 th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction) におけるホモロジーモデリング数 126 件
4. 課題全体の論文発表件数	12 件
5. 各機関の公開したツールやデータベース等の件数	41 件【詳細は各機関の成果報告票に記載】
6. 課題全体の特許出願件数	該当なし
7. 課題全体の成果を活用した共同研究数 (プログラム内外含む)	87 件 (プログラム内 66 件、プログラム外 21 件)【詳細は各機関の成果報告票に記載】
<p>8. 当初計画に対する達成度</p> <p>情報の共有と公開のために企画した情報システム群をプログラムの展開に即して構築・運用した。また、情報共有、情報公開ならびに情報検索・解析のシステムはコミュニティーにとって有用な情報資源となりつつある。事実、公開サイトの訪問者数が増加を続け、また、第3者が Wikipedia に記載したオートファジーDB はその後急速にアクセスが増加している。これは言い換えると、各システムの利用をのばす余地がまだまだ残っているとも言える。総合して、当初計画の 90% を達成したと判断している。</p>	
<p>9. 課題内の情報共有・連携体制</p> <p>構築運用の方向性、情報技術などの課題内情報共有と連携は、頻繁な電子メールのやりとりと、ビデオ会議も利用した情報 PF 班全体会議や情報 PF 運営委員会で実現した。</p> <p>【具体的連携の事例】事務局、日立製作所、遺伝研と共同で、研究課題と研究実施者の DB を構築運用；遺伝研と事務局が共同して公開ポータルサイトを構築運用；遺伝研によるチームタイム管理システムの構築と事務局による運用；シングルサインオンを含むポータルサイト構築、CASA、FUJI、AT Atlas ならびにチームライン情報システム構築における阪大との連携；旧課題 D1 と共同で TP リンクを整備し、また D1 の研究成果を公開ポータルサイトから公開。</p> <div data-bbox="815 981 1444 1317" style="float: right; border: 1px solid black; padding: 5px;"> <p>分担機関はそれぞれの専門性を活かしたサービスを開発提供</p> <p>遺伝研 研究成果データベースや網羅的データベースを構築提供 高速簡易な解析システムを開発提供</p> <p>阪大蛋白質研 実験プロトコルデータベースを開発提供</p> <p>事務局(東大・産) 各種イベント企画運営、広報資料作成、共有ポータルサイト運用など</p> <p>情報D1課題と連携 構造バイオインフォマティクス最先端技術の提供</p> <p>中核機関遺伝研の役割 ・中核・分担機関ならびに外部の情報資源を組合わせて利用できる情報環境を構築提供 ・プログラムの研究成果をアーカイブし各種指標を整備提供</p> <p>TP Atlas ターゲットタンパク質研究プログラム</p> <p>http://www.tanpaku.org/</p> </div>	
<p>10. 他の課題との情報共有・連携</p> <p>【詳細は各機関の成果報告票を参照】</p> <ul style="list-style-type: none"> ターゲットタンパク質研究ならびに技術開発研究の各課題と協力して各種データベースを構築した (例 TP Atlas、AT Atlas、PREIMS など)。 プログラム外の研究グループとの情報共有・連携によってツールやデータベースを構築した (例 オートファジーDB、PREIMS など) 多くのツール・データベースにおいてプログラム外のツール・データベースとの連携を実現した。 プログラム内外のモデル構造解析に協力した。 	
<p>11. 人材育成</p> <p>プログラム各課題の研究成果ならびに構造生物学一般の成果を集約して多様な利用者に提供するキュレーションあるいはタンパク質の構造と機能の解析に関する知識・技術・経験を持って、創薬等支援</p>	

情報C 1 代表機関のマネジメント

技術基盤プラットフォームに貢献することを期待できる。

1 2. 終了までの具体的な見通し

- ・ ターゲットタンパク研究課題からの研究成果開示が極大に達する最終年度において、情報共有と情報発信を円滑に進めるために、これまで以上に知的集約作業を行う。
- ・ 広く一般のタンパク質関連研究の情報基盤として、これまでに構築した DB のコンテンツと情報検索・解析システムの機能を更新・拡充する。
- ・ 学会発表や学術雑誌への発表によって情報 PF の利用拡大を図る。

1 3. 本プログラムにおける成果が我が国の科学技術発展及び産業応用等に与える効果について

- ・ 情報 PF は 2 つの面で科学技術発展と産業応用に貢献しうる。第 1 に、タンパク質研究にかかる最先端技術と研究成果を届けるメディアとして機能する。特に、当該分野への新規参入者や、当該分野の周辺領域の研究開発にとって有用である。第 2 に、タンパク質化合物相互作用探索や複合体構造予測などの情報解析の精度を上げていくことで、研究開発の効率化をもたらすことができる。
- ・ 情報 PF の仕組みならびに集約した公開可能なコンテンツは広く一般の利用に供する。一方で、クレジットを確保すべく必要に応じて Creative Commons のライセンス表示を活用する。

1 4. 特記事項：本稿で【平成 21 年度中間評価（外部評価）結果】への対応状況を述べる

【情報共有・公開についてプログラム全体で取り組むべき】推進委員会、運営委員会、交流会で議論
 【偏ったシステム開発を避けるべき・利便性の検証が必要】上記委員会や情報 PF 運営委員会での議論ならびに各種学会のポスター発表などでの情報収集を活かしながら開発を進めた。ターゲットタンパク研究の全ての課題についてアトラスを作成・公開し技術開発研究課題についても全課題についてアトラスを作成（年内公開予定）。各 DB や解析システムの訪問者数と訪問件数も見ながら機能向上を検討。

【予算規模の妥当性】平成 22 年度に班を再構成して予算圧縮。また、サービスの優先順位を検討して大規模データや高速解析に必要なサーバ増強を平成 21 年度で停止。

【PREIMS に対する指摘と技術利用支援の強化】一般の利用者が、SAIL 法等の技術開発研究領域の成果を検索・表示できるように、分かり易いシステムへと改良した。特に、日本語による検索を実現し、アニメーションを活用したマニュアルを作成。データ登録についても、実験研究者からのヒアリング結果に基づいて簡易化を進めた。利用支援については、PREIMS へのリンクを含む AT Atlas（年内公開予定）によって、入門的パンフレットと高度に専門的な PREIMS のギャップを埋めることを目指している。

【開発した DB などの開示】Creative commons のライセンスに則って開示を進めている。

【個別マネジメントシステム（LIMS）の実現】SPring-8 および Photon Factory において行われた X 線結晶構造解析実験の結果をネットワーク経由で取得した後に、標準フォーマットへと構造化してアーカイブし、オンラインで再利用可能とした。

1 5. 研究費一覧

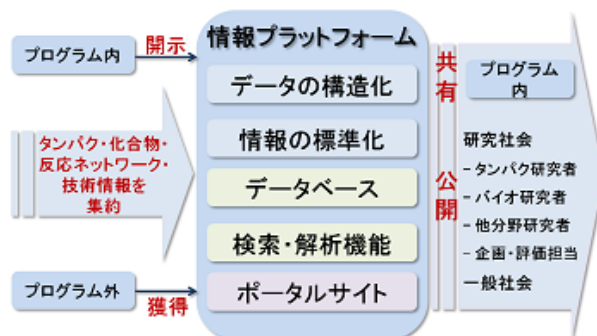
	19年度	20年度	21年度	22年度		23年度	計
設備備品費（千円）	69,060	16,625	10,791	2,238	物品費（千円）	2,249	/
試作品費（千円）	0	0	0	0	人件費・謝金（千円）	80,348	
人件費（千円）	59,444	90,135	82,201	84,850	業務実施費	122,320	
業務実施費（千円）	181,660	140,252	137,712	143,130			
間接経費（千円）	72,554	59,556	54,665	54,767			
合計（千円）	382,718	306,568	285,369	284,985	合計（千円）	204,917	

成果報告票(技術開発研究・各機関)

課 題 名	ターゲットタンパク研究情報プラットフォームの構築運用
機 関 名	大学共同利用機関法人情報・システム研究機構国立遺伝学研究所
代 表 研 究 者 名	菅原 秀明

1. 課題開始時における各機関の達成目標

国立遺伝学研究所（以下、遺伝研）は、ターゲットタンパク研究プログラム内のターゲットタンパク研究課題と技術開発3領域（生産・解析・制御）の研究拠点ならびに情報 PF 自身の成果情報を集約し、事務局機能と連動して、プロジェクト内での**情報共有**と、プログラム外への適切な**情報発信**を実現する情報プラットフォーム（以下、情報 PF）を構築運用する。右図に、情報 PF によるプログラム内外の研究情報の集約から提供までの流れを示す。



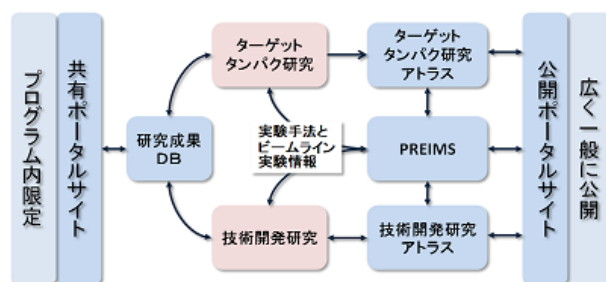
2. 平成23年5月末時点における各機関の事業計画に対する研究成果

(1) 成果概要

情報共有：非公開情報も含む情報共有を、共有ポータルサイト、研究成果 DB、ビームタイム予約システム、ビームライン情報システムによって実現した。また、評価のためのリストや指標を生成する機能を実現した（右下図 PREIMS は阪大の分担課題）。

情報発信：公開ポータルサイトと研究成果を構造化・標準化して集約したデータベース（以下、DB）群の構築運用、各種学会・公開シンポジウム・産学懇談会・講習会における講演、高校生を対象とするワークショップ実施、YouTube への TargetTanpaku チャンネルの設置などによって、情報発信に努めた。

タンパク質研究を支援する情報資源を情報 PF 自ら構築（右下図では「技術開発研究」のボックスに含まれる）：日米欧のアミノ酸配列 DB を集約した CASA db、アミノ酸配列のオンデマンド解析システム FUJI db、タンパク質-化合物相互作用の情報を集約した PCI db、ならびに有用なサイトを分類して短評を付けたリンク集（TP リンク）を構築・公開した。



(2) 技術開発の進捗及び成果（項目（1）の図を参照、テキスト中で【アルファベット】をふった事項については別紙に補足説明図または表を掲載）

1. 研究成果の集約

1) 研究の進捗を公開ポータルサイト (<http://www.tanpaku.org/>) 【A】に集約（H19年10月開始）

- ・ お知らせ・トピックス（H19年10月開始）：プログラムからのお知らせ累計90件掲載
- ・ TP 構造ギャラリー（H21年11月開始）：毎水曜日に PDB を確認してプログラム由来の PDB エント

リー累計 276 件を集約・速報。タンパク 3000 の成果については追跡調査をして PDB エントリー 3,677 件を構造ギャラリーに集約し、痛風・高尿酸血症治療薬に結実した複合体構造解析を含む重要な研究成果 53 件をリストアップ。また、生物種、酵素分類および解析方法の統計情報も表示【B】。

- ・ TP プレスリリース (H20 年 6 月開始) : 事務局への通知と新聞報道や拠点機関の Web サイトのチェック結果を併せて、プログラム由来のプレス発表 49 件を集約・速報
- ・ TP 論文オンライン (H20 年 12 月開始) : 各ジャーナルの電子版などを日々確認して、プログラム由来の発表論文累計 536 件を集約・速報。
- ・ 構造生物学に関連する論文や記事を日々集約・選択してタンパクニュースウオッチとして速報。
- ・ 論文、PDB エントリーならびに特許申請の動向を表示【C】

2) 非公開情報は共有ポータルサイト【D】(公開ポータルサイトからログイン) に集約

- ・ 研究成果 DB : 課題と研究実施者の DB に設定された利用者権限に応じて操作を制御することで、非公開情報も含めて各課題の成果を登録・蓄積・検索・閲覧・集計可能とするシステム群の総称【E】。H22 年度の自己点検の際には、研究成果 DB から出力された論文・PDB エントリー・特許出願のリストを各課題がチェックして報告するサイクルが動いた。
- ・ ビームタイム管理システム : SPring-8 ならびに Photon Factory と共同で、両施設のビームラインの一括申請・審査・承認システムを構築し平成 21 年 4 月の募集から運用した。H21 年度 77 件、H22 年度 164 件のビームタイム利用申請の効率的調整に貢献し、新規ビームライン追加にも対応した。
- ・ ビームライン情報システム : 前述両施設における実験情報を XML として構造化し、オンラインで研究成果 DB に登録しかつ分担機関阪大が開発運用している PREIMS に送信する LIMS を構築した。
- ・ ビームタイム申請からビームライン実験情報の取得までの情報の流れと関連システムの利用法を動画マニュアルとして表現し YouTube の TargetTanpaku チャンネルから公開【F】
- ・ セマンティック Web 技術を応用して、多様な切り口でのプロジェクト関係図を生成可能にした(年内に共有サイトで公開予定)【G】

3) 研究成果の詳細を DB に集約

- ・ eSOL (Solubility database of all *E.coli* proteins) : 再構築型の試験管内タンパク質合成系を用いて発現させた大腸菌全タンパク質の可溶率と合成量を集約(生産領域 PPC1 課題上田研究室と共同で開発)
- ・ オートファジー DB : 真核生物 81 種のオートファジーに関わる 7,444 タンパク質を収録し、種間の対応関係、配列・立体構造・機能・天然変性領域などの情報を付加し【H】、機能グループごとの閲覧を可能にした。プログラム外データとともに更新継続。(基本的生命 FBA3 課題の協力を得て開発)
- ・ トリパノソーマ DB : トリパノソーマに関するタンパク質 17 種と 38,117 編の論文書誌情報を集約。阻害剤の検索も実現。プログラム外データとともに更新継続。(医学薬学 MPA5 課題と共同で構築)
- ・ ネットワーク DB : ターゲットタンパク研究 35 課題の協力を得て構築した DB。各研究対象の反応経路を Cell Illustrator で描いた graphical summary、研究課題の概要をテキストで記述した general summary ならびに経路に登場するタンパク質と化合物の情報を統合した tabular summary で構成。平成 22 年 10 月からは ターゲットタンパク研究アトラス (Targeted Proteins Research Atlas、以下 TP Atlas) と呼ぶことにした。平成 23 年度には TP Atlas の動画マニュアルを作成して YouTube の TargetTanpaku チャンネルから公開した【I】。その後、同じ枠組みを技術開発研究課題に適用した 技術開発アトラス (Advanced Technologies Atlas、以下 AT Atlas) の構築も進めている【J】(年内公開)。

4) 統合検索システムの構築

平成 22 年度までに構築したサブシステム群を対象とする一括検索システム【K】の開発を進めてお

り、年内に一般公開する予定である。本システムには検索支援として、ライフサイエンス辞書を組み込むことで利用者が当初入力したキーワードを変更あるいは展開することを可能にした。

2. 研究成果の発信

公開ポータルサイトの利用は年々増加している。平成22年度の年間訪問者数（ユニークIP数）/訪問数が23,166人/81,828回に達し、平成23年5月の月間訪問者数/訪問件数は、計画停電期間の落ち込みから回復して過去最高となり3,500人/8,000件を超えた。プログラムの研究実施者数がおよそ700名であることから、プログラム外からのアクセスが多数を占めていると推定できる。また、IPアドレスから判断すると、その約60%が産業界など、40%程度がアカデミアや政府機関である。

3. 研究支援情報資源の整備

- ・ 配列DB (CASA: Comprehensive Amino Acids Sequences Annotated database) : UniRef、NCBI NR、DDBJ DAD などから網羅的に取得し非冗長化したアミノ酸配列にその時点で最新の参照DBを使ったアノテーションを付加。平成21年度版で1,300万件超のDBを、平成22年度版で2,100万件余りのDBを公開【L】。およそ300万件の新規配列を含む平成23年度版を年度内に公開予定。
- ・ 構造DB : 任意のアミノ酸配列に対して立体構造の情報のみならず立体構造を利用した機能予測も短時間で提供する類例のないオンデマンド解析 (FUJI: Functionally Annotated Japanese Protein Structural Information database) として平成21年3月に公開【M】。その後、単独で一定の立体構造を取る領域と、単独では決まった構造を取らない天然変性領域を高精度で分割する機能も付加。
- ・ 配列DB (CASA db) と構造DB (FUJI db) を融合させた CASA-FUJI db も開発し提供。融合dbでは、CASA db に登録されているアミノ酸配列に相当するタンパク質に対してはCASA dbからの情報を直ちに表示するとともに、FUJI db によって最新情報を付加しその結果を表示する。CASAS db 未登録のアミノ酸配列をもつタンパク質に対しては、FUJI db の解析結果を提供する。本システムは <http://sqpr-web.genes.nig.ac.jp/casafuji/casafuji.html> で公開しているが、年内に公開ポータルサイト情報検索・解析のメニューとして組込む。
- ・ TPリンク : タンパク質の配列や構造、解析技術等のDB 76件のサイトとタンパク質の構造や相互作用等を解析するツール提供サイト 170件を機能の観点で分類整理し、それぞれに短評を付けたリンク集。平成22年度には新規サイトの追加とともに、サイト全件を手動で再確認して、URLアドレスの変更や短評に必要な修正を加えた。
- ・ タンパク質-化合物相互作用DB (Protein-Compound Interaction (PCI) db) : 化合物ライブラリのDBに替えて、一般公開されている大規模あるいは特徴がある化合物関連DBからPCIの情報を集約して、タンパク質と化合物の双方向検索を可能とした。対象DBは、PubChem (化合物のBioActivityスクリーニングのアッセイ結果)、DrugBank (FDA認可済み薬剤などとターゲットタンパク質の情報)、CTD (Comparative Toxicogenomics Database : 化合物-遺伝子/タンパク質-疾患の関係を文献から専門家が抽出) ならびに ChEMBL (生理活性化合物のデータ) である【N】。また、KEGG DB (京都大学) と MeSH DB (米国NCBI) の情報も組込んだ。集約対象としている4種のDBのバージョンアップに対応して随時更新している。

3. タンパク質の構造解析件数

PDB 登録数	該当なし
PDB 登録の有無にかかわらず構造解析したタンパク質数	該当なし

情報C 1 - A

4. 各機関の論文発表件数	7 件
5. 各機関の公開したツールやデータベース等の件数	23 件
6. 各機関の特許出願件数	該当なし
7. 各機関の成果を活用した共同研究数（プログラム内外含む）	63 件（プログラム内 55 件、プログラム外 8 件）
8. 当初計画に対する達成度	<p>情報の共有と公開のために企画した情報システム群をプログラムの展開に即して構築・運用した。また、情報共有、情報公開ならびに情報検索・解析のシステムはコミュニティーにとって有用な情報資源となりつつある。事実、公開サイトの訪問者数が増加を続け、また、第 3 者が Wikipedia に記載したオートファジーDB はその後急速にアクセスが増加している。これは言い換えると、各システムの利用をのばす余地がまだまだ残っているとも言える。総合して、当初計画の 90% を達成したと判断している。</p>
9. 課題内の情報共有・連携体制	<p>構築運用の方向性、情報技術などの課題内情報共有と連携は、頻繁な電子メールのやりとりと、ビデオ会議も利用した情報 PF 班全体会議や情報 PF 運営委員会で実現した。</p> <p>【具体的連携の事例】事務局、日立製作所、遺伝研と共同で課題と研究実施者の DB を構築運用；遺伝研と事務局が共同して公開ポータルサイトを構築運用；遺伝研によるビームタイム管理システムの構築と事務局による運用；シングルサインオンを含むポータルサイト構築、CASA、FUJI、AT Atlas ならびにビームライン情報システム構築における阪大との連携；旧課題 D1 と共同で TP リンクを整備し、また D1 の研究成果を公開ポータルサイトから公開。</p>
10. 他の課題との情報共有・連携	<p>【プログラム内課題との情報共有・連携】</p> <ul style="list-style-type: none"> ・ SPring-8 ならびに Photon Factory と共同で、ビームタイム予約システムを構築運用し、また、ビームライン情報システムを構築して、X 線結晶解析実験結果データをオンラインで情報 PF へとアーカイブ可能に。 ・ ターゲットタンパク研究 35 課題と連携して、TP Atlas を構築運用 ・ 技術開発研究分野と連携して、生産領域が開発した 11 件の実験プロトコル、解析領域の 3 種類の実験手法ならびに制御領域の AT Atlas を構築し、AT Atlas と PREIMS の連携も実現（年内完成予定） ・ 生産領域 PPC1 課題上田研究室と共同で大腸菌全タンパク質の可溶率 DB である eSOL を構築運用 ・ 基本的生命 FBA3 課題と共同でオートファジーDB を構築運用 ・ 医学薬学 MPA5 課題と共同でトリパノソーマ DB を構築 ・ PCI の機能を株式会社ファイルデザイン（制御領域の分担機関）と連携しながら設定した。 <p>【プログラム外との情報共有・連携】</p> <ul style="list-style-type: none"> ・ 産業総合研究所の野口保副センター長の協力を得て、CASA db と Poodle の連携を実現 ・ DBCLS/NBDC の畠中秀樹研究員により、eSOL が TogoProt (http://lifesciencedb.jp/togoprot/) の検索対象に包含された。 ・ 前橋工科大学の福地研究室と共同で、天然変性領域と構造領域を正確に予測できる DICHOT を FUJI db に組込んだ。 ・ 理研サイネスの蛋白質統合データベースと情報 PF のタンパク 3000 構造ギャラリーなどとの連携を

情報C1-A

実現（年内公開予定）

- ・ オートファジーDB は、プログラム外の大阪大学吉森保研究室、東京医科歯科大学水島昇研究室、前橋工科大学の福地研究室とも共同して構築した。

1.1. 人材育成

プログラム各課題の研究成果ならびに構造生物学一般の成果を集約して多様な利用者に提供するキュレーションあるいはタンパク質の構造と機能の解析に関する知識・技術・経験を持って、創薬等支援技術基盤プラットフォームに貢献することを期待できる。

1.2. 終了までの具体的な見通し

- ・ ターゲットタンパク研究課題からの研究成果開示が極大に達すると最終年度において、情報共有と情報発信を円滑に進めるために、これまで以上に知的集約作業が行う。
- ・ 広く一般のタンパク質関連研究の情報基盤として、これまでに構築した DB のコンテンツと情報検索・解析システムの機能を更新・拡充。
- ・ 学会などでの発表や動画マニュアルの拡充と YouTube 掲載などによって情報 PF の利用拡大を図る。

1.3. 本プログラムにおける成果が我が国の科学技術発展及び産業応用等に与える効果について

- ・ 情報 PF は2つの面で科学技術発展と産業応用に貢献しうる。第1に、タンパク質研究にかかる最先端技術と研究成果を届けるメディアとして機能する。特に、当該分野への新規参入者や、当該分野の周辺領域の研究開発にとって有用である。第2に、構造・機能予測やタンパク質化合物相互作用探索などの情報解析の精度を上げていくことで、研究開発の効率化をもたらすことができる。
- ・ 情報 PF を先駆けとしてプロジェクト運用管理のデジタル化が進展することを期待する。学会発表申込や学術論文投稿についてはすでに、審査を経て発表・出版までインターネットを介して一貫してデジタル処理される場合が増えている。各種プロジェクトについても、PC で作成したデジタルファイルを印刷物で報告し、印刷物から情報を抽出してデータベース化するといった迂回をすることなく、募集・審査・開始・終了までの報告などをデジタルの世界で完結させることによって、プロジェクトの効率化と成果の迅速開示を実現できる。

1.4. 特記事項

- ・ 研究成果の新しい表現方法を提示した（TP Atlas, AT Atlas, プロジェクト関係図）
- ・ 大規模プロジェクトの運営支援システムを構築した（課題・実施者 DB、研究成果 DB など）
- ・ 任意のアミノ酸配列に対して構造・機能予測とともに新奇構造領域と天然変性領域の範囲を短時間で提供するシステムを構築した（CASA-FUJI db）

1.5. 研究費一覧

	19年度	20年度	21年度	22年度		23年度	計
設備備品費（千円）	50,493	13,818	9,975	1,520	物品費（千円）	2,000	
試作品費（千円）	0	0	0	0	人件費・謝金（千円）	0	
人件費（千円）	11,486	30,566	31,236	24,374	旅費（千円）	24,424	
業務実施費（千円）	44,175	54,986	48,786	88,106	その他（千円）	83,576	
間接経費（千円）	31,846	29,810	26,999	34,200	間接経費（千円）	0	
合計（千円）	138,000	129,180	116,996	148,200	合計（千円）	110,000	642,376

成果報告票（技術開発研究・各機関）

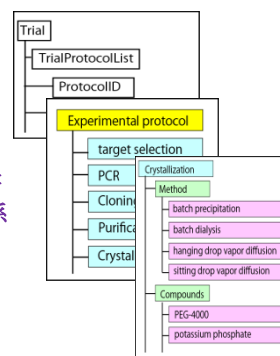
分担機関の課題名	ターゲットタンパク研究情報プラットフォームの構築運用 (情報マネジメントシステム: タンパク質実験情報の収集と品質管理)
分担機関	国立大学法人大阪大学
機関の代表研究者名	中村 春木

1. 課題開始時における各機関の達成目標

実験プロトコル・データに関してオントロジーを用いた標準化作業を実施し、基本的な実験プロトコルを収集・登録を行い、データベースを構築することを目的とする。このデータベースを検索・閲覧できるポータル (PREIMS: Protein Experimental Information Management System) の構築・公開を行う一方、ユーザがオフラインで簡易に実験プロトコル・データを登録・修正するためのツール (TPPE: Target Protein Protocol Editor) を開発する。こうして、ターゲットタンパク研究に携わる研究者が開発した実験プロトコルの登録を進め、研究プロジェクト内での実験プロトコル情報を共有するとともに、開示可能なものは外部へも発信する。一方、LIMS 開発として、遺伝研と協力し、自動化が進んでいる放射光 X 線結晶回折実験のログファイルを PREIMS へインポートする仕組みを開発し、SPring-8 と Photon Factory における実験結果データに対して、オントロジーで整理されたデータベース構築を行う。

タンパク質実験データ MS (申請時の説明図から)

- 1) 実験データ記述の標準化
 - 1-1) 発現、精製、構造決定、制御等の実験データの辞書
 - 1-2) オントロジーによる関係性の定義
- 2) データの検証 (Validation) による品質管理
- 3) データ単位のアクセス制御



こうして、ターゲットタンパク研究に携わる研究者が開発した実験プロトコルの登録を進め、研究プロジェクト内での実験プロトコル情報を共有するとともに、開示可能なものは外部へも発信する。一方、LIMS 開発として、遺伝研と協力し、自動化が進んでいる放射光 X 線結晶回折実験のログファイルを PREIMS へインポートする仕組みを開発し、SPring-8 と Photon Factory における実験結果データに対して、オントロジーで整理されたデータベース構築を行う。

2. 平成23年5月末時点における各機関の事業計画に対する研究成果

(1) 成果概要

① タンパク質実験プロトコルのオントロジーの開発と実験プロトコル・データの収集

タンパク質実験プロトコルを標準的に記述するためオントロジーを開発し、オントロジー・エディター法造および本プログラムで開発した実験プロトコル入力システム(TPPE)を利用して、本ターゲットタンパク研究によって開発された新規で有用な実験プロトコル・データを収集し、データベースとして整理した。また、生化学の知識を有するアノテータによって文献情報等の入力も継続的に行い、内容を充実させた。こうして、オントロジーによって整理したプロトコルを 261 件登録した。内訳は、クローニング 7, 発現 89, PCR 7, growth 1, 精製 47, NMR 実験 12, X 線結晶試料調整 8 結晶化 38, X 線結晶回折実験 43, 電子顕微鏡解析 2, 構造と機能の解析同定 7 件である。

② 実験プロトコル・データベースの開発と運営

実験プロトコル・データベース (PREIMS) とそのポータルを開発し、情報プラットフォームの他のデータベースやサービス等と連携して運用する一方、遺伝研ポータル (AT-Atlas) からの横断検索



実験プロトコル・データベースのポータル(PREIMS)のトップページ

を可能とした。実験プロトコル・データについては、261 件全てを本ターゲットタンパク研究のプロジェクト内だけでなく、国内外に対しても公開した。放射光施設における実験結果データ 1420 件は公開せず、実験者のみが閲覧できる仕組みとした。

③SPring-8 および Photon Factory での実験結果のデータベース化とポータルの開発

放射光施設 SPring-8 および Photon Factory において行われた X 線結晶回折実験の結果を、オントロジーを用いて整理した実験プロトコルのテンプレートを用い、データ・インスタンスをインポートする仕組みを開発した。この仕組みを使って、オントロジー化された実験データ 1420 件を登録した。

2) 技術開発の進捗及び成果

①タンパク質実験プロトコルのオントロジーの開発と実験プロトコル・データの収集

ターゲットタンパク研究を加速するため、これまでに開発されてきた困難なターゲットに対する優れた実験プロトコルをコミュニティで共有するため、データ項目やデータの意味とその間の関連性を明瞭に標準化された手法で定義し、精密に記載する必要がある。我々は、この標準化された記述を実現するためにオントロジーを採用し、[プロトコルサブ・プロトコルアクション]の階層構造で実験プロトコル記述を開発した。

実験プロトコルを、生化学実験における溶媒の緩衝液や培地の作成等での単位の手順にまで分解し、これらを様々な「action」として定義する。この「action」は、オントロジー辞書にて定義と説明がなされる。次に、この「action」が固有の順番でつながれたものが、比較的汎用的に利用される「サブ・プロトコル」と定義される。例えば、緩衝液の準備や培地の作成等などの作業を「サブ・プロトコル」(実際の名称は"buffer_preparation")として定義し、それら「サブ・プロトコル」をつなげたフローとして全体の「プロトコル」が記述される。この場合、作業自体は同一でも、対象の試薬または細胞や緩衝液の量、温度等は、当然、対象毎に異なるため、それらは「パラメータ」の値とし別途入力される。これらの「サブ・プロトコル」はライブラリとして登録され、それをダウンロードしてテンプレートとして利用し、パラメータを入力した後にその時系列に従うフローを作成する手順となる。

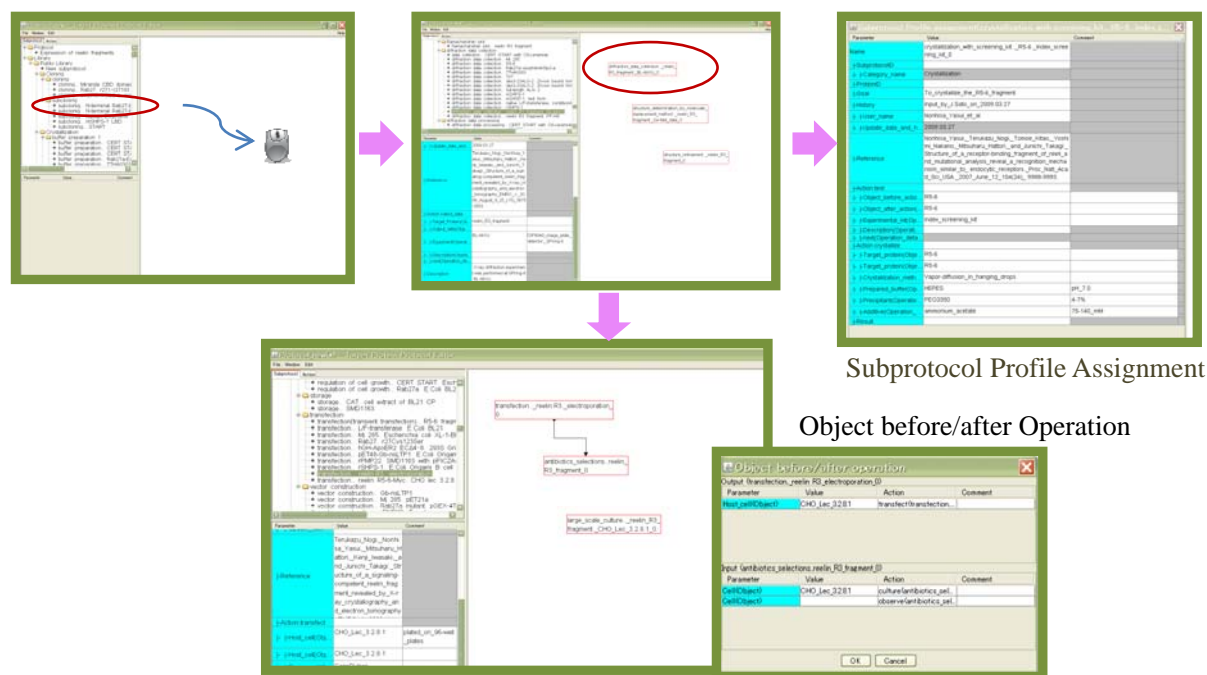


図1：PC 上でオフラインで稼動する実験プロトコル入力ツール (TPPE) によるプロトコル構築の流れ。左側にあるサブ・プロトコルのライブラリからサブ・プロトコルを選択し、そのプロファイル表に詳細を書き込んだり修正をし、その後にサブ・プロトコルを接続する。

手元の PC を用いてオフラインで実験プロトコルを簡便に入力するためのツール (TPPE) を開発し、一般の実験研究者が独自に実験プロトコルを構築し入力が行える支援ソフトとしてプロジェクトメンバーに対し公開した。利用者自身がプロトコルをすぐに記載できるように、単位のサブ・プロトコルをテキストで記述し、それらを矢印でつなげて一連のフローとして全体のプロトコルとする簡易版も構築・登録できる (図 1)。ターゲットタンパク研究に携わる研究者 (TPPE の利用者) の意見を加え初期の版の TPPE に対して修正をほどこし、オブジェクトの追加・削除を行ってプロトコルの修正に対して容易に利用者自らが対応できるようにした。

ターゲットタンパク研究で行われた実験プロトコルとして、ターゲット・タンパク研究に関わる研究者から、論文や実験データ資料等によって具体的な実験プロトコル例を提供していただき、オントロジーによって整理したプロトコルを 2011 年 5 月末までに 261 件登録した。内訳は、クローニング 7, 発現 89, PCR 7, growth 1, 精製 47, NMR 実験 12, X 線結晶試料調整 8, 結晶化 38, X 線結晶回折実験 43, 電子顕微鏡解析 2, 構造と機能の解析同定 7 件である。

②実験プロトコル・データベースの開発と運営

関係データベース(RDB)である PostgreSQL により実験 MS ポータル (PREIMS, <http://preims.pdbj.org/>) を開発し、そのポータルを用いて、利用者のアクセス制御を行いつつ、上記①にて収集し整理したプロトコルデータを公開した (図 2)。

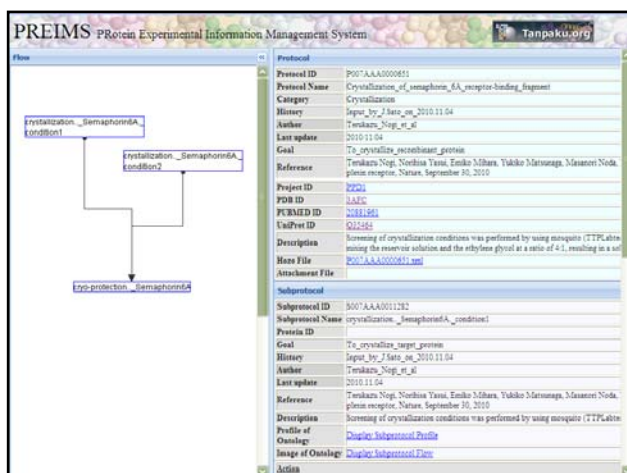


図 2 : 実験プロトコルデータベースのポータル (PREIMS) のプロトコル表示画面。この画面の左にある Profile of Ontology には、Image of Ontology がリンクされているおり、また、プロトコルのオントロジーで定義されている各プロセスの値も表として示される。

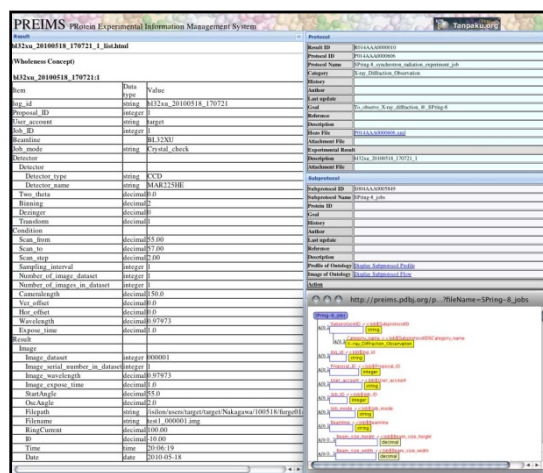


図 3 : PREIMS での放射光施設における X 線結晶回折実験結果が、左側の表で示される。右側は対応するプロトコル。測定結果のイメージ・ファイルは Filepath と Filename によって同定される (現在はリンクを張らない仕様としている)。

このポータルは、内容が英語で記載されているため基本的に全て英語による表示だが、トップページは日本語と英語の両方の表示のページを用意し、日本語から英語への翻訳機能により、日本語キーワード検索も可能とした。また、日本語によるオンラインの利用者マニュアルや動画による利用法の案内も作成した。このポータルを、情報プラットフォームの他のデータベースやサービス等と連携するため、遺伝研ポータル (AT-Atlas) からの横断検索も可能とした。、実験プロトコル・データについては、261 件全てを本ターゲットタンパク研究のプロジェクト内だけでなく、国内外に対しても公開した。

③SPring-8 および Photon Factory での実験結果のデータベース化とポータルの開発

<p>遺伝研と協力し、放射光施設 SPring-8 および Photon Factory から取得する X 線結晶回折実験の結果を、オントロジーを用いて整理しデータベース化した。まずこのデータベースへ記載する内容を整理し、XML スキーマを作成してそれぞれの実験ログ・ファイルを XML 化した。次に、対応する実験プロトコールのテンプレートを作成し、XML 化したログデータのインスタンスをインポートして、PREIMS データベースの一部として蓄積する仕組みを開発した。このインスタンスの内容を html 化し、表の形で PREIMS ポータルで表示する仕組みも構築した (図 3)。2011 年 5 月末の時点では、発表前の生の実験情報も多く含まれているため、データベース化した 1420 件全てをプロジェクト外へは公開せず、実験者のみが PREIMS で閲覧できる仕組みとしている。以上のシステムは LIMS ともなっており、この仕組みを拡張することにより実験室レベルでの LIMS へも拡張可能である。</p>	
<p>3. タンパク質の構造解析件数</p>	
PDB 登録数	該当しない
該当しない	該当しない
4. 各機関の論文発表件数	1 件
5. 各機関の公開したツールやデータベース等の件数	2 件
6. 各機関の特許出願件数	0 件 (うち国外 0 件)
7. 各機関の成果を活用した共同研究数 (プログラム内外含む)	9 件 (うち産業界との共同研究数 0 件)
<p>8. 当初計画に対する各機関の達成度</p>	
<p>A (95%)</p> <p>理由: 当初計画は、ほぼ全てを実現した。ただし、当初計画では日立製作所が実施することとなっていて具体的になっていなかった SPring-8 および Photon Factory での実験結果のデータベース化については、平成 22 年度から遺伝研との協力により LIMS 構築として開始した。そのため、SPring-8 の実験データに対するデータベースは構築できたが Photon Factory のデータに対するデータベース構築作業が平成 23 年度に残っており、また実験結果データの検索と表示方法についても改良の余地が残っている。</p>	
<p>9. 課題内の情報共有・連携体制</p>	
<ul style="list-style-type: none"> ・シングルサインオンの仕組みを代表機関 (国立遺伝学研究所) と分担者とで協力し、実現した。 ・PREIMS における横断検索、日本語キーワード検索について、代表機関と分担者とで協力して実施した。一方、代表機関の共有ポータル TanpakuWiki と PREIMS による新技術情報の統合については、新規研究参加者と専門家の双方に有用な情報資源となるよう、代表機関が実施する事業に協力した。 ・代表機関が開発・公開した網羅的配列データベース CASA db とオンデマンド構造解析 FUJI db のサービスについて、利用し易さの観点から改良する要望点を指摘し、その開発・拡充に協力した。 	
<p>10. 他の課題との情報共有・連携</p>	
<ul style="list-style-type: none"> ・他の課題 (ターゲットプログラム (生産) 課題総数 6 : PPC1, PPD1, PPD2, PPD3, SAC1, SAD2) において開発された実験プロトコールを、未発表の内容も含めてデータベース化した。 ・SPring-8 と Photon Factory の 2 つの機関の放射光 X 線結晶回折実験結果データをデータベース化した。(実際の image data file は、放射光実験施設の HDD 上におき、meta-data や file name を登録した。) ・実験プロトコールのオントロジー化は、世界的に見ても新たな試みであり、オントロジー研究の国際 	

的なパイオニアのグループ大阪大学産業科学研究所・溝口研究室の古崎晃司博士との連携により、データ内容が検証できるオントロジー・システムを構築した。

1.1. 人材育成

・具体的な個々の実験プロトコル作成を行うために雇用した研究者は、生化学の知識は持つものの、必ずしも情報科学やオントロジーについての知識は、当初は持っていなかった。しかし、オントロジーについての公開講習会や研究室での輪読会への参加、および on-the-job training によって、オントロジーを作成する十分な能力を持つ人材として育成できた。本プログラム終了後の進路は未定である。

1.2. 終了までの具体的な見通し

- ・本ターゲットタンパク研究プログラムで開発された有用でオリジナルな実験プロトコルについて、さらに多くのプロトコルをデータベース化し、内容が詳細で利用者にわかりやすいものとする。
- ・Photon Factory の放射光 X 線結晶解析実験結果データを、SPring-8 と同様のデータベースとして整理する。実験結果情報を、実験実施者、プロジェクト内、場合によってはプロジェクト外にも公開する。
- ・本プログラム終了後にも、大学スタッフによるデータの更新が可能な仕組みとなるよう整備しておく。

1.3. 本プログラムにおける成果が我が国の科学技術発展及び産業応用に与える効果について

- ・終了後に本プログラムで開発した実験プロトコルデータベースを陳腐化させないため、大阪大学蛋白質研究所において、国立遺伝学研究所と共同して、大学のスタッフが中心となり、継続的に本データベースを維持し、最新の知識を加えて発展させていく。
- ・タンパク質の発現・精製・結晶化等のプロトコル情報は、広く産業界にとっても重要な情報であり、論文では詳細が記述しきれていない内容を、データベースとして PREIMS ポータルから発信することは、本ターゲットタンパク研究の成果の橋渡しのツールとして大いに利用できる。

1.4. 特記事項

- ・本プログラムがスタートする以前には全く存在していなかった、タンパク質実験プロトコルについてのオントロジーを確立し、それを検索・表示するポータルを開発し公開した。
- ・オントロジーの標準化については、法造エディターで構築したライブラリーと辞書を他のオントロジー、例えば OBO foundry で公開されているオントロジーとのマッピングをすることで対応できる。
- ・構築した実験プロトコルのオントロジーに基づいて、SPring-8 や Photon Factory の放射光施設における X 線結晶回折実験の実験条件と実験結果であるイメージファイル情報を収集する LIMS を構築し、ポータルから利用者権限に応じて閲覧できる仕組みを開発し、運用を始めた。
- ・アウトリーチ：開発した実験プロトコル・データベースと、放射光施設における X 線結晶解析実験結果のデータベースの利用法については、アニメーションによる利用案内を作成し、公開した。

1.5. 研究費一覧

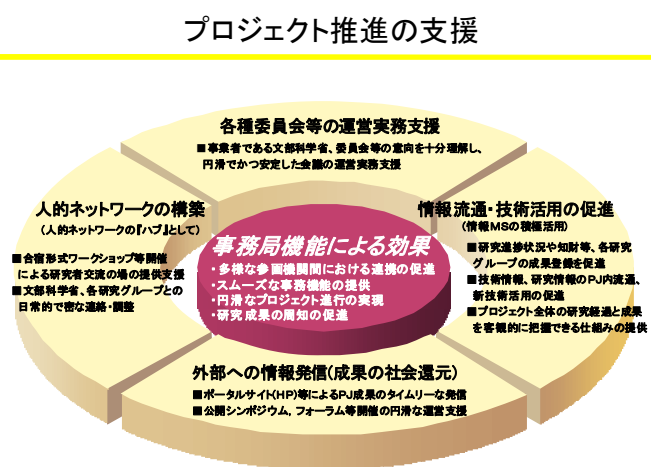
	19年度	20年度	21年度	22年度		23年度	計
設備備品費 (千円)	1,526	1,612	0	0	物品費 (千円)	0	
試作品費 (千円)	0	0	0	0	人件費・謝金 (千円)	4,407	
人件費 (千円)	2,857	4,401	4,360	4,372	旅費 (千円)	494	
業務実施費 (千円)	24,848	15,055	14,812	5,474	その他 (千円)	3,999	
間接経費 (千円)	8,769	6,320	5,751	2,954	間接経費 (千円)	0	
合計 (千円)	38,000	27,388	24,923	12,800	合計 (千円)	8,900	112,011

成果報告票（技術開発研究・各機関）

代表／分担機関の課題名	ターゲットタンパク研究情報プラットフォームの構築運用 (プロジェクト推進の支援)
分 担 機 関	国立大学法人 東京大学大学院農学生命科学研究科
機 関 の 代 表 研 究 者 名	大野 美恵

1. 課題開始時における各機関の達成目標

本プロジェクト全体における事業の円滑な運営と事業目標達成のため、研究等の支援業務、本プロジェクト内各種委員会など関係会議の開催実務、ホームページによる情報発信やシンポジウム開催等の広報業務等を実施する。具体には1) 参画機関のタイムリーな研究進捗調査、連絡・調整、及び参画研究機関間の交流の促進並びに事業者である文部科学省との連携を密に行うことにより、研究等の支援を行う。2) 本プロジェクト推進委員会及び各種委員会等関係会議の開催実務、並びにセミナー・ワークショップ等の毎年度の開催支援を行う。3) 「情報マネジメントシステム」及び「統合データベース」との連携のもと、プロジェクト情報発信に向けた広報用資料作成、及び公開シンポジウム等の開催、並びにプロジェクトに関する照会等実務の広報業務を行う。



2. 平成23年5月末時点における各機関の事業計画に対する研究成果

(1) 成果概要

推進委員会等の方針に沿って、本研究プログラムの事務局として、以下のようなプロジェクトの推進支援を行った。

- 1) 推進委員会等本研究プログラムに係わる各種委員会の開催にあたり、日程調整、会場設営、会議資料作成及び議事録作成等を行い、これら会議全般にわたりその円滑な運営を行った。
- 2) 毎年度開催される公開シンポジウムの開催に際し、日程調整及び会場設営、予稿集、ポスター、等の資料作成、当日の進行等、シンポジウム運営全般に係わる実務を行った。
また、本研究プログラムの成果を広く産業界に情報発信し、産業応用等につなげるために産業界との意見交換の場としての「産学懇談会」を日本製薬工業協会及びバイオインダストリー協会と提携して開催し、同協会等との連絡調整及び日程調整及び会場設営等運営全般を行った。
- 3) 参画研究機関間の情報交換、技術交流の場として、「全体交流会」（課題代表者）、分野ごとの「班会議」（若手研究者；合宿形式）及び特定トピックスに特化した「研究交流会」開催に際し、その全体の運営を担当した。
また、実施者によるネットワーク上の意見交換の場として、「共有ポータルサイト」を設置し、本研究プログラム内の情報発信等の運用に供した。
- 4) 参画研究機関の進捗状況・成果の調査及び同研究機関との連絡・調整等を行い、遺伝研チームと共同で本研究プログラムの公式サイトを運営し、広く社会への情報発信を行った。

また、本研究プログラムの研究内容あるいは研究成果を分かりやすく解説した2種類のパンフレットを作成・配布するとともに、タンパク研究の意義を高校生等非専門家にも理解できるようなリーフレットとしても発刊し、社会一般への広報を行った。

(2) 技術開発の進捗及び成果

以下の事項はいずれも当初目標として掲げた「プロジェクト推進の支援」の成果と考える。

1) 推進委員会等、本事業に関わる各種委員会の開催実務

毎年度開催された下記委員会について、日程調整及び会議資料作成、議事録作成等、会議全般にわたる実務の支援を行った（平成21年度に推進体制の改組があった。補足資料1）。

【平成19年度～平成20年度】推進委員会（6回）、プログラム連絡会（4回）、技術基盤委員会（2回）、ターゲットタンパク研究委員会（1回）、研究成果・情報プラットフォーム委員会（3回）、シンポジウム実行委員会（3回）

【平成21年度～平成23年度】推進委員会（5回）、運営委員会（5回）、技術開発検討ワーキンググループ、チームタイム利用調整ワーキンググループ、タンパク質研究検討会（以上各1回）

2) 本プロジェクトの成果を問う公開シンポジウム等の開催実務

①公開シンポジウム（補足資料2）

毎年度開催される公開シンポジウムの開催に際し、推進委員会等委員会の方針に沿って、日程調整及び会場設営、予稿集、ポスター等資料作成、当日の進行等、シンポジウム運営全般の実務を行った。開催にあたっては、顕著な研究成果を紹介するのみならず、海外に在住する研究者を講師として招聘し、海外の研究動向及びタンパク研究あるいは機能・構造解析研究の意義等についての講演も行った。下表に示すように企業及び一般からの参加者が毎年度40～50%を占めている。

参加者の所属	平成19年度	平成20年度	平成21年度	平成22年度*
大学関係	140(33%)	166(31%)	203(36%)	280(52%)
公的研究機関	73(17%)	81(15%)	61(11%)	49(9%)
一般企業	139(33%)	156(30%)	132(23%)	134(25%)
報道機関	9(2%)	6(1%)	6(1%)	3(1%)
その他一般	56(13%)	109(21%)	142(25%)	58(11%)
文科省、事務局	8(2%)	9(2%)	16(2%)	15(2%)
合計	425	527	564	539

*平成22年度は東北地方太平洋沖地震発生により会の途中で中止されたため、参加登録者数。

②産学懇談会

(1) 研究成果の産業界への応用（出口）展開の具体化 (2) タンパク質研究の社会的意義（産業応用、ライフサイエンス研究の基盤の必要性等）等の必要性から、本研究プログラムから産生されている研究成果を産業界に情報発信すること、及び産業界からの意見を聴取するための意見交換の場とすることを目的として、20～30名程度と比較的少人数の企業研究・技術者による「産学懇談会」を平成22年度開催した。具体には、関連する業界を束ねる日本製薬工業協会及びバイオインダストリー協会と提携して開催した。課題及び講師の選定等を含めて、同協会等との連絡調整及び日程調整及び会場設営等の運営等実務全般を行った。

	基本的な生命の解明	医学・薬学等への貢献	食品・環境等の産業利用	技術開発研究
テーマ数	2	3	2	7
企業からの参加者数	26	27	27	57

3) 参画研究機関間の情報交換、技術交流の場としてのワークショップ等の開催及び情報共有のためのポータルサイトの設置・運営等実務支援

①ワークショップ等の開催

本研究プログラム内での各種行事について、その開催に向けての日程調整及び会場設営、予稿集作成等の準備段階から当日の運営に至るまで行事全般にわたる実務支援を行った。

ア) 全体交流会：本研究プログラムは医・理・工・農等異分野の代表的研究者から構成されていることから、研究・技術交流による異分野融合を目指し、共同研究等更なる研究発展を期して開催したものである。本研究プログラム開始時より3年にわたり毎年度開催した。本交流会では、課題代表者が2日間の合宿形式をとることにより、夜を徹しての議論がなされた。(全3回)

イ) 成果発表会：上記「全体交流会」は研究者間の交流においては一定の成果をあげ、異分野の研究交流も個別に進められていることから、平成22年度は「全体交流会」を換骨奪胎し、PD・POによる内部評価のための成果発表の場とし、併せて、各研究課題の進捗状況や具体的研究上の隘路等についての研究者間の意見交換を可能ならしめることも目的として、3日間にわたり分野ごとに開催した。(1回)

ウ) 分野別班会議：本研究プログラムは3分野1領域からなるが、それぞれの分野・領域ごとに平成20年度から2年度にわたり、分野別の班会議を開催した。本会議は、実験に携わる若手研究者を主な対象者とし、実験上の隘路等を議論し、その問題解決を模索することを目的として開催したものである。このために、時間の制約を取り払って自由に議論しようということから始まったものであり、通常の会議終了後も夜を徹して議論ができるようにとの配慮から合宿形式にて行われた。夜2時過ぎまで有志が集まって議論を行った班会議グループもあった。(2回×4分野=全8回)

エ) 研究交流会：本研究交流会は、技術分野の特定のテーマに焦点を合わせ、その基礎から応用までの最新の技術についてのレビューを行うことを目的とした、非専門分野、とりわけ若手研究者に向けた研究セミナーである。平成19年度より毎年度、以下のようなテーマで開催した。(全4回)。
平成19年度：構造解析の基礎(116名)、平成20年度：化合物ライブラリーの活用(99名)
平成21年度：構造・機能解析の最新技術とその応用(94名)
平成22年度：最新の生産技術とその応用(157名)

なお、カッコ内数字は参加者数を示す。

②研究実施者向け情報共有ポータルサイトの設置及び運用

本ポータルサイトは、本研究プログラム内において(1)研究情報を共有すること、(2)意見交換を行うこと、(3)各種行事等の案内、(4)公式書類様式等のダウンロード、等を可能ならしめるために、平成19年度にサーバーを導入、並行してシステムの開発を行い、平成20年度6月より運用に供している。本ポータルサイトは本研究プログラム内部に限定されたものであり、機密保護の観点から、研究実施者及び推進委員会委員等個別にユーザーアカウント及びパスワードを連絡・付与し、統一されたセキュリティ・ポリシーの下、厳格なアクセス管理を行った。(アカウント数690)

これまで、ビームタイム枠の割当時間の案内等の「案内記事」120件、「ダウンロード資料」960件以上の記事掲載を行った。平成20年度6月の開設時より平成23年5月までの3年間の合計アクセス数約25,150件(月平均約700件)、月平均53名のユーザー訪問があった。

4) ホームページ及びパンフレット等の印刷物を通じた社会への情報発信

①本研究プログラムの公開サイト(ホームページ)運用による広報(補足資料3)

ターゲットタンパク研究プログラムにおける研究の意義及び研究内容を研究コミュニティ、マスコ

ミ、産業界及び一般に広くアピールし、さらに研究成果等の普及促進を目的として、本プログラムのポータルサイトを設置し、広報基盤とするために、平成19年10月に運用を開始した。

その後、遺伝研チームにおいて公開サイトのリニューアルが行われ、平成21年2月2日より運用を開始したが、その過程で開発に協力してきた。公開サイトの再構築にあたっては、公表された論文情報等は、従来前記「共有ポータルサイト」に掲載されていたが、これら研究情報は、むしろ積極的に外部に公開すべきとの判断から、公開しても問題のない研究情報を精選し、それらをできる限り公開サイトに移行して掲載することになった。また、研究成果の報道発表等顕著な成果の公表については、研究実施者からの情報提供を積極的に依頼し、情報発信してきた。これらは遺伝研チームとの密接な連携のもとに進められ、その後の運用についての主担当として推進してきた。

②パンフレット等の印刷物による情報発信

本研究プログラムの研究の意義や期待される成果等を紹介するパンフレットを本研究プログラム開始直後に、また、平成22年8月にそれまでに各課題で得られた成果を要約したパンフレットを作成し、研究機関・企業等に配布した。本パンフレットを作成するにあたっては、読者層として一般の方々にも理解しやすい内容にすることを目標とし、このために研究課題代表者の協力の下、サイエンスライターに作成委託を行った。また、本研究プログラムがなぜ必要になったか等、疾病との関連等卑近な例をあげながら、本研究の意義・背景を紹介することにも多くの稿をあてた。その結果、非常に分かりやすいパンフレットとなった。本パンフレットは公開シンポジウム等においても配布したが、そのほか各方面から提供の依頼を受けている。本パンフレットは公開サイトからもダウンロードで可能となっており、また電子ブックとしても閲覧できるようにしている。

また、本研究の社会的意義を広く周知するために、高校生等でも理解できるようなやさしい語り口で解説したリーフレットを作成し、シンポジウム等で配布してきた。本リーフレットは研究実施者等によるアウトリーチ活動の講義資料や各研究機関公開時の配布資料としてたびたび活用されており、現在までに関係部署への配布は14,000部に達している。(補足資料4)

また、平成22年度より遺伝研チームと協力して、本研究プログラムから産生された顕著な成果を特定のトピックスに絞ったニュースレターを発行した。第1号は植物ホルモンについての特集号を発行した。(補足資料5)

3. タンパク質の構造解析件数	
PDB 登録数	0件
PDB 登録の有無にかかわらず構造解析したタンパク質数	0件
4. 各機関の論文発表件数	1件
5. 各機関の公開したツールやデータベース等の件数	0件
6. 各機関の特許出願件数	0件
7. 各機関の成果を活用した共同研究数 (プログラム内外含む)	0件
8. 当初計画に対する各機関の達成度	

情報C1-C

当初目標としていた委員会運営、情報発信、本プログラム内の情報共有について、委員会委員及び研究実施者の積極的な協力体制を得ることができ、これらが円滑に実施できたことから、95%達成したと考えます。

9. 課題内の情報共有・連携体制

本文にて記載した通り、公開サイト及び共有ポータルサイトの運用に際しては、その機能強化及び記事の掲載等について、遺伝研及び日立製作所、大阪大学との日常的な連絡の下推進してきた。

10. 他の課題との情報共有・連携

当業務は、課題全般に係わるものであり、特定の課題との特別な関係を持たないが、当事務局の職務は、研究実施者から理解されており、報道発表や論文投稿の連絡等、実施者からの自主的な情報提供も得られるようになってきていること、また、当事務局からの情報提供はメールあるいはポータルサイト等の手段を活用することにより、情報共有をタイムリーに行っている。

11. 人材育成

本業務自体に若手人材は存在しないが、各研究機関にて従事している若手研究・技術者の育成を、事務局として支援するために、セミナーや勉強会等の開催運営に尽力してきた。

12. 終了までの具体的な見通し

引き続き、委員会等の円滑な運営及びタイムリーな広報活動を努めると同時に、最終年度にあたり、これまでの成果とりまとめ支援を行う。その一環として、特に本研究プログラムの知的財産取得のための支援及び調査業務を実施し、各研究機関での知的財産取得の進捗状況等を調査することとしたい。

13. 本プログラムにおける成果が我が国の科学技術発展及び産業応用に与える効果について

本文にも記した通り、産業界との意見交換の場としての「産学懇談会」は産業界の研究・技術者とのつながりを持つ機会を設けたが、今後各課題において具体的な産業移転につながるものとする。

14. 特記事項

当業務は事務局業務であることから、行政（文部科学省）、推進委員会、研究実施者との信頼関係が成立してはじめて有効に機能する。当事務局は本研究プログラムの「ハブ」の自覚をもって推進し、上記関係者との人的なつながりを何よりも重視しながら努めてきた。現在までに、当事務局は研究実施者からの第一次相談窓口としての役割を果たしており、また、実施者からの研究情報提供を積極的に呼び掛けていること等を通じて、当事務局の役割についての理解が十分得られており、相互の信頼関係は築かれているものと考えている。

15. 研究費一覧

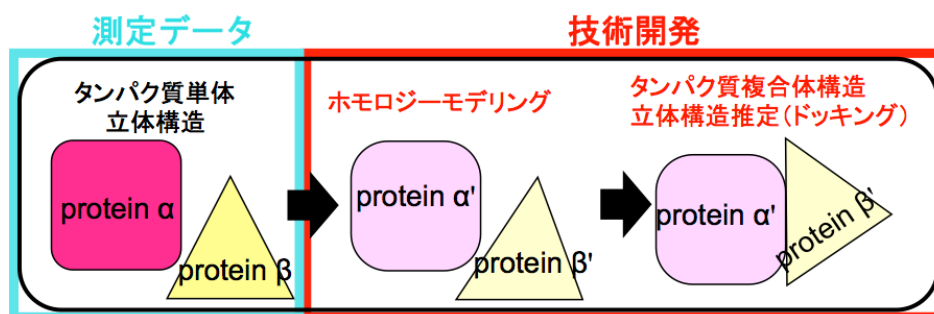
	19年度	20年度	21年度	22年度		23年度	計
設備備品費（千円）	6,027	500	300	518	設備備品費（千円）	249	/
試作品費（千円）	0	0	0	0	人件費（千円）	38,747	
人件費（千円）	27,130	37,160	37,291	37,695	業務実施費（千円）	19,162	
業務実施費（千円）	69,314	35,068	35,136	33,270			
一般管理費（千円）	10,247	7,273	7,273	7,148			
合計（千円）	112,718	80,000	80,000	78,631	合計（千円）	58,158	

代表／分担機関の課題名	ターゲットタンパク研究情報プラットフォームの構築運用（構造バイオインフォマティクス技術利用の推進）
代表／分担機関	国立大学法人お茶の水女子大学
機関の代表研究者名	由良 敬

1. 課題開始時における各機関の達成目標

(1) 生体高分子の高精度立体構造モデリング手法の研究開発：明らかになったタンパク質立体構造の情報を類縁タンパク質に拡大し、構造解析困難なタンパク質の立体構造を実験解析と同程度の精度で明らかにするために、従来よりも精度の高いタンパク質モデル構造を構築する。モデリング手法の要であるテンプレートとターゲットタンパク質のアラインメント法の改善と、モデル構造構築法の精度向上を達成する。

(2) タンパク質間相互作用部位予測法の研究開発：解析困難なタンパク質複合体の構造を明らかにし、産業応用などに利用するために、2つのタンパク質がどのような配向で相互作用するかを、それぞれのタンパク質の立体構造から推定する方法を研究開発する。



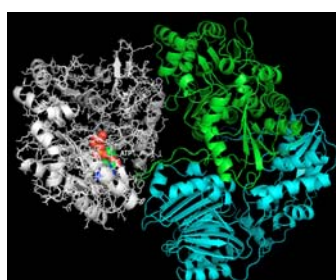
2. 平成23年5月末時点における各機関の事業計画に対する研究成果

(1) 成果概要

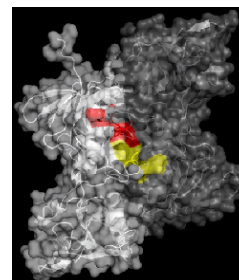
立体構造モデリングの基礎となる高精度アラインメント法の研究開発、およびタンパク質間相互作用の予測方法を研究開発し、さらにタンパク質相互作用界面に世界ではじめて新規特徴を見出すことに成功した。研究開発した技術はインターネットで公開し、プログラム内外の研究者が利用できるようにした。



高精度アラインメント法



相互作用の予測

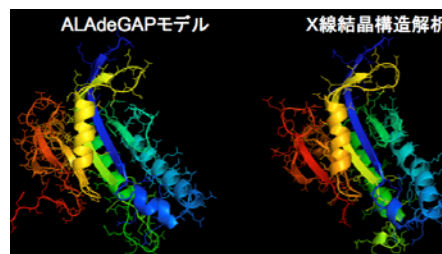


相互作用界面に新規特徴

(2) 技術開発の進捗及び成果

高精度アラインメント法の研究開発：ホモロジーモデリングを行う際に必要なアラインメントは、ターゲットとテンプレートの配列一致度が低い場合には、十分な精度が得られていなかった。当該研究開発において、アラインメントで高い精度を得るためには、ギャップペナルティを適確に定めることが重要であることを新たに見出し、ギャップペナルティの新規計算方法を考案した。共通祖先由来タンパ

ク質の立体構造を比較すると、ギャップ部位は挿入アミノ酸残基の溶媒露出度と指数関数的関係にあることがわかった。この関係から新しいギャップペナルティ算出法を考案し、従来よりも4~5%精度の高いアラインメントを構築できるようになり、モデル構造の精度を上げることに成功した (Hijikata et al., *PROTEINS*, 2011)。この技術を利用して、タンパク質立体構造予測とホモロジーモデリングの国際コンテスト CASP9 に参加し、他国の参加チームと互角の高精度ホモロジーモデル構造を構築できることを示せた。アラインメント法は ALAdeGAP と命名し、インターネットで公開しており、ターゲットタンパク研究プログラム内外の研究者が利用している。論文発表後 (2011年2月) からのアクセス数は、760件以上である。ALAdeGAP を利用したモデリング研究依頼が、プログラム外部からすでに3件届いている (東大、お茶大、千葉工大)。さらに高精度アラインメントをタンパク質の進化解析にも適用した (Yura, Go, *BMC Plant Biology*, 2008)。

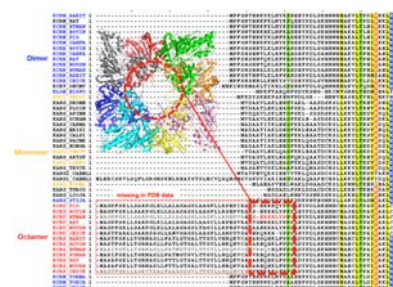


タンパク質間相互作用部位の予測方法の研究開発: タンパク質の

単体立体構造をドッキングさせて複合体構造を推定するために、単体構造をもとに、どのアミノ酸残基が複合体形成に関与するかを推定できるようにした。既知複合体タンパク質の界面における、アミノ酸残基出現頻度をもとにして、各アミノ酸残基種の界面形成度スコアを算出し、スコアの高いアミノ酸残基が局在する表面が、複合体界面であると予測した。その結果約60%の精度で、界面を正しく推定できることがわかった (論文準備中)。各種の予測法の精度と比べて、精度60%は相応の値である。この方法ももちいて、ターゲットタンパクプログラム内の東大田之倉研と共同でダイニンとマイクロチューブルの相互作用構造の推定を行った (論文準備中)。さらに複合体構造のデコイを生成するプログラム ZDOCK を併用し、複合体構造の予測ができるように研究開発を行った。この技術は ToSY と命名された統合サーバーから、インターネットで公開しており、ターゲットタンパク研究プログラム内外の研究者が利用できる。またプログラム外からも複合体構造の推定依頼が2件届いている (東大、千葉工大)。



タンパク質相互作用界面における新規特徴: タンパク質複合体構造の推定精度を上げるために、複合体界面がどのような構造特徴を持っているかを詳細に解析した。その結果、構造既知のタンパク質複合体界面の1/3には、複合体構造形成時のダイナミクスに関連する特殊な構造 (絡まり構造) があることがわかり、絡まり性 (Interwinding nature) というあたらしい概念を打ち出した。この構造は、複合体形成以前は安定な構造を取らず (天然変性部位)、複合体を形成した時点ではじめて安定な構造を取る部分と推定された (Yura, Hayward, *Bioinformatics*, 2009)。この概念は、国際学術雑誌に掲載されたタンパク質相互作用のレビューに取り上げられ、複合体形成部位と天然変性部位の関係を定量的に指摘した点において注目を集めている (Perkins, et al, *Structure*, 2010)。タンパク質複合体界面の予測は、界面には大きな構造変化があるために難しいと一般的に言われてきたが、実際には界面の2/3には、大きな構造変化はないことを、初めて定量的に示すことができた。相互作用面の絡まり性を同定する技術はインターネットに公開し、ターゲットタンパク研究プログラム内外の研究者に利用可能となっており、タンパク質複合体構造推定のための重要な情報を提供できるようにした。



<p>開発技術の利用推進：本課題において開発した技術はすべてをインターネットで利用可能とした。さらに技術の利用促進を図るために、開発技術とその適用例を論文、学会（国際学会を含む）、講演会、人材養成講義などで発表してきた。さらに技術利用をターゲットタンパク研究プログラム内で促進するために、技術講習会（平成22年度は2回、平成23年度も2回を予定）を開催した。</p>	
<p>3. タンパク質の構造解析件数</p>	
PDB 登録数	0 件
PDB 登録の有無にかかわらず構造解析したタンパク質数	126 件（CASP9 におけるホモロジーモデリング数）
4. 各機関の論文発表件数	4 件
5. 各機関の公開したツールやデータベース等の件数	10 件
6. 各機関の特許出願件数	0 件（うち国外 0 件）
7. 各機関の成果を活用した共同研究数（プログラム内外含む）	8 件 （うち産業界との共同研究数 0 件）
<p>8. 当初計画に対する各機関の達成度</p> <p>当初の計画を 100% 達成した。最終目標はタンパク質の高精度モデリングとタンパク質複合体構造のモデリングであった。タンパク質単体のモデリングの精度向上は、アラインメントの精度を向上することで達成した。精度向上はモデリングの国際コンテスト CASP9 で他グループと互角のモデルを構築できたことで証明された。タンパク質複合体構造のモデリングに必要な技術が完成し、実際の問題に適用できるようになった。</p>	
<p>9. 課題内の情報共有・連携体制</p> <p>課題内において、代表研究者と分担研究者は頻繁に連絡を取ると共に、進捗状況の確認を行った。情報の共有は、できるだけコンピュータネットワークを利用して効率的に行い、物理的な課題内の会合を年 2 回程度開催し、互いの進捗状況を確認した。特に、塩生氏のドメイン間の相互作用予測（共有結合があるドメイン間の相互作用予測）と、由良のタンパク質間相互作用予測（共有結合がないサブユニット間の相互作用予測）との間で情報共有を密に行い、手法と対象の独立性を明確にし続けた。平成 21 年度以降には、課題内の全技術を連携した統合サーバ（ToSY）を研究代表機関に構築し、またディザスタリカバリの観点から、各分担機関にもそれぞれの特色を生かしつつ代表機関のバックアップサーバとなるシステムを構築した。平成 23 年度には、関東圏の電力不足に対応し、課題内でコンピュータ資源配置の変更を行った。</p>	
<p>10. 他の課題との情報共有・連携</p> <p>以下のプログラム内課題と協力している。（1）<u>田之倉優グループ</u>（乾燥・高温ストレス耐性作物の開発に役立つ転写制御タンパク質の構造・機能解析）：タンパク質相互作用面予測結果の提供とタンパク質の動的構造変化の解析。</p> <p>また、以下のプログラム外研究者との共同研究を進めている。（1）<u>千葉工業大学河合剛太教授</u>：プリン合成系タンパク質群の複合体構造予測データ提供。（2）<u>お茶の水女子大学作田正明准教授</u>：酵素タンパク質のモデリングデータ提供。（3）<u>東京大学八木俊樹講師</u>：タンパク質複合体形成残基の</p>	

情報提供。

1.1. 人材育成

平成20年度にお茶の水女子大学生命情報学教育研究センターを創設し、由良がセンター長として学内の生命情報学の研究と教育拠点を構築した。本センターのもとで若手人材の育成を目的に、本課題で開発した高精度ホモロジーモデリング法、複合体界面推定法および絡まり性の概念および、計算生物学全般の講義と実習を、学部、大学院および学外の研究者に日本語と英語で行った。これらの活動は、生命情報学の女性人材を育てるプログラム（大学院GP採択）とともに運営し、外国からの講師による招聘講義も行った。毎回の講義には約10名の受講者が出席した。また最先端の生命情報学研究セミナーを、学内の学科や専攻を越え、外国人特任助教や外部からの参加者を含めて毎週英語で開催した。平成23年度には東京大学大学院と名古屋大学大学院におけるホモロジーモデリングに関する集中講義が予定されており、ライフサイエンス分野の若手研究者にタンパク質立体構造バイオインフォマティクスの解説を行う。韓国及び英国の大学院においても、2年間にわたりホモロジーモデリングの集中講義を行い、各集中講義に5名前後の大学院生が出席した。また、ターゲットタンパク研究プログラム内研究者向けに講習会を開催し、10名程度の受講者に対して利用方法の実習を行った。

講義実習とは独立に、タンパク質立体構造の理解を深めてもらうために、生物物理学会誌「生物物理」に「タンパク質立体構造散歩」を約2年間連載した（埋め草「タンパク質立体構造散歩」の経緯、*生物物理*, **51** (3), 144-145 に総括記事を掲載）。タンパク質の立体構造から、どのような考察を通して機能的情報を抽出するのかを、若手人材向けに記した。また、「タンパク質の立体構造入門：基礎から構造バイオインフォマティクスへ」を分担執筆し、構造バイオインフォマティクスの基礎を勉強する大学院生の教科書を日本語で記した。この教科書は韓国語への翻訳が始まっている。これら執筆と平行して「化学と生物」2010年1月号に、「タンパク質の構造から機能を推定する：構造バイオインフォマティクス」を掲載し、異分野の研究者に向けて、構造バイオインフォマティクスの重要性を記した。

平成20年度から2年間、物理学で博士号を取得した若手研究者を博士研究員として雇用し、異分野からの人材育成も行った。当該博士研究員によるタンパク質とATPの相互作用の研究は、現在論文執筆中である。

1.2. 終了までの具体的な見通し

ターゲットタンパク研究プログラム開始時点で、立体構造を決定するターゲットのリストにかかげられているにもかかわらず、現時点で構造が決定できていないタンパク質が約300個存在する。これらのタンパク質の立体構造を、研究開発した手法で可能な限りモデリングし、プログラム内に開示する。また相互作用することがわかっているタンパク質の複合体構造も可能な限り推定し、プログラム内に開示する。モデル構造の詳細や構造変化の解析が必要な場合は、共同研究によりデータを提供し、実験とバイオインフォマティクスとが協力して論文発表等を行う。

1.3. 本プログラムにおける成果が我が国の科学技術発展及び産業応用に与える効果について

タンパク質立体構造の高精度モデリング手法の開発により、実験的に決定が難しいタンパク質の予測構造を提供できる可能性を広げた。この技術は、タンパク質の構造に基づく機能発現機構の理解と、創薬分子ターゲット設計へ適用できる。

高精度モデリング手法を実際に適用することで、タンパク質のループ構造部分をモデリングすることが非常に難しいことが明らかになってきた。モジュール程度の長さの挿入や欠失（選択的スプライシング）があった場合にどのように対処するかは、まだ研究が進んでいない。病気の原因タンパク質において長い挿入や欠失があっても、タンパク質が立体構造と機能を維持する場合があることが明らかになっ

てきている。このようなタンパク質の立体構造すべてを実験的に決定するのは、時間とコストの両面で無理がある。長い挿入や欠失があるタンパク質のコンピュータモデリング法の研究と開発に進む必要がある。

タンパク質の複合体構造予測の技術開発によって、複合体構造の決定が難しい場合でも構造を提供できる可能性を広げた。この技術によりX線結晶解析が困難な複合体の構造を事前に推定し、この推定構造をもとに安定な複合体構造を得るために必要な変異をタンパク質に導入し、X線結晶解析で構造を決定することが可能となった。この新しい研究サイクルは各方面に応用可能である。

複合体構造予測の研究開発の過程で、「絡まり性」の概念を提案した。タンパク質が単体で存在するときは構造を取らない（天然変性）領域が、複合体形成を駆動する（他のサブユニットを呼びよせる）部位としてはたらし、複合体を形成した際には相手のサブユニットに絡まりつくようにして安定な立体構造を形成するモデルを提案した。この仮説を証明するためのシミュレーションと実験を進める必要がある。絡まり性の概念が正しければ、タンパク質の相互作用を制御するための理論的バックグラウンドを与えることができる。絡まり性を持つループの改変、またはループの動的構造を変化させる低分子との相互作用により、相互作用を制御できるようになり、この技術の産業転用が見えてくる。

1 4. 特記事項

新規開発高精度アランメント法 (ALAdGAP (Hijikata et al., *PROTEINS*, 2011)) にもとづくホモロジーモデリングによる、タンパク質立体構造予測国際コンテスト CASP9 で世界各国のトップレベルのモデリングチームと同レベルのタンパク質立体構造モデルを提供できた。「絡まり性」の概念提案は、タンパク質立体構造の専門誌レビュー (Perkins, et al, *Structure*, 2010) においてとりあがられるなど、ヨーロッパで注目を集め始めている。

1 5. 研究費一覧

	19年度	20年度	21年度	22年度		23年度	計
設備備品費 (千円)	0	0	0	721	物品費 (千円)	0	
試作品費 (千円)	0	0	0	0	人件費・謝金 (千円)	144	
人件費 (千円)	0	0	0	0	旅費 (千円)	1,439	
業務実施費 (千円)	0	0	0	13,356	その他 (千円)	8,817	
間接経費 (千円)	0	0	0	4,223	間接経費 (千円)	0	4,223
合計 (千円)	0	0	0	18,300	合計 (千円)	10,400	28,700

代表／分担機関の課題名	ターゲットタンパクプログラムの情報解析支援
代表／分担機関	分担機関 学校法人関西文理総合学園 長浜バイオ大学
機関の代表研究者名	塩生真史

1. 課題開始時における各機関の達成目標

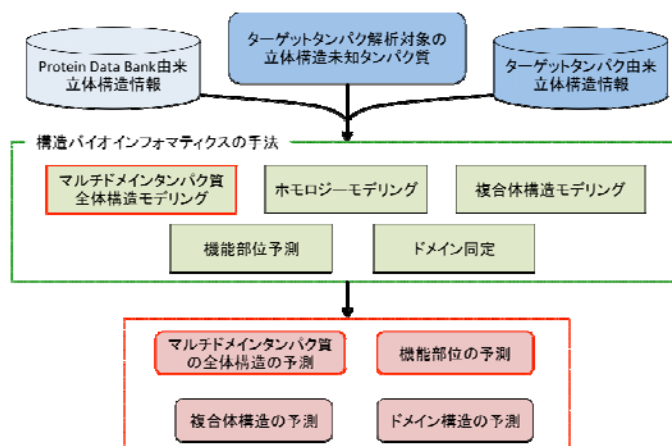
ヒトを含む高等生物由来のタンパク質の多くは多機能・マルチドメインである。それらは、各ドメインが機能を持つ他に、細胞接着に関わるインテグリンなどに見られるように、ドメインが共同して機能を発揮する場合が多い。一方で、マルチドメイン全体の立体構造を決定されたタンパク質は少なく、各ドメインを切り出して、その構造が決定される場合がほとんどである。そのようなマルチドメインタンパク質において、まだわかっていない機能の予測、および、機能の分子メカニズムの解明や機能改変のための論理的基盤として、それらがどのような全体構造を形成するかを予測することは非常に重要である。予測法は生命科学や創薬において希少価値をもつことから、タンパク質を対象とする生命情報研究者により研究開発されてはいるものの、創薬基盤技術としての重要性から非公開となっており、現在、自由に利用できる予測法は存在していない。そこでまず、

- 1) 既知の構造情報からドメイン間相互作用部位の特徴抽出を行う。
- 2) その情報をもとにドメイン間相互作用面を推定する方法の研究開発を行う。
- 3) さらに推定された相互作用面の情報に基づいてマルチドメインタンパク質の全体構造をモデリングする手法を研究開発する。

ことを目標とした。さらに、ターゲットタンパクプログラムから産出された立体構造データ、および、既存の立体構造データを用いて、

- 4) 開発したマルチドメインタンパク質のモデリング法を適用し、全体構造未知のマルチドメインタンパク質の構造予測を行う。
- 5) 開発中の機能部位予測法などの構造バイオインフォマティクス技術を適用し、マルチドメインタンパク質の機能推定を行う。

ことを目標とした。

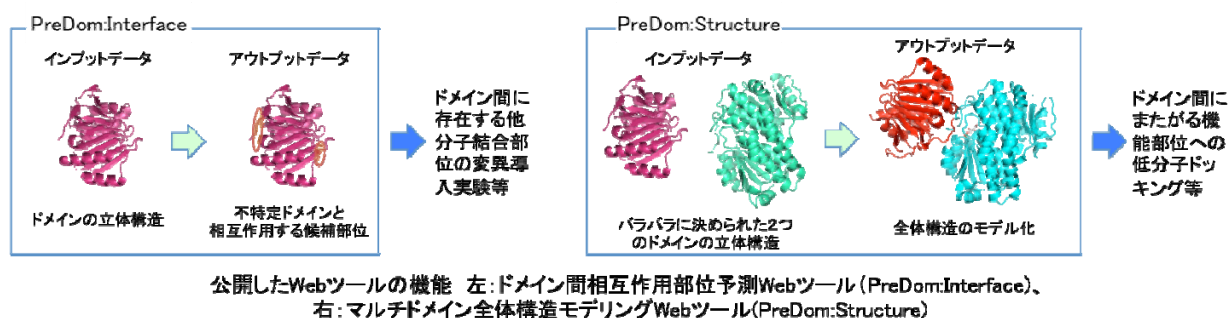


2. 平成23年5月末時点における各機関の事業計画に対する研究成果

(1) 成果概要

1)ドメイン間相互作用部位に特徴的に存在するアミノ酸残基の種類や、相同なドメインにおけるドメイン間相互作用部位の類似性について特徴抽出を行った。また、ターゲットタンパクプログラムにより得られた立体構造データについてドメインの範囲の同定を行い、既存のマルチドメインタンパク質とあわせてドメイン間相互作用部位のデータベースを構築した。このデータベースを **PreDom:DiD** として **Web** 上に公開した。2) この結果に基づくドメイン間相互作用部位の予測法を研究開発し、**Web** ツールとして公開した。3) 予測されたドメイン間相互作用部位やドメインリンカーに基づいて、マルチドメインタンパク質の立体構造をモデリングするシステムの研究開発を行い、**Web** ツールとして公開した。これらのドメイン間相互作用部位の予測法とそれによるモデリングシステムは共に、世界で自由に利用できる唯一の **Web** ツールである。 4) 愛知医科大学の渡辺秀人教授のグループと共同で、糖鎖合成酵素の

ドメインについてホモロジーモデリング法により構造予測を行い、その構造と機能の関係について解析した。



(2) 技術開発の進捗及び成果

1) ドメイン間相互作用部位の特徴抽出と、そのデータベースの公開

ドメイン間相互作用部位の情報抽出のために必要なドメイン間相互作用部位の同定と、ドメイン間相互作用部位に現れるアミノ酸残基種の特徴抽出に成功した。ドメイン間相互作用部位における各アミノ酸残基の出現頻度を求めると、疎水性のアミノ酸残基全般において出現傾向値が正の値をとり、親水性のアミノ酸残基の多くは負の値となることがわかった。これによりドメイン間も主に疎水性相互作用によりエネルギー的に安定になっていることが明らかになった。さらにドメイン間相互作用部位のうち、溶媒露出度が0.1以下であるコア相互作用部位は、疎水性アミノ酸残基全般がドメイン表面に比べて約2倍頻度が高く存在していることがわかった(由良ら、2009、54:1535-1541)。また、平均的なドメイン間相互作用部位は、コア相互作用部位として平均4残基程度の疎水性クラスターが存在し、コア相互作用部位の周辺にやや親水性が高い平均10残基程度のその他の相互作用部位が存在していることがわかった。これらの解析結果をもとにしてドメイン間相互作用部位の Web データベースを作成し、PreDom:DiDとして公開した(2010年度BMB2010で発表)。PreDom:DiDでは、ドメイン構造データベースを選択し、ドメイン間相互作用部位を閲覧したい立体構造データのPDB IDとそのデータに含まれるタンパク質鎖のIDを指定して実行すると、指定したタンパク質のドメインの範囲、ドメイン間相互作用残基、接触面積を調べることができる。

2) ドメイン間相互作用部位の予測法の研究開発

上記の解析結果から、ドメイン間相互作用部位におけるアミノ酸残基の出現傾向値を用いて、ドメイン間相互作用面の予測を行うアルゴリズム(IP法)を開発した(2009年度蛋白質科学会年会で発表)。

また、相同なドメインファミリーにおけるドメイン間相互作用部位の位置を解析したところ、相同ドメインにおけるドメイン間相互作用部位の位置情報が、ドメイン間相互作用残基を推定する際の良い指標になることがわかった。この解析結果に基づき、相同ドメインのドメイン間相互作用部位の位置情報をクエリーとするドメインに射影することでドメイン間相互作用部位を推定するアルゴリズム(KIP法)を開発した。さらに、ドメイン間相互作用残基

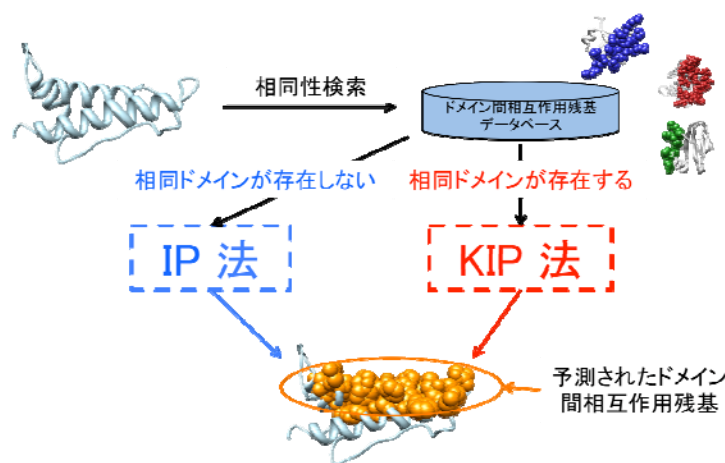


図1 PreDom:Interfaceのアルゴリズム

のデータベースにおいて、相同ドメインが検出できる場合にはKIP法を、検出できない場合はIP法を

適用することで、より多くのクエリーについてドメイン間相互作用部位推定が行えるようにしたアルゴリズム (図1) を、ドメイン間相互作用部位予測 Web ツール PreDom:Interface として公開した (2009年度の生物物理学会年会で発表)。

3) マルチドメインタンパク質構造モデリングシステムの研究開発

ドメイン間相互作用部位の予測結果、および、ドメインどうしの剛体ドッキングの結果、さらに、ドメイン間をつなぐリンカーの残基長に基づいて、マルチドメインタンパク質の全体構造をモデリングするシステムを開発した。この開発したアルゴリズムは、マルチドメインタンパク質の構造モデリングを行う Web ツール

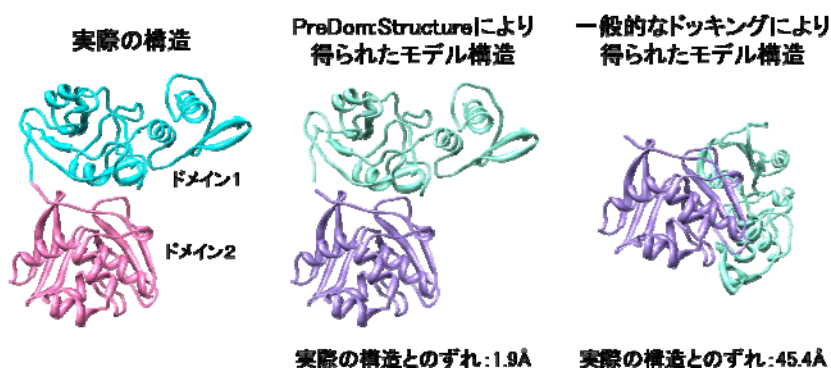


図2 PreDom:Structureにより得られたモデル構造の例

PreDom:Structure として実装し公開した (2010年度の蛋白質科学会年会・生物物理学会年会、BMB2010で発表)。この Web ツールで作成されるモデル構造の精度をブラインドテストにより調べたところ、実際の構造とのずれが平均 5.5 Åとなり、一般的なドッキング法によるモデル構造に比べて、ずれが平均して約 2.0 Å小さいことがわかった (図2)。また、情報プラットフォームにより公開されている TP Atlas を用いて、ターゲットタンパクプログラムの解析対象になっているタンパク質群のリストを取得し、タンパク質中の一部のドメインの立体構造が決定されているものを抽出し、全体構造のモデリングを開始した。このモデリング結果は平成23年度終了までに Web データベースとして公開していく。さらに、PreDom:Structure を基礎として、由良、および、藤と共同で、実験的に立体構造を決定することが困難なタンパク質の複合体構造、および、マルチドメインタンパク質の立体構造を推定する統合モデリングシステム ToSY を開発した。統合モデリングシステムでは、KIP 法により推定されるドメイン間相互作用部位を用いて、マルチドメインタンパク質のモデリングが行えるようにしている。ToSY では、2つのドメインの立体構造をクエリーとしてマルチドメインタンパク質の全体構造を推定するだけでなく、ドメインのアミノ酸配列をクエリーとした場合はホモロジーモデリングによりドメインの立体構造を推定し、その結果をもとに全体構造を推定できる。

4) コンドロイチン硫酸合成酵素の構造予測

愛知医科大学の渡辺秀人教授のグループと共同で、立体構造が未だ決定されていないコンドロイチン硫酸合成酵素の1種 (CSS2) について、ホモロジーモデリング法により構造予測を行い、その構造と機能の関係について解析した。マウスにおける CSS2 は選択的スプライシングにより2種類のタンパク質 (CSS2A と CSS2B) を生じることが知られており、それらについての酵素活性や細胞内での局在を解析した。その結果、CSS2B では、CSS2A に比べてグルクロン酸転位活性が低下していることが明らかとなった。CSS2B では N 末の 162 残基が欠失してはいるものの、酵素活性に重要な機能モチーフは CSS2A と同じであり、CSS2B の酵素活性の低下のメカニズムが不明であった。これを明らかにするために、CSS2 においてグルクロン酸転位活性を担う N 端側ドメインの構造予測を行い、スプライシングバリエーション間 (N-CSS2A、N-CSS2B) でそれらを比較した。この際、N-CSS2A、N-CSS2B とともに一般的な相同性検索では構造予測の鋳型となる構造既知の糖転移酵素が得られなかったため、既存のフォールド認識アルゴリズムを複数組み合わせ、それらのコンセンサスを取る手法を作成し、鋳型とできる

タンパク質を検索した。その結果、配列一致度が 10%程度しかないものの、有意にアミノ酸配列が類似した鋳型タンパク質を得ることができた。また、N-CSS2B と鋳型タンパク質の単純なアラインメントではドメインの立体構造形成に重要な疎水性コアが欠失してしまうため、ドメインの二次構造を考慮したアラインメントを新たに行い、それをもとに構造予測を行なった (図 3)。

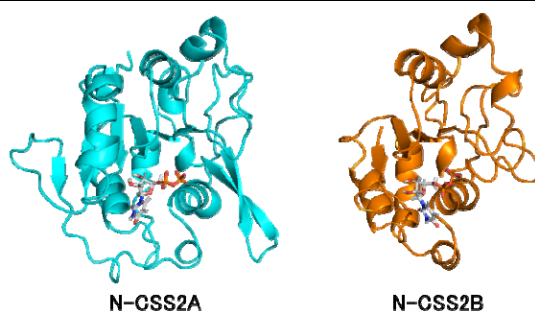


図3 推定したコンドロイチン硫酸合成酵素2の立体構造

その結果、N-CSS2A、N-CSS2B とともに構造評価ソフトウェアにより尤もらしい構造と評価されたものの、N-CSS2B では選択的スプライシングにより基質認識に関わる部位の構造が大きく変化していることがわかった。このことがグルクロン酸転位活性の低下の原因と考えられた。これらのことより、一般的な相同性検索では鋳型タンパク質が見つからないタンパク質であっても構造予測することができる例を示すことができた。また、コンドロイチン硫酸の合成は複数のタンパク質が複雑に相互作用しつつ行われるが、この研究により選択的スプライシングバリエーションがコンドロイチン硫酸合成の調節を行っているという初めての知見を示すことができた。

3. タンパク質の構造解析件数

PDB 登録数	0 件
PDB 登録の有無にかかわらず構造解析したタンパク質数	0 件

4. 各機関の論文発表件数

6 件

5. 各機関の公開したツールやデータベース等の件数

7 件

6. 各機関の特許出願件数

0 件 (うち国外 0 件)

7. 各機関の成果を活用した共同研究数 (プログラム内外含む)

5 件
(うち産業界との共同研究数 0 件)

8. 当初計画に対する各機関の達成度

課題開始時における達成目標であったドメイン間相互作用部位の特徴抽出、ドメイン間相互作用面推定法の研究開発、マルチドメインタンパク質の全体構造モデリング法の研究開発については、D1 課題終了年度である平成 21 年度までに達成することができた。一方で、C1 課題の活動として行うターゲットタンパク質の解析対象マルチドメインタンパク質の構造予測、および、その情報公開は、平成 23 年度いっぱいかかる予定であるため、達成度は 85%と自己評価する。

9. 課題内の情報共有・連携体制

年数回、不定期に各研究室や学会などで打ち合わせを行い、情報共有、進捗状況の把握に努め、互いの支援研究、個別解析の状況、また解析ツールのアップデートなどの情報を共有している。平成 23 年度には、関東圏の電力不足に対応し、コンピュータ資源を長浜バイオ大学に配置した。また、情報プラットフォームの会合には由良、藤、塩生のうちの誰かが必ず出席し、その内容の共有を行った。

10. 他の課題との情報共有・連携

本プログラム内では、菅原秀明教授のグループに協力し、タンパク質解析ツール群の調査を行った。

情報C1-F

本プログラム外では、愛知医科大学の渡辺秀人教授のグループとコンドロイチン硫酸合成酵素のモデリングにもとづく機能解析を連携して行った。さらに、長浜バイオ大学の水上民夫教授、長谷川慎准教授のグループとプロテアソームの立体構造情報にもとづく新規阻害剤の結合様式についての解析、同じく、長浜バイオ大学の中村卓講師のグループと好熱菌由来のハロ酸デハロゲナーゼファミリーに属する酵素についての機能予測を共同で行っている。

1.1. 人材育成

本課題の期間中、学部生34名、大学院生3名に対して、タンパク質の機能部位予測などの構造バイオインフォマティクスに関するテーマを与え、卒業・修士論文にまとめるまでの指導を行った。

また、愛知医科大学の渡辺秀人グループとの共同研究に際して、リモートホモログの取得やホモロジーモデリング法について習得してもらった。さらに、物理系の出身であるプログラム技術者を雇用し、人材が求められているバイオインフォマティクス分野に貢献する人材として育成している。

1.2. 終了までの具体的な見通し

現在、開発したマルチドメインタンパク質のモデリング手法について更なる精度向上を行うための改良を行いつつ、それらの結果にもとづく論文を平成23年度終了までに受理されるよう準備している。また、本プログラム内で解析対象となっているマルチドメインタンパク質の中で、現在も全体構造の決定がされていない解析困難なものについて、開発した手法により構造モデルを作成し、Web上で公開できるように解析を行っている。

1.3. 本プログラムにおける成果が我が国の科学技術発展及び産業応用に与える効果について

マルチドメインタンパク質の全体構造予測も含め、難構造解析のタンパク質に対してのバイオインフォマティクスの技術を用いた立体構造予測は、それらの構造情報を得るための唯一の方法である。開発した手法の精度をより高め、さらに、予測した立体構造を一般に公開することにより、それらの構造情報を用いたタンパク質機能解析や薬物設計などに資することができる。

1.4. 特記事項

塩生は本課題で開発した Web ツールの他にも、世界で最も早く開発・公開したタンパク質立体構造と低分子との相互作用のデータベース Het-PDB Naviの研究開発に携わった一人であり、現在は長浜バイオ大学からの公開・維持・更新を担当する責任者である。Het-PDB Naviの Web データベースでは、毎月1回データ更新をしており、本プログラムで立体構造が決定されたタンパク質と相互作用する低分子の情報も掲載している。国内だけでなく、アメリカなど海外からのアクセスも頻繁にある重要なデータベースとなっている。

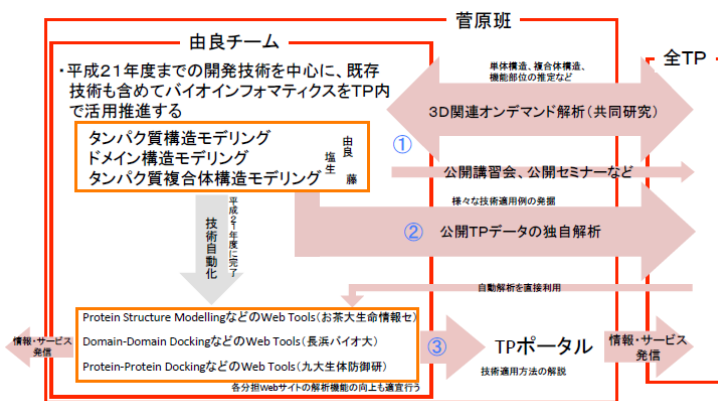
1.5. 研究費一覧

	19年度	20年度	21年度	22年度		23年度	計
設備備品費(千円)	0	0	0	0	物品費(千円)	0	
試作品費(千円)	0	0	0	0	人件費・謝金(千円)	0	
人件費(千円)	0	0	0	144	旅費(千円)	110	
業務実施費(千円)	0	0	0	3,856	その他(千円)	2,990	
間接経費(千円)	0	0	0	1,200	間接経費(千円)	0	-
合計(千円)	0	0	0	5,200	合計(千円)	3,100	8,300

代表／分担機関の課題名	「タンパク質の複合体構造を推定するための構造バイオインフォマティクス」(ターゲットタンパクプログラムの情報解析支援)
代表／分担機関	独立法人 産業技術総合研究所
機関の代表研究者名	藤 博幸

1. 課題開始時における各機関の達成目標

図 1



情報プラットフォーム内でお茶の水女子大学、産業技術総合研究所および長浜バイオ大学が協力して構造バイオインフォマティクスの先端技術を提供するプラットフォームを実現する。難解析タンパク質複合体の立体構造解析を支援するために開発された構造バイオインフォマティクス技術の共用利用の支援と技術向上を本事業目標とする。タンパク質の高精度立体構造モデリングや相互作用部位推定などの技術を、ターゲットタンパクプログラム内で

インターネットを介して共用し、オンデマンドで解析を支援する。

- ・ターゲットタンパク実験グループの支援

これまで研究支援をおこなってきたターゲットタンパクのメンバーとの情報解析支援を実施する。

- ・プログラム新規データの独自解析推進

ターゲットタンパクプログラムから産出された立体構造データを、過去3年間に本プログラムで開発したマルチドメインタンパク質のモデリング法や、ホモロジーモデリング法およびタンパク質複合体のモデリング法、既存のバイオインフォマティクス技術で解析することによってプログラム内に技術の適用例を示し、ターゲットタンパク研究チームの進展のための手がかりをプログラム内に開示する。

2. 平成23年5月末時点における各機関の事業計画に対する研究成果

(1) 成果概要

- ① ターゲットタンパク実験グループの支援 (上記図1の①に相当)

九州大学の住本教授から以前より依頼されていた、Nox系のペルオキシダーゼの触媒ドメインの分子系統解析に着手した。平成22年度は予備的な解析を行なったが、cyclooxygenase (COX)の中で二量体形成のインターフェースとなる部分について新規知見が得られた

- ② プログラム新規データの独自解析推進 (上記図1の②に相当)

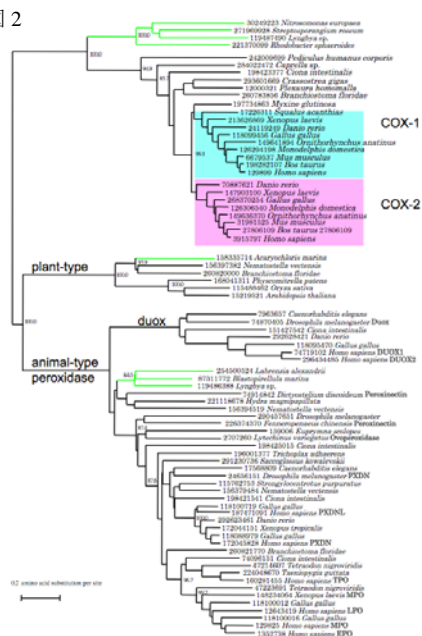
ターゲットタンパクの対象タンパク質の内、複数の実験グループの解析対象となっているユビキチンファミリーと、GPCRの一種であるケモカイン受容体について解析を行なった。ユビキチンについては、その低分子ハブとしての機能と表面電荷、disorderの関係を解析した。ケモカイン受容体については、生体内でシグナル伝達に関与するものの他に、構成的にシグナル伝達活性を有するウイルス受容体、またリガンド結合能のみを有するデコイ受容体など、機能的に異なるものの比較から、GPCRの作用機構に関する新規知見をえることができた。

(2) 技術開発の進捗及び成果

- ① ターゲットタンパク実験グループの支援

平成22年度は、九州大学の住本教授から以前より依頼されていた、Nox系のペルオキシダーゼの触媒ドメインの分子系統解析に着手した。データ数が多いことから、本年度は予備的な解析を行なったが、その中から新規の知見が得られた。以前にも、我々はこのファミリーの分子系統解析を行っており、cyclooxygenase (COX)の中で、ドメインとして注目されていなかったN末側の領域に

図 2

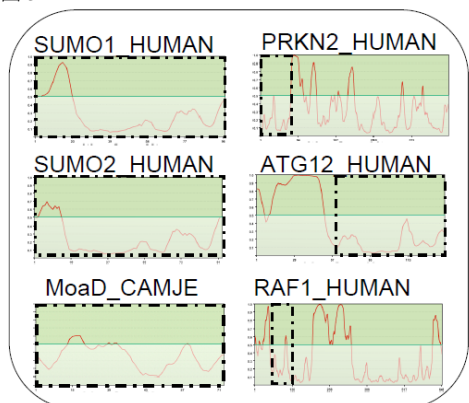


二量体形成のためのインターフェースとなる部分が存在することを明らかにした。今回、新たにデータベース検索を実施したところ、バクテリアに由来するcyclooxygenaseに近縁なタンパク質が検出された。下図は、COXとそのホモログの系統樹である。

緑色の枝が、真正細菌由来のCOX/ミエロペルオキシダーゼのホモログである。特にCOXに近縁な4種由来のタンパク質は、COXを特徴づけるEGF様ドメインや膜結合ドメインは持っていなかったが、触媒ドメインにおいては有意な類似性を示した。さらに、そのN末には触媒ドメインを越えてCOXと有意な配列類似性が観察された。EGF様ドメインも、その一部は二量体のインターフェースであることから、上記の領域だけではこれら真正細菌由来COX様タンパク質が二量体を形成するとはいえないが、実験等による検討が待たれる。一方、今回の解析から、ドメイン構造を完全に保持したCOX様タンパク質は刺胞動物から脊椎動物まで分布しているが、前口動物からは見つかっていない。今回検出された不完全なドメイン構造をもつCOXも、真正細菌全体で見いだされるのではなく、その分布は細菌内部でも非常に限られたものであり、その進化的な意味は、

今後のデータの増加を待たねばならないが、本ファミリーについて、進化的観点からばかりでなく、機能的な観点からも興味深い結果が得られたと考えている。今後、本ファミリーの解析をさらに進め、本ファミリーについて新たな知見を抽出していきたい。

図 3



② プログラム新規データの独自解析推進

ターゲットタンパクの研究促進あるいは研究成果の拡張を行なうことを目的として、ターゲットタンパクの解析対象となっているタンパク質に対して、情報の観点から独自に解析を実施した。そのため、由良、塩生と会合を開き分担を決めた。

藤は、分担タンパク質の内、複数の実験グループの解析対象となっているユビキチンファミリーと、GPCRの一種であるケモカイン受容体について解析を行なった。ユビキチンは、低分子ハブとして働き、そのハブとしての機能は表面電荷に依存することが中村らによって報告されている。今回、ファミリー全体で解析したところ、ユビキチンファミリーの中にも長いdisorder領域を持つものがあることがわかった。図2はユビキチンファミリーの代表的なメンバーに対してPOODLEを使用してdisorder予測を行なった結果である。SUMO1やPRKN2、またATG12などのN末にdisorder領域が存在することが示唆された。

そこで、ユビキチンファミリーの機能とこれらの構造的な特徴との関係をさらに調べることにした。まず、ユビキチンファミリーで立体構造が決定されているものを収集し、構造がないものについてはホモロジーモデリングで構造データをおこして解析に使用した。それらを、D, E, R, Kの頻度、疎水アミノ酸の表面での頻度、相互作用する相手の数、マルチドメインタンパク質の一部であるか否か、重合能、disorder領域の有無を要素とした9次元の特徴ベクトルで表現した。それらの特徴ベクトルを用いて主成分分析を行なったところ、まだ予備的な結果であるが、第三軸から相互作用の数は、Eの頻度に負の相関を示し、disorderとは正の相関を示した。また、第二軸からdisorderとマルチドメインであることに正の相関が、またdisorderとKやRなど正電荷を持つアミノ酸の頻度の間にも正の相関が見いだされた。今後データ数などもあわせてさらに検討していきたい。

ケモカイン受容体については、生体内でシグナル伝達に参与するもの他に、構成的にシグナル伝達活性を有するウイルス受容体、またリガンド結合能のみを有するデコイ受容体など、機能的に異なる近縁なものが多数存在しており、配列解析からそれら機能の差異をもたらすサイトを推測し、構造上に対応する残基をマッピングした。機能の差をもたらすサイトの推定には、自作のプログラムを用いたKullback-Leibler情報量による方法と、GuのDIVERGEを使用した方法の2種類を試みた。前者は、アラインメントを二つのグループに分割して、グループごとサイトごとにアミノ酸の組成を推定し、アミノ

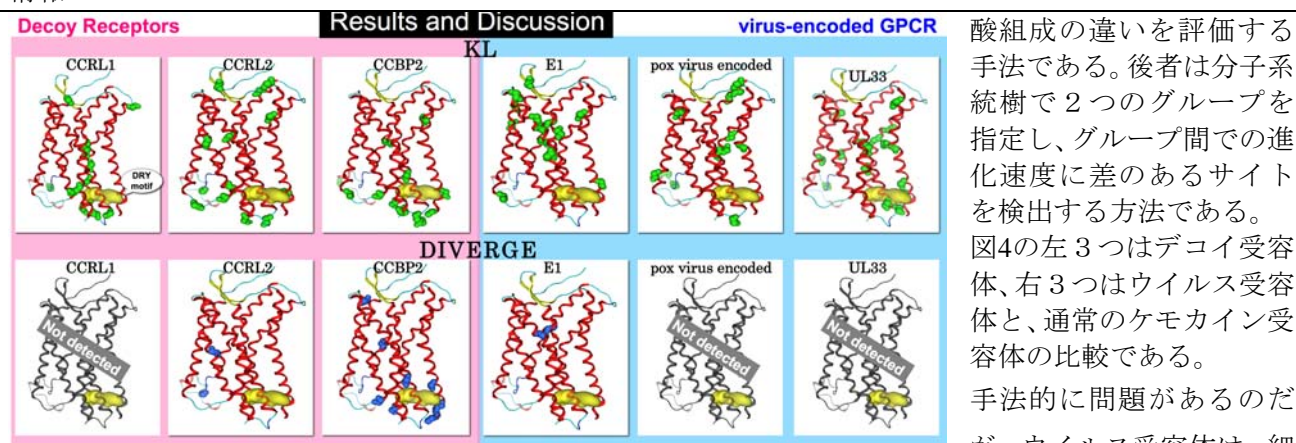


図 4

酸組成の違いを評価する手法である。後者は分子系統樹で2つのグループを指定し、グループ間での進化速度に差のあるサイトを検出する方法である。図4の左3つはデコイ受容体、右3つはウイルス受容体と、通常のケモカイン受容体の比較である。手法的に問題があるのだが、ウイルス受容体は、細胞外のリガンド結合に関与

すると思われる領域に、デコイ受容体については細胞内で受容体の活性/非活性のコンフォメーションの変化に関与すると考えられる DRY もチーフの近傍に、機能差を生じるサイトが集中する傾向が見いだされた。解析手法の開発も含めて、本予備的解析の結果の確認を行なっている。

また、図1の③のポータルサイトの充実のため、D1 班時代に開発したタンパク質間相互作用予測のさらなる改良を試みている。D1 班時代の開発手法は、サーバをたて公開している (ESPAER、<http://esper.cbrc.jp/~ESPER/form.php>、図5参照)。現在、震災に伴う節電のため停止中だが、由良、塩生と開発した複合体モデリングサーバーToSY に組み込まれている。最近、別プロジェクトでタンパク質の構造的制約と機能的制約を分離するために開発したプロファイル比較法により、インターフェイス構成残基を効率的に予測できることを偶然発見した (図6)。図6でオレンジ、シアンで色をつけたアミノ酸がプロファイル比較により機能的制約が働いている残基である。空間充填モデルで示された触媒残基に色がついていることがわかる。加えて、複合体の状態でインターフェイスを構成するアミノ酸も、多数色がつけられており、プロファイル比較により選択されていることがわかる。現在、この手法を導入することによる本システムのさらなるパフォーマンスの向上させる方法の開発を計画している。

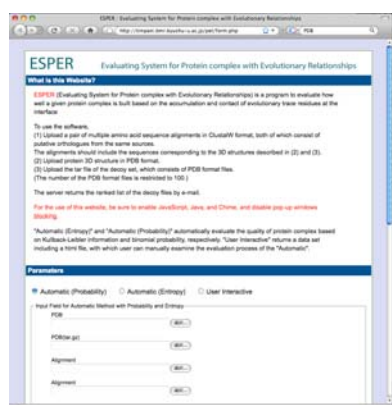


図5 ESPER トップページ

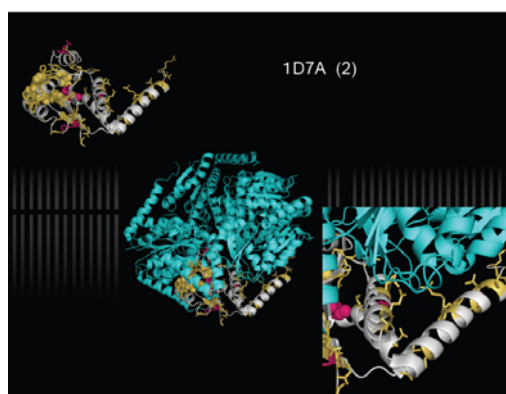


図6 プロファイル比較の結果

論文発表、公開したツールやデータベースの件数は以下の通りである。

- (1) 計算サーバ 1件

ESPER (<http://esper.cbrc.jp/~ESPER/form.php>) 震災後の節電のため停止中

- (2) 論文 6報 (別紙参照, D1班からの継続研究ためD1班時代の成果も含む)。

当初、設定していたターゲットタンパク内外の支援は着実に進行している。しかし、人的資源の不足のため、ターゲットタンパク全体を縦断する形での支援や個別解析は行なっていない。

論文や学会発表以外の外部への研究成果等の発信は以下の形で行なっている。

(1) D1 班時代に開発した手法は計算サーバ ESPER、統合モデリングサーバ ToSY、それらへの情報プラットフォームからのリンクの3つにより成果を発信している。

(2) 支援研究、個別解析の結果は、まとめ次第、情報プラットフォームを通じて、対応するターゲットタンパク内の実験研究者との情報共有をはかりたい。

本プログラムは C1 班になってからは、技術開発ではなく、情報からの研究支援を行なうことがミッションとして要求されているので、本プログラムからの技術基盤形成は本来ありえないが、D1 班時代の開発技術のアップデートを継続しており、その点が基盤整備と位置づけられる。藤の担当の複合体予測は、タンパク質の構造解析や複合体形成阻害による創薬などを支援する基盤技術を与える。上記のプロファイル比較手法の導入から、より高い性能の予測技術が開発できると考えている。

オンデマンドの解析については、当然のことながら結果を報告し、情報を共有してきた。一方、個別解析は平成 22 年度からのスタートであり、まだ予備的な段階であるため、情報の共有ははかられていない。まとめ次第、情報プラットフォームを通じて情報共有を進めていきたい。

3. タンパク質の構造解析件数

PDB 登録数	0 件
PDB 登録の有無にかかわらず構造解析したタンパク質数	0 件

4. 各機関の論文発表件数 6 件

5. 各機関の公開したツールやデータベース等の件数 1 件

6. 各機関の特許出願件数 0 件 (うち国外 0 件)

7. 各機関の成果を活用した共同研究数 (プログラム内外含む) 0 件 (うち産業界との共同研究数 0 件)

8. 当初計画に対する各機関の達成度

現時点での達成度ではあるが、60%程度であると考えている。ユビキチン系、ケモカイン受容体の系については、興味深い知見が得られてきており、基礎研究としてターゲット内の研究者との共同研究や、産業応用につながっていくのではないかと期待している。また、ターゲット内ばかりでなく、外部の研究者からの様々な要求に対応する形で、支援を行なっている。本プログラムの目的である実験グループの支援に加え、D1 班時代の成果である相互作用予測技術のアップデートについても検討しており、本プログラムとは異なる課題で研究していたプロファイル比較の手法の導入を計画している。

9. 課題内の情報共有・連携体制

本課題では、ポスドクを 1 名と大阪大学の研究協力者 1 名と協力して、ターゲットタンパク内外の研究者の実験支援を行なっている。ポスドクについては、進捗状況は、週 1 回のミーティング、月 1 回の monthly report で把握した。研究協力者については、主にメールで連絡をとりあい、情報を共有している。

年数回、不定期に代表の由良の研究室や学会などで打ち合せを行い、代表機関及び分担機関間の情報共有、進捗状況の把握に努め、互いの支援研究、個別解析の状況、また解析ツールのアップデートなどの情報を共有している。

10. 他の課題との情報共有・連携							
<p>ターゲットタンパク参加者である九州大学の住本英樹教授、九州大学 福井宣規教授、大阪大学 中村春木教授と実験支援（配列解析、モデリング）や共同研究を行なった。また、情報プラットフォームの運営会議に出席し、C1 班内での情報共有を行なった。</p> <p>本プログラム以外の研究者としては、東亜大学、渡部紀久子教授、岡山県立大学 高橋吉孝教授、九州大学 近藤久男教授、九州大学 横溝岳彦教授、九州大学 吉開泰信教授、大阪大学 森正樹教授、東京農工大 黒田裕准教授と共同研究を行なった。</p>							
1. 人材育成							
<p>ポスドクには週一回のミーティング、月報会で研究状況の把握に努めた。研究指導は、藤が主になって行なっている。ポスドクの本課題終了後の職などは未定。</p>							
12. 終了までの具体的な見通し							
<p>研究代表者の藤は、平成 22 年度より産業総合技術研究所生命情報工学研究センターへ異動し、そちらで C1 班として、個別解析、共同研究などを通じての支援を行なっている。サーバも生命情報工学研究センターに移設したが、現在震災後の節電対策のため停止している。研究成果の活用の推進は情報プラットフォームからのリンクを介して行なっている。支援のためには、開発した技術をアップデートしていく必要もあり、そのための取り組みも行なっている。先に述べたプロファイル比較などの手法の考慮は、相互作用予測システムのアップデートの一環である。研究支援、個別研究については、上に述べたものについては、既に予備的な解析は終了しており、これからまとめていく段階である。現在おこなっている3つの支援研究、個別解析の内、2つまでは23年度にまとめることができると考えている。</p>							
13. 本プログラムにおける成果が我が国の科学技術発展及び産業応用に与える効果について							
<p>ユビキチンの個別解析は、相互作用とネットワークという現在のタンパク質科学の研究の大きな流れの中にあるもので、タンパク質のハブ化について新たな知見をあたえてくれると考えている。また、ケモカイン受容体の解析は、ケモカイン受容体に留まらず、GPCR の作用機構についての一般的知見を与えてくれると期待している。また、D1 班時代からの相互作用解析のアップデートについては、プロファイル比較の導入によるさらなる性能改善を考えている。現在の予測システム、あるいは今後改良されるシステムは相互作用予測だけではなく、相互作用部位のキー残基が予測できるので、その情報を利用したプブチドミメティクスによる創薬の基盤を構成できる可能性がある。</p>							
14. 特記事項							
<p>Shimizu, K., Toh, H. <i>J. Mol. Biol.</i> 392, 1253-1265 (2009) は Faculty of 1000 Biology に選抜</p>							
15. 研究費一覧							
	19年度	20年度	21年度	22年度		23年度	計
設備備品費 (千円)				0	物品費 (千円)	0	
試作品費 (千円)				0	人件費・謝金 (千円)	6,257	
人件費 (千円)				6,236	旅費 (千円)	267	
業務実施費 (千円)				1,533	その他 (千円)	476	
間接経費 (千円)				2,331	間接経費 (千円)	0	2,331
合計 (千円)				10,100	合計 (千円)	7,000	17,100