|  |  |  |
| --- | --- | --- |
| **連携最適化による操業支援** | | |
| **抄 録** | 所属：イノベーションセンター　プロジェクトデザイン部 | 鎌田　健一 |
| 所属：イノベーションセンター　プロジェクトデザイン部 | 鵜飼　翔太 |
| 所属：イノベーションセンター　プロジェクトデザイン部 | 熊谷　渉 |
| 所属：イノベーションセンター　プロジェクトデザイン部 | 征矢　紘明 |
| 所属：イノベーションセンター　プロジェクトデザイン部 | 井本　量子 |
| 長期経営構想にて注目すべき事業領域と定められているバイオエコノミーにおいて、触媒タンパク質である酵素をベースに様々な有用化合物を生み出す「バイオ系物質生産」は中核領域だと考えている。ただし、任意の反応を触媒する酵素を人工設計する技術は未踏の領域であり、「タンパク質設計技術」は酵素の可能性を引き出す上で鍵を握る要素技術となりうる。一般的に、産業用酵素は天然のタンパク質、もしくはその点変異体が用いられることが殆どであり、天然タンパク質とは異なるが機能性を有する有用な配列を探し出すことは難しいとされている。本テーマでは、物理化学シミュレーションやAI/ML技術などの計算論的アプローチを駆使して、①*in silico*での構造・配列の生成、②*in silico*での機能性評価、③Wet実験での機能性評価、という要素技術からなるタンパク質設計戦略を想定し、効率よくタンパク質の改変・設計可能か検証した。研究試作にあたって、バイオマスリファイナリの重要酵素であるセルロース分解酵素の部分モジュールであるセルロース結合性タンパク質を題材とした。機械的・物理化学的に複雑な酵素を初手で扱うのは難しいと考え、まずはその部分モジュールである分子結合性タンパク質に焦点を絞った。設計プロトコルを通じて、題材とした天然タンパク質とはアミノ酸配列ベースで最大で33%異なるタンパク質変異体を獲得できた。一方で、当初想定したタンパク質設計戦略と戦略を実現する要素技術の一つである「設計タンパク質の機能性の机上評価とスクリーニング」について、現状では実現の目途がつけられなかった。これは、設計プロトコルを通じて獲得できたタンパク質変異体が偶然に任せた探索に対して十分効率の良く探索出来た結果なのか判断できないことを意味する。この結果を踏まえて、本テーマはLR2審査をもって研究テーマを中止した。 | | |
|  | | |

**目次**

[1. 研究開発の概要 4](#_Toc119585049)

[1.1 研究のスコープ 4](#_Toc119585050)

[1.2 研究の概要 4](#_Toc119585051)

[2. 研究開発のスケジュール 7](#_Toc119585052)

[3. 研究開発体制 8](#_Toc119585053)

[4. 開発工数 8](#_Toc119585054)

[5. 設備投資，使用経費 9](#_Toc119585055)

[5.1 設備投資リスト 9](#_Toc119585056)

[5.2 経費負担 9](#_Toc119585057)

[5.3 社外からの補助金，委託金等 9](#_Toc119585058)

[6. 要素技術ごとの検証内容と成果 10](#_Toc119585059)

[6.1 研究目的 10](#_Toc119585060)

[6.1.1 Rosettaによる立体構造最適化と側鎖改変 10](#_Toc119585061)

[6.2 モデリング技術 14](#_Toc119585062)

[6.2.1 非線型性と動特性に対応可能なモデリング技術の検証 14](#_Toc119585063)

[6.2.2](#_Toc119585064) [非線型性と制約抽出に対応可能なモデリング技術の検証 14](#_Toc119585064)

[6.2.3モデリング技術のまとめ 14](#_Toc119585064)

[6.3 最適化技術 15](#_Toc119585065)

[6.3.1 最適化技術の開発目標 15](#_Toc119585066)

[6.3.2 開発した最適化技術 15](#_Toc119585067)

[6.3.3 非線型性と有制約に対応可能な最適化技術の検証 17](#_Toc119585068)

[6.3.4 非線型性と混合変数に対応可能な最適化技術の検証 20](#_Toc119585069)

[6.3.5 最適化技術のまとめ 23](#_Toc119585070)

[6.4 まとめ 35](#_Toc119585084)

[6.5 参照文献 36](#_Toc119585085)

[7. 口頭発表，講演リスト 36](#_Toc119585086)

[8. 社外発表論文リスト 36](#_Toc119585087)

[9. 出願特許リスト 36](#_Toc119585088)

[10. 主要な関連社内報告書，LRの資料，議事録等のリスト 37](#_Toc119585089)

# 研究開発の概要

長期経営構想にて注目すべき事業領域と定められているバイオエコノミーにおいて、触媒タンパク質である酵素をベースに様々な有用化合物を生み出す「バイオ系物質生産」は中核領域である。ただし、任意の反応を触媒する酵素を人工設計する技術は未踏の領域であり、「タンパク質設計技術」は酵素の可能性を引き出す上で鍵を握る要素技術だと考えている。本テーマでは、計算機科学・合成生物学・遺伝子工学等を駆使した再現性のある酵素設計手法を模索し、天然物やその改変とは異なる人工酵素の創製とそのための要素技術獲得を目指した。

## 研究のスコープ

2019年度に行ったPhase1では、近い将来におけるバイオマス資源を活用した素材産業（狭義でのバイオエコノミー）の大きな進展を想定し、その根幹となるバイオマス分解を担う手段の一つであるセルロース分解関連酵素（以下、セルラーゼ）に焦点を当て、セルラーゼの人工設計の実現可能性を検証した。その検証の中で、触媒反応を含む化学反応メカニズム全体の人工設計は現状では解決すべき課題が多いことがわかってきた。そこで、Phase2では、セルラーゼの重要な構成要素の一つであるCBD（cellulose binding domain：セルロース結合ドメイン）に注目し、セルラーゼからCBDを単離した、セルロース結合性タンパク質の人工設計の実現可能性検証にスコープを絞って研究を開始した。

対象をセルラーゼのうちセルロース結合性タンパク質に絞った理由は、以下の3点である。

* 設計及びシミュレーションにおいて、量子化学的反応が支配的な触媒機能を扱うよりも、分子間力や静電相互作用の寄与が大きい結合機能を扱うほうが、技術的ハードルが低いと予想されたため
* 配列長が短く、設計、シミュレーション及び合成実験にかかる時間的、費用的コストを抑えられるため
* 評価実験系の構築が比較的容易であると予想されたため

本研究では、シミュレーションベースの設計プロトコルを実装したタンパク質改変のためのソフトウェアRosettaや、基質との結合シミュレーションを行うドッキング計算ソフトウェア、設計に関する最適化およびAI技術と、設計したタンパク質を実際に合成し、基質結合性を実験的に評価する手法を中心に調査を進め、実際的な機能性タンパク質設計のトライアルを行った。

## 研究の概要

機能性タンパク質の設計可能性検証を行うにあたり、本研究では、図 1のような設計プロトコルを想定した。

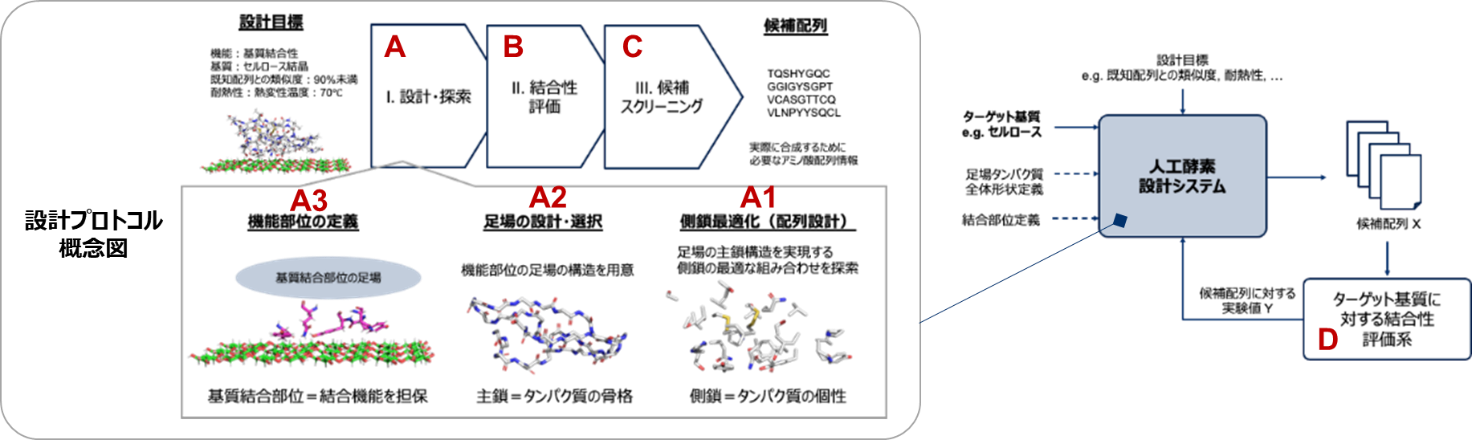


図 1 タンパク質設計プロトコル概念図

設計プロトコルを大きく分けると、タンパク質配列の候補生成（図 1中Aに相当）、設計タンパク質の機能性の机上評価とスクリーニング（図 1中BとCに相当）、Wet実験による機能性評価（図 1中Dに相当）の要素に分けられる。これらの要素技術について検証を進めることで、最終的に設計プロトコル全体を機能させることを目指した。

前提として、機能性タンパク質の人工設計を行うには、タンパク質分子という枠組みの中で目的の機能を持つと期待される候補を提案する技術が必要となる。そしてタンパク質分子の種類は、それを構成するアミノ酸配列によって定義されると考えられている。よって、タンパク質の候補を提案するためには、アミノ酸配列の候補を提案することができればよいと言える。

タンパク質は、数十から数千のアミノ酸がペプチド結合によって重合することにより構成されている分子であり、構成するアミノ酸は基本的には20種類であると言われている。これは、例え30残基という、タンパク質としては非常に小さな配列であったとしても、その組み合わせは20の30乗、すなわち約10の39乗通りという膨大なものになることを意味する。タンパク質の設計・改変を行うには、この膨大な配列空間の中から有望なものを探し出さなければならない。このような配列探索をすべて網羅的に行うことはまず不可能であり、人間の手で配列候補を見つけ出すにも限界がある。そこで、計算機的手法を用いて高速に大量の配列を評価し、配列空間を効率的に探索することが本研究の狙いである。

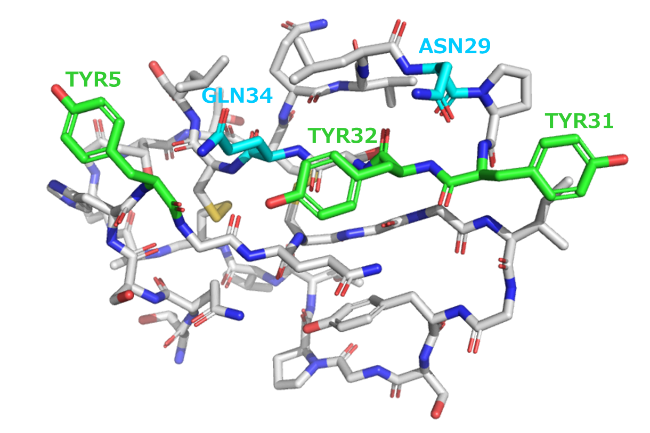
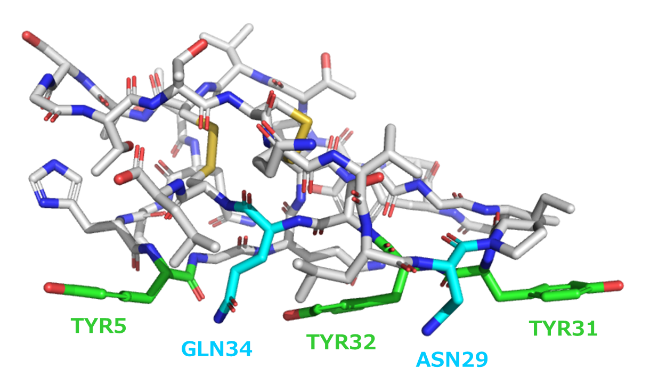
配列候補生成に取り組むにあたって、機能性タンパク質の改変・設計の技術には様々なレベルがあると想定されるが、本研究においては、最も基本的で容易であると考えられる、側鎖の改変に着手した。足場や機能部位については設計の難易度がより高くなるため本研究のスコープ外とし、天然タンパク質由来の既知のものを用いた。天然タンパク質としては、糸状菌Trichoderma reeseiの持つセルロース分解酵素TrCel7AのCBDである、TrCBM1（図 2）を主な対象とした。TrCBM1は36個という少数のアミノ酸残基からなる非常に小さなタンパク質であり、計算機上での取り扱いが比較的容易であると考えられる。



図 2 セルロース分解酵素TrCel7AのCBD である、TrCBM1（UniProt ID: P62694の478～513番目のアミノ酸残基）のアミノ酸配列 [1]。アミノ酸の種類は一般的な1文字表記で示している。結晶性セルロースに対する結合部位アミノ酸に下線を引いた。下線部のうち、緑色で示したものは芳香族アミノ酸、水色で示したものは親水性アミノ酸である。

計算機上でタンパク質の変異体を作り出すには、タンパク質配列の改変を行い、その配列に対応する立体構造を予測する技術が必要となる。本研究においては、タンパク質改変及び立体構造予測技術のベースとして、Rosettaを用いた。Rosettaは独自のエネルギー関数と探索アルゴリズムからなる、オープンソースのタンパク質モデリングソフトウェア群である。このソフトウェア群で実現されている技術のキャッチアップと実際にセルロース結合性タンパク質の変異体を探索するトライアルを目的に、立体構造の最適化計算、側鎖改変プロトコルを中心に論文・マニュアルの調査から実際の変異体候補探索までを実施した。Phase1では、酵素設計には量子化学的反応を考慮する必要があるためRosettaの適用は難しいとしていたが、結合性タンパク質の設計問題においてはRosettaが適用可能であると判断した。Rosettaには側鎖置換によるタンパク質の自由エネルギー変化の予測プロトコルであるCartesian DDG [2]や、Rosettaスコアが最適化されるように自動で配列探索を行うプロトコルであるFastDesignのほか、様々なプロトコルが実装されているが、本研究ではCartesian DDGとFastDesignの2つを配列の改変・探索に使用した。Rosettaのプロトコルを用いた改変については6.1に、FastDesignを用いた配列探索については6.2.1に記述する。

本研究ではTrCBM1の立体構造データをProtein Data Bank（PDB）から取得し（図 3）、Rosettaを用いて、TrCBM1の結合部位変異体、1～5点変異体、FastDesign設計、FastDesignから野生型配列へと戻していった変異体（BackMutationと命名）を作成した。結合部位変異体については、先行研究の情報をもとに、TYR5, ASN29, TYR31, TYR32, GLN34を結合部位とし（本文書内では、配列内のアミノ酸を略号と残基番号を並べた形で表すこととする。例えば5番目のチロシン残基であれば、アミノ酸3文字表記で「TYR5」あるいはアミノ酸1文字表記で「Y5」と表す）、基質結合性を失うと予想されるアラニンへの変異体31種と、維持あるいは強化されると予想されるチロシンから芳香族アミノ酸（トリプトファン、フェニルアラニン、ヒスチジン）への変異体63種を網羅的に作成した。1点変異体は36×19=684種、2点変異体は36×35÷2×19×19=227430種あり、これらを全てRosettaで作成し、スコアを評価した。3～5変異体については、Rosettaスコアをもとに、Rosetta外でのスコア最適化アルゴリズムによって配列探索を行った。3～5変異体についての詳細は6.2.2に記載する。FastDesignは4000種作成し、BackMutationはFastDesignのスコアの良いものを起点の配列として90種作成した。



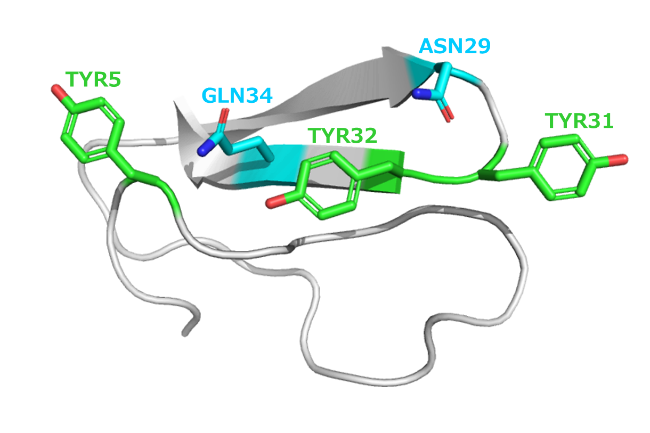
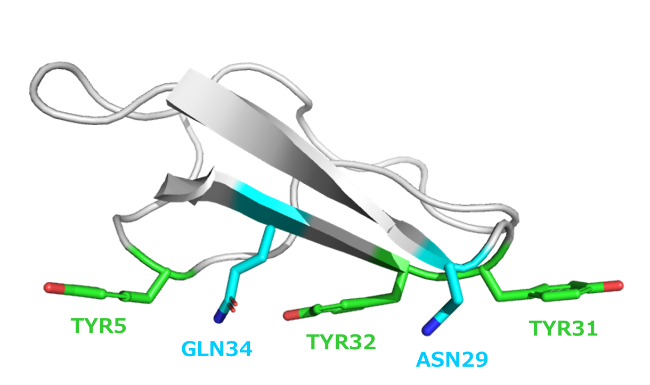


図 3 PDBから取得したTrCBM1の立体構造データ（PDB ID: 1CBH） [3]。アミノ酸の種類は3文字表記で示している。緑色と水色のアミノ酸残基は結晶性セルロースに対する結合部位であり、色は図 2と対応している。上段は水素原子を除く全原子を描画しており、下段は結合部位の側鎖以外をリボン表記で描画している。また、左列と右列は同じ立体構造を互いに異なる視点から見た図となっている。

候補配列の評価は、Rosettaスコアによるタンパク質安定性評価とAutoDock Vinaスコア（以下、Vinaスコアと呼ぶ）による基質結合性評価によって行った。AutoDock Vinaはタンパク質分子と基質分子の立体構造情報を入力として、複合体の結合の様式と結合親和性を予測するソフトウェアである。AutoDock Vinaを用いた評価の詳細は6.3.3に記述する。

Wet実験や計算時間のかかる机上評価に用いる配列は、TrCBM1の変異体の中から288種を選定した。スクリーニングは、机上評価によって適切なスコア付けを行うことができれば、スコアの良いものを残すことで実行可能だが、今回は机上評価の技術要素の検証を重視するため、結合部位変異体以外の変異体についてはRosettaスコアとVinaスコアが悪いものも含めて広範囲に分布するように選択した。また、TrCBM1以外のCBDも12種類評価対象に含めて、計300種をWet実験での評価対象とした。TrCBM1以外のCBDについてはPDB（Protein Data Bank）に登録されている配列から選択しており、机上評価は行っていない。

AutoDock Vinaの実行には時間がかかるため、代替手段として、より高速にスコアを予測可能な手法の実現が期待される。そこで本研究では、深層学習によるVinaスコアの学習・予測を試みた。この結果については6.3.4に記載する。また、RosettaやVinaによるスコアは物理化学的なエネルギーではない上、動的性質の評価はされていない。特にセルロース結晶との結合については、結晶表面での移動能力等の動的性質も重要であると想像されるため、本研究ではAmberを用いた分子動力学シミュレーションを実施し、物理化学的により正確な自由エネルギー計算手法で机上評価を行うことを試みた。この結果については6.3.5に記載する。また、近年登場したAlphaFold2では、配列からの構造予測の信頼度スコアを計算することができる。これを用いた評価については、6.3.2に記載する。

以上は既存のソフトウェアを用いた机上評価手法であるが、本研究では外部データベースからの特徴抽出や、実際の実験データを利用した特徴抽出も行った。PDBのデータによる分類モデルの作成や特徴抽出と、簡易評価実験のデータからの特徴抽出によって候補スクリーニングが可能かを試した。これについては6.4に記載する。

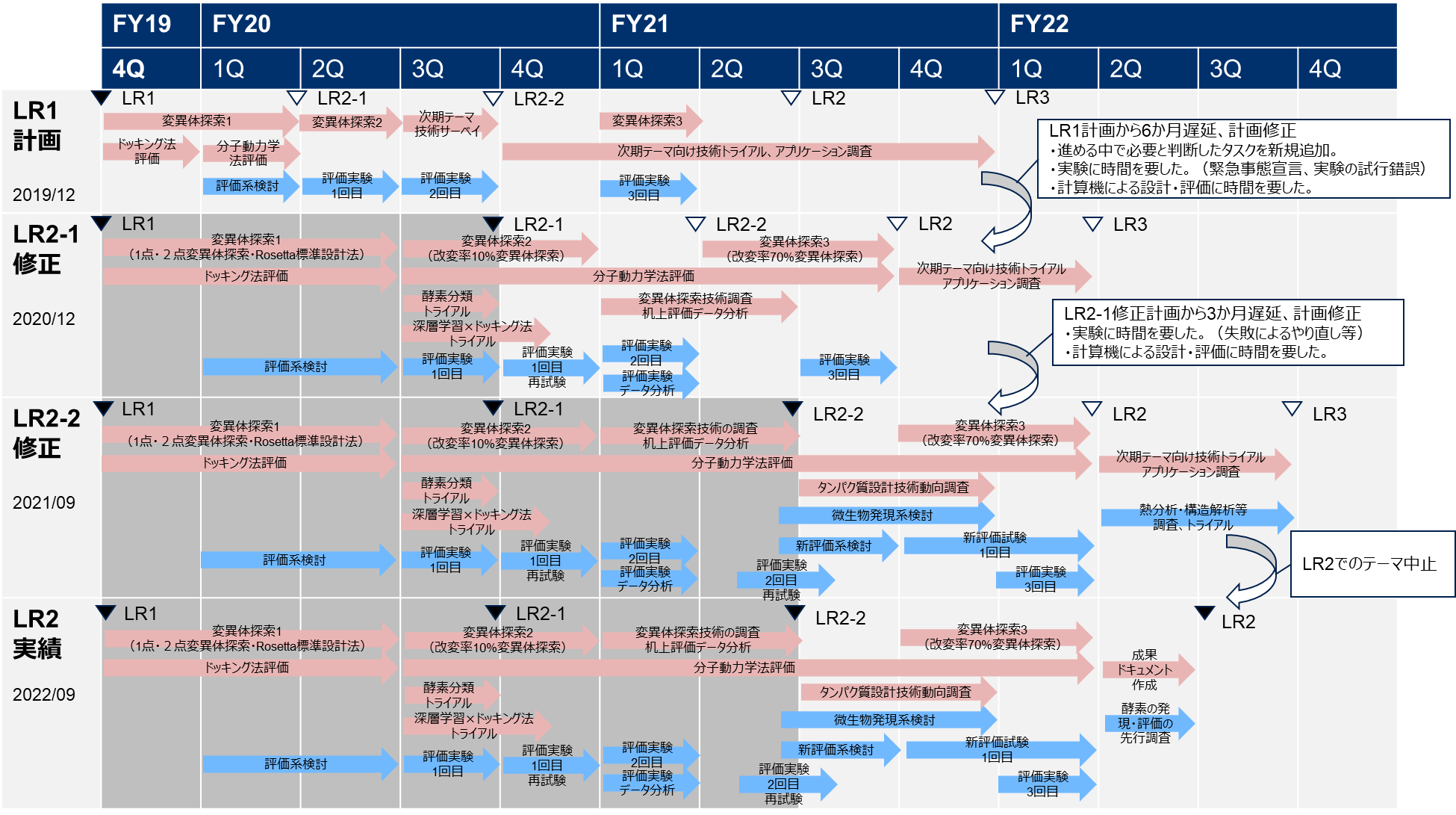
Wet実験によるスクリーニング評価に用いる発現系は、簡便で手間も少ないことから、コムギ胚芽無細胞合成系を採用した。また、今後NMR（Nuclear Magnetic Resonance: 核磁気共鳴分光法）を用いた実験等に取り組む可能性を考え、予備検討として、大腸菌およびメタノール資化酵母の発現系の検証も行った。

結合性の評価手法は、セルロースTLC（thin-layer chromatography：薄層クロマトグラフィー）プレートを用いた簡易的な方法と、タンパク質量を定量評価する方法を確立した。これらの実験的評価は、前述の300種の候補配列に対して行った。この結果については、6.6に記載する。

本テーマは、当初想定したタンパク質設計戦略と戦略を実現する要素技術の一つである「設計タンパク質の機能性の机上評価とスクリーニング」について、実現の目途がつけられなかったことから、LR2審査をもって研究テーマを中止した。

# 研究開発のスケジュール

|  |  |
| --- | --- |
| 実施日 | マイルストン |
| 2019年3月27日 | LR0報告 |
| 2019年12月25日 | LR1審査 |
| 2020年12月21日 | LR2-1中間報告 |
| 2021年9月29日 | LR2-2中間報告 |
| 2022年10月11日 | LR2審査、研究開発テーマの中止 |



# 研究開発体制

|  |  |  |
| --- | --- | --- |
| 役割 | 担当者 | 主な役割 |
| プロジェクト管理 | 鎌田　健一 | テーマ運営全般 |
| 要素技術開発 | 鵜飼　翔太 | ・モデリング技術の開発  ・モデリング技術の検証 |
| 熊谷　渉 | ・最適化技術の開発  ・最適化技術の検証 |
| 征矢　紘明 | ・モデリング技術の開発  ・モデリング技術の検証 |
| 井本　量子 | ・モデリング技術の検証 |
| 東京都立大学大学院  システムデザイン研究科  修士課程学生　数名 | ・最適化の要素技術となるアルゴリズムの開発と検証 |
| 社外アドバイザー | 東京都立大学大学院  システムデザイン研究科  安田　恵一郎　教授 | アドバイザー  （テーマ全般・アルゴリズムの検証に関する学生への指導） |

# 開発工数

|  |  |  |
| --- | --- | --- |
| 期 | 自部署(人月) | 他部署（人月） |
| FY19（上） |  | - |
| FY19（下） |  | - |
| FY20（上） |  | - |
| FY20（下） |  | - |
| FY21（上） |  | - |
| FY21（下） |  | - |
| FY22（上） |  | - |
| FY22（下） |  | - |
| FY23（上） |  |  |
| FY23（下） |  |  |
| 合計 | 13.9 | - |

# 設備投資，使用経費

## 設備投資リスト

自部署設備投資

|  |  |  |  |
| --- | --- | --- | --- |
| 設備名 | 目的 | 購入年月 | 購入価格 |
| 計算クラスタ CPU/GPUサーバ | Amber等の運用のため | 19年3月 | \22.4M |
| GPUサーバ | Amber等の運用のため | 20年6月 | \8.0M |
| Amber | 分子力学・分子動力学計算のため | 20年6月 | \3.0M |
| Reaction Plus | 量子化学計算用のため | 20年6月 | \2.5M |
| GPUサーバ拡張 | メモリ等の拡張のため | 21年9月 | \7.7M |
| 合計 | - | - | \43.6M |

他部署設備投資

|  |  |  |  |
| --- | --- | --- | --- |
| 設備名 | 目的 | 購入年月 | 購入価格 |
| - | - | - | - |
| 合計 | - | - | - |

## 経費負担

|  |  |  |  |
| --- | --- | --- | --- |
| 期 | R&D負担経費 | 事業部委託費 | 事業部負担経費 |
| FY19（上） | \5.1M | - | - |
| FY19（下） | \3.0M | - | - |
| FY20（上） | \6.3M | - | - |
| FY20（下） | \17.5M | - | - |
| FY21（上） | \9.3M | - | - |
| FY21（下） | \20.7M | - | - |
| FY22（上） | \7.0M | - | - |
| 合計 | \68.9M | - | - |

## 社外からの補助金，委託金等

補助金・委託金

|  |  |  |
| --- | --- | --- |
| 期 | 機関，団体 | 補助金，委託金額　（M\） |
| FY19（上） | - | - |
| FY19（下） | - | - |
| FY20（上） | - | - |
| FY20（下） | - | - |
| FY21（上） | - | - |
| FY21（下） | - | - |
| FY22（上） | - | - |
| 合計 | - | - |

# 要素技術ごとの検証内容と成果

セルロース結合性タンパク質の人工設計のため、本研究において取り組んだ要素技術の一覧を表 1に示す。

表 1 要素技術

|  |  |  |  |
| --- | --- | --- | --- |
| カテゴリ | 項 | 要素技術項目 | 関連する要素 |
| 配列候補生成 | 6.1 | タンパク質改変技術 | Rosetta |
| 6.2 | 配列探索技術 | Rosetta, 最適化 |
| 机上評価 | 6.3 | 安定性・基質結合性の机上評価技術 | Rosetta, AutoDock Vina, Amber,  AlphaFold2, 機械学習 |
| 6.4 | 構造・配列特徴抽出技術 | Protein Data Bank, 機械学習 |
| Wet実験評価 | 6.5 | タンパク質合成技術 | 無細胞合成系, 大腸菌発現系,  メタノール資化酵母発現系 |
| 6.6 | セルロース結合性評価技術 | セルロースTLCプレート,  定量評価系 |

本節では、各要素技術について、実施した検証内容とその成果を示す。

## 研究の背景【鎌田】

### Rosettaによる立体構造最適化と側鎖改変

Rosettaによるタンパク質設計・変異体探索は、側鎖の置き換え、主鎖・側鎖におけるねじれ自由度の中での構造変更、Rosetta独自のエネルギー関数を用いたエネルギー計算の3つの要素が主に組み合わって構成されている（図 4）。

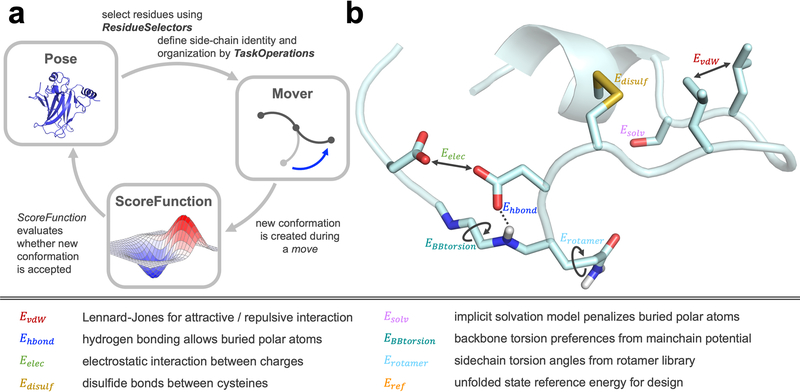


図 4 Rosettaの概念図 [4]

本共同研究においては、Rosettaに含まれる以下のプロトコルを使用して、タンパク質変異体探索や設計タンパク質の机上評価を行った。

* Score: Rosettaの独自スコアによる構造安定性の評価
* Relax: Rosettaスコアを最適化するように立体構造を緩和
* Cartesian DDG: 側鎖置換によるタンパク質の自由エネルギー変化の予測
* FastDesign: 主鎖構造を拘束してRosettaスコアを最適化するように側鎖改変の探索を実行

Rosetta Scoreは、入力したタンパク質立体構造を、Rosetta独自のスコア関数に基づいてスコアリングするプロトコルである。本研究においては、これをPDBで立体構造が提供されている野生型等の評価に用いた。Scoreに用いられるスコア関数は物理的に正確なエネルギーというわけではなく、開発者によってデータにフィットするようパラメータが最適化された関数となっており、年々改善が続いているものである。物理的に正確なエネルギーを用いないのは、計算量的が大きすぎることと、物理的に正確なエネルギーを静的構造に適用しても、自由エネルギーやタンパク質構造としての妥当性の予測はむしろ悪くなることが理由である。

Rosetta Relaxは、入力した立体構造をRosettaスコア上での安定構造に緩和させるプロトコルである。変異体探索の元となる野生型のタンパク質の立体構造には、PDBから取得したものを用いたが、PDBに登録されている構造は結晶構造であったりNMRによる構造の平均構造であったりするため、ほとんどの場合はRosettaスコアとしての安定構造ではない。また、側鎖を改変した場合はその配列での安定構造を探さなければ評価ができない。そこで本研究では、これらの外部データ由来の構造と設計によって得られる変異体をなるべく近い条件で比較するために、Relaxを用いた（図 5）。Relaxは、後述のCartesian DDGやFastDesignの内部でも使用されている。

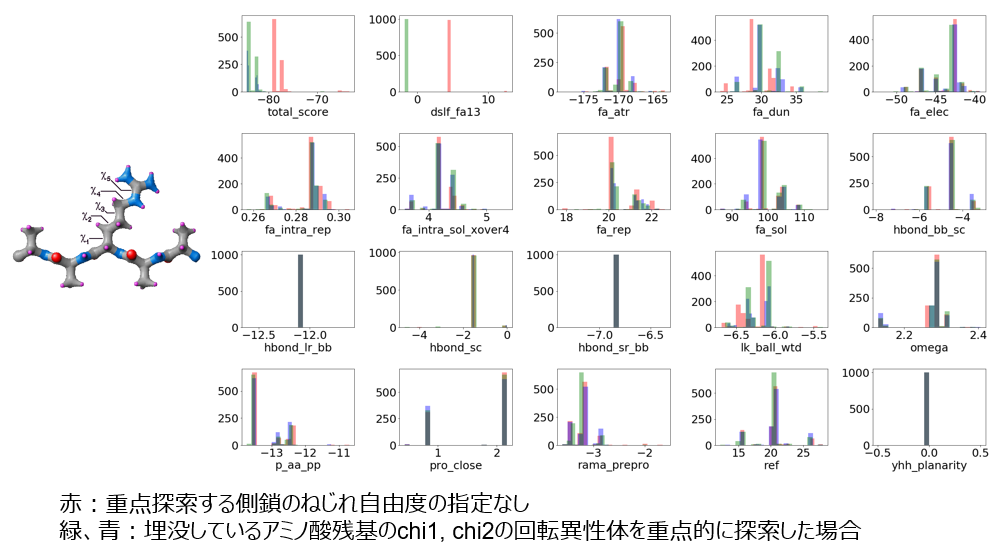


図 5 RelaxによるRosettaスコアの内訳の変化

Rosetta Cartesian DDGは、入力とするタンパク質の配列のうち、指定の部位を指定のアミノ酸に置換した際の構造と自由エネルギーを予測するプロトコルである。本研究では主に、1, 2変異体の網羅的な生成や、3～5変異体探索の際の候補生成に用いた。図 6と図 7に、1変異体684種、2変異体227430種を網羅的に生成した際の、改変による天然配列からの自由エネルギーΔGの変化量ΔΔGを計算したものを示す。ΔΔGの値は低いほどタンパク質として安定な構造であることを示す。

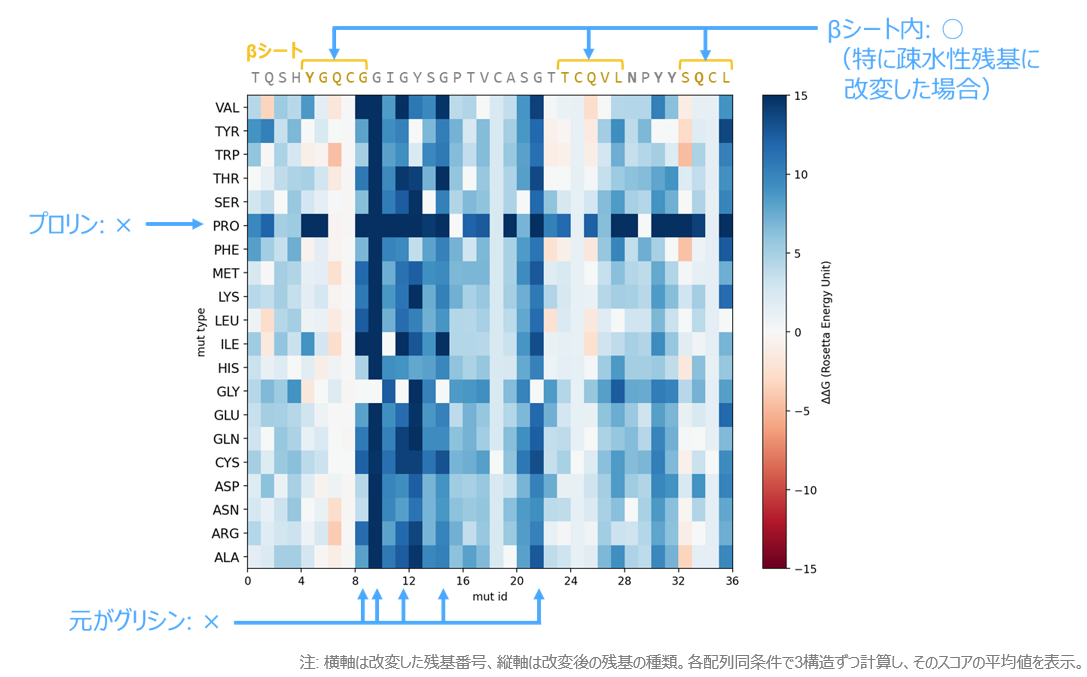


図 6 1変異体のΔΔG



図 7 2変異体のΔΔG

図 6図 7から、次のことがわかる。

* 野生型でグリシンとなっているアミノ酸を他のアミノ酸に置換すると不安定化する
* 野生型でプロリンでないアミノ酸をプロリンに置換すると不安定化する
* 野生型でβシート内のアミノ酸の置換は比較的安定化しやすい（特に疎水性アミノ酸に置換すると良い場合が多い）

また、2変異体のΔΔGから、各変異の1変異体のΔΔGの和を引いたものを計算することで、2つの変異が共同的に作用して安定化する可能性があるかの検証を行った。図 8がその結果である。この図から、特に第10残基と第33残基について、変異の組み合わせによって構造安定性が極端に変化することがわかる。2つの側鎖は野生型において互いに向かい合っていることから、この結果は、空間的に近接する側鎖を同時に改変することで側鎖間の相互作用によって構造安定性を変化させられる可能性があることを示していると言える。

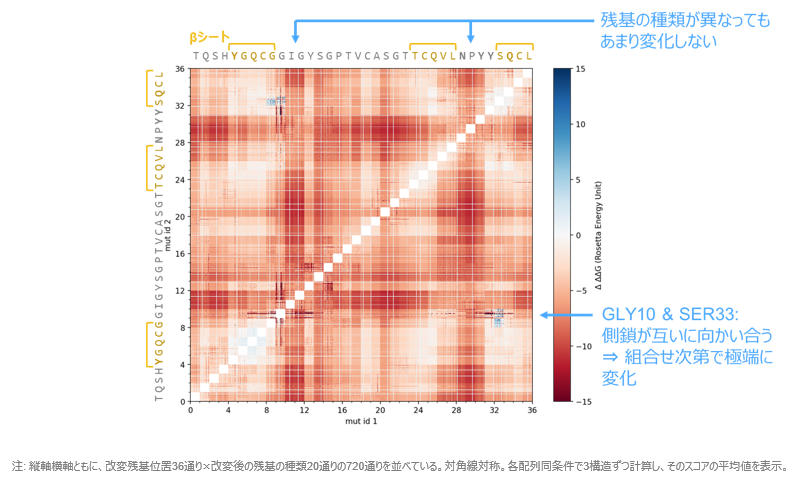


図 8 2変異体のΔΔGから、各変異の1変異体のΔΔGの和を引いた値

以上、本研究で用いた、RosettaのScoreによるスコア付け、Relaxによる構造緩和、Cartesian DDGによる配列改変のプロトコルを紹介した。これらは計算機上でのタンパク質改変を行う場面ではよく用いられている手法であるが、検証の中で課題点も見つかった。主な課題点としては、Rosettaスコアがタンパク質の安定性を予測する上で必ずしも精度の高いスコアとは限らない点、構造緩和の出力が入力構造に強く依存しており局所最適解に陥る可能性がある点、Cartesian DDGによる配列改変後はRelaxによって最終的な構造が予測されるため、配列のフォールド可能性を十分に評価できない点が挙げられる。

## モデリング技術【鵜飼・熊谷・鎌田】

### 非線型性と動特性に対応可能なモデリング技術の検証【鵜飼】

Rosetta FastDesignは、入力したタンパク質の主鎖構造を固定して、Rosettaスコアを最適化するように側鎖改変の探索を実行するプロトコルである。出力される配列候補は、全体の構造を変えないまま、改変率が非常に高くなることが多く、本研究では、結合部位の配列を固定した探索を行うことで、改変率と結合性の高さを両立した配列候補が得られるかという検証のため利用した。図 9に、FastDesignにより設計された変異体配列のシーケンスロゴを示す。図 9のシーケンスロゴ中で大きく書かれている文字と元の配列を見比べると、FastDesignによって、結合部位のアミノ酸を保存する制約下で、入力した野生型の配列からの改変率が高い配列を得ることができたことがわかる。

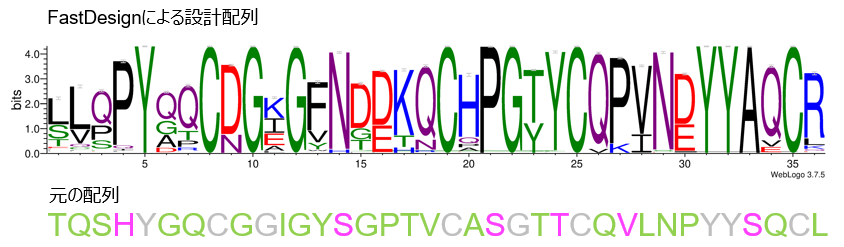


図 9 FastDesignにより設計された変異体配列のシーケンスロゴ [5]。元の配列において灰色で示しているアミノ酸は、重要な結合部位として機能する部位や全体の構造への寄与が大きいと考えられた部位であるため、変異を行わないように指定した。下段の配列における文字色は、FastDesignにより生成した変異体に占める該当位置のアミノ酸の割合が、元と異なる1つの種類が支配的な位置はピンク色で、プロトコル上の制限により元と同じ種類のみの位置は灰色で、その他の位置は緑色で示した。

ただし、本研究で目指す設計プロトコルを構成する上では課題点もある。FastDesignはRosettaスコアに基づく探索プロトコルであるため、基質結合性の維持については、結合部位の配列の保存をすること以上の操作ができず、基質結合性を担保した探索が困難である。また、探索アルゴリズムはプロトコルで定義されている物を使用することとなるため、より複雑な制約の下での探索や異なるアルゴリズムの検証が難しい。そのため、本研究においては、FastDesignのみを用いた配列生成は参考にとどめ、FastDesignにより得られた配列の改変率の高さは本研究の成果としていない。

FastDesignによって得られた配列のWet実験による基質結合性検証については、6.6の結果に含める。

### 非線型性と制約抽出に対応可能なモデリング技術の検証【熊谷】

本項では、カーネルPCAの検証方法と結果を記載する。変異数制約を満たすWild-Typeの変異体候補を探索・生成する方法を確立した。具体的には、Rosettaの設計プロトコルに変異数制約を加えた上位最適化を組み込むことで、全探索やランダム探索よりも効率的に変異体を生成した。Rosettaの設計プロトコルは、Cartesian DDGを使用した。Cartesian DDGは、Wild-Typeのアミノ酸配列の変異位置・残基を入力すると、それに整合するように最小化されたエネルギースコア（Rosettaスコア）と構造データを出力する。しかしながら、指定した変異数内での候補配列を得るには、全探索やランダム探索によって与える方法は非効率である。よって、候補生成の効率を上げるために、Cartesian DDGに上位最適化を組み込んだ技術を検討した。

図 10に今回検討した技術の概要を示す。指定した変異数内の変異位置・残基の組を、上位最適化の最適化変数に設定し、目的関数値であるエネルギースコア（Rosettaスコア）を最小化するような、最適化問題を構築した。最適化アルゴリズムはLocal Searchを使用したため、局所解のいずれかに収束する。

図 11に、1CBHのWild-Typeに三点変異を指定した場合の探索で得た、変異体の立体構造例を示す。三点変異体のアミノ酸残基に変異した部位を紫色で示している。Wild-TypeのRosettaスコアは-127.9だが、三点変異体のRosettaスコアは-157.8であるため、大きく改善した変異体が得られていることが確認できる。

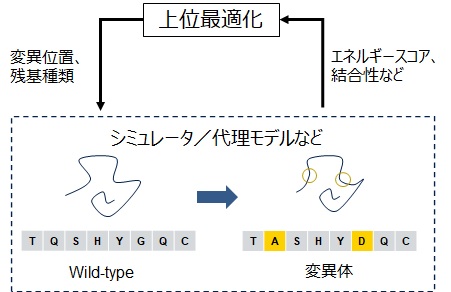


図 10 変異体探索プロトコルの概要

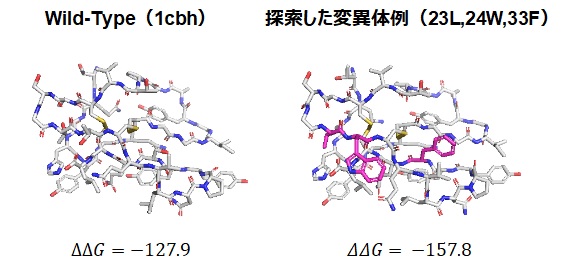


図 11 1CBHの三点変異体の主鎖構造（紫色は変異部分）

### モデリング技術のまとめ【鎌田】

本目では、カーネルPCAの検証方法と結果を記載する。変異数制約を満たすWild-Typeの変異体候補を探索・生成する方法を確立した。具体的には、Rosettaの設計プロトコルに変異数制約を加えた上位最適化を組み込むことで、全探索やランダム探索よりも効率的に変異体を生成した。Rosettaの設計プロトコルは、Cartesian DDGを使用した。Cartesian DDGは、Wild-Typeのアミノ酸配列の変異位置・残基を入力すると、それに整合するように最小化されたエネルギースコア（Rosettaスコア）と構造データを出力する。しかしながら、指定した変異数内での候補配列を得るには、全探索やランダム探索によって与える方法は非効率である。よって、候補生成の効率を上げるために、Cartesian DDGに上位最適化を組み込んだ技術を検討した。

## 最適化技術【熊谷】

### 最適化技術の目的

本節では、本研究テーマにおける最適化技術の開発や検証方法、検証結果について述べる。本研究テーマでは、プラント操業計画問題を解くための最適化技術を開発することを目指した。図 11に、その手順を示す。まず、プラント内の各設備のOn/Off状態を示すバイナリ変数を導入する、数理モデル等を使って各変数間の関係式を求めるなど、プラントモデルを構築する。次に、目的関数（操業コスト）と制約条件（上下限制約等）を決め、最適化問題として定式化する。なお、設備特性の式は、前述のプラントモデルから自動的に反映される。最後に、その問題を最適化アルゴリズムで解き、操作計画と操業コストを算出する。

以上の問題を解くためには、有制約・混合変数・大域的最適化が可能な技術が必要となるため、それを開発することを目的とした。図 12に示す通り、本研究テーマの最適化技術を実現するために、各アプローチを調査したところ、2つのアプローチが候補として上がった。一つ目（O1）は、非線型・有制約・連続変数の最適化技術を、混合変数へ拡張する技術と併用することで達成するアプローチである。O1の領域には、有制約大域的最適化技術が該当しており、主にメタヒューリスティクスと制約対処法を組み合わせることで達成される技術である。二つ目（O2）は、非線型・無制約・混合変数の最適化技術を、有制約に拡張することで達成するアプローチである。O2の領域には、凸緩和・分枝限定法が該当している。なお、線型・有制約・混合変数から非線型へ拡張するアプローチ（O3）について調査したが、該当技術は無いと判断し、対象から除外した。このため、O1、O2の二つのアプローチを開発・検証の対象とした。

最適化技術の目標は、有制約・混合変数・大域的最適化の問題性質に対応できることに加えて、15分以内に実行可能解（制約条件を満たす解）を得られることとした。これは、オンラインでプラント操業計画算出を支援するときに、1時間ステップの操業計画について、約30分～1時間間隔の最適化計算で求めて提出するというケースを想定したためである。

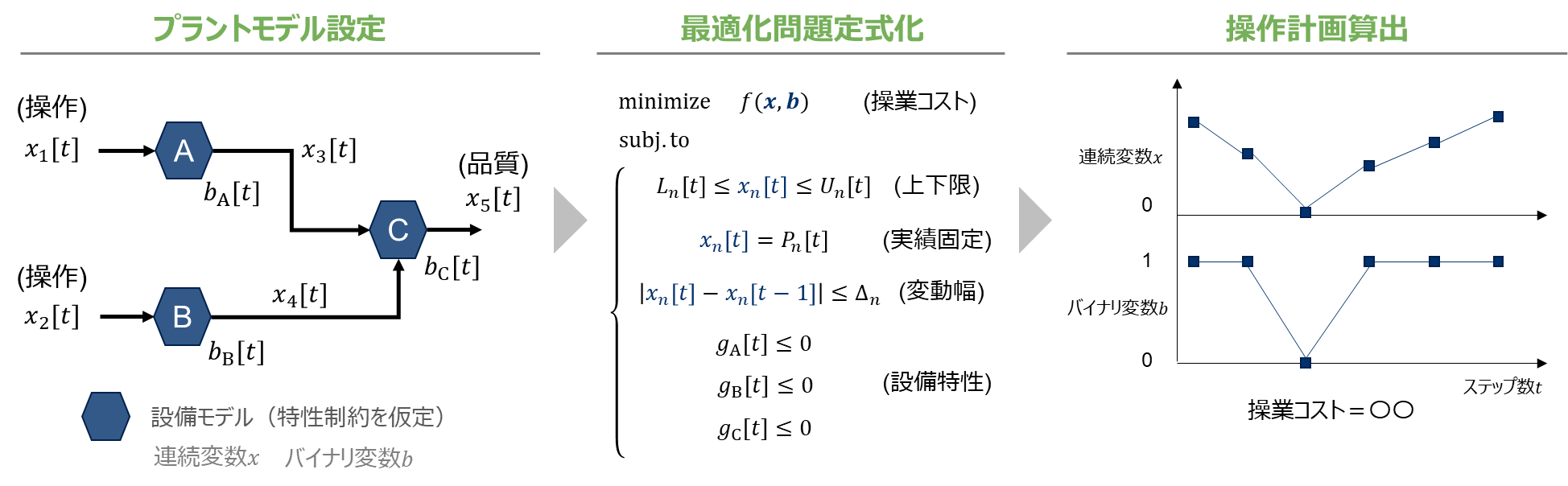


図 11 プラント操業計画問題を解く手順

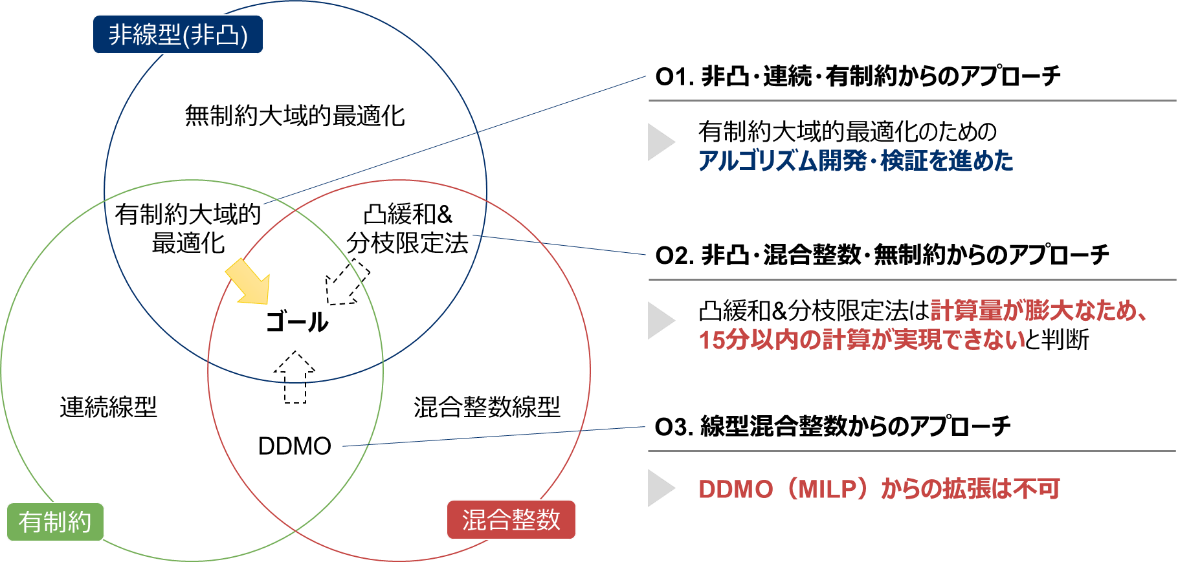


図 12 有制約・混合整数・大域的最適化へのアプローチ

### 開発した最適化技術

本項に、本研究テーマで開発・検証した最適化技術について簡潔に述べる。なお、Adaptive Weight MOEA/D、重み調整付きMOEA/Dにおける正規化法、ハイブリッド制約対処法、バイナリ分離法、冗長な変数・制約削減法などのアルゴリズムを、共同研究先の東京都立大学と協力して開発した。

* 最適化アルゴリズム：メタヒューリスティクスと制約対処法

メタヒューリスティクスとは、自然現象などのアナロジーに基づいて構築された、無制約最適化アルゴリズムである [6]。目的関数の勾配を使わない（直接探索）、複数の探索点で探索する、などが特徴として挙げられ、代表的な手法として、Genetic Algorithm、Particle Swarm Optimization、Differential Evolutionなどが挙げられる。

また、制約対処法は、メタヒューリスティクスの適用範囲を有制約最適化問題へ拡張する方法である。無制約対象のメタヒューリスティクスは、探索過程の解の比較・選択において、目的関数値を比較して解の優劣を決定するが、制約対処法は、目的関数値だけでなく制約条件を違反する度合（制約違反量）の両方を評価・比較することで、解の優劣を決定する。これにより、制約条件を満たしながら目的関数値が小さな解が得られる。

* 最適化アルゴリズム：Feasibility Rule（FR）

FRは、制約対処法の一つで、二つの解の比較の際に、(1)互いの制約違反量vが同じならば、目的関数値fを比較して、優劣を決定する、(2)そうでなければ、制約違反量vを比較して、優劣を決定するというルールである。これにより、目的関数値よりも制約違反量の削減を優先することになる。

* 最適化アルゴリズム：Adaptive Weight MOEA/D（AW-MOEA/D）

　AW-MOEA/Dは、制約対処法の一つで、単一目的有制約問題を、目的関数fと制約違反量vの二目的無制約問題に変換して、多目的最適化アルゴリズム（MOEA/D）で解く方法である。MOEA/Dは、目的関数fと制約違反量vの空間（(f, v)空間）を重みで分割し、各領域の重みとスカラ化関数で優劣を決定する。通常のMOEA/Dは、(f, v)空間のパレートフロンティアを広く求めるために、探索過程で重みを固定するが、AW-MOEA/Dは、パレートフロンティアの実行可能領域の部分だけを求めるために、探索過程で重みを適応的に調整するルールが備わっている。これにより、目的関数fと制約違反量vのトレードオフ関係を考慮し、制約境界付近の目的関数値が小さい領域を早期に探索することが期待できる。

* 最適化アルゴリズム：重み調整付きMOEA/Dにおける正規化法（未定）

　正規化法は、AW-MOEA/Dなどの重み調整付きMOEA/Dにおいて適用する方法である。MOEA/Dは、目的関数値fと制約違反量vのスケール差が大きいとき、(f, v)空間の重み分割の間隔が不均等となり、パレートフロンティアの一部だけしか獲得できなくなる。

この簡潔な対処法として、単なる正規化が考えられる。例えば、スカラ化関数で評価するときに、目的関数値fと制約違反量vをそのまま用いるのではなく、各探索で得た範囲で正規化されたf’, v’を用いることで実現できる。しかし、実行可能解を得た後は、目的関数値fと制約違反量vの絶対量が少ないため、スケール差が激しく変化し、探索に悪影響を及ぼす。

一方で、この正規化法は、実行可能解を得るまでは、解の評価に正規化されたf’, v’を用いて、実行可能解を得た後は、正規化に使用するスケーリング範囲を固定し続ける。これにより、f, vのスケール差を抑制して問題分割すると同時に、探索過程におけるスケール差の激しい変化を抑制することが期待できる。

* 最適化アルゴリズム：ハイブリッド制約対処法（未定）

　図 13に、Feasibility Rule（FR）とAdaptive Weight MOEA/D （AW-MOEA/D）の違いを示す。FRは、制約違反量の削減を優先するため、実行可能解を早期に得られる、また、AW-MOEA/Dは目的関数と制約違反量のトレードオフを考慮するため、制約境界付近の目的関数値が小さい領域を早期に探索できるなど、これらは性質が大きく異なる制約対処法である。言い換えると、問題性質が変わると、得意・不得意が変わってしまうため、対象問題に応じてアルゴリズムを選択する必要があるが、その選択は専門知識を持たないユーザにとって難しい。

そこで、このハイブリッド制約対処法は、FRとAW-MOEA/Dの両方の選択方法を組み込んだ制約対処アルゴリズムである。このため、互いの不得意を打ち消し、アルゴリズムの選択が容易になることが期待できる。

* 最適化アルゴリズム：凸緩和＋分枝限定法

　分枝限定法は、離散変数を含む問題の解法として良く知られており、データ空間で、整数変数をある値に仮定した領域に細かく子問題に分割していき、子問題を取捨選択しながら順に解くことで、元の問題の解を求める方法である。分枝限定法は、連続変数とバイナリ変数の混合変数問題にも適用でき、バイナリ変数の部分だけを分枝限定法によって子問題を生成した後、各子問題は連続問題なので非線形最適化アルゴリズムで解くことが可能だが、凸最適化と組み合わせるのが一般的である。

そこで、非凸かつ混合変数の問題を対象とする場合は、非凸部分に凸緩和や凸関数合成を適用し、バイナリ変数に分枝限定法を適用することで、大域的最適解を求めることが可能できる。

* 最適化アルゴリズム：バイナリ分離法

　メタヒューリスティクスは、連続問題か離散問題専用のアルゴリズムであるため、連続変数とバイナリ変数の混合変数問題にはそのまま適用できない。バイナリ分離法は、図 14に示す通り、探索過程の候補生成において、解を実数部とバイナリ部に分離し、各部位だけで近傍生成を実施した後、各部位の近傍解を合体し、混合変数の近傍解を構成する。これにより、離散変数部分の離散制約を強制的に満たすことが可能になり、別の制約条件だけを考慮すれば良い。

* 最適化問題実装の工夫：冗長な変数・制約削減法

　これは、冗長な変数の個数を削減する方法を開発した。図 14に示す通り、スケジューリング形式と解の形式の間は変換が必要である。

また、冗長な制約違反量の個数を削減する方法も開発した。L<=x<=Uという挟み撃ち形式の不等式の場合、通常は下限L<=xと上限x<=Uの二つの制約条件とみなす。しかし、実際にはこれらを同時に満たすことはないため、上下限のどちら側に近いのかを判定し、片方の制約に対する違反量を計算すれば十分である。よって、挟み撃ち形式の制約の場合は、探索過程で中央の(L+U)/2とxを比較し、片方の制約違反量のみを総和の対象とする。これにより、挟み撃ち形式の制約条件の違反量の個数は半数となる。

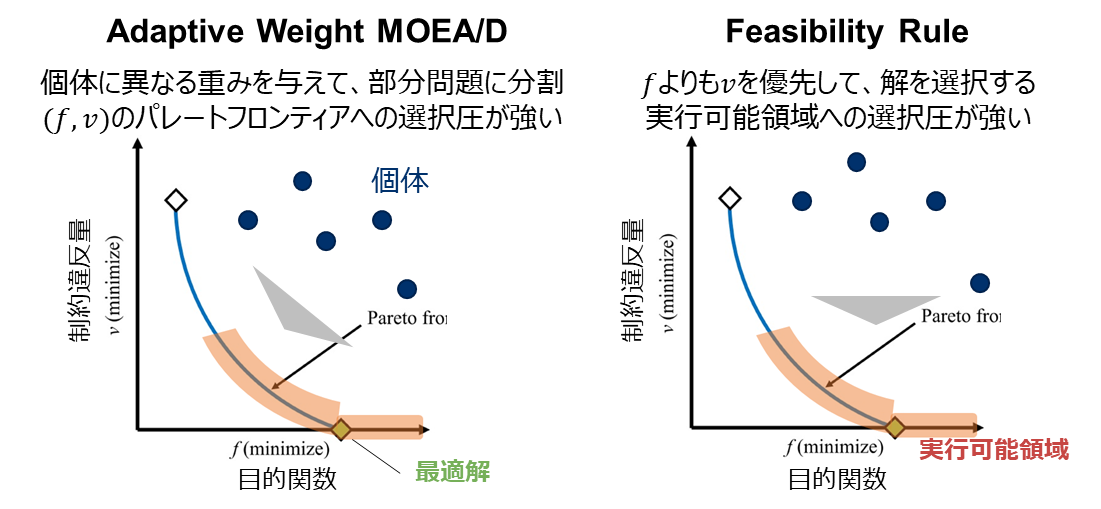


図 13 Feasibility RuleとAdaptive Weight MOEA/Dの概要

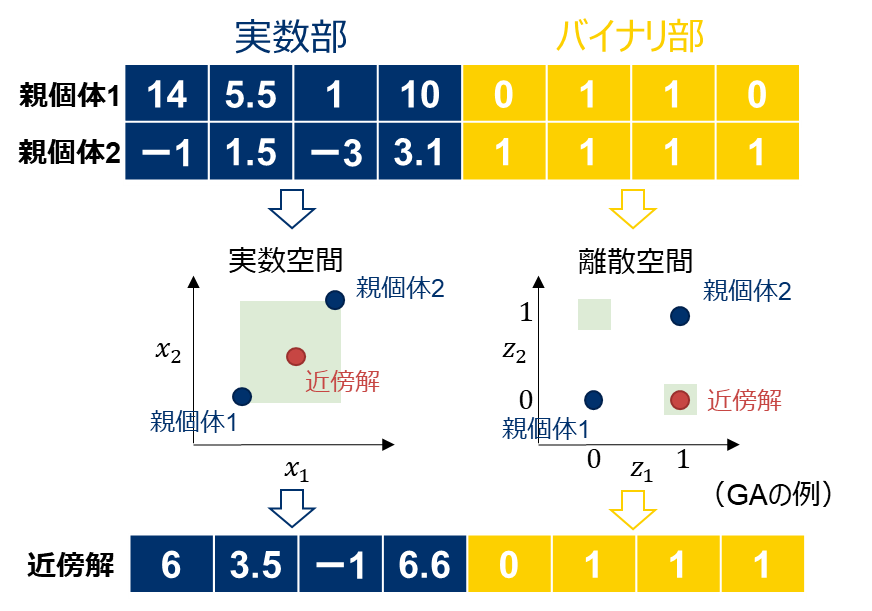


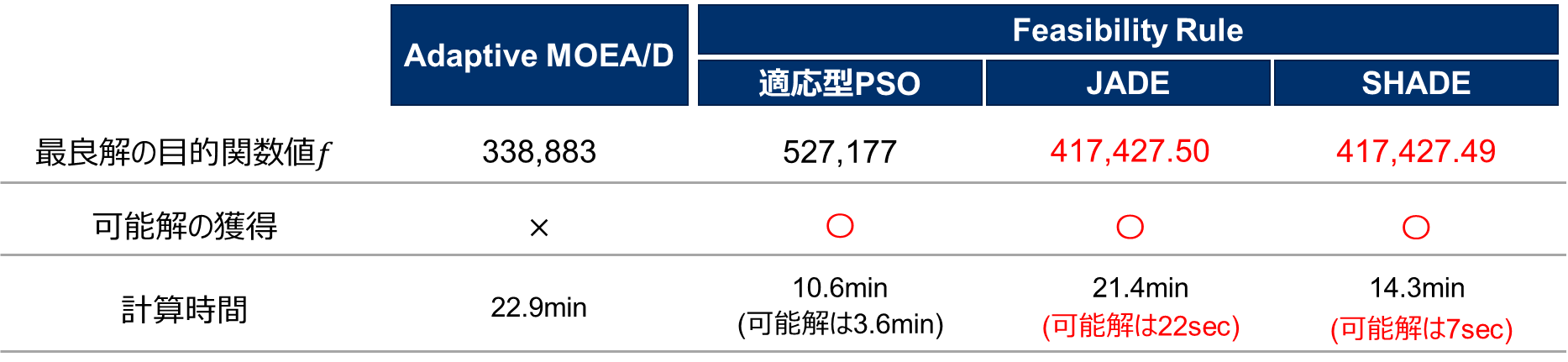
図 14 バイナリ分離法の概要

### 非線型性と有制約問題に対応可能な最適化技術の検証

本項では、最適化技術の核となる、有制約大域的最適化の性能検証に向けて、。

* 紙パルプ蒸解工程運転計画最適化問題における性能検証

表 1 蒸解工程操作計画問題における最適化結果



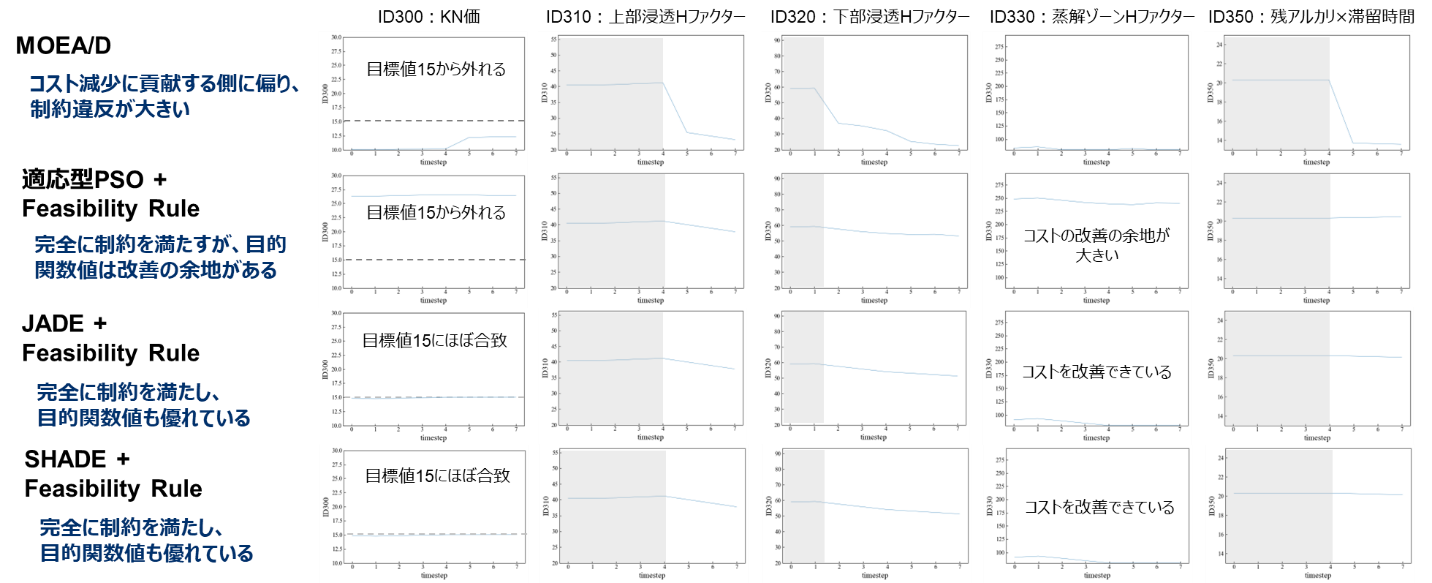


図 14 蒸解工程における最適操作計画

* 再生水RO運転計画最適化問題における性能検証

　図15に下水飲用化（再生水）プロセス概要を示す。下水処理施設から流入する水をUF膜、RO膜、UV/AOPなどで、処理・浄化して飲用水に再生するプロセスである。特にRO膜は、水イオン以外をほぼ除去できるが、膜の内部に詰まりが発生する、または洗浄や薬液で劣化するなど、管理が重要である。このため、ROプロセスでは、RO膜の状態（水質基準、膜の詰まり、膜洗浄、膜寿命など）を考慮しながら、薬液コストを最小化することが求められる。本目では、ROプロセスの実データを用いてモデルを構築し、1か月の操業計画を計算した。

検証条件を述べる。まず、1分データを1時間データに前処理した。1か月のRO操業計画を得るために、1週間を学習し、1週間を最適化するという計算を、時期を移動させながら5週間実施し、それらの結果を繋げた。今回は、水質基準として、RO透過導電率（EC）と水中有機物（TOC）の除去率を一定以上に保つ制約条件と、UF前の消毒剤と膜詰まり防止剤（Scaling防止剤）のコストを目的関数として設定した。なお、RO膜の水質予測モデルは、線形重回帰で構築し、制約条件に埋め込んだ。最適化アルゴリズムは、SHADEとFeasibility Ruleを組み合わせたアルゴリズムを使用し、5000反復、100個体に設定した。

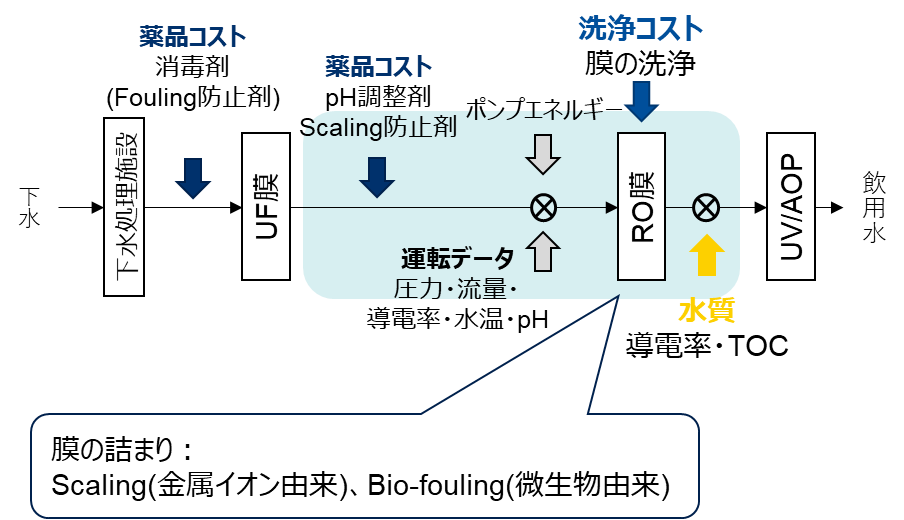


図 15 下水飲用化（再生水）プロセス概要

図16に結果を示す。

1週間計画算出に、約10分、問題規模は336変数、1859制約である。様々な制約条件を満たしながら、コストを削減する操業計画が得られた。

各週で特性が異なるのはモデルに起因する。

より滑らかにするには、変動幅制約を厳しくする必要があるが、15分以内の実行可能解獲得が困難となる。

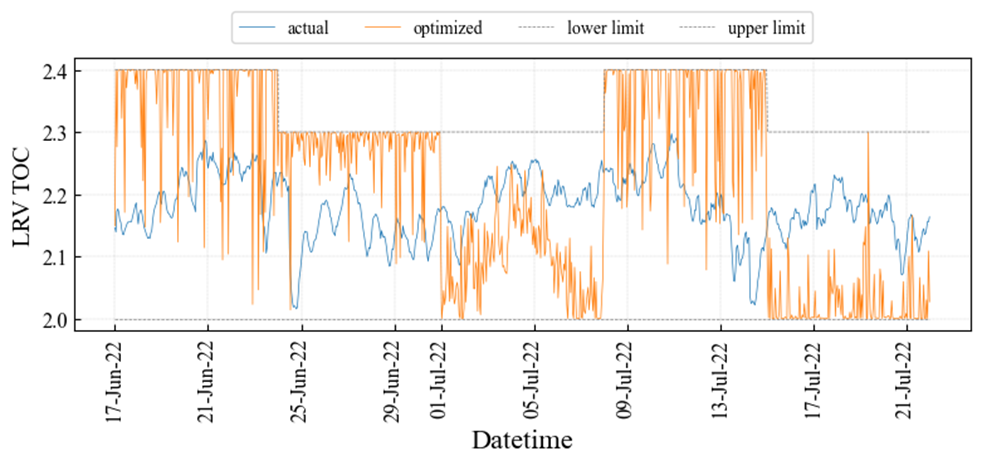


図 16 (a) TOC除去率の結果



図 16 (b) スケール防止剤の操作計画

図 16 最適化計算で求めたRO操業計画

* ブラックボックス制約問題における性能検証

　人工データから抽出したブラックボックス（非凸）な制約を課した問題に適用し、性能を検証した。図17に最適化問題の概要を示す。図17(a)のように、3次元球殻状に人工データを生成し、AutoEncoderでデータの特性式を抽出し、それを制約として課した最適化問題を構築した。目的関数は、原点との距離に比例する関数とした。また、図18にAutoEncoderによる特性抽出原理を示す。学習データと、AutoEncoderからの復元データの距離がε=10^-2以内に含まれる領域を特性式とする。なお、この特性式はニューラルネットワークの一種であるAutoEncoderに基づくモデルであるため、ブラックボックスかつ非凸である。

非凸な制約でも、実行可能な準最適解を得た。条件1は100反復／10個体、条件2は100反復／100個体である。

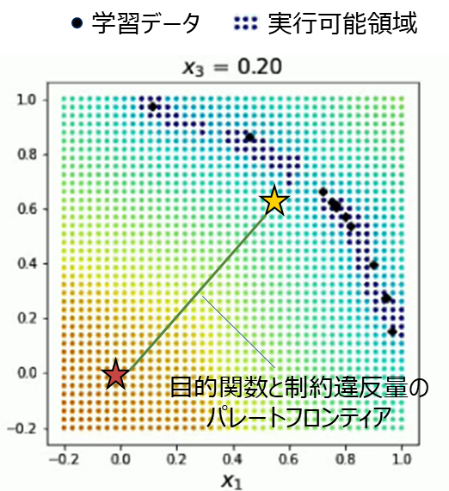
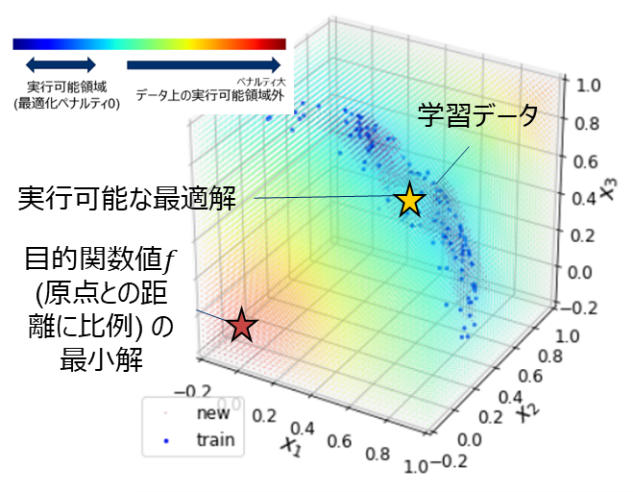


図 17 (a) 3次元空間　　　　　　　　　　　図 17 (b) 2次元断面（x1-x2）

図 17 人工データの特性式をAutoEncoderで抽出した制約違反量の景観

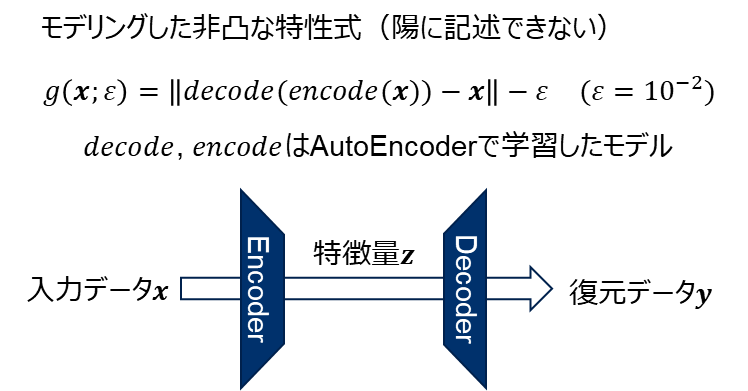
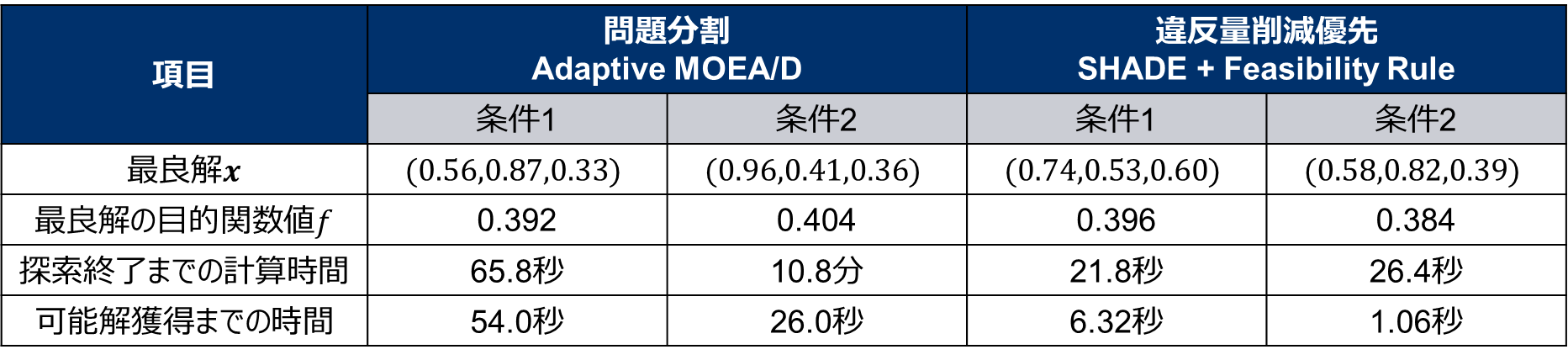


図 18 AutoEncoderによる特性抽出の原理

表 2 ブラックボックス制約問題における最適化結果



* (f, v)空間のパレートフロンティアが狭い問題における性能検証

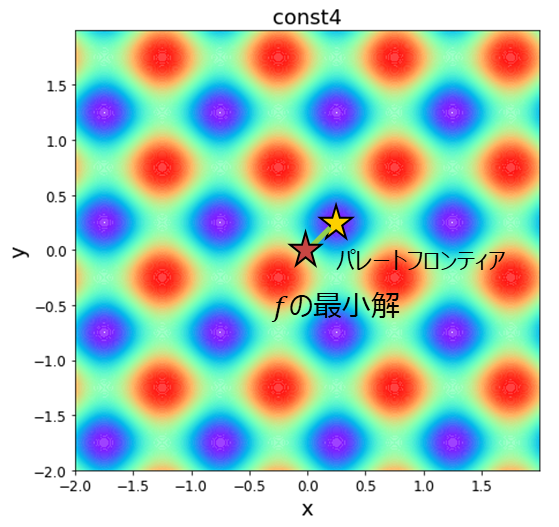
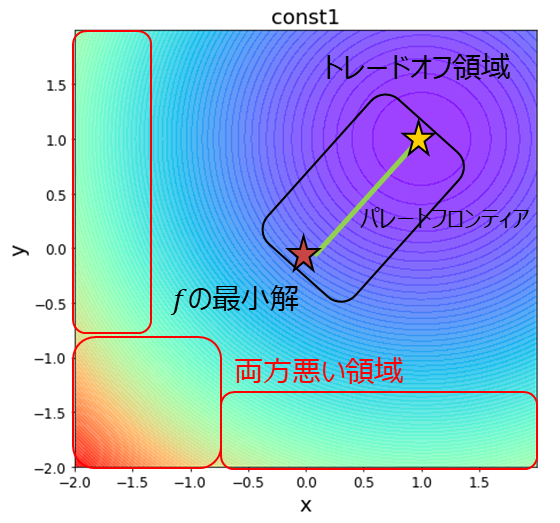


図 19 (a) Prob.1　　　　 　　　　図 19 (b) Prob.4

図 19 パレートフロンティアが狭いベンチマーク問題における制約違反量の景観

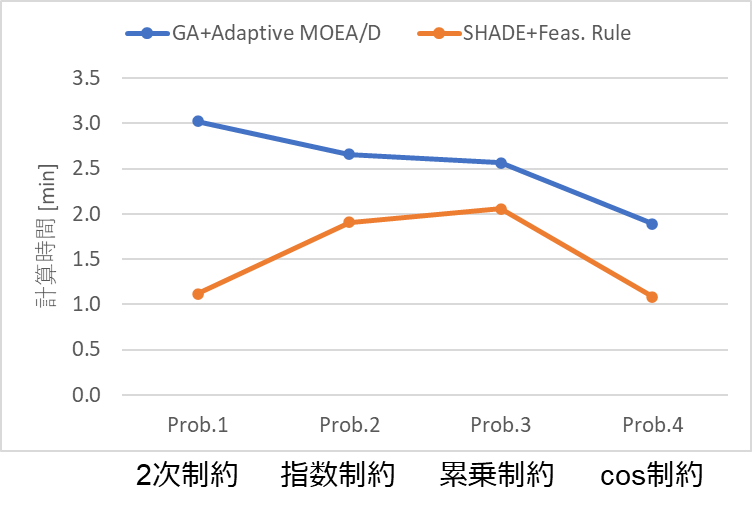
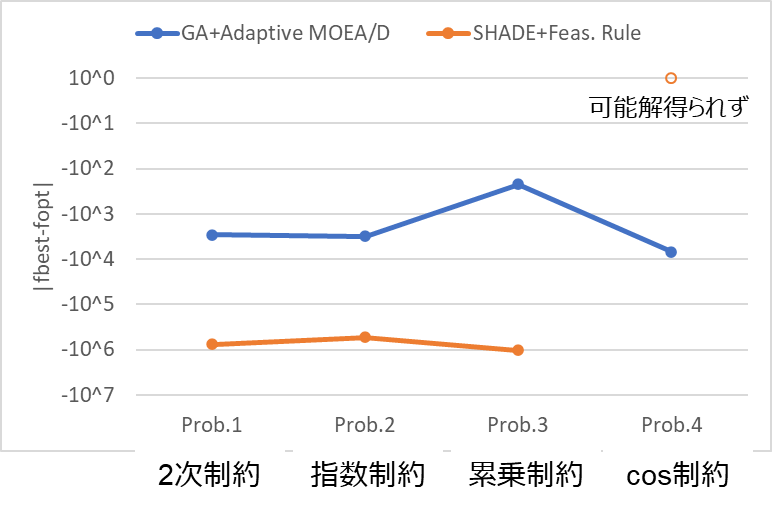


図 20 (a) 最適値と最良値の差　　　　 　　　 　図 20 (b) 計算時間

図 20 パレートフロンティアが狭いベンチマーク問題における最適化結果

* (f, v)のスケール性の影響分析（未定）
* (f, v)空間のパレートフロンティアの広さと実行可能領域の非連結性の影響分析（未定）

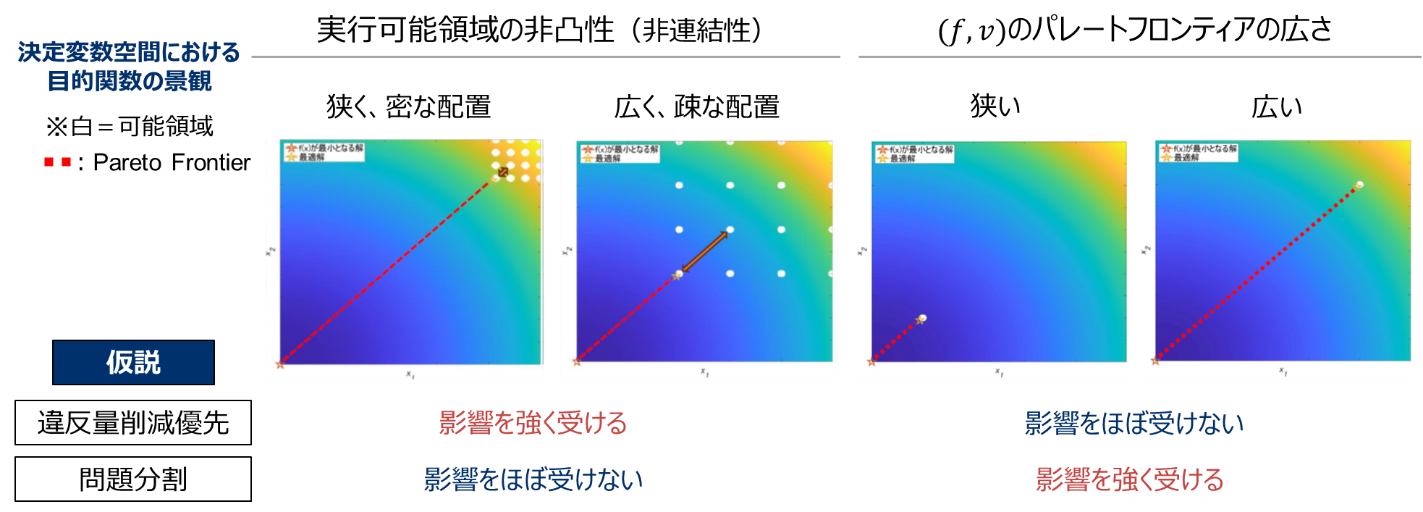


図 20 実行可能領域の非連結性とパレートフロンティアの広さを変更可能なベンチマーク問題の制約違反量の景観

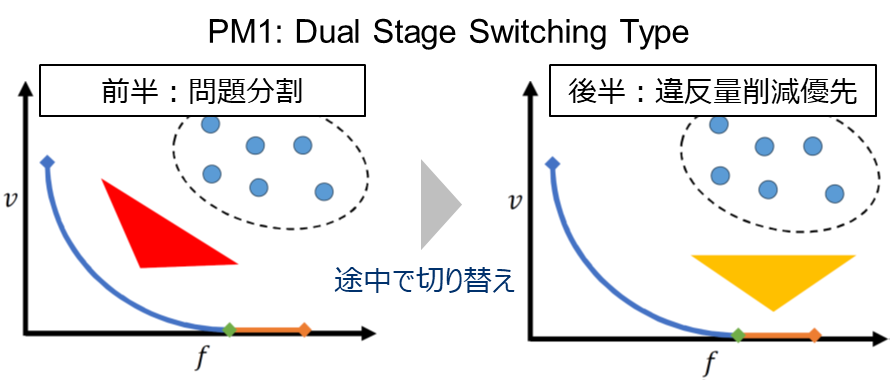


図 20 (a) ハイブリッド手法Proposed Method1 (PM1)

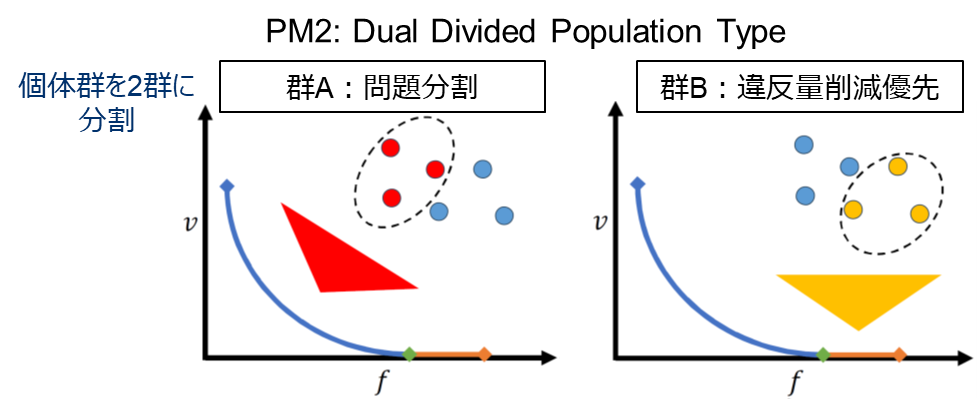


図 20 (b) ハイブリッド手法Proposed Method2 (PM2)

図 20 ハイブリッド手法の概要

ハイブリッド手法を開発し、相反する性質に対するロバスト性を両立できた。

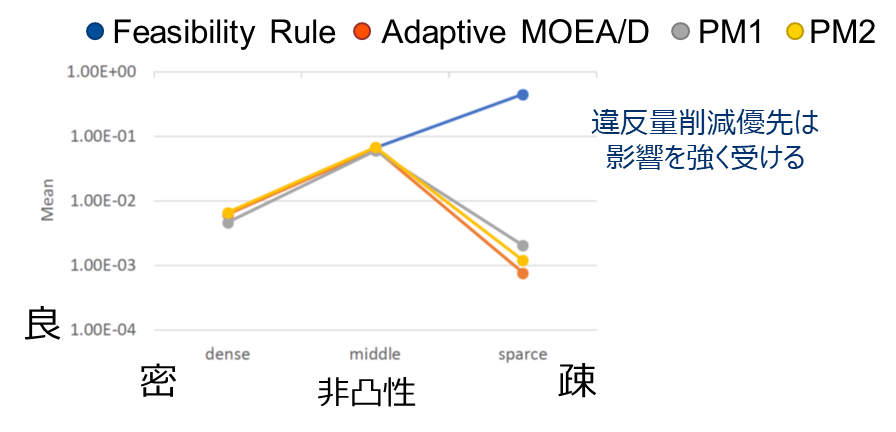


図 20 (a) 実行可能領域の非連結性を変更したときの最適化結果



図 20 (b) パレートフロンティアの広さを変更したときの最適化結果

図 20 数値実験結果

### 非線型性と混合変数に対応可能な最適化技術の検証

　本項では、非線型性と混合変数の問題に適用し、その性能を確認した。検証対象の手法は、凸緩和＋分枝限定法と、有制約大域的最適化とバイナリ分離法を組み合わせた手法である。

　【実験する必要あり】

次元数を変えた場合、バイナリ変数の割合を増やした場合、で解の精度と計算時間がどうなるのかを確認する。

### 最適化技術のまとめ

表 3 各設計CBDの設計情報と評価結果のまとめ



Sample#：番号（20-272）は設計CBDの通し番号を示す。

mut type：変異タイプを示す。例えば、Y5は5番目のチロシン（Y）を示す（補足：アミノ酸表記）。

## まとめ

本共同研究では、セルラーゼの重要な構成要素の一つであるセルロース結合性タンパク質の人工設計に向けて、アミノ酸配列改変・設計技術、計算機上でのタンパク質評価技術、Wetでの配列合成・結合性評価技術についての、適用可能性検証と課題抽出を行ってきた。

一般に、タンパク質のアミノ酸配列の組合せの数は網羅的探索を行うにはあまりにも膨大であるため、『所望の機能を有する新規配列のタンパク質を獲得する』という目的の達成には、効率的な配列提案手法が必要となる。

本研究ではこれを、まず機械学習的手法を用いて実現することを考えた。この機械学習的手法の実現には非常に多数のタンパク質評価データが必要となるが、十分な数のデータをWet実験によって短期間で得ることは、現状の技術では難しい。そのため外部データベースを用いた学習を試みたが、公になっているタンパク質の立体構造データが少なく、機械学習の適用は難しかった。

そこで我々は、既存のタンパク質改変プロトコルに上位の最適化手法を組み合わせることで、効率的な配列提案の実現を目指した。最終的に、Rosettaの変異体生成とスコア計算の機能を用いた最適化によって配列探索を行うことで、天然タンパク質から最大13%異なるタンパク質変異体の候補を獲得できた。そして、その候補に対して計算機上での安定性と基質結合性評価を行い、Wetでの合成と結合性評価を実施することで、想定した設計プロトコルの検証を行うことができた。設計プロトコルによるスクリーニング性能は期待したほど高くはならなかったものの、検証を通じて各要素技術についての課題を明確にすることができたことの意義は大きかったと思われる。

今後については、これらの課題点を踏まえたうえで、設計プロトコルの全体像の変更も含め、関連する分野への技術展開を検討したいと考えている。

## **参照文献**

|  |  |
| --- | --- |
| [1] | “cbh1 - Exoglucanase 1 - Hypocrea jecorina (Trichoderma reesei) | UniProtKB | UniProt,” [オンライン]. Available: https://www.uniprot.org/uniprotkb/P62694/entry.. [アクセス日: 16 11 2022]. |
| [2] | “cartesian\_ddg application (rosettacommons.org),” [オンライン]. Available: https://www.rosettacommons.org/docs/latest/cartesian-ddG.. [アクセス日: 16 11 2022]. |
| [3] | “RCSB PDB - 1CBH: DETERMINATION OF THE THREE-DIMENSIONAL STRUCTURE OF THE C-TERMINAL DOMAIN OF CELLOBIOHYDROLASE I FROM TRICHODERMA REESEI. A STUDY USING NUCLEAR MAGNETIC RESONANCE AND HYBRID DISTANCE GEOMETRY-DYNAMICAL SIMULATED ANNEALING,” [オンライン]. Available: https://www.rcsb.org/structure/1CBH. [アクセス日: 16 11 2022]. |
| [4] | J. K. Leman et al., "Macromolecular modeling and design in Rosetta: recent methods and frameworks," Nature Methods, 2020. |
| [5] | T. D. Schneider , R. M. Stephens, “Sequence logos: a new way to display consensus sequences,” Nucleic Acids Research, 1990. |
| [6] | J. Jumper, “Highly accurate protein structure prediction with AlphaFold.,” Nature, 2021. |
| [7] | T. Kosugi, “Solubility-Aware Protein Binding Peptide Design Using AlphaFold.,” Biomed, 2022. |
| [8] | M. Mirdita, “ColabFold: making protein folding accessible to all.,” Nat Methods, 2022. |
| [9] | J. Eberhardt, “AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings.,” J Chem Inf Model., 2021. |
| [10] | O. Trott, “AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.,” J. Comput. Chem., 2010. |
| [11] | T. C. F. Gomes, “Cellulose-Builder: A toolkit for building crystalline structures of cellulose.,” J Comput Chem., 2012. |
| [12] | R. Salomon-Ferrer, “An overview of the Amber biomolecular simulation package.,” Wiley Interdiscip Rev Comput Mol Sci, 2013. |
| [13] | D. A. Case, “The Amber biomolecular simulation programs.,” J. Comput. Chem., 2005. |
| [14] | N. Sugimoto, “Functional analysis and application of fungal cellulose-binding domains,” 2014. |
| [15] | M. Linder, “The difference in affinity between two fungal cellulose-binding domains is dominated by a single amino acid substitution.,” FEBS Lett, 1995. |

# 口頭発表，講演リスト

|  |  |
| --- | --- |
| [1] | 安田・熊谷・田村・安田：「メタヒューリスティクスのための制約対処法に関する解析とその課題」、令和2年電気学会全国大会、3-018，pp.24-25（2020.03.11） |
| [2] | 安田・熊谷・田村・安田：「メタヒューリスティクスにおける加重和に基づく制約対処法の基礎検討」、計測自動制御学会 システム・情報部門学術講演会 2020，GS1-3-1，pp.256-261（2020.11.17） |
| [3] | 安田・熊谷・田村・安田：「有制約最適化問題のためのMOEA/Dに基づく制約対処法のパラメータ解析」、2021年 電気学会 電子・情報・システム部門大会，GS12-5，pp.1226-1231（2021.09.18） |
| [4] | 安田・熊谷・田村・安田：「適応的重み調整を用いたMOEA/Dによる有制約最適化」、計測自動制御学会 システム・情報部門学術講演会 2021，GS5-2-1，pp.252-257（2021.11.22）【SSI優秀論文賞】 |
| [5] | Y. Yasuda, W. Kumagai, K. Tamura, K. Yasuda: “Feasibility-based Weighted MOEA/D in Constrained Optimization”, 電気学会C部門大会2022, SS1-1 (2022.08.31) |
| [6] | 安田・熊谷・田村・安田：「制約条件を目的関数に変換する有制約メタヒューリスティクスの提案」、2022年 電気学会 電子・情報・システム部門大会、GS10-2（2022.09.02） |
| [7] | 小嶋・熊谷・田村・安田：「有制約最適化のためのMOEA/Dに基づく制約対処法の検討」、2022年 電気学会 電子・情報・システム部門大会、GS10-3（2022.09.02） |
| [8] | 佐藤・熊谷・安田・田村・安田：「有制約最適化のためのDifferential Evolutionの基礎検討」、2022年 電気学会 電子・情報・システム部門大会、GS10-4（2022.09.02） |
| [9] | Y. Yasuda, W. Kumagai, K. Tamura, K. Yasuda: “MOEA/D with Feasibility-based Weight Adjustment for Constrained Optimization”, IEEE Symposium Series on Computational Intelligence 2022（2022.12.7）【査読有】 |
| [10] | Y. Yasuda, H. Kojima, W. Kumagai, K. Tamura, K. Yasuda: “Analysis and Switching of Normalization in the Decomposition-based Constraint Handling Technique for Constrained Optimization”, 2023年 電気学会 電子・情報・システム部門大会，SS2-2-7，pp.1912-1913 (2023.08.30) |
| [11] | Y. Sato, W. Kumagai, Y. Yasuda, K. Tamura, K. Yasuda: “Constrained Differential Evolution with Superior Infeasible Solutions for Mutation”, 2023年 電気学会 電子・情報・システム部門大会，SS2-2-4，pp.1906-1907 (2023.08.30) |
| [12] | 宇津本，安田，熊谷，田村，安田：「有制約最適化のためのMOEA/DとFeasibility Ruleの比較検討」，2023年 電気学会 電子・情報・システム部門大会，PS6-5（2023.08.30） |
| [13] | 佐藤，熊谷，安田，田村，安田：「外部アーカイブを用いて優れた実行不可能解を活用する有制約Differential Evolutionの比較検討」，2023年 電気学会 電子・情報・システム部門大会，GS11-5，pp.1468-1473（2023.09.01） |
| [14] | Y. Yasuda, H. Kojima, W. Kumagai, K. Tamura, K. Yasuda: “A Switching Normalization Method in Decomposition-based Constraint Handling for Constrained Optimization”, SICE Annual Conference 2023（2023.09.09）【査読有】 |
| [15] | Y. Sato, W. Kumagai, Y. Yasuda, K. Tamura, K. Yasuda: “Differential Evolution Using Superior Infeasible Solutions for Constrained Optimization”, IEEE Conference Systems, Man, Cybernetics 2023（2023.10.04）【査読有】 |
| [16] | 宇津本，安田，熊谷，田村，安田：「有制約最適化のためのFeasibility RuleとMOEA/Dに基づくハイブリッド制約対処法の検討」，進化計算シンポジウム2023，P1-02（2023.12.22） |
| [17] | 佐藤，熊谷，安田，田村，安田：「有制約Differential Evolutionにおける外部アーカイブの活用法の解析」，進化計算シンポジウム2023，P1-16（2023.12.22） |

# 社外発表論文リスト

|  |  |
| --- | --- |
| [1] | 安田・熊谷・田村・安田：「MOEA/Dの有制約最適化への拡張と適応的重み調整に関する基礎検討」、電気学会 C部門誌（2022.03） |
| [2] | 安田・熊谷・田村・安田：「有制約最適化のための制約条件の目的関数化と適応的重み調整を用いたMOEA/D」、電気学会 C部門誌（2023.03） |
| [3] | 安田・小嶋・熊谷・田村・安田：「有制約最適化のための分割に基づく制約対処法における正規化の検討」、電気学会 C部門誌（2024）【査読中】 |

# 出願特許リスト

なし

# 主要な関連社内報告書，LRの資料，議事録等のリスト

|  |  |
| --- | --- |
| 文書 | Document No. |
| LR0審査　資料 | SMM-SCRR-048 |
| LR0審査　議事録 | SMM-SCRR-018 |
| LR1審査　資料 | SMM-BD18-INV-10R-001 |
| LR1審査　議事録 | SMM-BD18-INV-10G-004 |
| LR2-1中間報告　資料 | SMM-BD18-INV-10R-002 |
| LR2-1中間報告　議事録 | SMM-BD18-INV-10G-007 |
| LR2-2中間報告　資料 | SMM-BD18-INV-10R-003 |
| LR2-2中間報告　議事録 | SMM-BD18-INV-10G-013 |
| 2022年5月報告会　資料 | SMM-BD18-INV-10R-004 |
| 2022年5月報告会　議事録 | SMM-BD18-INV-10G-016 |
| LR2審査　資料 | SMM-BD18-INV-10R-005 |
| LR2審査　議事録 | SMM-BD18-INV-10R-018 |
| 人工酵素設計FY21技術動向調査報告 | SMM-BD18-INV-10D-001 |
| HPCシステムズ社委託計算報告書：セルロース分解酵素によるβ-1,4グリコシド結合の加水分解 | SMM-BD18-INV-10H-005 |
| 報告書：タンパク質候補の机上生成 | SMM-BD18-INV-10H-022 |
| 報告書：分子シミュレーションによる機能性の机上評価 | SMM-BD18-INV-10H-024 |
| 報告書：構造的特徴抽出 | SMM-BD18-INV-10H-023 |
| テクノプロ社委託試験報告書：1案件管理番号YD-416-2　コムギ無細胞発現系で合成したCBD-EGFPのCBD評価 | SMM-BD18-INV-10H-004 |
| テクノプロ社委託試験報告書：案件管理番号YD-416-1　コムギ無細胞発現系/ピキア酵母発現系のユニバーサルカセットデザイン | SMM-BD18-INV-10H-003 |
| テクノプロ社委託試験報告書：案件管理番号YD-639-1　TLCプレートを用いたダイナミックレンジの広いCBD結合能の評価条件と評価指標の決定 | SMM-BD18-INV-10H-006 |
| テクノプロ社委託試験報告書：案件管理番号YD-640-2　変異CBD-EGFPのコムギ胚芽無細胞タンパク質合成・精製 | SMM-BD18-INV-10H-009 |
| テクノプロ社委託試験報告書：案件管理番号YD-639-2　TLCプレートを用いたダイナミックレンジの広いCBD結合能の評価条件と評価指標の決定（2） | SMM-BD18-INV-10H-008 |
| テクノプロ社委託試験報告書：案件管理番号YD-640-1　変異CBD-EGFPのコムギ胚芽無細胞タンパク質合成用ベクター構築 | SMM-BD18-INV-10H-007 |
| テクノプロ社委託試験報告書：案件管理番号YD-849-2　セルロース結合性評価試験② | SMM-BD18-INV-10H-011 |
| テクノプロ社委託試験報告書：案件管理番号YD-849-1 セルロース結合性評価試験① | SMM-BD18-INV-10H-010 |
| テクノプロ社委託試験報告書：案件管理番号YD-960-1　コムギ胚芽無細胞タンパク質合成系を用いたCBD-EGFPアミノ酸変異体301種類のタンパク質調製 | SMM-BD18-INV-10H-013 |
| テクノプロ社委託試験報告書：案件管理番号YD-960-2　TLCプレートを用いたCBD-EGFPアミノ酸変異体301種類のセルロース結合性評価 | SMM-BD18-INV-10H-014 |
| テクノプロ社委託試験報告書：案件管理番号YD-851-1　変異CBD-EGFPのコムギ胚芽無細胞タンパク質合成用ベクター構築（２） | SMM-BD18-INV-10D-012 |
| テクノプロ社委託試験報告書：案件管理番号YD-970-1　大腸菌発現系を用いたCBDペプチドの調製 | SMM-BD18-INV-10H-025 |
| テクノプロ社委託試験報告書：案件管理番号YD-970-2　Pichia酵母発現系を用いたCBDペプチドの調製 | SMM-BD18-INV-10H-026 |
| テクノプロ社委託試験報告書：案件管理番号YD-1145-1 コムギ胚芽無細胞タンパク質合成による2種類のタンパク質調製および輸送による影響確認試験 | SMM-BD18-INV-10H-019 |
| テクノプロ社委託試験報告書：案件管理番号YD-1145-2 コムギ胚芽無細胞タンパク質合成による3種類のタンパク質調製 | SMM-BD18-INV-10H-020 |
| テクノプロ社報告書 案件管理番号YD-1145-3 コムギ胚芽無細胞タンパク質合成による2種類のタンパク質追加調製 | SMM-BD18-INV-10H-021 |
| 社内実験報告書：結合率による設計セルロース結合ドメインの定量的評価 | SMM-BD18-INV-10H-016 |
| 社内実験報告書：CBD結合能の定量的評価系の検討 | SMM-BD18-INV-10H-017 |
| 社内実験報告書：CBD結合能の定量的評価系の確立 | SMM-BD18-INV-10H-018 |
| 東京大学実験報告書：酵母を用いた2種類のセルロース分解酵素の発現確認実験 | SMM-BD18-INV-10H-015 |
| FY2019 東京大学共同研究 活動報告書 | SMM-BD18-INV-10H-002 |
| FY2020-2022　東京大学共同研究　活動報告書 | SMM-BD18-INV-10H-027 |