



Title	タンパク質配列情報を活用したシアノバクテリア Synechocystis sp. PCC 6803の代謝改変に関する研究
Author(s)	西口, 大貴
Citation	大阪大学, 2025, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/103169
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

タンパク質配列情報を活用した
シアノバクテリア *Synechocystis* sp. PCC 6803
の代謝改変に関する研究

提出先 大阪大学大学院情報科学研究科

提出年月 2025 年 7 月

西口 大貴

学位取得に関わる発表論文

学術雑誌

- [1] Hiroki Nishiguchi, Teppei Niide, Yoshihiro Toya, Hiroshi Shimizu, Modification of intracellular metabolism by expression of a C-terminal variant of phosphoribulokinase from *Synechocystis* sp. PCC 6803, Bioscience, Biotechnology, and Biochemistry, 2025,, zba013, <https://doi.org/10.1093/bbb/zba013>
(博士論文第 2 章)

- [2] Hiroki Nishiguchi, Teppei Niide, Yoshihiro Toya, Hiroshi Shimizu, Efficient selection of pyruvate decarboxylase sequences from database for high ethanol productivity in *Synechocystis* sp. PCC 6803, Journal of Bioscience and Bioengineering, 2025, ISSN 1389-1723, <https://doi.org/10.1016/j.jbiosc.2025.05.011>.
(博士論文第 3 章)

国際学会

- [1] Hiroki Nishiguchi, Teppei Niide, Yoshihiro Toya, Hiroshi Shimizu., “Selection for useful sequences for chemical production in cyanobacteria from enormous sequences in the database”, 1st Asia & Pacific Bioinformatics Joint Conference 2024, October 22-25 (2024), Okinawa, Japan
(博士論文第 3 章)

内容梗概

気候変動対策の一環として、CO₂を原料とした物質生産への関心が高まっている。従来は廃棄物と見なされていた CO₂を、有用な化学物質やエネルギー源の製造原料として活用する試みが進展している中、光合成によって高効率に CO₂を固定する微細藻類の利用が注目されている。一方で、微細藻類の産業利用にはコスト面の課題が依然として存在する。その解決策として、カーボンプライシング制度などの経済的インセンティブの導入と、微細藻類自体の物質生産性の向上に向けた研究開発が進められている。

そのような中で、シアノバクテリアは微細藻類の中でも増殖が速く、形質転換効率も高いため、有望な宿主生物として注目されている。シアノバクテリアの代謝を理解し、有用物質生産に適した状態へと改変することは代謝工学の目的のひとつである。代謝を改変するために、好ましい機能を有する代謝酵素を発現することは有効なアプローチである。このアプローチは対象とする機能に関わる酵素が宿主に内在する場合と、内在しない場合に大別される。対象とする機能に関わる酵素が宿主に内在している場合には、その酵素にアミノ酸変異を導入して機能を改良することが可能であり、この方法は宿主の代謝に与える悪影響が比較的小さい点で有利である。一方で、宿主が目的とする機能に関与する内在性酵素を持たない場合には、最適な異種酵素を探索し、それを発現させる必要がある。アミノ酸変異による機能改良と異種酵素の探索は相補的に重要な技術であるといえ、両方の場合における方法論を確立することが、代謝工学のために重要である。本研究では、タンパク質配列情報をアミノ酸残基の保存度と有用遺伝子情報のリソースとして活用し、アミノ酸変異による機能改良と異種酵素の探索によってシアノバクテリアの代謝を改変するための方法論を示すことを目的とした。

本学位論文は、第1章から第4章により構成される。第1章では本研究の背景と目的について述べた。

第2章では、内在性酵素の合理的改変のための方法論として、①アミノ酸残基の保存度を指標に、機能に重要かつ改変に適した残基を予測し、②酵素活性など維持したい機能に関わる領域から離れた残基に変異を導入するという戦略を提案した。本方法論の有効性を実証するための適用例として、カルビンベンソン回路の律速酵素である Phosphoribulokinase (PRK) の活性制御緩和に取り組んだ。その結果、変異 PRK を発現する株において、細胞内 PRK 活性が野生株に対して 1.5 倍に向上し、下流の有機酸の蓄積量が最大 1.8 倍に増加した。これは本方法論が、内在性酵素の機能を効果的に改変し、物質生産に有利な代謝状態へと変化させたことを示している。本方法論は、特に酸化還元制御を受ける酵素が多く存在するシアノバクテリアや植物において、保存性の高いシステイン残基を標的とすることで広く適用可能であると期待される。

第 3 章では、有用な異種酵素を効率的に探索するための方法論として、①系統的に多様な配列群の評価による配列空間の「広い探索」と、②有望空間内での天然・非天然配列を含む「密な探索」からなる 2 段階の探索戦略を提案した。本方法論の適用例として、シアノバクテリアに内在しない pyruvate decarboxylase (PDC) の探索を行った。その結果、従来良く用いられてきた *Zymomonas mobilis* 由来 PDC を上回るエタノール生産性 (1.12 倍に向上) を示す *Gluconobacter oxydans* 由来 PDC を同定することに成功した。これは本方法論が、広大な配列空間から既存酵素よりも優れた性能を持つ酵素を効率的に見出せることを示している。本方法論は、機能がアノテーションされ、ベンチマークとなる酵素が存在する多くのタンパク質ファミリーに適用可能であり、様々な物質生産経路の最適化に貢献できる。

第 4 章では、本研究で得られた知見をまとめ、配列情報を活用したシアノバクテリアの代謝工学について将来の展望を述べ、結論とした。

目次

学位取得に関わる発表論文	1
学術雑誌	1
国際学会	1
内容梗概	1
目次	3
第 1 章 諸論	7
1.1 研究背景	7
1.1.1 CO ₂ を原料としたモノづくりを取り巻く環境	7
1.1.2 シアノバクテリアによる CO ₂ からの有用物質生産	8
1.1.3 タンパク質工学による代謝改変のアプローチ	9
1.1.4 タンパク質配列データの拡充と有用配列の探索	9
1.1.5 タンパク質におけるアミノ酸残基の保存度と機能	10
1.1.6 本論文における表記に関して	11
1.2 本論文の目的及び構成	12
第 2 章 C 末端変異 Phosphoriblokinae の発現によるシアノバクテリアの代謝 改変 14	
2.1 緒言	14
2.2 実験方法	17
2.2.1 PRK 系統樹の作成	17
2.2.2 ラン藻類由来 PRK の C 末端領域の配列保存性確認	17
2.2.3 PCC6803 由来 PRK の立体構造予測	17
2.2.4 アミノ酸残基の露出度計算	18
2.2.5 活性中心からの距離の計算	18
2.2.6 ラン藻類由来 PRK の配列アライメント	18
2.2.7 使用菌株	18
2.2.8 培地組成	19
2.2.9 培養条件	20
2.2.10 菌体濃度及び培地中グルコース濃度の測定	20
2.2.11 PRK 追加過剰発現用 DNA 断片の構築	21
2.2.12 シアノバクテリアの形質転換	21
2.2.13 シアノバクテリアの粗酵素液調製	22

2.2.14	野生型及び変異型 PRK の大腸菌での調製.....	22
2.2.15	PRK タンパク質の精製.....	23
2.2.16	SDS-PAGE	23
2.2.17	PRK 活性測定	23
2.2.18	酵素反応速度及び活性値の算出	24
2.2.19	中間代謝物質の定量分析	25
2.2.20	中間代謝物質の誘導体化法	28
2.3	結果.....	29
2.3.1	PRK 配列の系統解析.....	29
2.3.2	シアノバクテリア由来 PRK の各残基の保存度計算	30
2.3.3	PCC6803 由来 PRK の立体構造を用いた変異残基絞り込み	31
2.3.4	PCC6803 由来 PRK の C 末端ループ変異体設計	34
2.3.5	PRK の C 末端ループ変異体の発現・精製	36
2.3.6	PRK 変異体の <i>in vitro</i> 活性測定	37
2.3.7	変異 PRK 発現株の細胞内 PRK 活性測定 (Photoautotroph).....	38
2.3.8	変異 PRK 発現株の細胞内 PRK 活性測定 (Photomixotroph)	40
2.3.9	Photomixotroph 条件における変異 PRK 発現株の培養挙動.....	42
2.3.10	変異 PRK 発現株の代謝状態変化の解析.....	42
2.4	考察.....	45
2.5	結言.....	47
第 3 章	シアノバクテリアのエタノール高生産化に寄与する異種 pyruvate decarboxylase の探索	49
3.1	緒言.....	49
3.2	実験方法.....	52
3.2.1	配列的に多様な PDC 配列の系統解析 (第 1 段階：広い探索)	52
3.2.2	バクテリア由来 PDC 配列の系統解析 (第 2 段階：密な探索)	53
3.2.3	PROSS による <i>Zm</i> PDC 変異体の設計	54
3.2.4	ASR によるバクテリア由来 PDC の祖先型配列推定	54
3.2.5	使用菌株.....	55
3.2.6	培地組成.....	56
3.2.7	培養条件.....	56
3.2.8	菌体濃度及び培地中グルコース濃度の測定.....	57
3.2.9	エタノール生合成経路発現用 DNA 断片の構築	57
3.2.10	シアノバクテリアの形質転換	60
3.2.11	シアノバクテリアの粗酵素液調製	60

3.2.12	PDC 活性測定	60
3.2.13	酵素反応速度及び活性値の算出	61
3.2.14	SDS-PAGE	61
3.3	結果.....	62
3.3.1	配列的な多様性に富んだ PDC 配列の系統解析 (第 1 段階：広い探索) ...	62
3.3.2	多様な PDC 配列を発現する株のエタノール生産性	63
3.3.3	バクテリア由来 PDC の系統解析 (第 2 段階：密な探索)	64
3.3.4	PROSS による <i>Zm</i> PDC と配列相同性の高い非天然タンパク質の設計	67
3.3.5	ASR による <i>Zm</i> PDC と配列相同性の高い非天然タンパク質の設計	69
3.3.6	バクテリア由来 PDC 発現候補の配列相同性解析	71
3.3.7	バクテリア由来 PDC 発現株のエタノール生産性	71
3.3.8	細胞内 PDC 活性の測定	72
3.3.9	SDS-PAGE による PDC の発現確認	73
3.4	考察	75
3.5	結言	78
第 4 章	結論	80
4.1	結果のまとめ	80
4.2	本研究の成果と方法論の適用範囲	81
4.3	タンパク質配列情報の代謝工学への活用の展望	82
	参考文献	84
	付録	95
	謝辞	113

第1章 諸論

1.1 研究背景

1.1.1 CO₂を原料としたモノづくりを取り巻く環境

化石燃料の大量消費と大気中二酸化炭素濃度の上昇による気候変動への対策として微生物発酵を利用した物質生産の重要性が強く認識されている (Zhang *et al.*, 2013)。近年では特に、CO₂を単なる廃棄物として捉えるのではなく、多様な製品を製造するための貴重な原料と見なす考え方が注目されている (Aresta *et al.*, 2007)。その中で、微細藻類は光合成を通じて CO₂を吸収し、バイオマスや様々な有機化合物を生成する独自の能力を持つ (Li *et al.*, 2024)。微細藻類は、陸上植物と比較して CO₂固定効率が 10 倍から 50 倍も高いとされており (Man *et al.*, 2012)、バイオ燃料だけでなく、食品、栄養補助食品、医薬品、化粧品など、多岐にわたる高付加価値製品の生産に利用可能であると期待されている。

日本においても、微細藻類を活用した CO₂ 固定・利用に関する研究開発、実証実験、そして事業化に向けた取り組みが活発に進められている。一例として、株式会社 IHI は、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) との共同研究プロジェクトで、微細藻類 (*Botryococcus Braunii*) 由来のバイオジェット燃料の安定生産技術開発を進めてきた (Prussi *et al.*, 2021)。この技術で製造されたバイオジェット燃料は、国際標準規格である ASTM D7566 に適合しており、既存のジェット燃料と混合して航空機の運航に使用できる (Prussi *et al.*, 2021)。2021 年には、日本航空 (JAL) と全日本空輸 (ANA) の定期国内便に、この微細藻類由来バイオジェット燃料が供給され、試験飛行が実施された (IHI Corporation, 2021)。

しかしながら、微細藻類による CO₂ 固定・利用技術は、微細藻類の培養、収穫、加工にかかるコストが依然として高いという経済性の観点で課題を有する (Bernardo *et al.*, 2021)。この課題を解決すべく、大きく分けて 2 つの取り組みがなされている。1 つ目はカーボンプライシング制度などの経済的インセンティブである (Mengesha *et al.*, 2025)。カーボンプライシング制度は、CO₂ 排出にコストを課すことで、排出量の削減を促すことを目標に導入された制度である。微細藻類由来の物質生産は、生産過程での CO₂ 排出量が少ないことに加え、炭素クレジットの取得といった恩恵を受ける可能性がある (Ana *et al.*, 2023)。また、カーボンプライシングによる化石燃料のコスト上昇は、微細藻類を原料とした代替製品の相対的な経済性を向上させることも予想できる。

経済性の課題を解決する 2 つ目の取り組みとしては研究開発による微細藻類の物質生産性向上である。これまでに微細藻類の遺伝子改変 (Choi *et al.*, 2016) や培養条件の改良 (Velmurugan *et al.*, 2019) によって生産性を数倍向上させた例も報告されている。微細藻類の物質生産性向上に向けた検討は生産コストの改善に直結するため極めて重要である。しかしながら、大腸菌や酵母のような従属微生物と比較して、微細藻類は菌株

構築や培養に時間を要することが、検討を進めるうえでのボトルネックのひとつになっていると言える。

1.1.2 シアノバクテリアによる CO₂ からの有用物質生産

シアノバクテリアは、微細藻類の中でも形質転換効率が高く、増殖速度が速いことが知られている (Vioque, 2007)。大腸菌や酵母には及ばないものの、微細藻類研究のボトルネックのひとつを緩和した宿主であると言える。また、シアノバクテリアは増殖速度の速い単純な宿主であるものの、酸素発生型光合成の機構やカルビンベンソン回路 (CBB 回路) のような炭素固定代謝は植物や他の微細藻類とも共通しており、得られた知見を他の宿主へ応用できるケースがある (McFarlane *et al.*, 2019)。

シアノバクテリアは食品や色素の生産においてすでに産業的に利用されている (Fernandes *et al.*, 2023)。近年ではシアノバクテリアが天然に合成する脂質や糖といった高分子物質、色素等だけでなく、細胞が本来生産しないエタノール (Deng and Coleman, 1999)、ブタノール、イソブタノール (Atsumi *et al.*, 2009) といった様々な有用物質の生産へと可能性が広がってきた。中でもエタノールはバイオ燃料としての汎用性が高い、生合成経路が単純といった理由から広く研究されてきた。1999 年に Deng and Coleman が初めてシアノバクテリア *Synechocystis sp.* PCC 6803 に *Zymomonas mobilis* 由来の pyruvate decarboxylase 遺伝子 (*pdc*) と alcohol dehydrogenase 遺伝子 (*adh II*) を導入しエタノール生産に成功し (Deng and Coleman, 1999)、4 週間の培養を行ったところ、培養液中のエタノール濃度はおよそ 0.23 g/L となった。2012 年の Gao らの研究では、*Synechocystis sp.* PCC 6803 株に *Z. mobilis* 由来の *pdc* と *Synechocystis sp.* PCC 6803 株由来のアルコールデヒドロゲナーゼ遺伝子 (*slr1192*) を導入し、26 日間培養を行うことで 5.50 g/L (212 mg L⁻¹day⁻¹) という高い培地中のエタノール濃度が得られた (Gao *et al.*, 2012)。このような物質生産性向上に向けた検討は、エタノール以外の物質生産でも広く行われている。例として、代謝酵素の過剰発現と制御因子の破壊の組み合わせ (Kanno *et al.*, 2017)、短縮型フィコビリソームの発現 (Sengupta *et al.*, 2023)、PSI/PSII 比の低減 (Moore *et al.*, 2024)、培地への金属添加による NADPH 再生 (Velmurugan *et al.*, 2019)、CBB サイクルを構成する酵素の過剰発現 (Nishiguchi *et al.*, 2019)、生合成経路発現への強発現プロモーターの利用 (Zhou *et al.*, 2014) などが挙げられる (Fig. 1-1-1)。加えて、光の照射と共に培地にグルコースを添加した条件 (Photomixotroph 条件) において様々な有用化合物の生産性が向上することも報告されている (Matson and Atsumi, 2017)。Photomixotroph 条件では CO₂ の取り込み速度が向上することも知られており、社会実装に向けた重要な技術であると考えられている (Kanno *et al.*, 2017)。

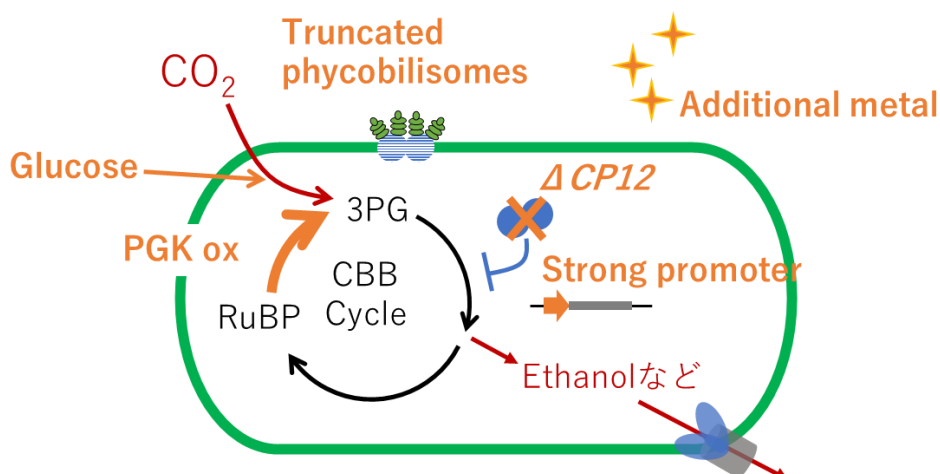


Fig. 1-1-1 シアノバクテリアの有用物質生産向上に寄与した因子の例

代謝酵素の過剰発現と制御因子の破壊の組み合わせ (Kanno *et al.*, 2017)、短縮型フィコビリソームの発現 (Sengupta *et al.*, 2023)、培地への金属添加による NADPH 再生 (Velmurugan *et al.*, 2019)、CBB サイクルを構成する酵素 (PGK) の過剰発現 (Nishiguchi *et al.*, 2019)、生合成経路発現への強発現プロモーター (*cpc560p*) の利用 (Zhou *et al.*, 2014) について記載した。

1.1.3 タンパク質工学による代謝改変のアプローチ

物質生産性を向上させることを目的とした代謝工学において、タンパク質工学的手法によって機能を改変した代謝酵素の活用は重要である。代謝酵素の機能改変と代謝工学での活用は、アロステリック阻害の緩和や基質特異性の変更などを目的として、古くから行われてきた。特に大腸菌や酵母といったモデル生物では、ランダム変異導入と高速スクリーニングを組み合わせた進化分子工学的手法や、構造情報に基づく合理的設計により、生産性を向上させた事例が多数報告されている (Pleiss, 2011; Nevoigt, 2008)。しかし、シアノバクテリアにおいては、これらの手法の適用は限定的である。その理由として、大腸菌などに比べて形質転換や培養に時間を要し、ハイスループットな評価系を構築しにくいことや、利用可能な遺伝子改変ツールが限られてきたことが挙げられる (Hitchcock *et al.*, 2019)。そのため、シアノバクテリアでは、より少ない試行回数で効果的な変異体を得るための、合理的な酵素設計戦略の確立が特に重要となる。

1.1.4 タンパク質配列データの拡充と有用配列の探索

タンパク質の配列情報は UniProt などのタンパク質配列データベースに集積されており、アミノ酸配列、機能アノテーション、構造情報、進化的な関係性などを参照する

ことが可能となっている (Apweiler *et al.*, 2004)。UniProt に登録されているタンパク質配列情報の数は著しく増加しており、2016 年頃には 6000 万件程度であった UniProt の配列数が (UniProt Consortium, 2017)、2024 年には 2 億 4600 万件を超えるまでに拡大している (UniProt Consortium, 2025)。配列情報から得られる知見として、タンパク質の進化的関連性や局在情報、相互作用情報、アミノ酸残基の保存度などが挙げられる。近年では構造予測の精度向上により、配列情報と立体構造も関連付けられつつある (Yang *et al.*, 2023)。

タンパク質配列数の増加に伴い、研究者は大量の利用可能なデータから自身の目的に適したタンパク質配列を選択することが可能となった。しかしながら、有用なタンパク質配列の探索は容易ではない。タンパク質の配列空間は、 20^n (n = アミノ酸残基数。PDC の場合、 $n = 500$) であり非常に膨大である。そのため、従来は機能的な配列である可能性が高い天然の配列について評価を進め、有用なタンパク質配列が同定されてきた (Yang *et al.*, 2024)。しかしながら、天然の配列数も増加してきており (PDC の場合、20 万配列以上 (Paysan-Lafosse *et al.*, 2023))、適切な配列の選抜が必要となる。また、天然には存在しない配列が有用である可能性もあるため、好ましくは非天然の配列も探索すべきであるといえる。

1.1.5 タンパク質におけるアミノ酸残基の保存度と機能

膨大なタンパク質の配列情報から抽出できる情報のひとつとして保存度が挙げられる。保存度はアミノ酸残基の機能に対する重要性や変異導入の指標になり得る。タンパク質改変の際には保存度が低く、変異の許容性が高い残基に変異を導入することで機能を向上したタンパク質変異体を獲得するというアプローチがしばしば取られている。これまでに、保存性の低い残基への変異導入によってタンパク質の安定性を向上した例 (Zheng *et al.*, 2022) や、基質特異性を狭くすることに成功した例 (Gerth *et al.*, 2007) などが報告されている。

しかしながら、保存度の低い残基への変異導入は、タンパク質の機能や構造にどのような影響を与えるかを予測することが困難であり (Tungtur *et al.*, 2010)、目的の機能変化をもたらす変異を特定するためには変異体ライブラリーの構築とスクリーニングが必要となる可能性がある (Xu *et al.*, 2020)。また、保存性の低い残基はファミリー内での機能の多様性を生む部分でもあるため、特定のタンパク質で得られた知見を同一ファミリー内で横展開することが困難な可能性も考えられる。

保存度の高い残基への変異導入は機能に破壊的な影響を与える可能性があるため避けられる傾向にある。保存度の高い残基が担う機能の例としては、酵素の活性中心、リガンド結合部位 (Perozich *et al.*, 1998)、フォールディングの促進 (Chikunova *et al.*, 2022) などが挙げられる。しかしながら、保存度の高さが意味する機能的な重要性は、酵素活性のように「維持すべき必須機能」を担う残基と、活性制御のように「改変の対

象となりうる制御機能」に分けることが出来るはずである。後者の様な残基に戦略的に変異を導入することが出来れば、酵素の基本機能を維持したまま、特定の制御のみを解除するという、大きな機能改変を効率的に達成できる可能性がある。

1.1.6 本論文における表記に関して

本論文では、生物種名はイタリック体で表記する (例: *Zymomonas mobilis*)。遺伝子名は小文字のイタリック体 (例: *prk*)、タンパク質・酵素名は立体 (例: PRK, Pyruvate decarboxylase) で表記することを原則とする。

1.2 本論文の目的及び構成

シアノバクテリア *Synechocystis* sp. PCC 6803 は高い増殖性と形質転換効率からモデル生物として広く研究されているラン藻類の一種であり、酸素発生型光合成により CO_2 を固定し、様々な有用物質を生産可能である。シアノバクテリアの代謝を理解し、有用物質生産に適した状態へと改変することは代謝工学の目的のひとつである。代謝を改変するために、好ましい機能を有する代謝酵素を発現することは有効なアプローチである。このアプローチは対象とする機能に関わる酵素が宿主に内在する場合と、内在しない場合に大別される。対象とする機能に関わる酵素が宿主に内在している場合には、その酵素にアミノ酸変異を導入して機能を改良することが可能であり、この方法は宿主の代謝に与える悪影響が比較的小さい点で有利である。一方で、宿主が目的とする機能に関与する内在性酵素を持たない場合には、最適な異種酵素を探索し、それを発現させる必要がある。アミノ酸変異による機能改良と異種酵素の探索は相補的に重要な技術であるといえ、両方の場合における方法論を確立することが、代謝工学のために重要である。近年、遺伝子解析技術の進歩により UniProt を始めとするタンパク質配列データベース上の配列数は急速に増加しており、これらの配列情報にはシアノバクテリアの代謝を改良するための有益な情報が含まれていると考えられる。本研究では、タンパク質配列情報をアミノ酸残基の保存度と有用遺伝子情報のリソースとして活用し、アミノ酸変異による機能改良と異種酵素の探索によってシアノバクテリアの代謝を改変するための方法論を示すことを目的とした (Fig. 1-2-1)。

本学位論文は、第1章から第4章により構成される。第1章では本研究の背景と目的について述べた。

第2章では、内在性酵素の機能改変の方法論について述べた。ここでは、タンパク質ファミリー内でのアミノ酸の「保存度」を、アミノ酸残基の機能的な重要性を見積もる指標として活用した。機能的に重要で、活性中心から離れた露出残基を変異導入の対象とすることで、酵素の活性を損なうことなく、特定の制御機能のみを合理的に改変する戦略を提案した。本方法論の適用例として、シアノバクテリアに内在し、かつ活性制御の解除が物質生産性向上に繋がると予測される Phosphoribulokinase (PRK) を対象とした。提案した方法に基づいて設計した変異 PRK を発現する株では CO_2 とグルコースを添加した Photomixotroph 条件でのみ、変異 PRK 発現株が野生型 PRK 発現株よりも高い細胞内 PRK 活性を示すことが明らかとなった。さらに、変異 PRK 発現株ではピルビン酸やクエン酸サイクル中の代謝物濃度が増加しており、これらを前駆体とする有用物質生産に適した形質を示すことが確認された。このような保存度の高いアミノ酸残基に対する変異導入という戦略は、タンパク質の機能改変と代謝状態改変に有用であることが確認された。

第3章では、異種酵素探索のための方法論について述べた。ここでは、広大な「配列空間」を効率的に探索するため、系統的に多様な配列の評価で有望な領域にあたり

をつける「広い探索」と、その領域を天然・非天然配列を含めて評価する「密な探索」からなる2段階の戦略を提案した。本方法論の適用例として、エタノール生産に必要なもののシアノバクテリアに内在しない pyruvate decarboxylase (PDC) を選択した。シアノバクテリアにおけるエタノール生産では、経験的にテキーラの醸造菌である *Zymomonas mobilis* (Zm) 由来 PDC が良く用いられてきた。近年では BRENDA 上に Zm PDC よりも高い k_{cat}/K_M を有する複数の酵素が登録されていることに加え、配列データベース上の PDC 配列の数も増加してきており、データベース上の多様な PDC 配列を探索すればよりシアノバクテリアでのエタノール生産に適した PDC 配列を同定できる可能性がある。シアノバクテリアでのエタノール生産において、PDC 活性がボトルネックになる例が報告されており (Luan *et al.*, 2015)、最適な PDC の獲得により、エタノールの生産性が向上すると期待される。探索の結果、従来良く用いられてきた Zm PDC よりも高いエタノール生産性を示す *Gluconobacter oxydans* PDC を同定することに成功し、データベース上の多数の配列から有用遺伝子を探索する方法を示すことができた。

第 4 章では、本研究で得られた知見をまとめ配列情報を活用したシアノバクテリアの代謝工学について将来の展望を述べ、結論とした。

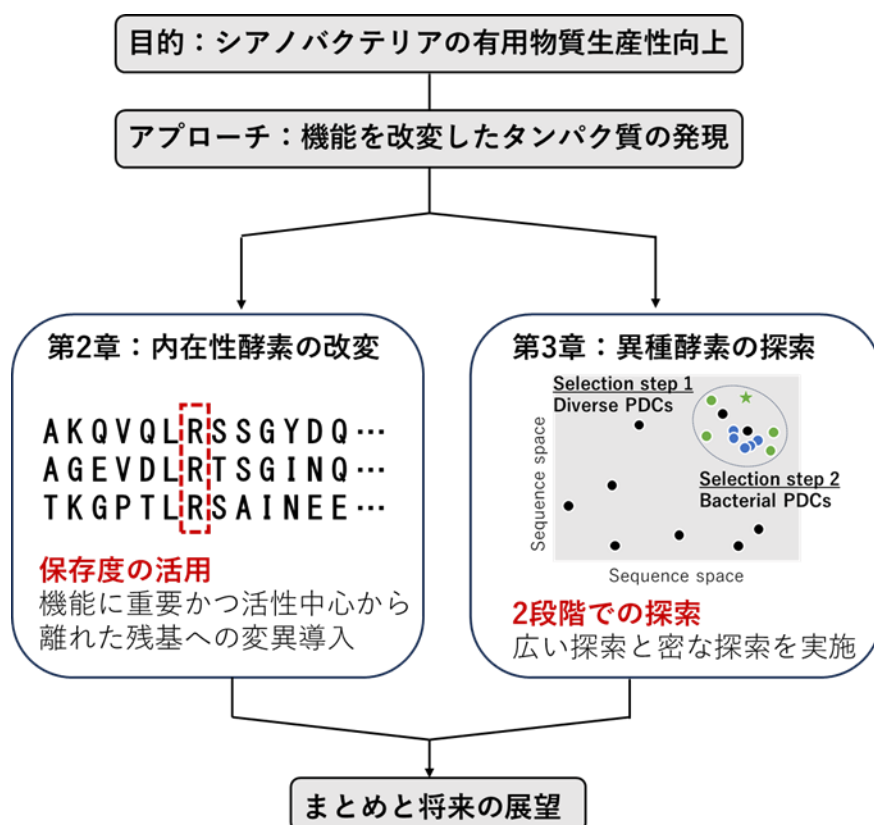


Fig. 1-2-1 本論文の目的及び構成

第2章 C 末端変異 Phosphoriblokinae の発現によるシアノバクテリアの代謝改変

2.1 緒言

微生物による物質生産性の向上には、好ましい機能を有する代謝酵素の発現が有効である。第1章で述べた通り、シアノバクテリアのようなハイスループット評価が困難な生物種では、少数の変異体セットで効果的な改変を行うための合理的な設計が求められる。そこで本章では、内在性酵素の合理的改変のための方法論として保存度を活用した変異導入戦略を提案し、その有用性を検証することを目指した。

アミノ酸残基の保存度は変異体を設計する際の有用な指標の一つである。一般的に保存度の高いアミノ酸残基は重要な機能を担っていることが知られているため (Sharir-Ivry *et al.*, 2021)、該当残基への変異導入はタンパク質の機能を改変する可能性が高い。タンパク質配列データベースに含まれる配列数の増加に伴い、多様なタンパク質ファミリーごとのアライメントを作製し、保存度を参照することが可能になってきている。保存度の高いアミノ酸残基に対して戦略的に変異を導入するという方法は様々なタンパク質設計にも応用できる可能性がある。

本方法論の有用性を検証する対象には、シアノバクテリアの炭素固定を担うカルビンベンソン回路の重要酵素 Phosphoribulokinase (PRK) を用いた。PRK は CO₂ を固定するカルビンサイクルを構成する重要な代謝酵素であり (Fig. 2-1-1)、Metabolic Control Analysis において、この酵素の活性が増加すると、解糖系下流に位置する PYK 反応のフラックスは同程度の割合で増加する (PRK 活性増加率の 1.08 倍) ことが予測されている (Nishiguchi *et al.*, 2019)。そのため、CP12 と呼ばれる小さなタンパク質との複合体形成や (Blanc-Garin *et al.*, 2022 ; Lucius *et al.*, 2022)、C 末端側ループにおける分子内ジスルフィド結合形成による活性抑制 (Fukui *et al.*, 2022) などによってその活性は精緻に制御されている。この抑制制御を部分的にでも解除し、細胞内の PRK 活性を向上させることができれば、シアノバクテリア細胞内の代謝状態が大きく変化し、有用化合物の生産に適した状態になることが期待できる。

これまでに、PRK の抑制制御に関与する CP12 の破壊株での細胞内 PRK 活性や代謝状態の変化については複数報告されている (Blanc-Garin *et al.*, 2022 ; Lucius *et al.*, 2022 ; McFarlane *et al.*, 2019)。CP12 はシアノバクテリア細胞内で PRK と Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) と可逆的に結合し、活性を抑制していることが報告されている (McFarlane *et al.*, 2019)。PCC6803 株においては CP12 による活性抑制機構は Photoautotroph では働いておらず、Photomixotroph 条件や暗所で機能していることが報告されている (Blanc-Garin *et al.*, 2022 ; Lucius *et al.*, 2022)。Photomixotroph 条件での培養によってシアノバクテリアの有用化合物の生産性が大きく向上する例が複数報告されており (Matson and Atsumi, 2017)、この条件での培養可能性は産業応用において重要であると考えられる。PRK 活性を抑制している CP12 を破壊することで Photomixotroph 条

件における PRK 活性を増加させ、解糖系下流や TCA サイクルへの炭素の引き抜きを増大させるアプローチが有効であると考えられるが、CP12 破壊株では Photomixotroph 条件での生育が困難になることが報告されている(Blanc-Garin *et al.*, 2022 ; Lucius *et al.*, 2022)。このことは、Photomixotroph 条件下で CP12 破壊株を利用することが困難であることを示している。

近年では CP12 以外の因子による PRK 活性の制御機構についても研究が進められている(Fukui *et al.*, 2022 ; Nishiguchi *et al.*, 2020 ; Wilson *et al.*, 2019)。先行研究において、*Anabaena* sp. PCC 7120 の PRK 分子内において C 末端ループ上の 2 つのシステイン(Cys230 と Cys236)によるジスルフィド結合形成が PRK の活性を抑制することが報告されている(Fukui *et al.*, 2022)。また、*Synechococcus* sp. strain PCC 6301 の Arg231 が相互作用界面の負電荷キャビティにはまることが PRK の活性抑制に寄与することも示唆されている(Wilson *et al.*, 2019)。しかしながら、シアノバクテリアの C 末端ループ上のアミノ酸残基の役割については *in vitro* での検証と考察にとどまっている。こうした *in vitro* での活性測定は細胞内で遭遇する条件とはかけ離れた条件下で行われるため、得られた知見の生理学的意義には疑問が残る(Castaño-Cerezo *et al.*, 2024)。実際に、アロステリック阻害を解除したタンパク質の発現によっても狙ったフェノタイプの変化が起こらない例(Sedivy *et al.*, 1986)や、*in vitro* と *in vivo* でアミノ酸変異による効果が異なる例などが報告されている(Davidi *et al.*, 2016; Comino *et al.*, 2017)。こうした乖離がある場合には変異体構築を含む実験全体が徒労に終わる可能性があるため、慎重に変異対象とするアミノ酸残基を選定する必要がある。

そこで本研究ではアミノ酸の保存度を指標とした細胞内での機能性推定と、保存性の高い残基への戦略的な変異導入を試みた。最初に、シアノバクテリア由来 PRK の各アミノ酸残基の保存度を計算し、保存度が高く、機能に重要な残基を抽出した。そのうち、活性中心の周辺残基やタンパク質のパッキングに関与する埋没残基を変異対象から除くことで、変異候補を絞り込んだ。得られた候補残基のうち、C 末端ループに位置するアミノ酸残基が変異導入の候補として特に有望であると考えられたため、C 末端側ループ上にアミノ酸変異を導入した変異 PRK を細胞内で発現させたところ、Photomixotroph 条件での生育を低下させずに細胞内 PRK 活性を向上することに成功した。変異 PRK 発現株についてメタボローム解析を行ったところ、細胞内の有機酸濃度が向上していることが明らかとなり、この株は有機酸を前駆体とした有用化合物生産に適している可能性が示唆された。本研究の成果はシアノバクテリア PRK の C 末端側ループのジスルフィド結合形成による活性抑制の解除がシアノバクテリアの代謝工学に有用であることを示している。加えて、本研究で採用した保存度を指標としたアミノ酸変異導入は、機能的に重要な残基を予測し、効率的に設計を行うための手法になりうるといえる。

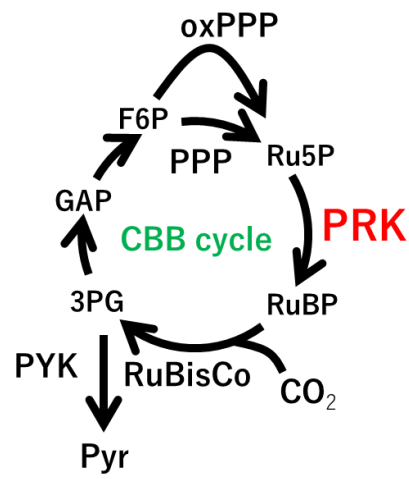


Fig. 2-1-1 カルビンサイクル模式図

2.2 実験方法

2.2.1 PRK 系統樹の作成

ラン藻類由来 PRK の C 末端領域の配列保存度と許容されるアミノ酸変異を解析すべく、PRK の系統樹作成を行った。UniProtKB reference proteomes + Swiss-Prot (Apweiler *et al.*, 2004) をデータベースとして PCC6803 由来 PRK をクエリ配列にした検索を行い、E-value 上位 5000 配列を取得した (2022 年 11 月 5 日検索)。これらの配列を閾値 0.8 で CD-hit (Li *et al.*, 2006) によってクラスタリングを行い、得られた配列群の配列長中央値よりも 30%以上長い又は短い配列を除去することで 694 配列を得た (配列長 228 aa ~ 424 aa の配列のみを抽出)。CD-hit によるクラスタリングで同一クラスター内の最も配列長の長い配列が選択されることで PCC6803 由来 PRK が除かれたため、PCC6803 由来 PRK を配列群に再度加え、計 695 配列とした。MSA は MAFFT (Katoh *et al.*, 2013)を用いて作成し、Trimal (Capella-Gutiérrez *et al.*, 2009) の gt メソッドで閾値を 0.9 として保存領域を抽出したのちに IQTree v2 を用いて最尤法により系統樹を推定した (Minh *et al.*, 2020)。進化モデルには ModelFinder を用いたベイズ情報量基準スコアによって LG+R9 を選択した (Kalyaanamoorthy *et al.*, 2017)。推定されたクレードの信頼性は、UFBoot2 (Hoang *et al.*, 2018) を用いて 1000 回の反復を行い、ブートストラップ法で評価した。

2.2.2 ラン藻類由来 PRK の C 末端領域の配列保存性確認

得られた系統樹においてラン藻類由来 PRK クレードに含まれる 31 配列を抽出し、それぞれの配列が含まれるクラスターの全配列を再回収することで 779 配列を得た。これらの配列から、配列長 228 aa ~ 424 aa の配列のみ (759 配列) を抽出し、multiple sequence alignment (MSA) を MAFFT を用いて作成した。PCC6803 由来 PRK の各アミノ酸残基の保存度 (出現頻度) を、各アミノ酸の出現回数を配列数 (759) で割ることによって算出した。

2.2.3 PCC6803 由来 PRK の立体構造予測

PCC6803 由来の野生型 PRK ホモダイマー構造は ColabFold v1.5.5: AlphaFold2 using MMseqs2 (Mirdita *et al.*, 2022) を用いて予測した。テンプレート構造として *Synechococcus* sp. strain PCC 6301 由来 PRK のホモダイマー構造である 6HZK (PDB) を入力し、num_recycles = 6 に設定した。その他はデフォルトパラメーターで立体構造を予測した。5 つの複合体構造を予測した中で pTM 値が最良であったモデルを選択した (pLDDT=91.5 pTM=0.899)。この構造を元に、一方の Chain を削除することでモノマー立体構造を得た。立体構造及び相互作用の図示は Discovery Studio Visualizer v21.1.0.20298 (Biovia, 2019) を使用した。

2.2.4 アミノ酸残基の露出度計算

溶媒露出度の計算は Discovery Studio Visualizer v21.1.0.20298 (Biovia, 2019) を用いて行った。溶媒露出度はタンパク質を構成する原子やアミノ酸残基が、水分子などの溶媒にどの程度表面を露出しているかを示す指標であり、溶媒アクセス可能表面積 (Solvent Accessible Surface Area: SASA) として数値化され、原子のファンデルワールス半径に溶媒プローブ (通常 1.4 Å) の半径を加えたプローブ半径を用いて、プローブが原子表面に沿って描く表面積が、グリッドベースの数値計算アルゴリズムにより算出される。溶媒露出度は、特定の理想的な伸長コンフォメーションを持つアミノ酸残基 (Ala-X-Ala トリペプチドの X) の SASA を 100% の基準としているため、解析対象の残基がこの基準コンフォメーションよりもさらに溶媒に露出した場合、計算値が 100% を超えることがある。

2.2.5 活性中心からの距離の計算

活性中心からの距離の計算は Discovery Studio Visualizer v21.1.0.20298 (Biovia, 2019) を用いて行った。ADP 又は RuBP と相互作用すると推定される Ser17, Lys21, Ser22, Asp43, Arg50, Arg53, Tyr89, His91, Arg164 (Yu *et al.*, 2020) を選択し、いずれかの残基から 8 Å 以内に含まれるアミノ酸残基を変異導入を行わない残基として抽出した。

2.2.6 ラン藻類由来 PRK の配列アライメント

シアノバクテリア由来 PRK の配列は UniprotKB データベース (Apweiler *et al.*, 2004) から取得した (*Synechococcus* sp. strain PCC 6301 : Q31PL2、*Anabaena* sp. PCC 7120 : Q8YPR9 (UniprotKB)、*Synechocystis* sp. PCC 6803 : P37101 (UniprotKB))。アライメントは CLUSTAL 2.1 (Larkin *et al.*, 2007) を用い、デフォルトパラメーターで実施した。

2.2.7 使用菌株

本研究では、*Synechocystis* sp. PCC 6803 グルコース耐性 (GT) 株 (PCC 6803) と、新たに構築した野生型及び変異型 PRK を追加過剰発現した株を検討に用いた (Table 2-1、株の構築については 2.2.2 に記載)。

Table 2-2-1 使用菌株

Strain	Genotype	reference
PCC 6803	<i>Synechocystis</i> sp. PCC 6803 glucose-tolerant	Williams., 1988
PCC 6803-PRK _{WT}	PCC6803 + <i>slr0168::PpsbA2-prk-Smr</i>	This work
PCC 6803-PRK _{CCSS}	PCC6803 + <i>slr0168::PpsbA2-prk(Cys229Ser Cys235Ser)-Smr</i>	This work
PCC 6803-PRK _{R231T}	PCC6803 + <i>slr0168::PpsbA2-prk(Arg231Thr)- Smr</i>	This work

slr0168 はニュートラルサイトとして知られている。ストレプトマイシン耐性遺伝子を *Smr* と記載した。

2.2.8 培地組成

シアノバクテリアの前培養と本培養には共に 25 mM となるように NaHCO₃ を添加した BG11 培地を用いた (Table 2-2-2)。

Table 2-2-2 BG11 培地組成

	(mM)
EDTA 2Na	2.7×10^{-3}
NaNO ₃	1.8×10
H ₃ BO ₃	4.6×10^{-2}
HEPES	2.0×10
Na ₂ MoO ₄ · 2H ₂ O	1.6×10^{-3}
K ₂ HPO ₄	2.2×10^{-1}
MgSO ₄ · 7H ₂ O	3.0×10^{-1}
CaCl ₂	2.6×10^{-1}
MnCl ₂ · 4H ₂ O	9.1×10^{-3}
ZnSO ₄ · 7H ₂ O	7.7×10^{-4}
CuSO ₄ · 5H ₂ O	3.2×10^{-4}
Co(NO ₃) ₂ · 6H ₂ O	1.7×10^{-4}
FeCl ₂ · 4H ₂ O	1.6×10^{-2}
NH ₄ Cl	2.5×10^{-2}

2.2.9 培養条件

・ Photoautotroph 条件

前培養

20 mL の 25 mM NaHCO₃ BG11 培地入り 100 mL 容三角フラスコにシアノバクテリアを植菌し、30 °C、150 rpm、40 μ mol/m² sec で 96~120 時間培養した (バイオシェーカー BR-43FL (LC-LED450W; TAITEC, Saitama, Japan))。遺伝子組換え細胞の選択マーカーとして必要に応じて終濃度 20 μ g/mL となるようにストレプトマイシンを添加した。

本培養

20 mL の 25 mM NaHCO₃ BG11 培地入り 100 mL 容三角フラスコに、前培養から初期 OD₇₃₀ =0.05 となるようにシアノバクテリアを植菌し、30 °C、150 rpm、40 μ mol/m² sec で 48 時間培養した。

・ Photomixotroph 条件

前培養

20 mL の 25 mM NaHCO₃ 及び 5 mM グルコースを含む BG11 培地入り 100 mL 容三角フラスコにシアノバクテリアを植菌し、30 °C、150 rpm、30 μ mol/m² sec で 96~120 時間培養した。Photomixotroph 条件では光強度約 40 μ mol/m² sec において野生型の PRK 追加過剰発現株の増殖が確認されなかったため、光強度を約 30 μ mol/m² sec となるように弱く設定した。

遺伝子組換え細胞の選択マーカーとして必要に応じて終濃度 20 μ g/mL となるようにストレプトマイシンを添加した。

本培養

20 mL の 25 mM NaHCO₃ 及び 5 mM グルコースを含む BG11 培地入り 100 mL 容三角フラスコに、前培養から初期 OD₇₃₀ =0.05 となるようにシアノバクテリアを植菌し、30 °C、150 rpm、30 μ mol/m² sec で 48 時間培養した。

2.2.10 菌体濃度及び培地中グルコース濃度の測定

菌体濃度分析

分光光度計 (UVmini-1240 (島津製作所)) を用いて培養液の 730 nm の吸光度 (OD₇₃₀) を測定した。培養液の濃度が高い場合は必要に応じて希釈して測定した。

培地中グルコース濃度の測定

適宜培養液を 1.0 mL 採取し、遠心分離した (15,000 rpm, 10 min, 4°C) 上清を分析サンプルとした。培地中のグルコース濃度は Glucose Assay Kit-WST (Dojindo Laboratories, Kumamoto, Japan) を用いてプロトコル通りに測定した。

2.2.11 PRK 追加過剰発現用 DNA 断片の構築

シアノバクテリアのニュートラルサイトである *slr0168* に光化学系 II のプロモーター *psbA2p* の下に野生型及び変異型 *prk* 遺伝子を結合したもの、およびストレプトマイシン耐性遺伝子を組み込んだ株を構築するための DNA 断片を構築した。本研究で用いたプライマーの配列を Table 2-2-3 に示した。すべての遺伝子操作は先行研究 (Nishiguchi *et al.*, 2019 ; Yoshikawa *et al.*, 2015) に記載された方法に従って行った。先行研究で構築された Et-g + Slr0168::PpsbA2-prk-Smr (Nishiguchi *et al.*, 2019) のゲノムを鋳型として、Primer 1 と 2 を用いた PCR によってニュートラルサイトである *Slr0168* サイト (Yoshikawa *et al.*, 2015) に *PpsbA2* プロモーター、PCC6803 野生型 PRK (PRK_{WT})、ストレプトマイシン耐性遺伝子を組み込むための DNA 断片を得た。さらに、同じゲノムを鋳型として Primer 1 と 3、Primer 2 と 4 でそれぞれ PCR を行い、得られた PCR 産物を Primer 1 と 2 による SOE-PCR によって結合することで、PRK_{WT} に Cys229Ser と Cys235Ser を導入した変異体 (PRK_{CCSS}) 追加過剰発現用の DNA 断片を得た。同様に、Primer 4 の代わりに Primer 5 を用いて PRK_{WT} に Arg231Thr を導入した変異体 (PRK_{R231T}) 追加過剰発現用の DNA 断片を得た。

Table 2-2-3 使用プライマー

Number	Name	Sequence
1	slr0168_up_F	CCAAACCACCGCTGATTGGAG
2	slr0168_down_R	CAGGTCACTGAAGCGGTCTAAC
3	Prk_int_R	GGGACGCCAATCAATGGTGG CGAAGGTTCCACCATTGATTGGCGTCCCTCAGGT
4	Prk_CCSS_F	CGGAAGCTGACCTCAACCTATCCTGGCATCAAGATG CGAAGGTTCCACCATTGATTGGCGTCCCTGTGG
5	Prk_R231T_F	TACAAAGCTGACCTGTACCTATCCTGGCATCAAGATG
6	Prk_CCSS_pet_F	ATTGATTGGCGTCCCTCAGGTCGGAAGCTGACCTCAACCTATCCTGGCATCAAGATG
7	Prk_R231T_pet_F	ATTGATTGGCGTCCCTGTGGTACAAAGCTGACCTGTACCTATCCTGGCATCAAGATG

2.2.12 シアノバクテリアの形質転換

シアノバクテリアは自然形質転換能を持つため、ゲノム DNA との相同領域をもつ DNA 断片を菌体と混合することで形質転換を行った。5 mM のグルコースを含む BG11

液体培地に PCC6803 のコロニーを植菌し、30 °C、150 rpm、30 μ mol/m² sec で 48~96 時間培養した。OD₇₃₀=0.05 となるように 5 mM のグルコースを含む BG11 培地に植え継ぎ、30 °C、150 rpm、30 μ mol/m² sec で 24 時間培養した菌体を 50 mL ファルコンに移し、室温 3000 g×5 min で集菌し、グルコースを含まない BG11 に再懸濁した。菌液を 200 μ l ずつ分注し、2-10 μ g 程度の DNA 断片を混合した。30 °C、30 μ mol/m² sec で 6 時間インキュベートしたのち、BG11 プレート (薬剤無し) にニトロセルロース膜 (ポアサイズ ϕ 0.22 μ m, 直径 5.0 μ m, Millipore, MF Mixed Cellulose Ester) を載せ菌液をスポットし、コンラージ棒で塗り広げた。24 時間蛍光灯照射下でインキュベートした後、10 μ g/mL のストレプトマイシン (Sm) を含む BG11 プレートにニトロセルロース膜を移した。1~2 週間、30 °C の蛍光灯照射下で培養することで、目的の DNA 断片が形質転換された変異株を選択した。得られたコロニー数個を独立に新しい BG11 プレート 20 μ g/mL Sm に植え継いだ。ゲノム上の元々の遺伝子が目的の DNA 断片に全て置き換わるまで植え継ぎを繰り返した。目的遺伝子カセットの組み込みはコロニー PCR で確認した。形質転換が完了したコロニーを終濃度 20 μ g/mL となるようにストレプトマイシンを添加した BG11 液体培地に植菌し、OD₇₃₀=1 程度になるまで培養した。その後フリーズストックを作製した。

2.2.13 シアノバクテリアの粗酵素液調製

培養開始 48 時間目の培養液から遠心分離 (10 分、3000 g、室温) によって菌体を回収し、プロテアーゼ阻害剤 (Roche, Switzerland, 1 錠/25 mL) を含む 20 mM Tris-HCl 緩衝液 (pH 8.0) に再懸濁した。この懸濁液を Bioruptor (Sonicbio, Knanagawa, Japan) を用いて 10 回超音波処理した (30 秒の超音波処理、処理間隔は 30 秒)。細胞抽出物を遠心分離し、上清を使用まで -80 °C で保存した。タンパク質濃度は Bradford Protein Assay Kit (Takara Bio Inc., Shiga, Japan) を用いて定量した。

2.2.14 野生型及び変異型 PRK の大腸菌での調製

先行研究で構築した pET28-a(+)/prk (Nishiguchi *et al.*, 2020) を鋳型として Table 2-2-3 の Primer3 と 6、Primer 3 と 7 でそれぞれ PCR を行い、得られた断片で大腸菌 DH5a 株を形質転換することで変異型 PRK 発現用プラスミド pET28-a(+)/prk (CCSS) 及び pET28-a(+)/prk (R231T) を得た。得られたプラスミドで大腸菌 BL21 (DE3) を形質転換した。得られたコロニーを標的タンパク質を高発現する大腸菌として取得し、グリセロールストックとして -80 °C で保存した。

30 μ g/mL のカナマイシンを含む 1 mL の LB 培地で試験管による前培養をおこなった後、30 μ g/mL のカナマイシンを含む 40 mL の LB 培地を入れた 200 mL 容バツフルフラスコに初期 OD=0.05 になるように植え継いだ。OD=0.5 になったときに終濃度 0.1 mM

となるように IPTG を添加し、24 時間目の菌体を 10 mL 回収し、遠心分離 (3000 rpm, 5 min, 4 °C) することによって培地を除いた。

2.2.15 PRK タンパク質の精製

得られた大腸菌菌体を 3 mL の BUGBUSTER(R) MASTER MIX (Novagen, Madison, WI, USA) を用いて溶菌させ、クルードサンプルとした。次いで、クルードサンプル全量を用いて Capturem™ His-Tagged Purification Miniprep kit (TaKaRa) によってタンパク質を精製した。Amicon Ultra Filter Units (10 kDa, Millipore, Schwalbach, Germany) を用いて 20 mM Tris-HCl 緩衝液 (pH 8.0) へとバッファー交換を行った。精製タンパク質の純度は SDS-PAGE によって確認した。タンパク質濃度は、Bradford Protein Assay Kit (Takara Bio Inc., Shiga, Japan) を用いて定量した。

2.2.16 SDS-PAGE

標的タンパク質の発現確認を行うため SDS-PAGE を行った。サンプルと 2 X sample buffer の混合物を 99 °C で 5 分間加熱した。長寿ゲル (Oriental Instruments Co., Ltd., Kanagawa, Japan) を泳動槽にセットし、標準 running buffer (0.25 mM Tris, 0.192 mM Glycine, 0.1 % SDS) を泳動槽に注いだ。サンプルを 5 µL をウェルに添加し、40 mA で 40 分間電気泳動を行った。泳動後のゲルを CBB R250 染色液 (BIO RAD, CA, USA) で染色した後、脱染色してタンパク質のバンドを確認した。分子量マーカータには Precision Plus Protein™ All Blue Prestained Protein Standards (BIO RAD, CA, USA) を用いた。

2.2.17 PRK 活性測定

PRK の酵素活性測定は以下の反応式に示すように、PRK 反応と PYK 及び LDH 反応をカップリングすることによって行った (Wadano *et al.*, 1998)。PYK と LDH には Rabbit 由来の精製酵素を使用した (Sigma Aldrich, MO, USA)。PRK 反応によって生成した ADP が PYK 反応を進行させ、それによって生じる PYR が LDH 反応を進行させ、NADH を NAD へと酸化する。PRK の活性は NADH の酸化による 340 nm の吸光度の変化として検出することができる。PRK 活性測定溶液を Table 2-2-4 の組成で調製し、シアノバクテリア粗酵素液と精製 PRK の活性をそれぞれ下記に示す方法で測定した。

$\text{D-Ribulose 5-Phosphate} + \text{ATP} \xrightarrow{\text{PRK}} \text{D-Ribulose 1,5-Bisphosphate} + \text{ADP}$ (PRK 反応)

$\text{ADP} + \text{PEP} \xrightarrow{\text{PYK}} \text{Pyruvate} + \text{ATP}$ (PYK 反応)

$\text{Pyruvate} + \beta\text{-NADH} \xrightarrow{\text{LDH}} \text{Lactate} + \beta\text{-NAD}$ (LDH 反応)

・シアノバクテリア粗酵素液

タンパク質濃度が 60 mg/L となるようにシアノバクテリア粗酵素液を 20 mM Tris-HCl 緩衝液 (pH 8.0) で希釈した。希釈サンプル 16 µL と 100 mM DTT 4 µL 又はミリ Q 水

を混合し、室温で 20 分間インキュベートすることで還元及び非還元の粗酵素液を作製した。得られた粗酵素液 10 μL と PRK 活性測定溶液 80 μL を混合し、酵素反応を開始させた。反応開始から 20 分間の 340 nm の吸光度をプレートリーダー (SYNERGY, BioTec, Green Mountains, VT) によって測定した。酵素反応速度及び活性値は 340 nm の吸光度のプロット直線部分の傾きから算出した。

・精製 PRK

タンパク質濃度が 20 mg/L となるように精製 PRK を 20 mM Tris-HCl 緩衝液 (pH 8.0) で希釈した。希釈サンプル 16 μL と 100 mM DTT 4 μL 又はミリ Q 水を混合し、室温で 60 分間インキュベートすることで還元及び非還元の酵素液を作製した。得られた酵素液 8 μL と PRK 活性測定溶液 80 μL を混合し、酵素反応を開始させた。反応開始から 20 分間の 340 nm の吸光度をプレートリーダーによって測定した。酵素反応速度及び活性値は 340 nm の吸光度のプロット直線部分の傾きから算出した。

Table 2-2-4 PRK 活性測定溶液

Ru5P	2 mM
PEP	2.5 mM
ATP	2 mM
NADH	0.3 mM
MgCl ₂	10 mM
PYK + LDH	4 U
50 mM Tris-HCl pH8 に溶解	

2.2.18 酵素反応速度及び活性値の算出

酵素反応速度及び活性値はそれぞれ以下の式に従って算出した。 $\Delta A_{340\text{nm}}$ は 20 分間の吸光測定において、吸光度のプロットに直線性が見られた7点を利用して算出した。NADH のモル吸光定数には 6300 L/mol/cm を利用した。光路長は反応液の全体積が 100 μL であることと、使用した 96 穴プレートの底面積が 0.32 cm^2 であることから算出した。

$$\text{酵素反応速度 (mM/h)} = \frac{\Delta A_{340\text{nm}} (\text{/min}) \times 60 \times 10^3}{\text{光路長 (cm)} \times \text{モル吸光係数 (L/mol/cm)}}$$

$$\text{酵素活性 (Unit/mL)} = \frac{\Delta A_{340\text{nm}} (\text{/min}) \times \text{全反応容量 } (\mu\text{L}) \times 10^3}{\text{光路長 (cm)} \times \text{モル吸光係数 (L/mol/cm)} \times \text{酵素容量 } (\mu\text{L})}$$

2.2.19 中間代謝物質の定量分析

サンプリング

サンプリングは対数増殖期中である培養開始 48 時間目に行った。まず、減圧濾過用フィルターホルダーにフィルターをセットした (PTFE メンブレンフィルター; 孔径 0.5 μm 、フィルター径 90 mm (ADVANTECH))。培養液を $\text{OD}_{730} \times \text{サンプリング液量 (mL)} = 20$ となるようにサンプリングし、フィルター上に培養液を流し込んだ。アスピレーターにより減圧濾過を行い、フィルトレーションを行った。フィルトレーション後、フィルターごとクエンチ溶液である 1.6 mL のメタノールに浸漬し、 -80°C で保存した。クエンチ溶液はサンプリング中も氷上に静置しておいた。

細胞内代謝物質の抽出

採取したサンプルに Milli-Q 水 640 μL 、クロロホルム 1.6 mL を加えて超音波を用いて懸濁し攪拌した後、遠心した (3700 g, 4°C , 20 min)。遠心後に 4°C で静置した後、上層を 250 μL ずつ 5 本のエッペンチューブに分注し、濃縮遠心機 (Centrifugal Concentrator CC-105, TOMY, Tokyo, Japan) を用いて常温で減圧乾燥し、乾燥後のサンプルは -80°C で保存した。LC-MS/MS での分析の際には乾燥後のサンプルを Milli-Q 水 50 μL に再溶解した。GC-MS での分析の際には乾燥後のサンプルに誘導体化の処理を行った。

分析条件

分析には液体クロマトグラフィー・質量分析計 (LC-MS/MS) (液体クロマトグラフィー; HPLC (Agilent 1100series; Agilent Technologies, CA, USA), 質量分析部; MS/MS (API 2000TM LC-MS/MS システム; AB SCIEX, YTO, Canada) 及び、ガスクロマトグラフィー・質量分析計 (GC-MS) (GC-MS-QP2020 NX (Shimadzu, Kyoto, Japan)) を使用した。代謝物のピークは標準物質のクロマトグラフと保持時間を比較して同定し、ピーク面積の算出にはそれぞれ Analyst Software Version 1.6.2 (AB SCIEX, YTO, Canada) 及び、GC-MS solution (Shimadzu, Kyoto, Japan) を用いた。Table 2-2-5 に HPLC の分析条件、Table 2-2-6 にイオン源パラメータ、Table 2-2-7 に MS/MS の分析条件を示し、解糖系とペントースリン酸経路の中間体代謝物質の Multiple reactions monitoring (MRM) トランジションと MRM パラメータを Table 2-2-8 に示す。さらに、Table 2-2-9 に GC-MS の分析条件、Table 2-2-10 に GC-MS で分析した代謝物の m/z を示す。

Table 2-2-5 HPLC の分析条件

カラム	ProteCol-P C18 HQ103 / 長さ:150 mm, 内径:2.1 mm, 粒径:3 μm
溶離液	(A) 15 mM トリブチルアミン, 10 mM 酢酸水溶液, (B) メタノール
グラジエント	A:100% (~24.0 min) ~A:10%, B:90% (~24.1 min) ~A:100% (~30 min)
流速	200 mL min ⁻¹
注入量	3 μL
カラム温度	35°C

Table 2-2-6 イオン源パラメータ

Ion source	Turbo spray
Curtain Gas	40 psi
Collision Gas	8 psi
Ionspray Voltage	-4500 V
Temperature	500°C
Ion Source Gas1	70 psi
Ion Source Gas2	60 psi
Interface Heater	Off

Table 2-2-7 MS/MS の分析条件

Mass 分析モード	負イオンモード
Gas temp	300°C
Gas flow	10 L min ⁻¹
Nebulizer	55 psi
Sheath gas temp	380°C
Sheath gas flow	11 L min ⁻¹
Capillary pos/neg	3500/3500 V
Nozzle voltage pos/neg	1000/1000 V
Resolution Q1/Q3	Wide/Wider
EMV 電圧	2000 V

Table 2-2-8 各代謝物質の MRM トランジションと MRM パラメータ

Compounds	Precursor ion	Product ion	Precursor ion	Product ion	Decustering Potential	Entrance Potential	Collision Energy	Collision Cell Exit Potential
	m/z for non-	m/z for non-	m/z for full	m/z for full				
	labelled	labelled	¹³ C labelled	¹³ C labelled				
	metabolites	metabolites	metabolites	metabolites				
	(Number of ¹³ C)	(Number of ¹³ C)	(Number of ¹³ C)	(Number of ¹³ C)	(V)	(V)	(V)	(V)
G6P/F6P	259 (0)	79 (0)	265 (6)	79 (0)	-31	-10	-22	-5
G1P	259 (0)	79 (0)	265 (6)	79 (0)	-60	-4	-64	-4
FBP	339 (0)	97 (0)	345 (6)	97 (0)	-37	-10	-28	-5
PEP	167 (0)	79 (0)	170 (3)	79 (0)	-22	-10	-20	-5
6PG	275 (0)	79 (0)	281 (6)	79 (0)	-37	-10	-66	-5
R5P/Xu5P/Ru5P	229 (0)	97 (0)	234 (5)	97 (0)	-31	-10	-22	-5
S7P	289 (0)	97 (0)	296 (7)	97 (0)	-21	-10	-22	-5

Table 2-2-9 GC-MS の分析条件

カラム	DB-5MS+DG (Agilent Technologies) 長さ:30 m, 内径:0.25 mm, 膜厚:0.25 μm
気化室温度	250°C
スプリット比	Split (Split ratio 50.1)
オープン温度	60°Cで 3.5 min 保持後、10°C/min で昇温、325°Cで 10 min 保持
カラム流量	ヘリウムガス 1.14 mL/min
ページ流量	3.0 mL/min
インターフェイス	250°C
イオン源	200°C
マスレンジ	<i>m/z</i> 100-950
スキャンスピード	3333 u/sec
検出器電圧	チューニング相対値 0 kV

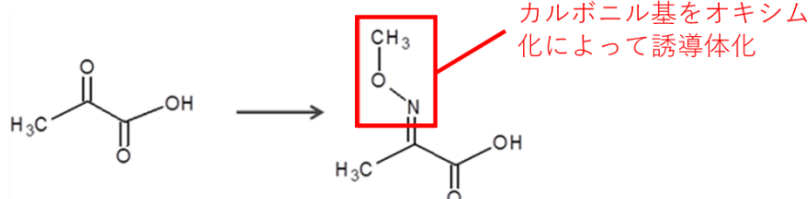
Table 2-2-10 GC-MS で分析した代謝物の m/z

	m/z		m/z		m/z
Pyr	174	Mal	233	Ribitol	217
Suc	247	Isocit	245	Cit	273
Fum	245				

2.2.20 中間代謝物質の誘導体化法

誘導体化には Methoxyamine hydrochloride によるオキシム化と N-Methyl-N-trimethylsilyltrifluoroacetamide + 1% trimethylchlorosilane (MSTFA+1%TMCS) によるトリメチルシリル化(TMS 化)を用いる。ピルビン酸を例とした概略を Fig. 2-2-1 に示した。乾燥したサンプルに 40 mg/mL Methoxyamine hydrochloride in pyridine 10 μ L 加え、Deep Well Maximizer (TAITEC, Saitama, Japan) で 30 $^{\circ}$ C、90 min 攪拌したのち、MSTFA+1% TMCS 40 μ L を加えて 37 $^{\circ}$ C、30 min 攪拌した。アルミホイルで遮光した後、室温で 2 時間以上安置し、上清を GC-MS 分析に供した。

Methoxyamine hydrochlorideによるオキシム化



MSTFAによるTMS化

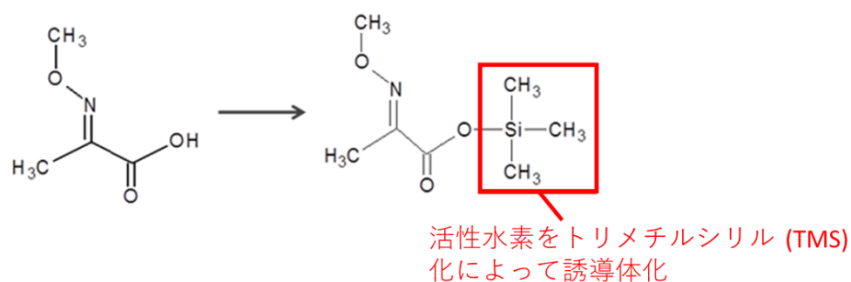


Fig. 2-2-1 Methoxyamine hydrochloride と MSTFA による誘導体化の概略

2.3 結果

2.3.1 PRK 配列の系統解析

UniProtKB reference proteomes + Swiss-Prot (Apweiler *et al.*, 2004) をデータベースとして PCC6803 由来 PRK をクエリ配列にした検索を行うことで得られた 5000 配列を出発配列とした。これらの配列から冗長性の削除などを行うことで得られた 695 配列を用いて系統解析を行った (Fig. 2-3-1)。系統解析によると、植物型 PRK が明瞭なクレードを形成し、その中にシアノバクテリア PRK のクレードが形成されていることが確認された。このシアノバクテリア PRK クレードに属する 31 配列を保存性の解析へと進めることとした。また、植物由来 PRK の他にも古細菌型 PRK や光合成細菌型 PRK も配列として多く存在していることが確認された。

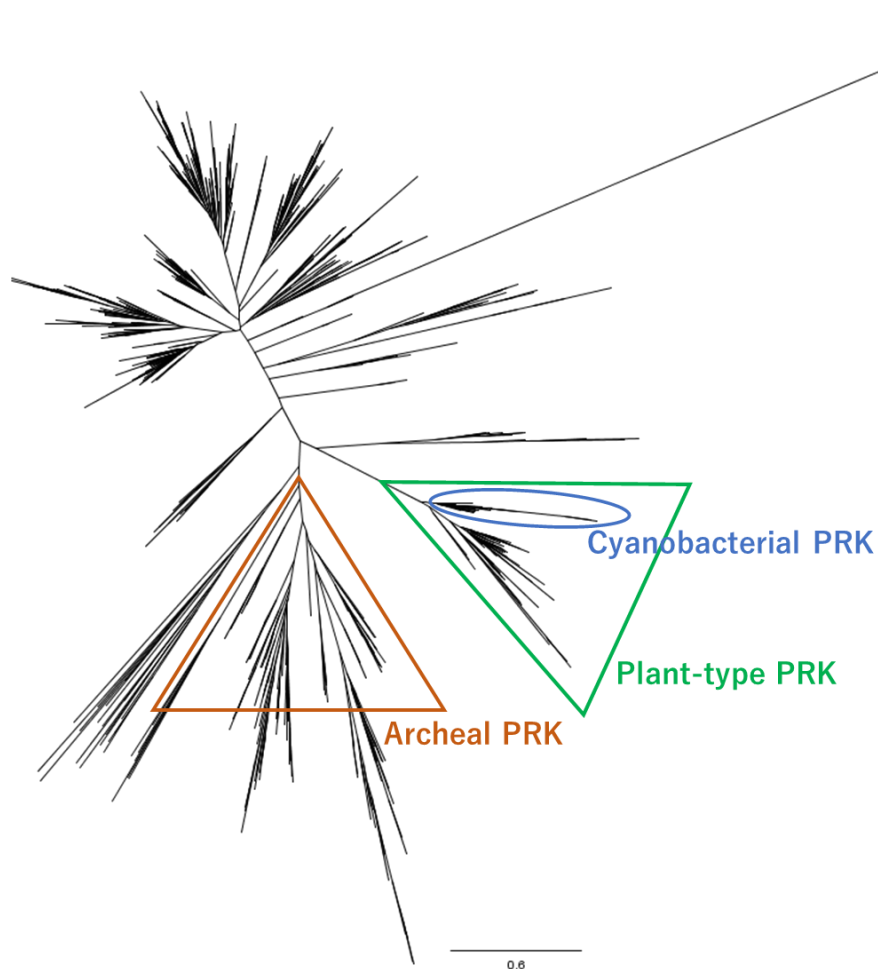


Fig. 2-3-1 PRK 系統樹

植物型 PRK と古細菌型 PRK が構成するクレードをそれぞれ緑とオレンジで囲んだ。その他のクレードには光合成細菌型 PRK が含まれている。

2.3.2 シアノバクテリア由来 PRK の各残基の保存度計算

シアノバクテリア PRK クレードに属する 31 配列を出発配列とし、冗長性を削除する前の配列を再回収することで、シアノバクテリア PRK の 779 配列を得た。これらの配列から配列長が長すぎる又は短すぎる一部の配列を削除して得られた 759 配列を元にシアノバクテリア由来 PRK の各位置のアミノ酸残基の保存度を計算し、Fig. 2-3-2 に示した。この結果によると、シアノバクテリア由来 PRK という比較的配列多様性の小さい集団に対して解析を行ったことで、保存度の高い残基が多いことが分かる。本研究ではその中でも特にシアノバクテリア PRK の機能に重要な残基を変異対象に選定するため、保存度 95%以上の残基を変異の対象とすることにした。PCC6803 由来 PRK の全長が 332 アミノ酸であるのに対し、保存度 95%以上の残基は 158 アミノ酸であり、全体の 40%程度であることが明らかとなった。これらの保存性が高い残基の中には、PRK 活性の維持に重要な残基、構造維持に重要な残基、制御機構に重要な残基などが含まれていると考えられる。

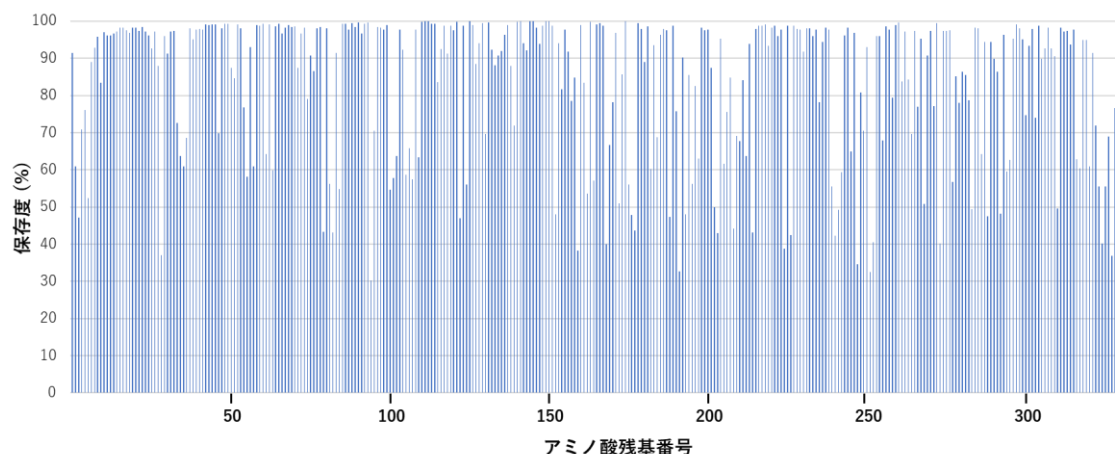


Fig. 2-3-2 シアノバクテリア由来 PRK の各残基の保存度

アミノ酸残基の番号付けは PCC6803 由来 PRK に従った。横軸は 50 残基ごとに番号を振った。

2.3.3 PCC6803 由来 PRK の立体構造を用いた変異残基絞り込み

保存度の高い残基のうち、PRK 活性の維持に重要な残基、構造維持に重要な残基は変異によって PRK の機能を損ねる可能性が高いため、変異対象から除くべきである。そこで、アミノ酸残基の露出度と、立体構造上の活性中心からの距離を指標に対象残基をさらに選抜することとした。Colabfold による構造予測で得られた PCC6803 由来 PRK のモノマー構造を元に、各アミノ酸の溶媒露出度と活性中心からの距離を算出した (Fig. 2-3-3)。

タンパク質の内部に位置するアミノ酸残基の精密な構造的配置と密なパッキングは、タンパク質の折りたたみや安定性を維持するために極めて重要であり、埋没残基への変異はタンパク質を極めて不安定化することが知られている (Bajaj *et al.*, 2005)。文献では露出度が 25%以下である場合に残基が埋没していると考えるのが一般的であり (Wu *et al.*, 2017)、PCC6803 由来 PRK の場合、332 アミノ酸のうち 169 アミノ酸が埋没していることになる。

酵素の活性中心や、その近接残基への変異導入は酵素活性を著しく低下させる可能性が高い (Sunden *et al.*, 2015)。ある特定の残基に対する近接残基の定義は文献ごとに大きく異なっており、5-10 Å 程度までさまざまである (Jiao *et al.*, 2013)。本研究では活性中心を構成するいずれかのアミノ酸残基から 8 Å 以内を近接残基として定義し、活性中心を広く覆う形で変異導入を避けることとした。この定義に基づくと、PCC6803 由来 PRK の場合、332 アミノ酸のうち 91 アミノ酸が活性中心に近接していることとなる。

2.3.2 項に示した保存度の高い残基から埋没残基と活性中心近傍の残基を除くことで変異導入候補として 31 残基が得られた (Fig. 2-3-4)。これらの残基は立体構造上で大きく分けて 4 つの領域に位置していた (Fig. 2-3-4c)。シアノバクテリア由来の代謝酵素は様々な酸化還元制御を受けることが知られており (Michelet *et al.*, 2013)、酸化還元制御に関与することが多い残基としては Cys や Met が知られている (Lim *et al.*, 2019; Klomsiri *et al.*, 2011)。PRK も酸化状態では大きく活性が低下することが報告されており、酸化還元制御の解除は PRK 活性向上につながると期待される。このことから、領域 C または D が変異導入の有望な候補であると考えられる。中でも領域 D に含まれる Cys 残基は *Anabaena* sp. PCC 7120 の PRK において酸化還元制御への関与が *in vitro* で報告されており (Fukui *et al.*, 2022)、特に有望な候補であるといえる。そこで本研究ではこのシステイン残基に着目し、変異体設計を進めることとした。

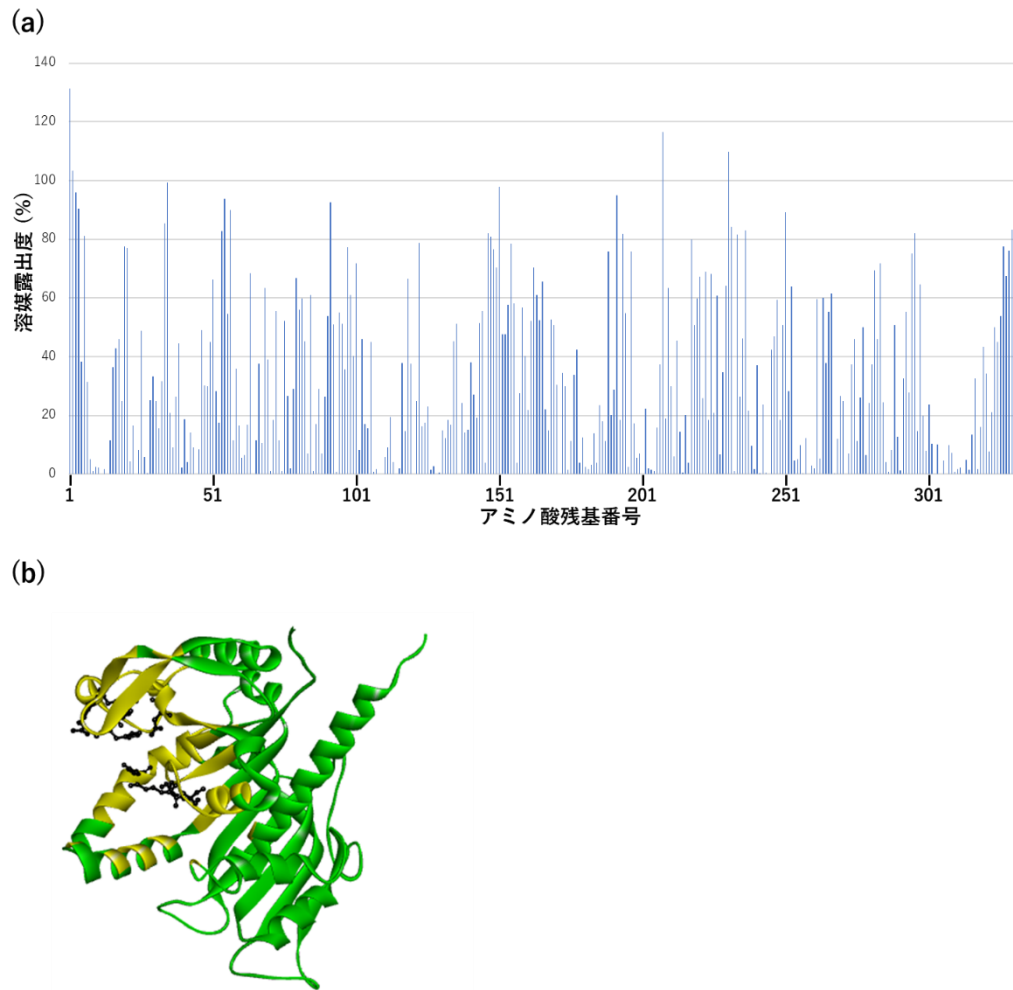


Fig. 2-3-3 PCC6803 由来 PRK の溶媒露出度と活性中心近傍残基

(a) PCC680 由来 PRK 各残基の溶媒露出度。溶媒露出度は、特定の理想的な伸長コンフォメーションを持つアミノ酸残基 (Ala-X-Ala トリペプチドの X) の SASA を 100% の基準としているため、解析対象の残基がこの基準コンフォメーションよりもさらに溶媒に露出した場合、計算値が 100% を超えることがある。(b) ADP 又は RuBP と相互作用すると推定される残基 (Yu et al., 2020) から 8 Å 以内に含まれるアミノ酸残基を黄色で示した。

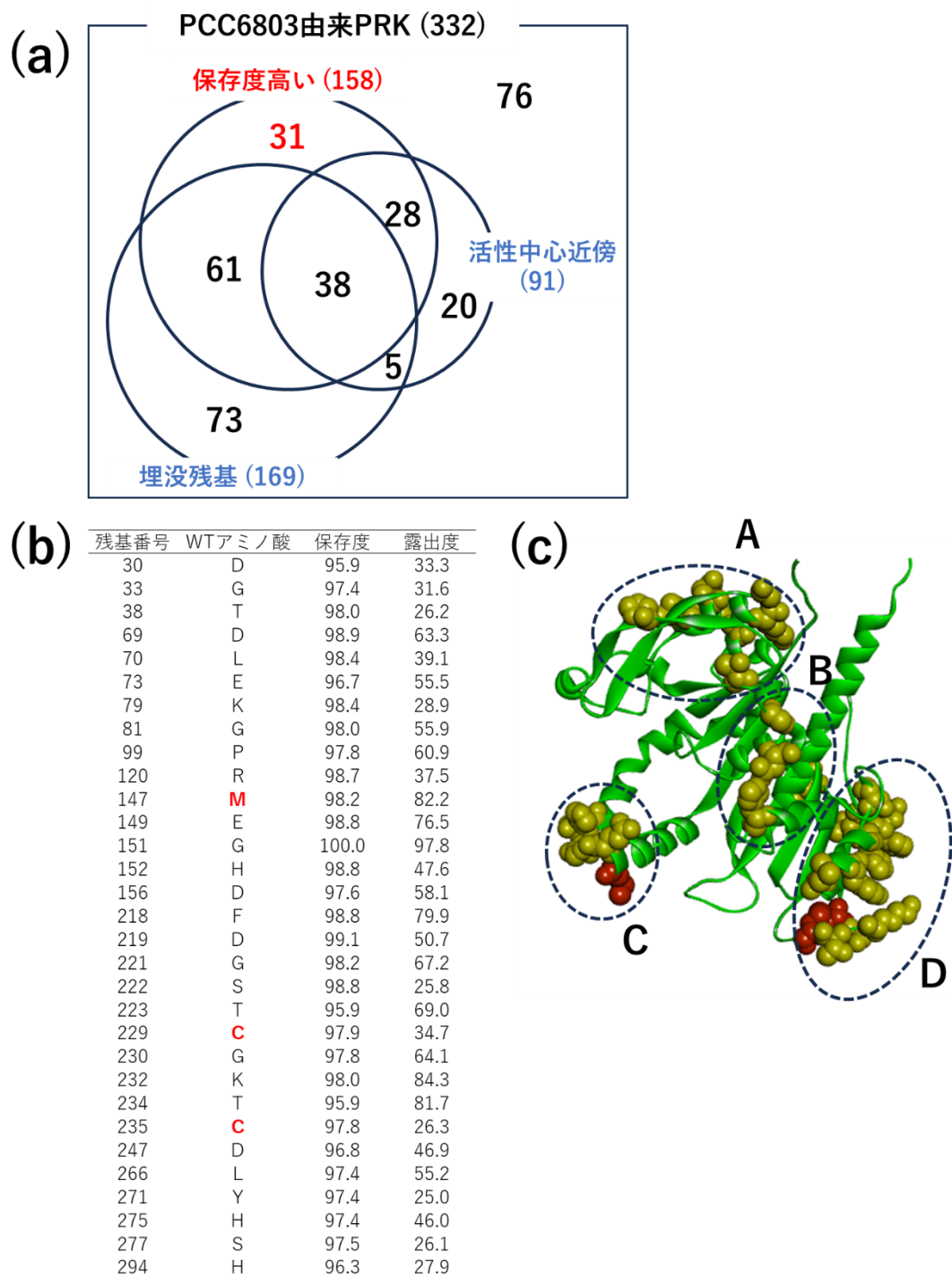


Fig. 2-3-4 抽出した変異候補残基

(a) 保存度の高い残基、埋没残基、活性中心近傍残基の包含関係。赤で示す 31 残基が変異導入候補として抽出された。(b) 抽出した残基一覧。酸化還元制御に寄与することの多い Cys 及び Met を赤で示した。(c) 抽出した残基を CPK で示し、特に Cys 及び Met を赤の CPK で示した。抽出した残基が集中する領域に A-D の名称を付した。

2.3.4 PCC6803 由来 PRK の C 末端ループ変異体設計

Fig. 2-3-4c に示す領域 D は立体構造上で C 末端の位置に近い C 末端ループ領域に位置する。この C 末端ループ上の各アミノ酸残基の保存度を、PCC6803、*Anabaena* sp. PCC 7120、*Synechococcus* sp. strain PCC 6301 由来 PRK 配列のアライメントと共に Fig. 2-3-5 に示した。*Anabaena* sp. PCC 7120 由来 PRK は C 末端ループ上のジスルフィド結合形成で PRK 活性が低下することが確認されている配列 (Fukui *et al.*, 2022)、*Synechococcus* sp. strain PCC 6301 由来 PRK はホモダイマーの立体構造が解かれている配列 (Wilson *et al.*, 2019) であるため、参考として PCC6803 とのアライメントを作製した。

Fig. 2-3-5b によると、C 末端ループ上の各残基の保存度は比較的高く、このループが生体内でも重要な機能を担っていることが示唆された。C 末端ループ上の Cys 残基は *Anabaena* sp. PCC 7120 の PRK において酸化還元制御への関与が *in vitro* で報告されていることに加え (Fukui *et al.*, 2022)、*Synechococcus* sp. strain PCC 6301 の Arg231 がホモダイマー相互作用界面の負電荷キャビティにはまることが PRK の活性抑制に寄与することも示唆されており (Wilson *et al.*, 2019)、先行研究の知見も C 末端ループの制御への重要性を支持している。Colabfold による PCC6803 由来 PRK の 2 量体構造予測によると、*Synechococcus* sp. strain PCC 6301 由来 PRK と同様に C 末端ループ上でのジスルフィド結合と、Arg231 の負電荷キャビティへの挿入が確認された (Fig. 2-3-6)。そこで、保存度の高さを指標として選抜した Cys229 と Cys235 に加え、同じ C 末端ループに位置する Arg231 も変異導入の対象とした。

変異設計の際には、変異後にどのようなアミノ酸残基を導入するかも重要である。変異後のアミノ酸残基には、物理化学的性質の近いアミノ酸残基または、保存度解析で一定の割合で出現するアミノ酸残基が有利であると考えられる。そこで、1 つ目の変異体として PCC6803 由来 PRK の Cys229 と Cys235 を置換した変異体では、Cys と物理化学的性質 (Krigbaum & Komoriya, 1979) の近い Ser への置換を検討することとした (PRK_{CCSS})。 *Anabaena* sp. PCC 7120 の PRK において同様の変異 (Cys230Ser と Cys236Ser) によって酸化条件においても活性が維持されることが報告されており (Fukui *et al.*, 2022)、PCC6803 由来 PRK でも同様の効果が期待できる。そして 2 つ目の Arg231 の変異体については、231 位に 2 番目に多く出現する (3.3%)Thr へと置換した変異体を検討することとした (PRK_{R231T})。 *Synechococcus* sp. strain PCC 6301 において Arg231 が相互作用界面の負電荷キャビティにはまることが PRK の活性抑制に寄与することが示唆されており (Wilson *et al.*, 2019)、Arg の正電荷を失わせ、相互作用を弱くすることで活性抑制が緩和できると期待できる。

(a)

CLUSTAL 2.1 multiple sequence alignment

```

ana7120_PRK      MTTKPERVYVLIQVAGDSGCGKSTFLRRLIDLFGEFMTVICLDDYHSLDRKQKQKGTGITA
syn6301_PRK      -MSKPDYVYVLIQVAGDSGCGKSTFLNRLADLFGTELMTYICLDDYHSLDRKGRKEAGVTA
syn6803_PRK      MTTQLDRYVYVLIQVAGDSGCGKSTFLRRLTDLFGEEFMTVICLDDYHSLDRQGRKAAGVTA
                  :::*****:*****:*****:*****:*****:*****:*****:*****:*****
ana7120_PRK      LDPRANNFDLMYEQIKALKEGQTINKPIYNHETGLIDPPEIVKPNHIVVVEGLHPLYDER
syn6301_PRK      LDPRANNFDLMYEQVYKALKNGETIMKPIYNHETGLIDPPEKIEPNRIIVIEGLHPLYDER
syn6803_PRK      LDPRANNFDLMYEQIKTLKSGQSIMKPIYNHETGLIDPPEKVEPNKYVVEGLHPLYDER
                  *****:*****:*****:*****:*****:*****:*****:*****:*****
ana7120_PRK      VRSLLDFSYYFDISDEVKIAWKIQRDMAERGHRYEDVLAANSRKPDFQKYIEPQREFAD
syn6301_PRK      VRELLDFSYYLDIDDEVKIAWKIQRDMAERGHSEYEDVLAISIEARRPDFKAYIEPQRGHAD
syn6803_PRK      VRELVDYGVYLDISEEVKINWKIQRDMAERGHYEDVLAISINARKPDFTAYIEPQKQYAD
                  *:***:***:***:***:***:***:***:***:***:***:***:***:***:***:***:***
ana7120_PRK      VYLQVLPTNLIKDDTERKVLRYVRLQREGKEGFEPAYLFDEGSTINWTPCGRKLTCSYPG
syn6301_PRK      IYIRYMPQLIPNDTERKVLRYVRLQREGKGFEPAYLFDEGSTIQWTPCGRKLTCSYPG
syn6803_PRK      VYIQVLPTRLIE-DKESKLLRVLYQKEGVKFFEPAYLFDEGSTIDWRPCGRKLTCTYPG
                  :*:***:***:***:***:***:***:***:***:***:***:***:***:***:***:***
ana7120_PRK      MQLYYGSDVYYGRYVSVLEVVGQFONLEEVYIETHLSNTSTKYQGELTQLLLQHREYPG
syn6301_PRK      IRLAYGPDYYGHEVSVLEVVGQFONLEEMIYVEGHLSTDTQYYGELTHLLQHKDYPG
syn6803_PRK      IKMYYGPDNFMGNEVSLLEVVGQFONLEEMIYVENHLSKGTGKYYGEMTELLKHKDYPG
                  :::***:***:***:***:***:***:***:***:***:***:***:***:***:***:***
ana7120_PRK      SNNGTGFFQVLTGLKMRAYERLTKEAKLAVQV
syn6301_PRK      SNNGTGLFQVLTGLKMRAYERLTSQAAPYAASV
syn6803_PRK      TDNGTGLFQVLTGLKMRKYVEQLTAEAKYPAASY-
                  :::***:***:***:***:***:***:***:***:***:***:***:***:***:***:***

```

(b)

Position	228	229	230	231	232	233	234	235	236	237	238
Rank 1	P	C	G	R	K	L	T	C	S	Y	P
Frequency (%)	98.7	97.9	97.8	91.7	98.0	98.0	95.9	97.8	78.1	94.5	98.3
Rank 2	-	-	N	T	-	F	S	N	A	F	S
Frequency	1.2	2.1	0.1	3.3	0.9	1.1	1.4	1.1	12.0	1.6	0.4

Fig. 2-3-5 ラン藻類 PRK のアライメントと C 末端ループ配列の保存度

(a) PCC6803、*Anabaena* sp. PCC 7120、*Synechococcus* sp. strain PCC 6301 由来 PRK 配列のアライメント。本研究で着目する C 末端ループに相当する残基 (PCC6803 配列で 229 位から 238 位に相当する残基) に赤の線を付した。(b) C 末端ループ上の各位置について、最も出現頻度の高い残基と 2 番目に出現頻度が高い残基をその出現頻度と共に示した。本研究で変異の対象として選抜した残基をオレンジ又は緑で示した。

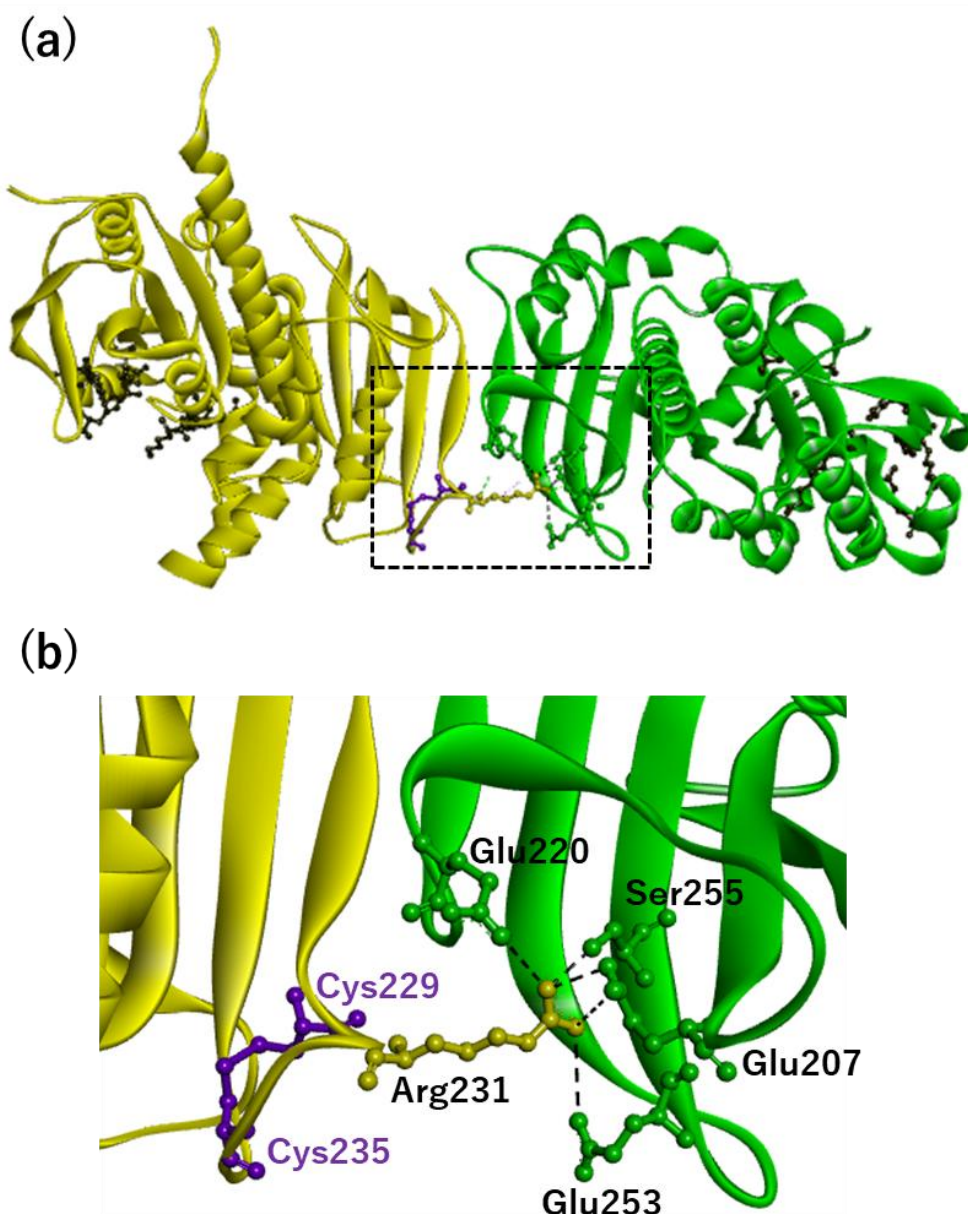


Fig. 2-3-6 ColabFold による PCC6803 由来 PRK のダイマー構造予測結果

(a) Chain A を黄色、Chain B を緑のリボンモデルで示した。ADP 又は RuBP と相互作用すると推定される残基 (Yu *et al.*, 2020) を黒の Ball and stick で示した。ATP 及び Ru5P を基質とする反応においても近い領域が相互作用に関与すると考えられる。(b) Chain A の C 末端側ループ付近の拡大図。本研究で変異の対象とした Cys229, Cys235, Arg231 及び、Arg231 と相互作用する負電荷キャビティを形成する残基を Ball and stick で示し、主要な相互作用を黒の点線で示した。

2.3.5 PRK の C 末端ループ変異体の発現・精製

PCC6803 由来 PRK 野生型 (PRK_{WT}) と 2 つの変異体 (PRK_{CCSS} 及び PRK_{R231T}) を大腸菌で発現・精製し、PRK 活性を維持しているかを *in vitro* での活性測定によって確認する

ことを試みた。野生型及び変異型 PRK を組み込んだ pET28-a(+)ベクターを構築し、タシバク質生産宿主である大腸菌 BL21 (DE3) に導入した。この菌株を IPTG 誘導下で培養することで野生型及び変異型 PRK を生産させ、His-tag 精製とバッファー交換を行った後に SDS-PAGE によって PRK の純度を確認した (Fig. 2-3-7)。5~7 レーン目を見ると、3 種の PRK いずれについても高い純度で精製できていることが分かった。

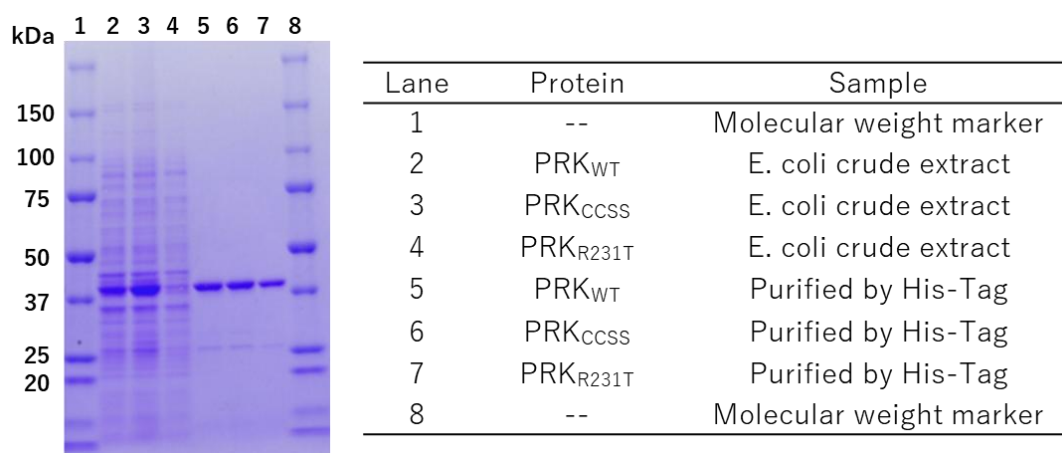


Fig. 2-3-7 組み換え PRK の SDS-PAGE

2.3.6 PRK 変異体の *in vitro* 活性測定

シアノバクテリアの PRK は大腸菌における発現・精製の過程で C 末端のループに存在するシステインがジスルフィド結合を形成することが報告されており (Wilson *et al.*, 2019)、活性測定の前に 60 分間還元処理を行った場合と、行わない場合で活性を比較した (Fig. 2-3-8)。還元処理を行った後の活性測定のデータによると、PRK_{CCSS} は PRK_{WT} と比較して PRK 活性が低下していないことが分かる。Anabaena sp. PCC 7120 由来の PRK では同様の変異で PRK 活性が 75% 程度に低下することが報告されており (Fukui *et al.*, 2022)、PCC6803 由来 PRK とは変異効果が異なることが明らかとなった。また、PRK_{R231T} 変異体では PRK_{WT} と比較して 31% にまで活性が顕著に低下していた。

その一方で、事前の還元処理を行っていない条件を見ると、全ての PRK で事前の還元処理を行った条件と比較して活性が低下しているものの 2 つの変異体は共に PRK_{WT} よりも活性が高くなっていることが分かる。このことから、2 つの変異体は酸化による活性抑制を部分的に解除した変異体であることが示唆された。この結果を受けて、変異 PRK をシアノバクテリア細胞内での発現検討へと進めることとした。

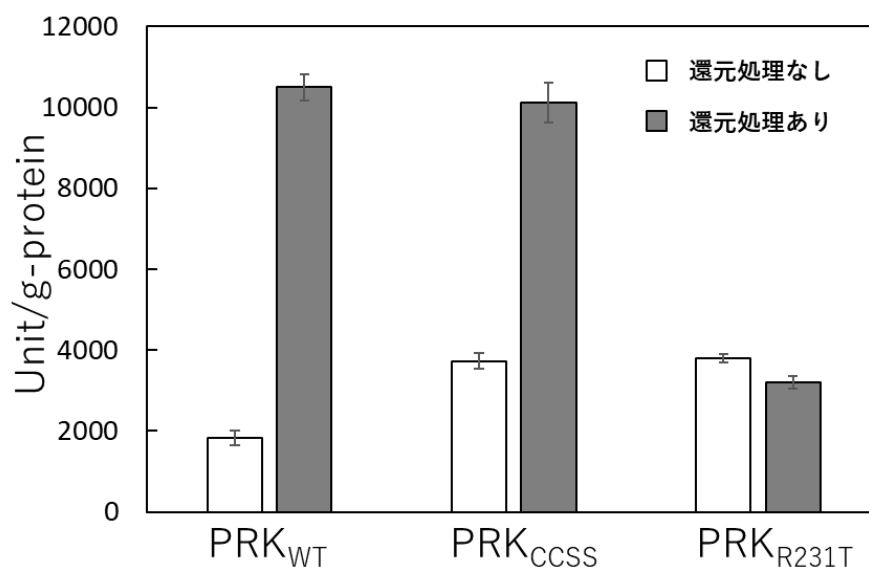


Fig. 2-3-8 野生型及び変異型 PRK の精製酵素アッセイ

終濃度 20 mM の DTT を添加し、60 min の還元反応を行った後に測定する条件と DTT の添加直後に測定する条件で精製酵素の PRK 活性を測定した。すべてのデータは平均値±SD で示した (n=3)。60 min の事前還元処理を行う条件における PRK_{CCSS} と PRK_{WT} の活性には有意な差が見られなかった (両側 t 検定、p 値=0.42)。

2.3.7 変異 PRK 発現株の細胞内 PRK 活性測定 (Photoautotroph)

prk はシアノバクテリアの生育に必須な遺伝子であり、内在性の *prk* 破壊や置換が困難であると予想されたため、ニュートラルサイトである *Slr0168* サイト (Yoshikawa *et al.*, 2015) に *PpsbA2* プロモーター制御下の野生型又は変異型 *prk* 遺伝子を追加導入したシアノバクテリア株を構築し、細胞内の PRK 活性を評価することとした。PCC6803 を親株として PRK_{WT}、PRK_{CCSS}、PRK_{R231T} をそれぞれ追加過剰発現するシアノバクテリア株を構築し、まずは CO₂ のみから生育する条件である Photoautotroph 条件における細胞内 PRK 活性を測定した (Fig. 2-3-9)。その際、事前の還元処理を行わない細胞内の PRK 活性測定に加え、還元処理を行った後の活性測定を行うことで細胞内における PRK の酸化制御の大きさを活性比として得た。

Fig. 2-3-9a によると、PCC6803 と比較して細胞内の PRK 活性が向上した株は見られなかった。このことは、シアノバクテリア細胞内において PRK 活性は高度に抑制制御されていることを示しているといえる。

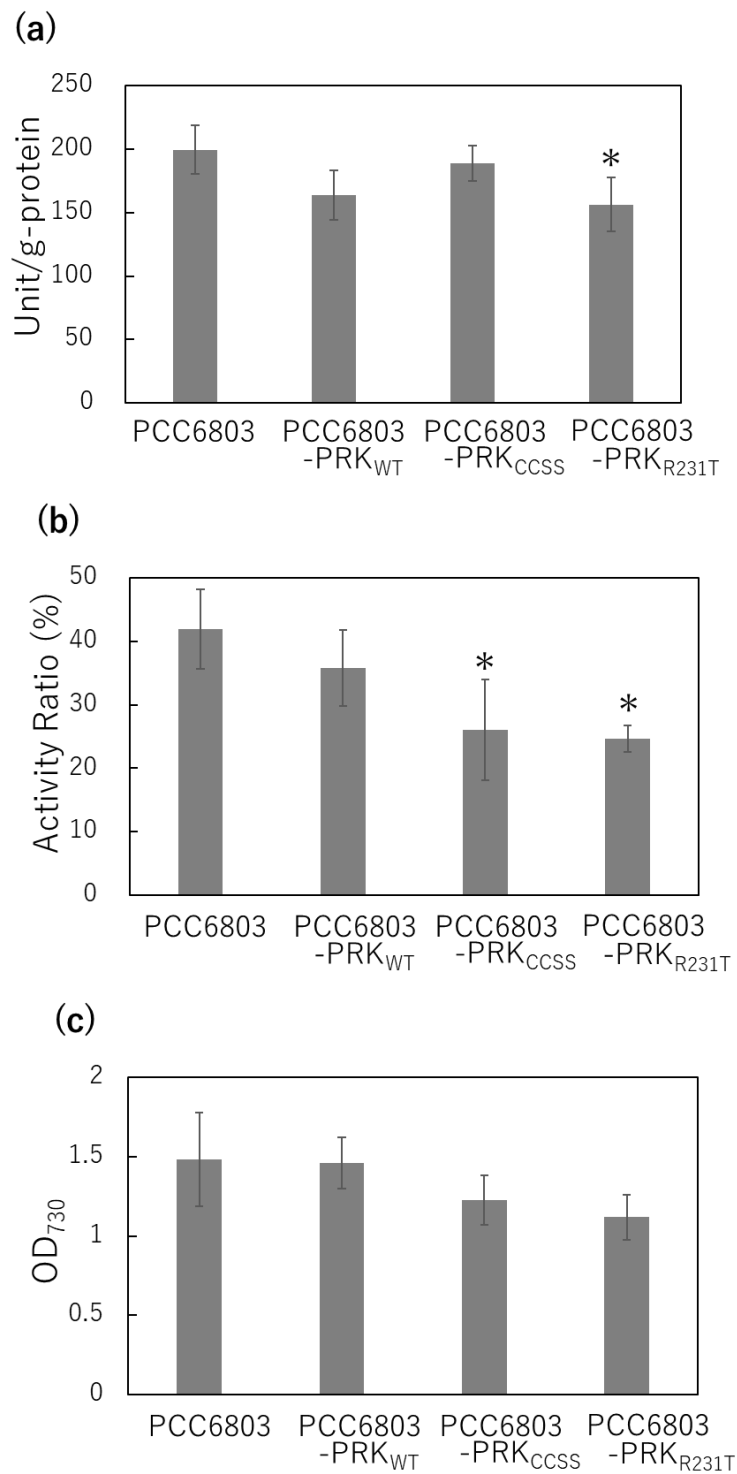


Fig. 2-3-9 PRK 追加過剰発現株の細胞内 PRK 活性 (Photoautotroph)

すべてのデータは平均値±SD で示した (n=3)。PCC6803 に対して $p < 0.05$ (片側 t 検定) のデータについて * を付した。Photoautotroph 条件での培養 48 時間目における (a) 非還元状態のクルードの PRK 活性、(b) 還元状態のクルードの PRK 活性に対する非還元状態のクルードの PRK 活性比、(c) 菌体密度 (OD₇₃₀) をそれぞれ表している。

2.3.8 変異 PRK 発現株の細胞内 PRK 活性測定 (Photomixotroph)

PCC6803 を親株として PRK_{WT}、PRK_{CCSS}、PRK_{R231T} をそれぞれ追加過剰発現するシアノバクテリア株について、CO₂ とグルコースによって生育する Photomixotroph 条件での細胞内 PRK 活性測定も行うこととした。Photomixotroph 条件では NADPH に対する NADP の比率が増加し細胞内が酸化的な条件となることが知られており (Nishiguchi *et al.*, 2019)、細胞内 PRK 活性の向上に C 末端ループ領域の酸化制御を緩和した変異 PRK の発現が有効になると期待できる。栄養条件以外は 2.3.7 章と同一の条件で細胞内の PRK 活性を測定したデータを Fig. 2-3-10 に示した。

Fig. 2-3-10a によると、PCC6803 と比較して PCC6803-PRK_{CCSS} 株で細胞内 PRK 活性を 1.5 倍に向上することに成功した (p 値 = 0.017)。PCC6803-PRK_{CCSS} 株の活性比 (Fig. 2-3-10b) については、PCC6803-PRK_{WT} 株に対して有意に向上し (1.66 倍、p 値 = 0.024)、PCC6803 と比較するとわずかに低下しているものの (0.93 倍、p 値 = 0.013)、PRK 活性の抑制が改善していることが確認された。このことは当初の設計通り、C 末端ループ上への変異導入により、酸化による活性抑制が緩和されたことで細胞内の PRK 活性が向上したことを示している。また、PCC6803-PRK_{R231T} 株では細胞内 PRK 活性は向上しなかった。これは、酸化による活性抑制を部分的に解除した変異体ではあるものの、PRK 活性が低下している (Fig. 2-3-8) ために十分に効果を発現できなかったことによる可能性が考えられる。

Fig. 2-3-10b において PCC6803-PRK_{WT} の活性化割合が大きく低下した要因については詳細は不明であるが、一時的な PRK 活性の増加によってカルビンサイクルが亢進し、NADPH が消費されることで細胞内が酸化的な条件に偏り、活性化割合が低下した可能性などが考えられる。また、Fig. 2-3-9 の独立栄養条件における細胞内 PRK 活性のデータと比較して、混合栄養条件の細胞内 PRK 活性は低下していることが分かる。混合栄養条件では光合成活性が低下することで NADPH に対する NADP の比率が増加し、細胞内が酸化的な条件へと傾くことが報告されており (Nishiguchi *et al.*, 2019)、CP12 による活性抑制や C 末端ループ上のシステイン残基による酸化制御が促進されることで PRK の活性が低下していると考えられる。

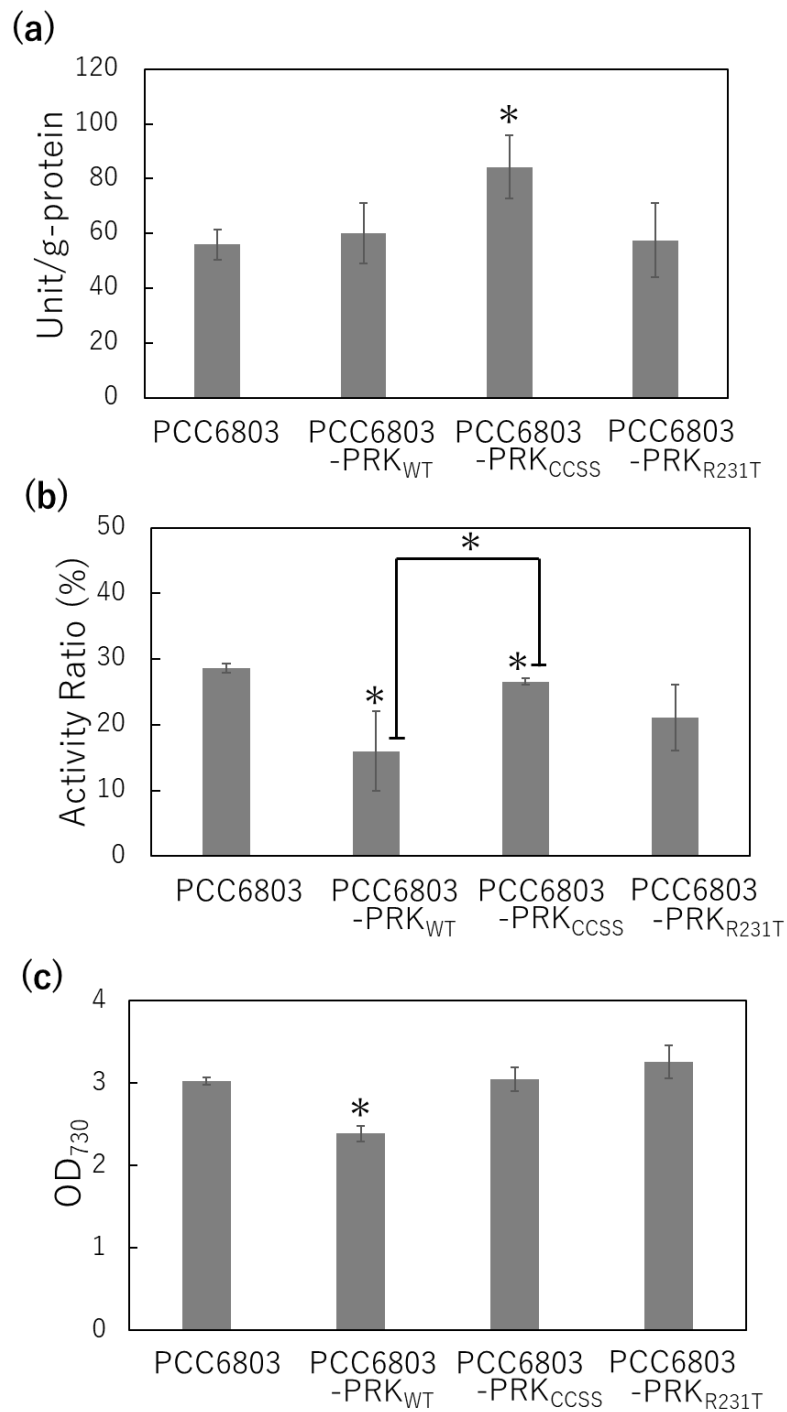


Fig. 2-3-10 PRK 追加過剰発現株の細胞内 PRK 活性 (Photomixotroph)

すべてのデータは平均値±SD で示した (n=3)。PCC6803 に対して $p < 0.05$ (片側 t 検定) のデータについて * を付した。(a) Photomixotroph 条件での培養 48 時間目における非還元状態のクルードの PRK 活性、(b) 還元状態のクルードの PRK 活性に対する非還元状態のクルードの PRK 活性比、(c) 菌体密度 (OD₇₃₀) をそれぞれ表している。

2.3.9 Photomixotroph 条件における変異 PRK 発現株の培養挙動

細胞内 PRK 活性の向上が確認された Photomixotroph 条件において、変異 PRK の追加過剰発現による培養挙動の変化を確認した (Fig. 2-3-11)。継時の培養挙動を見ると、PCC6803-PRK_{WT} 株でのみ増殖の立ち上がりが遅くなり、グルコース消費も遅延していることが分かる。先行研究において、PRK と GAPDH の抑制制御を担う CP12 の遺伝子を破壊した PCC6803 株において Photomixotroph 条件下での生育が低下することが報告されており (Blanc-Garin *et al.*, 2022 ; Lucius *et al.*, 2022)、これと同様に PRK_{WT} の追加過剰発現によって PRK、GAPDH、CP12 のタンパク質量の比率が変化し、制御機構に影響を与えたことで、生育に悪影響を及ぼしたことが示唆される。その一方で、変異 PRK の追加過剰発現株は PCC6803 と培養挙動に差が見られなかった。この要因の詳細は不明であるが、酸化による活性抑制を受けにくい変異 PRK の存在によってカルビンサイクルが駆動し、そこで消費される還元力 (NADPH) を酸化的ペントースリン酸経路を介した NADPH 生成で補うためにグルコース取り込みが促進されている可能性が考えられる。

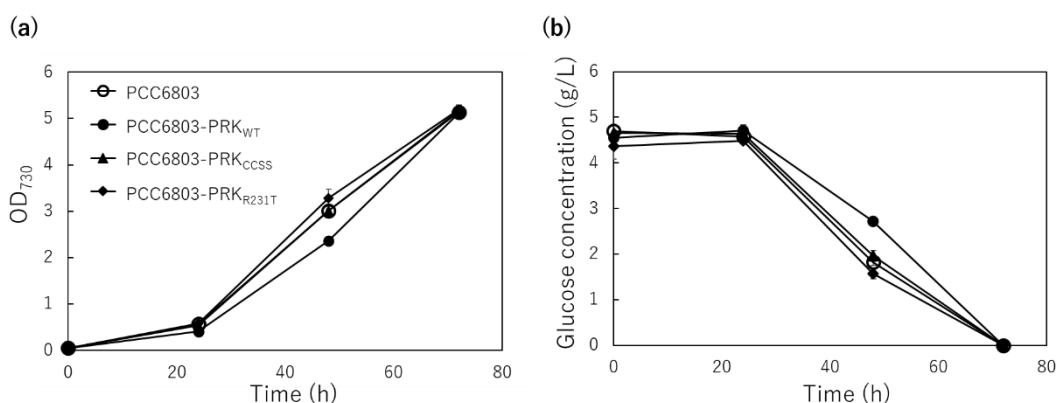


Fig. 2-3-11 Photomixotroph 条件における増殖とグルコース消費挙動

すべてのデータは平均値±SD で示した (n = 3)。(a) 菌体密度 (OD₇₃₀)。 (b) 培養液中のグルコース濃度。

2.3.10 変異 PRK 発現株の代謝状態変化の解析

Photomixotroph 条件における変異 PRK の追加過剰発現がどのような代謝状態の変化につながるか、及び、どのような化合物生産に有用かを確かめるべく、対数増殖期中のシアノバクテリア細胞を回収し、細胞内代謝物の濃度を定量した (Fig. 2-3-11)。本研究で着目した代謝酵素である PRK の基質である Ru5P についてみると、PCC6803 と比較して PCC6803-PRK_{WT} 株において有意に濃度が低下していることが分かる。この株はグルコース消費速度が減少していることから (Fig. 2-3-11)、Ru5P 濃度の減少は酸化的

ペントースリン酸経路を介した基質の供給が少ないことによると考えられる。また、PCC6803-PRK_{CCSS}株でも Ru5P 濃度は減少傾向 (p 値 = 0.064) であった。これは細胞内 PRK 活性が向上したことで基質である Ru5P の消費速度が向上したことによると考えられる。

解糖系上流の代謝物濃度をみると、PRK の追加過剰発現による代謝物濃度の変化は比較的小さく、かつ G1P や F6P といったいくつかの代謝物については濃度が低下していることが分かる。このことから、PRK の追加過剰発現株は解糖系上流の代謝物を前駆体とする化合物の生産宿主としては適していないことが示唆された。その一方で、PRK の追加過剰発現株は解糖系下流や TCA サイクル中の有機酸 (Pyr, Cit, IsoCit, Suc, Fum, Mal) 濃度が増加しており、これらの代謝物を前駆体とする化合物の生産宿主として適していると期待される。特に濃度の増加度合いが大きい代謝物の 1 つである Mal では、PCC6803-PRK_{WT}, PCC6803-PRK_{CCSS}, PCC6803-PRK_{R231T} 株それぞれで親株である PCC6803 株と比較して 1.5, 1.8, 1.7 倍にまで細胞内濃度が向上していた。驚くべきことに、Fig. 2-3-10 に示す PRK 活性測定では活性の増加を検出することが出来なかった PCC6803-PRK_{WT} においても有機酸の濃度は増加していた。細胞内の酵素活性値は時間で動的に変化するものであり、その蓄積として代謝物濃度の変化が起これば考える、PCC6803-PRK_{WT} においても培養開始後のいずれかの時間には細胞内の PRK 活性が増加しているのかもしれない。しかしながら、PRK 活性の増加により、カルビンサイクルが亢進すると、還元力として NADPH が消費されるため、細胞内が酸化的条件に偏り、PCC6803-PRK_{WT} では C 末端ループ上の酸化による活性抑制によって細胞内の PRK 活性がすぐに低下してしまう可能性が考えられる。また、PCC6803-PRK_{CCSS} 株の有機酸濃度は特に高くなっており、PCC6803-PRK_{WT} 株よりも特に有用であると考えられる。エタノールや 2,3-ブタンジオールの前駆体である Pyr については、PCC6803-PRK_{CCSS} 株でのみ PCC6803 株と比較して 1.5 倍の有意な濃度増加が確認された。

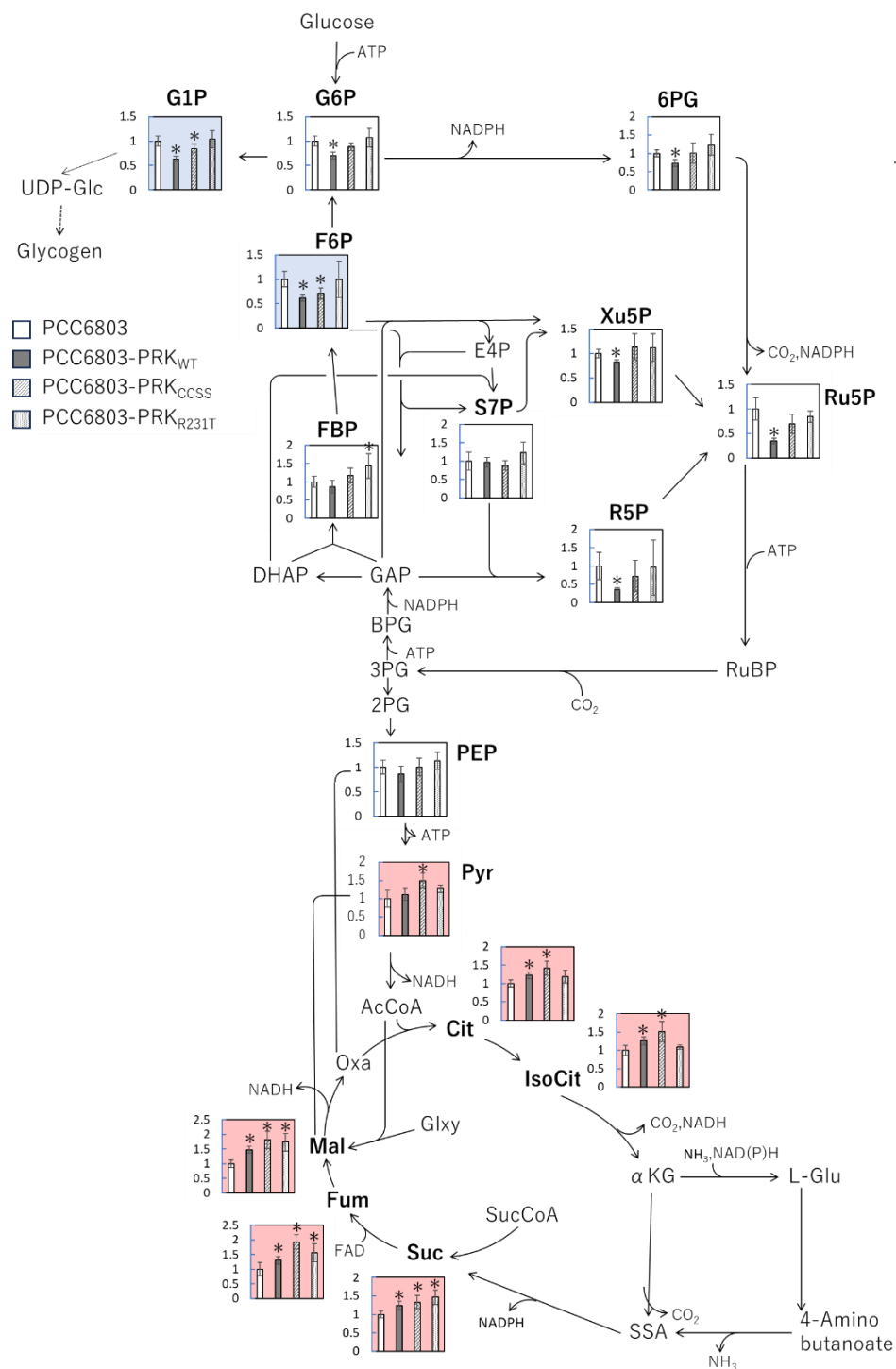


Fig. 2-3-11 Photomixotroph 条件における細胞内代謝物濃度

培養 48 時間目における細胞内代謝物濃度の相対値。すべてのデータは平均値±SD で示した (n=4)。PCC6803 に対して $p < 0.05$ (片側 t 検定) のデータについて * を付した。PCC6803 に対して PCC6803-PRK_{CCSS} で有意に増加した代謝物のグラフを赤で、有意に減少した代謝物のグラフを青で色付けた。

2.4 考察

内在性酵素の合理的改変のための方法論として保存度を活用した変異導入戦略を提案し、その有用性を検証することを目指した。最初にシアノバクテリア由来 PRK の保存度を解析すべく、多様な PRK 配列の系統解析を行った (Fig. 2-3-1)。これによると、データベース上には植物型 PRK だけでなく、多様な古細菌由来 PRK や光合成細菌由来 PRK が登録されていることが明らかとなった。古細菌由来 PRK は V_{max} が植物由来 PRK よりも低いことが知られており (Kono *et al.*, 2017)、シアノバクテリア細胞内で発現を行っても細胞内 PRK 活性向上の期待値は小さいと考えられる。その一方で、光合成細菌由来 PRK の一部は比較的高い V_{max} や基質への親和性を有することが報告されており (Kono *et al.*, 2017)、これらの配列のシアノバクテリア細胞内での発現は細胞内 PRK 活性向上につながる可能性がある。異種タンパク質の発現は細胞毒性や発現性の低さ (Kaur *et al.*, 2018) など考慮すべき要素が多く、かつ PCC6803 は内在性の PRK を有し、その変異体で細胞内 PRK 活性向上の可能性が考えられたため、本研究では光合成細菌由来 PRK の発現は検討しなかったが、今後の研究課題としては非常に興味深いものである。

保存度が高く、細胞内で重要な機能を担っていることが示唆された残基から活性中心付近の残基と、埋没残基を除くことで変異対象として4つの領域が抽出された。そのうち、C末端ループ上の残基は先行研究において制御への関与が示唆されており、本研究の変異対象として選抜された。実際に変異体構築を進めたところ、*Anabaena sp.* PCC 7120 由来 PRK の例ではC末端のシステイン対を Ser 残基に置換することによって PRK 活性が低下することが報告されている (Fukui *et al.*, 2022) のに対し、PCC6803 由来 PRK では同様の変異導入によっても PRK 活性の低下が確認されなかった (Fig. 2-3-8)。このことは PCC6803 由来 PRK がC末端側のループへの変異寛容性が高いことを示しているのかもしれない。同様に保存度が高く、二量体界面で重要な機能を担っていると考えられる Arg231 (Wilson *et al.*, 2019) への変異導入によっても活性は低下したものの PRK 活性を保持できたことは驚きに値する。これは、231 位におけるアミノ酸の出現頻度の解析において、わずかながら出現することを確認した Thr への変異を採用したことで変異が許容された可能性が考えられる。シアノバクテリアの他の種における Arg231 を含む C 末端側ループの変異寛容性については今後の研究が必要である。

野生型及び変異型 PRK を追加過剰発現した株の細胞内 PRK 活性は栄養条件によって異なっていた。本研究での Photoautotroph における PRK 活性値と活性比の測定結果は先行研究 (Blanc-Garin *et al.*, 2022) と近い値を示しており、測定系の妥当性が確認された。Photoautotroph 条件では変異型 PRK の追加過剰発現によっても細胞内 PRK 活性は変化しなかったにも関わらず、Photomixotroph 条件では PCC6803-PRK_{CCSS} 株において野生株と比較して有意な PRK 活性の向上が確認された。Photomixotroph 条件では細胞内の PRK 反応のフラックスが Photoautotroph 条件と比較して低下することが知られて

いる(Nakajima *et al.*, 2014 ; Nakajima *et al.*, 2017)。本研究のデータから、Photomixotroph 条件のような PRK 活性が低い条件では細胞内の主要な PRK 活性制御機構である CP12 との複合体形成による活性抑制 (McFarlane *et al.*, 2019) に加え、C 末端側ループでのジスルフィド結合形成による PRK 活性の抑制制御が働いていることが示唆された。CP12 破壊株では Photomixotroph 条件での生育が困難になることが報告されているため (Blanc-Garin *et al.*, 2022 ; Lucius *et al.*, 2022)、それとは異なる手法で細胞内 PRK 活性の向上が可能となることは有益である。また、Photomixotroph 条件における PCC6803-PRK_{CCSS} 株の PCC6803 に対する PRK 活性向上は 1.5 倍となっていたが、本研究では変異型 PRK の追加過剰発現株においても内在性の野生型 PRK が発現しており、こちらは従来と同様の活性制御を受けることになる。追加過剰発現では無く、内在性の野生型 PRK を変異型 PRK に置換したシアノバクテリア株を作製することでより大きな PRK 活性の向上効果が得られる可能性がある。

PRK を追加過剰発現した株は PCC6803 と比較して細胞内の有機酸濃度が向上していた。有機酸の蓄積は特に、本研究で構築した PCC6803-PRK_{CCSS} 株で顕著であり、例えば細胞内のピルビン酸濃度は野生株と比較して 1.5 倍にまで向上していた。このことは、PCC6803-PRK_{CCSS} 株が有機酸を前駆体とした有用化合物生産に適している可能性がある。また、Metabolic Control Analysis において、PRK の活性が増加すると、解糖系下流に位置する PYK 反応のフラックスは同程度の割合で増加する (PRK 活性増加率の 1.08 倍) ことが予測されており (Nishiguchi *et al.*, 2019)、PCC6803-PRK_{CCSS} 株の野生株に対する PRK 活性増加率とピルビン酸濃度の増加率がおおよそ一致することは先行研究での予測と一致する。また、変異 PRK の追加過剰発現株では PCC6803 株と比較してグルコース取り込み速度が変化していないにもかかわらず有機酸濃度が向上しており、CO₂ やグルコースに由来する炭素を効率よく利用できていると言える。有機酸のうち、ピルビン酸の濃度は変異 PRK を発現する株において親株と比較して 1.5 倍にまで向上していた。先行研究においてピルビン酸供給量の増加がエタノール生産性向上に重要であることが示されており、本研究の成果はエタノールの生産性向上につながることが期待される。また、生産の対象となる化合物としてはエタノールだけでなく、2,3-ブタンジオールやコハク酸なども挙げられる (Knoot *et al.*, 2018)。先行研究において、Photomixotroph 条件で CP12 の破壊や PRK の追加過剰発現を含む複数の遺伝子改変を施した *Synechococcus elongatus* PCC 7942 株によって 2,3-ブタンジオールの生産性を向上した例が報告されており (Kanno *et al.*, 2017)、有機酸を前駆体とした有用化合物の生産において PRK 活性の向上が有効であることが確認されている。また、PCC6803-PRK_{WT} 株では Photomixotroph 条件において増殖が遅延しているのに対し、変異型 PRK を追加過剰発現した株では遅延が見られず、より短期間での有用物質生産が可能になることが期待される。本研究の成果のひとつはシアノバクテリア由来 PRK の C 末端側ループのジスルフィド結合形成による活性抑制の解除がシアノバクテリアの代謝工学に有用であることを示している。さらに、本研究で採用した保存度を指標と

したアミノ酸残基の機能性推定と変異導入は、効率的に変異体設計を行うための手法となり得る。特に、保存度の高い残基は機能に重要である可能性が高く、変異による機能改変の成功確率を高める。本研究に示す PRK の例においては、*in vitro* での先行研究が存在したものの、本手法は先行研究における知見が少ない場面でも適用可能であると考えている。近年では配列データベースの拡充に伴い、様々なタンパク質について保存度の解析が可能になっており、本手法と同様の変異導入戦略が適用可能であると考えられる。

2.5 結言

本研究では内在性酵素の合理的改変のための方法論として保存度を活用した変異導入戦略を提案し、その有用性を検証することを目指した。PRK を対象として保存度が高く、機能に重要な残基に変異を導入したところ、その残基が担う酸化制御を解除して細胞内の PRK 活性を向上することに成功した。この事象は、変異導入に際して保存度の高い残基を避けることの問題点を示しているといえる。保存度の高い残基の中には、触媒活性などの必須な機能を破壊する変異点と、必須ではない機能を破壊する変異点が存在し、構造上の位置や相互作用の有無などを元に変異の可否を慎重に判断すべきであるといえる。本研究で有用な変異体を構築することに成功したアミノ酸残基は立体構造上で活性中心から離れており (Fig. 2-3-6a)、表面に露出していたため、酵素活性を維持しながら酸化制御を解除することに成功した。このように、維持したい機能 (本研究では PRK 活性や酵素の折りたたみ) を有するアミノ酸残基から離れた位置に保存された残基が存在する場合、機能を改変する有用な変異点となる可能性がある。また、変異導入の位置を決定した後には、どのようなアミノ酸残基への置換を行うかを検討する必要がある。変異後のアミノ酸残基には、物理化学的性質の近いアミノ酸残基または、保存度解析で一定の割合で出現するアミノ酸残基が有利であると考えられる。先行研究において、連続進化によって Phosphoenolpyruvate carboxylase (PPC) に I829S または R849S 変異が導入されることで *in vivo* での L-アスパラギン酸によるアロステリック阻害の解除に成功した例が報告されている (Tokuyama *et al.*, 2018)。PPC について各アミノ酸残基の保存度を計算すると (計算方法は付録)、有効な変異点となったアミノ酸残基のうち、849 位は保存度が 100%であり、高度に保存されていることが明らかとなった。このように、保存度が高いアミノ酸残基には細胞内で機能的な制御に関連する残基が含まれている可能性があり、変異点として有効であるといえる。

また、本研究で対象とした PRK ではシアノバクテリアや一部の緑藻類・植物にも保存されているような非常に保存度の高い残基を変異対象として選択したことで、本研究の知見を他の宿主の PRK にも展開できる可能性も高いと考えている。例えば、2.4 項で触れた、複数の遺伝子改変を組み合わせた 2,3-ブタンジオール高生産 *Synechococcus elongatus* PCC 7942 株 (Kanno *et al.*, 2017) に C 末端変異 PRK を過剰発現させることでさ

らに生産性を向上することが可能かもしれない。実際に、*Synechococcus elongatus* PCC 7942 株由来 PRK も PCC6803 と同じ位置にシステイン対を有していた。このように保存度の高いアミノ酸残基を変異の対象とすることは得られた知見を他の生産宿主に応用する際にも利点がある。

本研究ではデータベースから得られた保存度を指標としてアミノ酸残基の機能性を推定し、保存度の高いアミノ酸残基へ変異を導入するという代謝酵素の設計法を示した。代謝酵素はアロステリック制御や酸化還元制御など、細胞内で様々な活性制御を受けている。保存度が高いアミノ酸残基には細胞内で機能的な制御に関連する残基が含まれている可能性があり、変異点として有効であると考えられる。しかしながら、保存度が高いアミノ酸残基には酵素の活性中心や酵素の折りたたみに関連する残基も含まれるため、変異点の慎重な選択が必要になる。その際、活性中心に近接するアミノ酸残基や埋没残基を除くアミノ酸からの選択が有効であると考えられる。実際に、本研究で対象とした PRK では変異点としても活性中心から離れた露出残基を選択した。このように代謝酵素変異体の設計として最初のステップで系統的な保存度の解析によって機能に重要なアミノ酸残基を絞り込み、その中から、維持したい機能に重要な領域 (活性中心など) とは離れた残基を選択し変異を導入することで効果的に機能を改変した酵素を獲得できると考えられる。しかしながら、この変異候補選抜では変異対象残基を 1 つにまで絞り込むことは出来ず、複数の候補が抽出されることになる。PCC6803 由来 PRK の例では、アミノ酸配列全長の約 9% に相当する 31 残基が抽出された。この中から先行研究の知見などを元にして実際に変異対象とするアミノ酸残基を決定することが必要となる。シアノバクテリアや植物に由来する酵素は細胞内で様々な酸化還元制御を受けていることが知られている (Michelet *et al.*, 2013)。こうした酸化還元制御に関与するアミノ酸残基には Cys や Met が多いとされるため (Lim *et al.*, 2019; Klomsiri *et al.*, 2011)、抽出された変異対象に Cys や Met が存在する場合は特に有望な変異候補になりうる。アロステリック部位には疎水性相互作用や特異性を担う荷電残基が重要とされており、これらの残基も変異の対象となり得る (Omage *et al.*, 2024)。本研究で示した変異設計のためのアプローチはシアノバクテリアや植物の物質生産性向上に向けた研究開発に大きく貢献しうると期待される。

本研究のように対象とする機能に関わる酵素が宿主に内在している場合には、その酵素にアミノ酸変異を導入して機能を改良することが可能であり、この方法は宿主の代謝に与える悪影響が比較的小さい点で有利である。その一方で、代謝工学研究においては宿主に目的とする価値の高い物質を生産させるべく、異種の遺伝子を発現させることも多い (Deng and Coleman., 1999 ; Atsumi *et al.*, 2009)。このようにより広い配列空間を探索する必要がある際の配列データベース活用や探索手法については第 3 章で論じる。

第3章 シアノバクテリアのエタノール高生産化に寄与する異種 pyruvate decarboxylase の探索

3.1 緒言

モデルラン藻であるシアノバクテリア *Synechocystis* sp. PCC 6803 (PCC 6803) は、酸素発生型光合成によって CO₂ から有用化合物を生産可能な宿主として期待されている (Lai *et al.*, 2015 ; Kanno *et al.*, 2017)。代謝工学研究においては宿主に目的とする価値の高い物質を生産させるべく、異種の遺伝子を発現させることも多い (Deng and Coleman., 1999 ; Atsumi *et al.*, 2009)。このケースでは、2章に示したタンパク質配列情報を活用した内在性酵素の改変ではなく、配列データベースからの目的物質生産に適した異種酵素の探索が必要となる。しかしながら、データベース上に登録されている酵素配列は膨大であり、その中から有用な酵素を絞り込むことは容易ではない。

シアノバクテリアによるエタノール生産はよく研究されており、これまでに代謝改変、培養条件や遺伝子発現量の最適化などによって生産性を向上した例が複数報告されている (Velmurugan *et al.*, 2019 ; Luan *et al.*, 2015 ; Roussou *et al.*, 2021)。ピルビン酸からアセトアルデヒドを介してエタノールを生産する生合成経路において、PCC6803 はアセトアルデヒドをエタノールに変換可能な内在性の alcohol dehydrogenase (EC 1.1.1.1) を有しており、内在性の遺伝子であることに加え、活性が高いことから、多くの研究では同遺伝子が利用されている (Velmurugan *et al.*, 2019 ; Luan *et al.*, 2015 ; Roussou *et al.*, 2021)。その一方で、PCC6803 はピルビン酸を脱炭酸し、アセトアルデヒドに変換する酵素である pyruvate decarboxylase (PDC、EC 4.1.1.1) を持たないため、ピルビン酸からエタノールを生産するためには異種の PDC を細胞内で発現させる必要がある。エタノール生産の先行研究の多くは *Zymomonas mobilis* (Zm) 由来の同一の PDC 配列を利用している (Luan *et al.*, 2015; Roussou *et al.*, 2021 ; Choi *et al.*, 2016)。一部の研究では *Saccharomyces cerevisiae* (*S. cerevisiae*) 由来や、 K_M 値が Zm PDC よりも小さいことが知られている *Zymobacter palmae* 由来 PDC も利用されているものの (Velmurugan *et al.*, 2019 ; Quinn *et al.*, 2019)、エタノールの生産性を向上する PDC を系統的に探索した例は報告されていない。他の有用物質生産において、生合成経路の酵素をより最適な酵素に置き換えることで生産性を向上した例が報告されており (Miao *et al.*, 2017 ; Ethan *et al.*, 2013)、エタノール生産においても様々な PDC 配列を探索・評価することで、より生産に適した酵素が得られる可能性がある。

PDC は植物や真菌に広く見られ (König *et al.*, 1998)、InterPro の pyruvate decarboxylase/indolepyruvate decarboxylase-like (IPR012110) ファミリーに分類されている配列だけでも 1 万を超えている (Paysan-Lafosse *et al.*, 2023)。一方で、Zm PDC のようなバクテリア由来の PDC は比較的まれであり、これらの PDC は至適 pH が酸性側で、至適温度も 50 度以上の高温側にあることが多く (Van *et al.*, 2014)、細胞内で 4 量体を形成している (Pohl., 1997)。また、データベース上に pyruvate decarboxylase として登録され

ている酵素以外にも、indolepyruvate decarboxylase (IPDC) や branched-chain keto acid decarboxylase (KDC) についても Pyruvate を基質として反応し Acetaldehyde を生成することがあり (Yep *et al.*, 2006 ; Schütz *et al.*, 2003)、これらも含めると利用可能な酵素の多様性は増大する。

さらに、近縁種の配列情報を元に機能を向上した非天然の酵素変異体を設計する protein repair one-stop shop (PROSS) (Goldenzweig *et al.*, 2016 ; Peleg *et al.*, 2021) や ancestral sequence reconstruction (ASR) (Spence *et al.*, 2021 ; Foley *et al.*, 2022) といった有用酵素の探索手法も利用可能である。PROSS は、立体構造と近縁種に由来する multiple sequence alignment (MSA) を元に複数の安定化変異を導入した変異体を提案するアルゴリズムであり (Goldenzweig *et al.*, 2016)、この手法で設計された変異体は耐熱性の向上と異種での生産性が改善することが報告されている (Peleg *et al.*, 2021)。ASR は系統解析の結果を元に、系統樹のノードに相当する祖先型配列を推定する手法であり (Spence *et al.*, 2021)、当初は酵素の進化解析のためのツールであったが (Gaucher *et al.*, 2003)、本手法によって設計される酵素が高い耐熱性や耐酸性、基質特異性の広さを有するという有用な形質が多いため、有用酵素設計の手法として用いられることも多い (Spence *et al.*, 2021)。これらの手法の活用により、現存する配列が少ない配列空間についても効率的に有用酵素探索が可能となる。

本研究では、多様なタンパク質配列空間を 2 段階に分けて探索することで目的に合った酵素を同定することを目指した。具体的には、データベースに存在する多様な酵素配列を探索し、現在最もよく用いられる *Zm* PDC よりもシアノバクテリアで発現した際のエタノール生産性が向上する PDC を獲得することを試みた。生産性向上のためにはシアノバクテリア細胞内で発現し、活性を持つことが必要であり、本研究ではこれらを包括した上で有用な配列を探索した。第 1 段階では、配列的に多様な PDC 配列を複数評価し、どのような配列がシアノバクテリアのエタノール生産にとって有用であるかを特徴づけすることを試みた。この探索は先行研究においてよく利用される *Zm* PDC よりも好ましい配列群が存在する可能性を考慮して実施した。続く 2 段階目ではシアノバクテリアのエタノール生産に適していることが確認されたバクテリア由来 PDC について、密に配列評価を行った。この評価では、バクテリア由来 PDC クレードの中からシアノバクテリアでのエタノール生産性向上に有用な PDC 配列を同定することを目指した。このような 2 段階の探索によってエタノールの生産性が向上する PDC を獲得することに成功した (Fig. 3-1-1)。本研究で実施した広い配列空間の探索・評価と、絞り込んだ配列空間での非天然配列を含めた探索・評価というステップワイズな戦略は、様々な有用酵素の選抜や利用可能な酵素配列の拡張に有用であると考えられる。

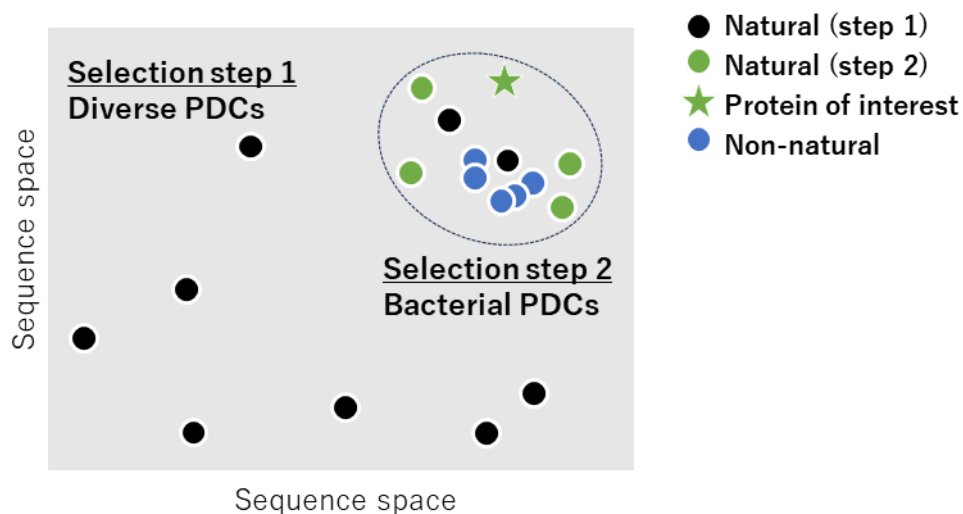


Fig. 3-1-1 シアノバクテリアに適した PDC 配列の探索概念図

2段階での PDC 探索を行った。1段階目では黒のプロットで示すように多様な天然 PDC 配列の評価を行い、バクテリア由来 PDC がシアノバクテリアでのエタノール生産に適していることを明らかにした。2段階目ではバクテリア由来 PDC の周辺配列に絞り、緑で示す天然配列と青で示す人工的に設計した非天然配列を評価し、目的の配列を獲得することに成功した。

3.2 実験方法

3.2.1 配列的に多様な PDC 配列の系統解析 (第 1 段階：広い探索)

シアノバクテリアのエタノール生産に適した PDC 同定の第 1 段階として、配列的に多様な PDC 配列を発現候補として選抜し、エタノールの生産性を検証することとした。配列的に多様な PDC を選抜するために必要な解析として、多様な PDC 配列を収集し、系統解析を試みた。多様な PDC 配列の系統解析の際には、BRENDA (Jeske *et al.*, 2019)、Cyanobase (Nakamura *et al.*, 2000)、InterPro (Paysan-Lafosse *et al.*, 2023) から配列情報を収集した。系統解析の大まかな流れを Fig. 3-2-1 に示した。BRENDA からは EC 4.1.1.1 (pyruvate decarboxylase) として情報が登録されている酵素のうち、高い k_{cat}/K_M value や小さい K_M 値といった好ましい性質を有している酵素の配列情報を抽出した (計 11 配列)。Cyanobase からは、pyruvate decarboxylase 又は indole pyruvate decarboxylase としてアノテーションされている計 6 配列を取得した。InterPro からは Entry IPR012110 (Pyruvate decarboxylase/indolepyruvate decarboxylase-like) に属する配列のうち、Swiss-Prot にもエントリーされているレビュー済みの 50 配列を取得した。得られた計 67 配列から、配列長の中央値よりも 20%以上長い又は短い配列を削除し (4 配列が除去された)、Cd-hit (Li *et al.*, 2006) を用いて閾値 99% でクラスタリングを行うことで同一配列を削除した (計 47 配列となった)。MSA は MAFFT (Kato *et al.*, 2013) を用いて作成し、Trimal (Capella-Gutiérrez *et al.*, 2009) の Gappyout メソッドで保存領域を抽出したのちに IQTree v2 を用いて最尤法により系統樹を推定した (Minh *et al.*, 2020)。進化モデルには ModelFinder を用いたベイズ情報量基準スコアによって LG+R5 を選択した (Kalyaanamoorthy *et al.*, 2017)。推定されたクレードの信頼性は、UFBoot2 (Hoang *et al.*, 2018) を用いて 1000 回の反復を行い、ブートストラップ法で評価した。

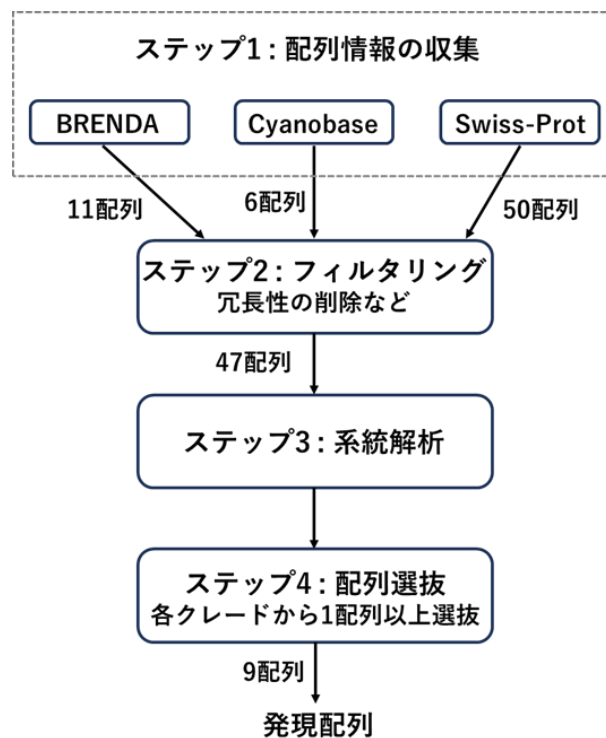


Fig. 3-2-1 多様な PDC 配列の系統解析と配列選抜フロー

3.2.2 バクテリア由来 PDC 配列の系統解析 (第 2 段階：密な探索)

PDC 同定の第 2 段階として、第 1 段階の探索で有望とされたバクテリア由来 PDC クレードをより詳細に探索することとした。バクテリア由来の PDC 配列の系統解析の際には、UniProtKB reference proteomes + Swiss-Prot (Apweiler *et al.*, 2004) をデータベースとして *Zm* PDC をクエリ配列にした検索を行い、E-value 上位 50 配列を取得した。*Zm* PDC よりも 20%以上長い又は短い配列が含まれないことを確認したのち、閾値 0.9 で CD-hit によってクラスタリングを行うことで 47 配列を得た。ここに、真核生物由来 PDC である *Zea mays* 及び *Arabidopsis thaliana* 由来 PDC の配列 (外群として利用する配列) と、CD-hit によるクラスタリングで除かれた配列のうち、先行研究 (Van *et al.*, 2014) において中性付近の pH でも比較的高い PDC 活性を有することが報告されている *Acetobacter pasteurianus* 由来 PDC の配列を加え、計 50 配列とした。MSA は MAFFT を用いて作成し、Trimal の Gappyout メソッドで保存領域を抽出したのちに IQTree v2 を用いて最尤法により系統樹を推定した。進化モデルには ModelFinder を用いたベイズ情報量基準スコアによって LG+R5 を選択した。推定されたクレードの信頼性は、UFBoot2 を用いて 1000 回の反復を行い、ブートストラップ法で評価した。真核生物に由来する 3 配列を Re-root し、outgroup とすることで最終的な系統樹を得た。

3.2.3 PROSS による *Zm* PDC 変異体の設計

本研究では、Jonathan らによって提供されている PROSS2 server (Weinstein *et al.*, 2021) を用いて変異体のデザインを行った。PROSS では入力配列のホモログのアライメントを元に、許容される変異が特定される。それらの変異のうち、エネルギー計算によって安定化変異と予測された複数の変異を導入した多重変異体が提案される (Goldenzweig *et al.*, 2016)(Fig. 3-2-2)。デフォルト設定では、安定化変異と定義するエネルギー変化の閾値が異なる 8 つの変異体配列が出力される。安定化変異と定義するエネルギー変化の閾値が厳しい変異体ほど少数の変異しか導入されないこととなる。

変異体設計の制約条件として、*Zm* 由来 PDC のホモ 4 量体立体構造である PDB ID : 2wva (Pei *et al.*, 2010) の V 鎖を入力構造とし、変異を導入しない固定残基として基質や補因子など (TPU 600 V, PYR 602 Z, MG 601 V, TPU 600 Z, PYR 602 V) の周辺とホモテトラマー形成時の相互作用領域を選択した。許容される変異を推定するための MSA には PROSS2 server のデフォルト設定で得られるものまたは、*Zm* PDC 近傍のクレードの配列から作成される MSA を使用した。デフォルト設定で得られた MSA を用いた配列デザインでは、*Zm* PDC に対する配列相同性が 94.8, 93.3, 92.9, 91.7, 91.0, 89.8, 89.6, 87.4% の配列が設計され、そのうち 94.8% と 87.4% の配列相同性を有する配列をそれぞれ PROSS 1、PROSS 2 と名称付け、シアノバクテリアでのエタノール生産検討に用いた。また、*Zm* PDC 近傍のクレードの配列から作成される MSA を用いた配列デザインでは、配列相同性が 96.4, 95.7, 95.3, 95.1, 95.1, 95.3, 94.0, 93.3, 91.7, 84.0% の配列が設計され、そのうち 96.4% の配列相同性を有する配列を PROSS 3 と名称付け、シアノバクテリアでのエタノール生産検討に用いた。

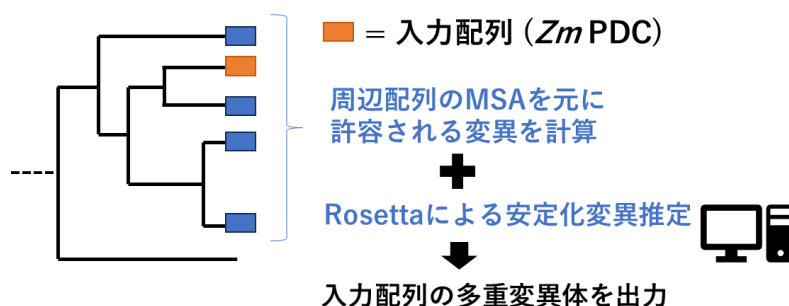


Fig. 3-2-2 PROSS による変異体設計の概要

枝の長さを考慮しない系統樹の模式図。変異を導入したい入力配列をオレンジの四角で、その周辺のホモログ配列を青の四角で示した。

3.2.4 ASR によるバクテリア由来 PDC の祖先型配列推定

バクテリア由来 PDC の祖先型配列の設計には GRASP server を使用した (Foley *et al.*, 2022)。祖先型配列設計は系統解析の結果を元にした配列比較から系統樹の分岐に相当

する祖先型配列を推定する手法である (Fig. 3-2-2)。3.2.2 項で得られた系統樹と、系統樹作成の過程で作成したアライメントを入力とし、進化モデルには LG モデルを指定して設計を行った。推定された祖先型配列のうち、*Zm* PDC の祖先型配列に相当する配列を選定し、発現検討へと進めた。

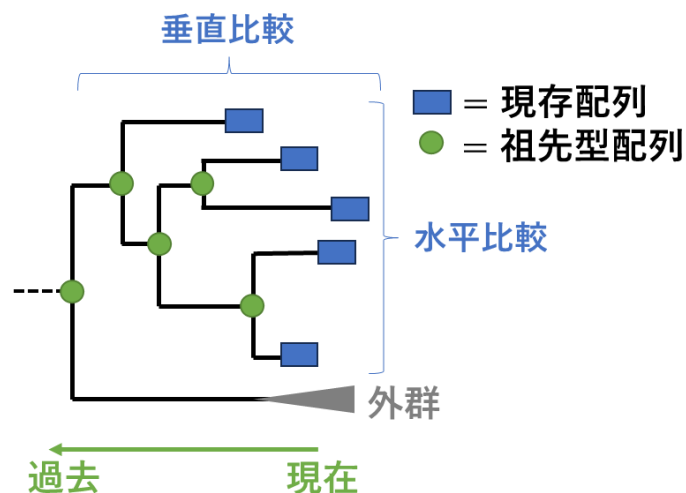


Fig. 3-2-3 ASR による変異体設計の概要

枝の長さを考慮した系統樹の模式図。現存する天然のタンパク質配列を青の四角で、推定の対象となる祖先型配列を緑のまるで示した。系統樹の樹形を元に、天然配列間の垂直比較と水平比較によって祖先型配列が推定される。

3.2.5 使用菌株

本研究では、*Synechocystis* sp. PCC 6803 グルコース耐性 (GT) 株 (PCC 6803) と、新たに構築したエタノール生産株を検討に用いた (Table 3-2-1、株の構築については 3.2.9 に記載)。

Table 3-2-1 使用菌株

Strain	Genotype	reference
PCC 6803	<i>Synechocystis</i> sp. PCC 6803 glucose-tolerant	Williams., 1988
Et-Zm	PCC6803 + <i>Slr0168::PnblA-pdc(Zymomonas mobilis)-adhII-Smr</i>	This work
Et-Gd	PCC6803 + <i>Slr0168::PnblA-pdc(Gluconacetobacter diazotrophicus)-adhII-Smr</i>	This work
Et-Zmays	PCC6803 + <i>Slr0168::PnblA-pdc(Zea mays)-adhII-Smr</i>	This work
Et-At	PCC6803 + <i>Slr0168::PnblA-pdc(Arabidopsis thaliana)-adhII-Smr</i>	This work
Et-Cs	PCC6803 + <i>Slr0168::PnblA-pdc(Crocospaera subtropica)-adhII-Smr</i>	This work
Et-Ca	PCC6803 + <i>Slr0168::PnblA-pdc(Clostridium acetobutylicum)-adhII-Smr</i>	This work
Et-syn7803	PCC6803 + <i>Slr0168::PnblA-pdc(Synechococcus</i> sp. (strain WH7803)) <i>-adhII-Smr</i>	This work
Et-Ao	PCC6803 + <i>Slr0168::PnblA-pdc(Aspergillus oryzae)-adhII-Smr</i>	This work
Et-Sc	PCC6803 + <i>Slr0168::PnblA-pdc(Saccharomyces cerevisiae)-adhII-Smr</i>	This work
Et-Go	PCC6803 + <i>Slr0168::PnblA-pdc(Gluconobacter oxydans)-adhII-Smr</i>	This work
Et-AP	PCC6803 + <i>Slr0168::PnblA-pdc(Acetobacter pasteurianus)-adhII-Smr</i>	This work
Et-Zp	PCC6803 + <i>Slr0168::PnblA-pdc(Zymobacter palmae)-adhII-Smr</i>	This work
Et-Ar	PCC6803 + <i>Slr0168::PnblA-pdc(Ancylobacter rudongensis)-adhII-Smr</i>	This work
Et-Av	PCC6803 + <i>Slr0168::PnblA-pdc(Acetobacter vaccinii)-adhII-Smr</i>	This work
Et-ASR1	PCC6803 + <i>Slr0168::PnblA-pdc(ASR1)-adhII-Smr</i>	This work
Et-ASR2	PCC6803 + <i>Slr0168::PnblA-pdc(ASR2)-adhII-Smr</i>	This work
Et-PROSS1	PCC6803 + <i>Slr0168::PnblA-pdc(PROSS1)-adhII-Smr</i>	This work
Et-PROSS2	PCC6803 + <i>Slr0168::PnblA-pdc(PROSS2)-adhII-Smr</i>	This work
Et-PROSS3	PCC6803 + <i>Slr0168::PnblA-pdc(PROSS3)-adhII-Smr</i>	This work

slr0168 はニュートラルサイトとして知られている。ストレプトマイシン耐性遺伝子を *Smr* と記載した。

3.2.6 培地組成

2.2.6 項参照

3.2.7 培養条件

本章ではエタノール生産性の評価を Photomixotroph 条件で行った。

前培養

20 mL の 25 mM NaHCO₃ 及び 50 mM グルコースを含む BG11 培地入り 100 mL 容三角フラスコにシアノバクテリアを植菌し、30 °C、150 rpm、30 μ mol/m² sec で 96~120

時間培養した。遺伝子組換え細胞の選択マーカーとして必要に応じて終濃度 20 μ g/mL となるようにストレプトマイシンを添加した。

本培養

20 mL の 25 mM NaHCO₃ 及び 50 mM グルコースを含む BG11 培地入り 100 mL 容三角フラスコに、前培養から初期 OD₇₃₀ = 0.05 となるようにシアノバクテリアを植菌し、30 °C、150 rpm、30 μ mol/m² sec で 120 時間培養した。

3.2.8 菌体濃度及び培地中グルコース濃度の測定

菌体濃度分析

分光光度計 (UVmini-1240 (島津製作所)) を用いて培養液の 730 nm の吸光度 (OD₇₃₀) を測定した。培養液の濃度が高い場合は必要に応じて希釈して測定した。

培地中グルコース濃度の測定

適宜培養液を 1.0 mL 採取し、遠心分離した (15,000 rpm, 10 min, 4°C) 上清を分析サンプルとした。培地中のグルコース濃度は Glucose Assay Kit-WST (Dojindo Laboratories, Kumamoto, Japan) を用いてプロトコル通りに測定した。

3.2.9 エタノール生合成経路発現用 DNA 断片の構築

シアノバクテリアのニュートラルサイトである *slr0168* に *nblAp* プロモーターの制御下に *pdc* 及び *adhII* 遺伝子を連結したもの、およびストレプトマイシン耐性遺伝子を組み込んだ株を構築するための DNA 断片を構築した (Fig. 3-2-3)。 *nblAp* プロモーターは先行研究においてエタノール生産関連遺伝子を発現させた実績のあるプロモーターの一つである (Yoshikawa *et al.*, 2015)。本研究で用いたプライマーの配列を Table 3-2-2 に示した。PDC の DNA 断片は Invitrogen (Thermo Fisher Scientific, Waltham, MA) で人工合成し、その配列を Table S1 に示した。PDC の DNA 配列はすべてコドン最適化ツール (Integrated DNA Technologies, Coralville, USA) を利用して PCC6803 に最適化している。

すべての遺伝子操作は先行研究 (Yoshikawa *et al.*, 2015; Nishiguchi *et al.*, 2019) に記載された方法に従って行った。先行研究で構築された Et-g + Slr0168::PpsbA2-prk-Smr (Nishiguchi *et al.*, 2019) のゲノムを鋳型として、Primer 1 と 2、Primer 3 と 4 でそれぞれ PCR を行い、ニュートラルサイトである Slr0168 サイト (Yoshikawa *et al.*, 2015) に目的の遺伝子と、ストレプトマイシン耐性遺伝子を組み込むための DNA 断片を得た。PCC6803 株のゲノムを鋳型として、Primer 5 と 6、Primer 7 と 8 でそれぞれ PCR を行い、PCC6803 内在性の *adhII* 遺伝子と *rbc* 遺伝子のターミネーターを得た。さらに、同じゲノムを鋳型として Primer 9 と 10 で PCR を行い、*PnblA* プロモーターを得た。これらの PCR 産物と PCC6803 のコドンに最適化して人工合成した *Zm* 由来 *pdc* 遺伝子を Primer 11 と 12 で増幅したものを混合し、Primer 1 と 4 による SOE-PCR によって結合

することで *Slr0168* サイトにストレプトマイシンを選択マーカーとして *PnblA* プロモーターがドライブする *Zm pdc-adhII* オペロンを組み込むための DNA 断片を得た。

上記で得られた DNA 断片を用い、3.2.10 項に記載の方法でシアノバクテリアを形質転換することで Et-Zm 株を得た。Et-Zm 株のゲノムを鋳型として Primer 1 と 10、4 と 5 でそれぞれ PCR を行い、*pdc* 遺伝子のみを変更するための DNA 断片を得た。これらの DNA 断片と、PCC6803 のコドンに最適化して人工合成した様々な *pdc* 遺伝子を鋳型として Primer 12 と Primer 13-30 の組み合わせで PCR を行うことで得られた DNA 断片を混合し、Primer 1 と 4 による SOE-PCR によって結合することで多様な *pdc* 遺伝子を含むエタノール生合成経路組み込み用の DNA 断片を得た。

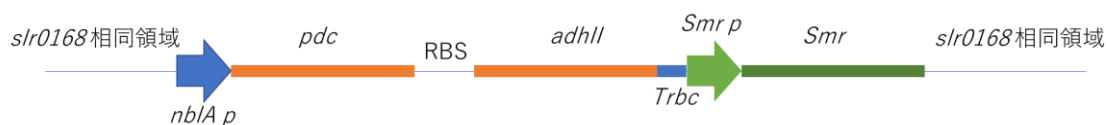


Fig. 3-2-3 エタノール生合成経路発現用 DNA 断片

RBS は ribosome binding site を表し、先行研究 (Heidorn *et al.*, 2011) で設計された tactagagta gtggagggtta ctag という配列を用いた。*Trbc* は *rbc* に由来するターミネーター配列、*Smr p* はストレプトマイシン耐性遺伝子発現用のプロモーターを表している。

Table 3-2-2 使用プライマー

Number	Name	Sequence
1	slr0168_up_F	CCAGCACAAATTATTGGGGTCAG
2	slr0168_up_nblA_R	CGAGGAGGAGGAGGCTAAACTGTTCTTCCCATATAACCATCAAA GCCATAGTTGG
3	Sm_F	GTAGCGACCGAGTGAGCTAG
4	Slr0168_down_R	GTGGTTAAATCCGTGGCTTGG
5	adhII_F	ATGATTAAAGCCTACGCTGC
6	adhII_R	CTAATTTTTACTATGGCTGAGCACTACC
7	Trbc_F	GGTAGTGCTCAGCCATAGTAAAAATTAGtaaGTTACAGTTTTGGC AATTACTAAAAAACTGAC
8	Trbc_R	CAAATAGCTAGCTCACTCGGTCTGCTACTTCCCCACTTAGATAAA AAATCCGGTCAG
9	PnblA_F	GAACAGTTTAGCCTCCTCCTCC
10	PnblA_R	AGCTGTTGCCCTCCAAGGCGACTAC
11	pdc_Zm_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGAGTTATACAGTTGGG ACG
12	pdc_R	GCTTCCAGGGCAGCGTAGGCTTTAATCATctagtaacctccactactctag
13	pdc_Gd_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGACCTACACGGTTGGT CG
14	pdc_Zmays_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGGAAACACTGCTCGC GGGTAATCCAGC
15	pdc_At_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGGACACGAAAATTGG GTCC
16	pdc-Cs_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGGAAAATAATAAGAGT ATTGGCAAC
17	pdc_Ca_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGAAGAGTGAGTATACC ATTGGACG
18	pdc_syn7803_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGCCCCCTAGCGTAGTT ACATACG
19	pdc_Ao_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGGCGACAGATATCGCT ACTCG
20	pdc_Sc_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGAGTGAGATTACACTG GGTAAG
21	pdc_Go_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGACTTATACCGTAGGC CAC
22	pdc_Ap_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGACGTACACAGTTGG CATG
23	pdc_Zp_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGTATACGGTCGGCATG TACC
24	pdc_Ar_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGACATATACAGTGGGT AGC
25	pdc_Av_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGACTTACACCGTCGGC AC
26	pdc_ASRI_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGACCTATACTGTCTGGG AC
27	pdc_ASRI_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGACTTATACGGTAGGC ACG
28	pdc_PROSS1_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGTCTTATACCGTGGGA ACG
29	pdc_PROSS2_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGTCCTATACGGTGGGG AC
30	pdc_PROSS3_F	GGTAGTCGCCTTGGAGGGCAACAGCTATGTCTTACACTGTAGGA ATG

3.2.10 シアノバクテリアの形質転換

2.2.10 項参照。

3.2.11 シアノバクテリアの粗酵素液調製

培養開始 96 時間目の培養液から遠心分離（10 分、3000 g、室温）によって菌体を回収し、プロテアーゼ阻害剤（EDTA フリー, Roche, Switzerland, 1 錠/25 mL）を含む 20 mM Tris-HCl 緩衝液（pH 7.0）に再懸濁した。この懸濁液を Bioruptor（Sonicbio, Knanagawa, Japan）を用いて 10 回超音波処理した（30 秒の超音波処理、処理間隔は 30 秒）。細胞抽出物を遠心分離し、上清を使用まで -80°C で保存した。タンパク質濃度は Bradford Protein Assay Kit (Takara Bio Inc., Shiga, JapanTakara) を用いて定量した。

3.2.12 PDC 活性測定

PDC の酵素活性測定は以下の反応式に示すように、PDC 反応と ADHII 反応をカップリングすることによって行った (Dörte *et al.*, 2009)。ADHII には *Saccharomyces cerevisiae* 由来の酵素を使用した (Sigma Aldrich, MO, USA)。PDC 反応によって生成した Acetaldehyde が ADHII 反応を進行させ、NADH を NAD へと酸化する。PDC の活性は NADH の酸化による 340 nm の吸光度の変化として検出することができる。PDC 活性測定溶液を Table 3-2-3 の組成で調製した。タンパク質濃度が 60 mg/L となるようにシアノバクテリア粗酵素液を 20 mM Tris-HCl 緩衝液（pH 7.0）で希釈し、希釈サンプル 10 μ L と PDC 活性測定溶液 100 μ L を混合し、酵素反応を開始させた。反応開始から 20 分間の 340 nm の吸光度をプレートリーダー（SYNERGY, BioTec, Green Mountains, VT）によって測定した。酵素反応速度及び活性値は 340 nm の吸光度のプロット直線部分の傾きから算出した。

Pyruvate \Rightarrow Acetaldehyde + CO₂ (PDC 反応)

Acetaldehyde + β -NADH \Rightarrow Ethanol + β -NAD (ADHII 反応)

Table 3-2-3 PDC 活性測定溶液

Sodium Pyruvate	5 mM
ThDP	0.1 mM
NADH	0.25 mM
MgCl ₂	5 mM
ADH	10 U
20 mM Tris-HCl pH7 に溶解	

3.2.13 酵素反応速度及び活性値の算出

酵素反応速度及び活性値はそれぞれ以下の式に従って算出した。 ΔA_{340nm} は 20 分間の吸光測定において、吸光度のプロットに直線性が見られた 13 点を利用して算出した。NADH のモル吸光定数には 6300 L/mol/cm を利用した。光路長は反応液の全体積が 110 μ L であることと、使用した 96 穴プレートの底面積が 0.32 cm^2 であることから算出した。

$$\text{酵素反応速度 (mM/h)} = \frac{\Delta A_{340nm} \text{ (/min)} \times 60 \times 10^3}{\text{光路長 (cm)} \times \text{モル吸光係数 (L/mol/cm)}}$$

$$\text{酵素活性 (Unit/mL)} = \frac{\Delta A_{340nm} \text{ (/min)} \times \text{全反応容量 } (\mu\text{L}) \times 10^3}{\text{光路長 (cm)} \times \text{モル吸光係数 (L/mol/cm)} \times \text{酵素容量 } (\mu\text{L})}$$

3.2.14 SDS-PAGE

PDC の発現確認を行うためシアノバクテリア粗酵素液の SDS-PAGE を行った。サンプルと 2 X sample buffer の混合物を 99 $^{\circ}$ C で 5 分間加熱した。長寿ゲル (Oriental Instruments Co., Ltd., Kanagawa, Japan) を泳動槽にセットし、標準 running buffer (0.25 mM Tris, 0.192 mM Glycine, 0.1 % SDS) を泳動槽に注いだ。サンプルを 10 μ L をウェルに添加し、40 mA で 40 分間電気泳動を行った。泳動後のゲルを CBB R250 染色液 (BIO RAD, CA, USA) で染色した後、脱染色してタンパク質のバンドを確認した。分子量マーカーには Precision Plus ProteinTM All Blue Prestained Protein Standards (BIO RAD, CA, USA) を用いた。

3.3 結果

3.3.1 配列的な多様性に富んだ PDC 配列の系統解析 (第 1 段階：広い探索)

第一段階として、機能的で、シアノバクテリアでの発現可能性が高く、配列多様性に富んだ PDC 配列の選抜を目指した。BRENDA は酵素の分子および生化学情報が掲載されたデータベースであり (Jeske *et al.*, 2019)、PDC 活性が実験的に確認されている酵素の情報を収集できると考えられる。そこで、BRENDA に EC 4.1.1.1 (pyruvate decarboxylase) として登録され、実験的な PDC 活性が確認されている酵素のうち、特に K_{cat}/K_M の値が高い又は低い K_M 値を有する 11 配列を収集した。加えて、近縁種に由来する酵素の発現可能性は高いと考え、本研究におけるエタノール生産宿主であるシアノバクテリア PCC6803 株と近縁の宿主に存在する PDC 遺伝子の情報を獲得するため、Cyanobase を活用することとした (Nakamura *et al.*, 2000)。Cyanobase はシアノバクテリアの総合的なゲノム情報データベースであり、遺伝子機能も提供されている。Cyanobase において、pyruvate decarboxylase 又は indole pyruvate decarboxylase に分類されている遺伝子を検索することで 6 配列を獲得した。最後に、InterPro Entry IPR012110 (Pyruvate decarboxylase/indolepyruvate decarboxylase-like) に属する配列のうち、Swiss-Prot にもエントリーされているレビュー済みの 50 配列を収集した。

得られた計 67 配列を出発配列として系統解析を行い、収集した PDC 配列の分類とシアノバクテリアで発現する酵素の選定を試みた (Fig. 3-3-1)。得られた系統樹によると、植物由来 PDC や真核微生物由来 PDC が主要なクレードを形成し、それとは別に IPDC や KDC もクレードを形成しており、多様な PDC 配列を収集できていることが分かる。IPDC や KDC はそれぞれ indolepyruvate と α ケト酸を基質として脱炭酸反応を起こす酵素であるものの、構造が類似している Pyruvate への反応性を有する酵素も知られているため (Yep *et al.*, 2006; Schütz *et al.*, 2003)、配列選抜でも候補に加えることとした。シアノバクテリアで発現する酵素としては、BRENDA 及び Cyanobase から抽出した配列を優先的に選抜し (*Zm* PDC を含む 7 配列)、多様な配列を評価するために、Swiss-Prot から得た配列も選抜した (2 配列)。選抜した 9 配列は系統樹の主要なクレードから選抜された 1 つ以上の配列を含んでおり、配列多様性を担保できていると考えられる。異なるクレードから選抜した PDC 配列はアライメントが困難 (配列相同性 40%以下) であるほど配列的に異なっていた。

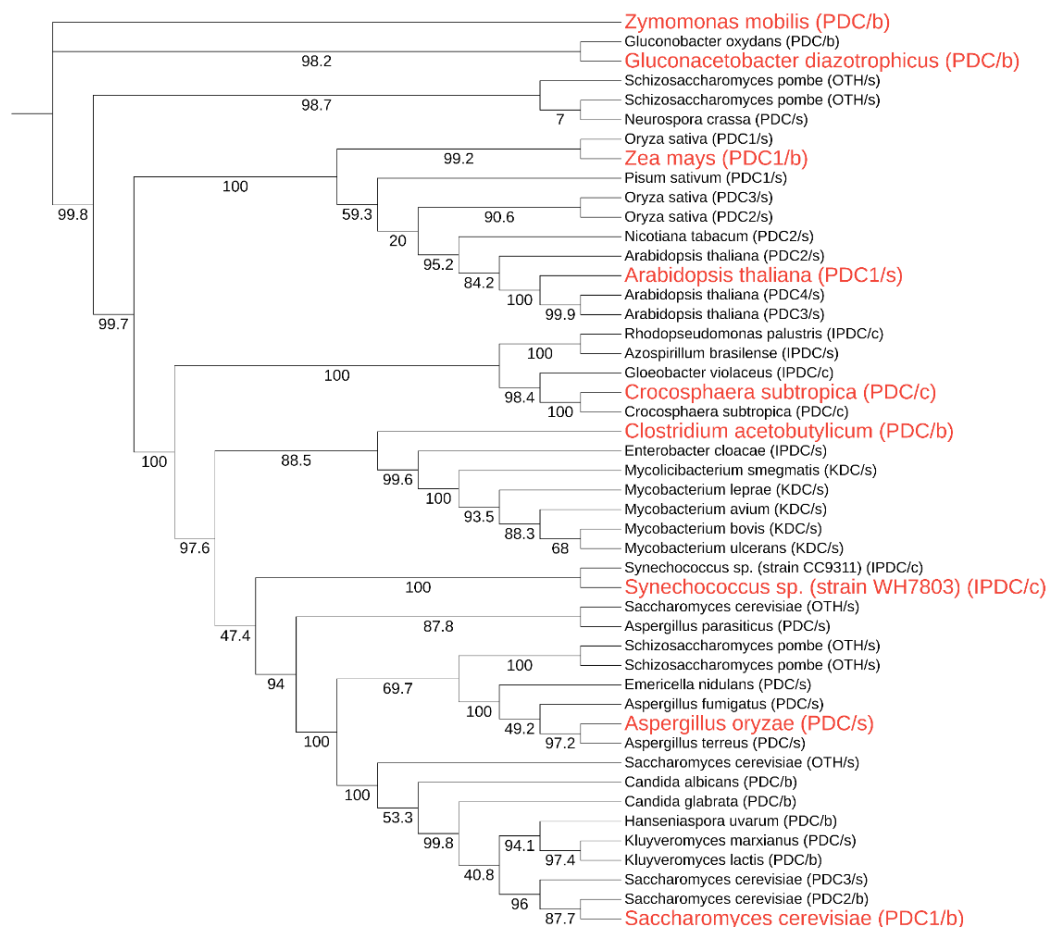


Fig. 3-3-1 多様な PDC 配列の系統樹

枝の長さを考慮しない系統樹を作成した。カッコ内にタンパク質名と由来するデータベースの略称を示した。各分岐には bootstrap probability を表示した。シアノバクテリアでの発現を行う酵素として選抜したものを赤で示した。タンパク質名とデータベースの略称、PDC1 : Pyruvate decarboxylase isozyme 1, OTH : Other, b : BRENDA, c : cyanobase, s : Swiss-prot

3.3.2 多様な PDC 配列を発現する株のエタノール生産性

選抜した9種のPDC及びPCC6803由来ADHII遺伝子をPnblAプロモーターがドライブするオペロンカセットとしてPCC6803株のゲノムに導入することでエタノール生産シアノバクテリア株を構築した(Table 3-2-1)。多様なPDCを発現する株の比較にはエタノールの生産性が高い条件が好ましいと考えられるため、本研究ではシアノバクテリアの有用化合物の生産性が大きく向上する例が複数報告されている Photomixotroph 条件での評価を実施することとした(Matson and Atsumi, 2017)。エタノール生産シアノバ

クテリア株を Mixotrough 条件で 120 時間培養し、培養上清中のエタノール濃度を評価した (Fig. 3-3-2)。その結果、*Zm* PDC を発現する Et-*Zm* 株の生産性が最も高く、その他 8 種の PDC に関しては *Gluconobacter diazotrophicus* (*Gd*) PDC を発現する Et-*Gd* 株でのみエタノール生産に成功した。エタノール生産に成功した PDC はいずれもバクテリア由来であり、PCC6803 株におけるエタノール生産には *Zm* PDC 配列周辺のバクテリア由来 PDC が適していることが示唆された。

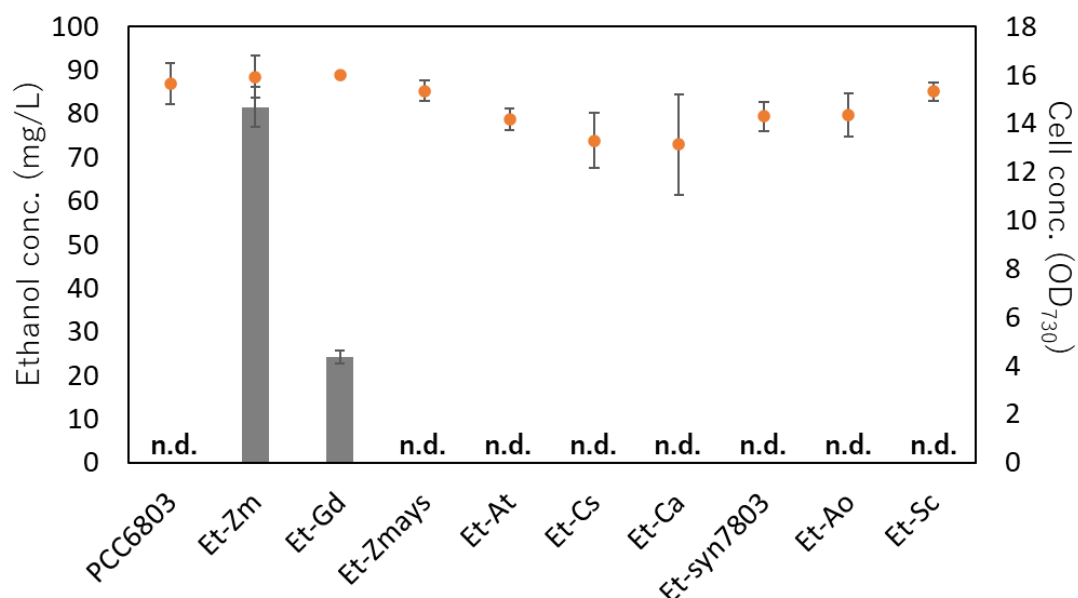


Fig. 3-3-2 多様な PDC を発現する株の培養上清中エタノール濃度

120 時間の培養終了時点における培養上清中のエタノール濃度 (左軸、棒グラフ) と菌体濃度 (右軸、散布図)。すべてのデータは平均値±SD で示した (n=3)。上清中にエタノールが検出されなかった株は n.d. で示した。

3.3.3 バクテリア由来 PDC の系統解析 (第 2 段階：密な探索)

バクテリア由来の PDC がシアノバクテリアにおけるエタノール生産に適していることが明らかとなったことを受け、探索の第 2 段階として、バクテリア由来 PDC クレードの詳細な評価を行うこととした。3.3.2 項においてシアノバクテリアにおけるエタノール生産への適合性が示唆された *Zm* PDC 配列と配列的に類似した Bacteria 由来 PDC 配列の中から新たな発現候補を選抜し、同クレード内を密に探索すべく、*Zm* PDC をクエリとした BLAST sequence similarity search によって Uniprot から 50 配列を取得した (Apweiler *et al.*, 2004)。3.3.2 項の結果から、Et-*Zm* 株のエタノール生産性は高く、*Zm* PDC と類似した配列において高い生産性が期待できるにも関わらず、Uniprot から得られた配列のうち、最も *Zm* PDC と類似した配列は *Gd* PDC であり、その配列相同性は 62.1% しかなかった。また、別のデータベースである non-redundant NCBI protein

sequences database (Pruitt *et al.*, 2006) に対する BLAST 検索においても *Zm* PDC と最も類似した配列は *Gd* PDC であった。そこで、*Zm* PDC 周辺の配列を評価すべく、3.3.4 項において人工配列設計を行うこととした。

Uniprot から得た配列に真核生物由来配列として 2 配列を加え、系統解析を行い、真核生物由来配列を外群とすることで Fig. 3-3-3 に示す系統樹を得た。系統樹に示す配列のうち、*Gluconobacter oxydans* (*Go*) PDC、*Acetobacter pasteurianus* (*Ap*) PDC、*Zymobacter palmae* (*Zp*) PDC は pH 6.5 又は 7 という中性付近の pH でも比較的高い PDC 活性を有することが報告されており (Van *et al.*, 2014)、シアノバクテリア細胞内の中性付近の条件でも活性発現することが期待できるため、新たな発現配列として選抜した。また、選抜した 3 配列と別のクレードに属する配列のうち、保存領域の欠損が比較的小さい *Acetobacter vaccinii* 及び *Ancylobacter rudongensis* に由来する 2 配列についても発現候補として選抜した。これら 2 配列は実験的に PDC 活性が確認された例は報告されていない。

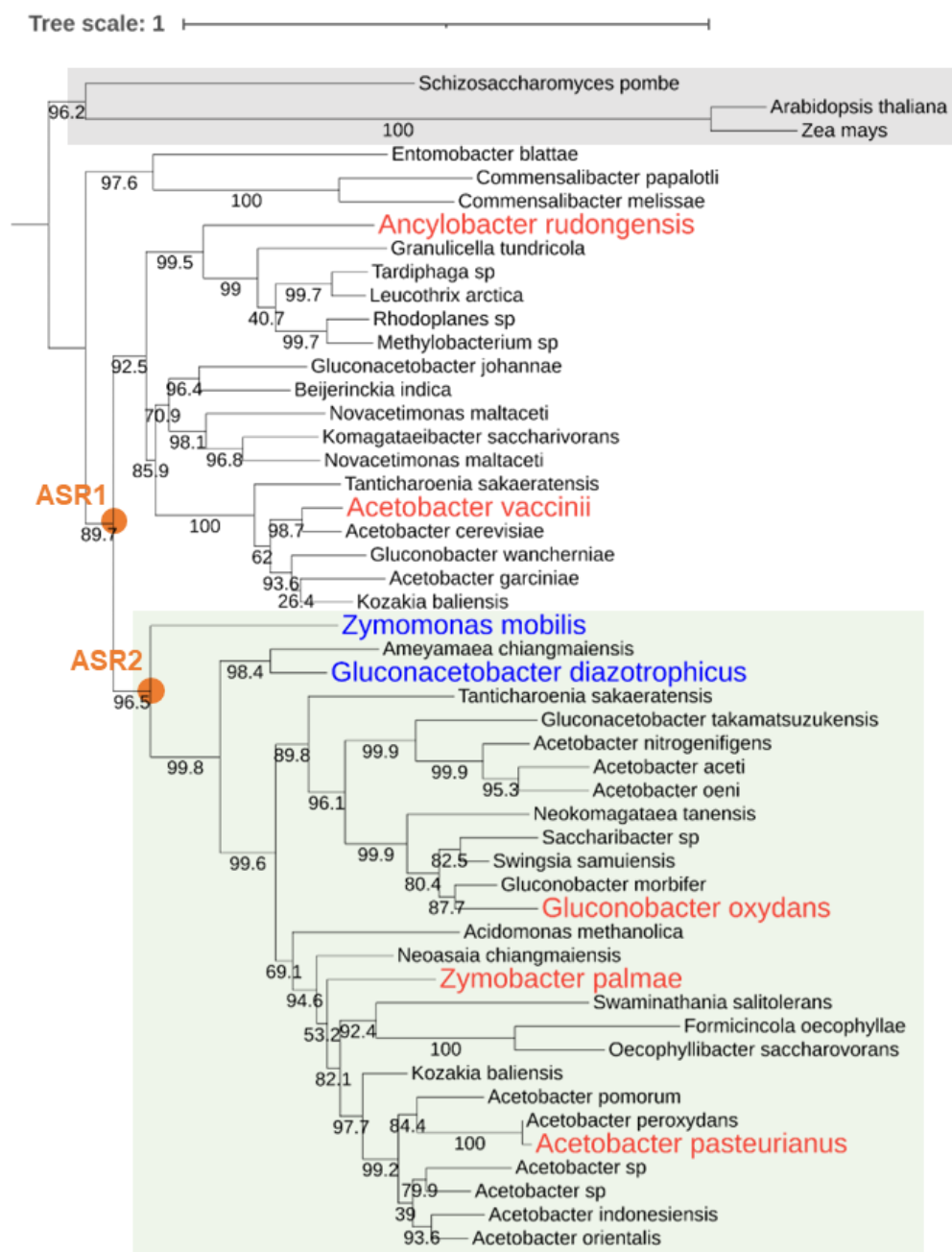


Fig. 3-3-3 バクテリア由来 PDC の系統樹

枝の長さを考慮した系統樹を作成した。各分岐には bootstrap probability を表示した。シアノバクテリアでの発現を行う酵素として新たに選抜したものの種名を赤で示し、3.3.2 項においてエタノールの生産が確認された 2 つの酵素的種名を青で示した。PROSS における MSA 作成に使用したクレードを緑で色付けし、外群とした真核生物由来の配列が属するクレードをグレーで色付けた。また、ASR1 と ASR2 の配列に対応するノードをオレンジで示した。

3.3.4 PROSS による *Zm* PDC と配列相同性の高い非天然タンパク質の設計

シアノバクテリアにおけるエタノール生産への適合性が示唆された *Zm* PDC と配列相同性の高い酵素がデータベース上に存在しなかったため、人工的に *Zm* PDC と類似性の高い配列を創出し、評価対象に加えることを目指した。PROSS は、立体構造と MSA を元に複数の安定化変異を導入した変異体を提案するアルゴリズムであり (Goldenzweig *et al.*, 2016)、この手法で設計された変異体は異種での生産性が改善することが報告されている (Goldenzweig *et al.*, 2016; Peleg *et al.*, 2021)。

本研究では、PROSS2 server (Weinstein *et al.*, 2021) を用い、*Zm* PDC の立体構造 (PDB ID : 2wva (Pei *et al.*, 2010)) を入力として、デフォルト設定で作成した MSA または、Fig. 3-3-2 の系統樹において緑で示す *Zm* PDC を含むクレードの配列情報のみで作成した MSA を用い、変異体の設計を行った。デフォルト設定の MSA には、*Zm* PDC との配列相同性が 40% 以下の配列も含まれているのに対し、系統解析を元にした MSA は 40% 以上の配列相同性を有する配列のみを含むため、MSA を元にした許容変異の計算においてより少数の変異しか許容されず、最終的に出力される変異体にも少数の安定化変異のみが導入されと考えられる。PROSS2 server での計算の結果、デフォルト設定による MSA では、*Zm* PDC との配列相同性 94.8 ~ 84.0% (29~84 個の変異を含む) の配列が得られ、系統解析による MSA では配列相同性 96.4 ~ 91.7% (20~45 個の変異を含む) の配列が得られた。デフォルト設定による MSA を入力として得られた配列のうち、*Zm* PDC との配列相同性が 94.8 % と 87.4% の配列をそれぞれ PROSS_design 1、PROSS_design 2 として、系統解析による MSA を入力として得られた配列のうち配列相同性が 96.4% の配列を PROSS_design 3 としてシアノバクテリアでの発現候補に加えた。*Zm* PDC と PROSS で設計した発現候補 3 配列のアライメントを Fig. 3-3-4 に示した。

[illegible]

PROSS による設計のクエリとした *Zm* PDC と設計された 3 配列のアライメント。配列全長に
変異が導入されていることが分かる。

3.3.5 ASR による *Zm* PDC と配列相同性の高い非天然タンパク質の設計

3.3.4 項と同様に人工的に *Zm* PDC と類似性の高い配列を創出することを目指し、ASR による非天然タンパク質の設計を行った。ASR は系統解析の結果を元に、系統樹のノードに相当する祖先型配列を推定する手法である (Spence *et al.*, 2021; Foley *et al.*, 2022)。祖先的な生物種は少数の酵素が多様な機能を担っており、1 つ 1 つの酵素の発現量が高かったと考えられているため (Gomez-Fernandez *et al.*, 2018)、ASR で設計された祖先型配列は異種での生産性が改善することが期待できる。

本研究では、Fig. 3-3-2 に示した系統樹を asr.GRASP server (Foley *et al.*, 2022) に供し、*Zm* PDC の祖先に相当するノード 1 と 2 の配列を ASR1, 2 (Fig. 3-3-3) としてシアノバクテリアでの発現候補に加えた。これらの配列の *Zm* PDC に対する相同性はそれぞれ 71.5% と 76.2% であった。*Zm* PDC と ASR で設計した発現候補 2 配列のアライメントを Fig. 3-3-5 に示した。

[illegible]

Zm PDC の祖先型配列として設計された 2 配列とのアライメント。PROSS とは異なり、ASR では *Zm* PDC と配列長の異なる配列が設計されていることが分かる。

3.3.6 バクテリア由来 PDC 発現候補の配列相同性解析

発現候補として新たに選抜した配列の多様性を評価すべく、*Zm* PDC、*Gd* PDC 及び系統樹を元に選抜した天然配列 5 配列と PROSS または ASR で設計した非天然配列 5 配列の相同性行列を作成した (Fig. 3-3-6)。これによると、*Zm* PDC を含む天然配列間の相同性は 45~68% であり、バクテリア由来 PDC の中では配列的に離れたものが評価対象となっていることが分かる。一方で、人工的に設計した非天然配列は *Zm* PDC との相同性が 72~96% であり、PROSS 及び ASR の活用によって、データベース上の配列よりも *Zm* PDC と類似した配列を発現候補に加えることができていたことが分かる。また、PROSS と ASR で設計した配列間の相同性は最大でも 75.9 % であり、2 つの方法の活用によって多様な非天然配列を発現候補に加えることができていた。

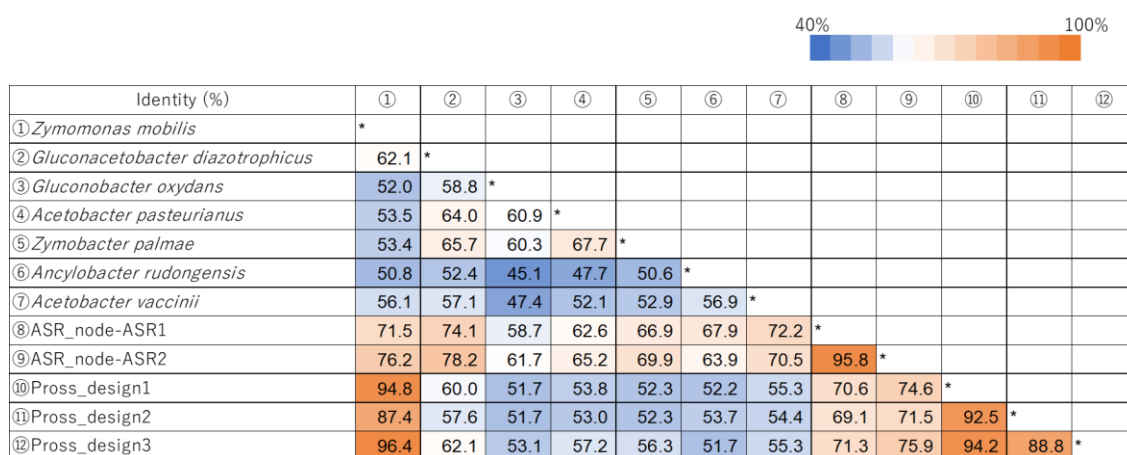


Fig. 3-3-6 Identity matrix of bacterial PDCs and non-native sequences

バクテリア由来 PDC と PROSS または ASR で設計した配列間の相同性行列。配列相同性が低いペアほど青に、高いペアほどオレンジになるように色付けした。

3.3.7 バクテリア由来 PDC 発現株のエタノール生産性

3.3.2 項と同様に、新たに選抜した 10 種の PDC 及び PCC6803 由来 ADHII 遺伝子を発現するシアノバクテリア株を構築し、そのエタノール生産性を評価した (Fig. 3-3-7)。その結果、*Go* PDC を発現する Et-*Go* 株のエタノール生産性が最も高くなっていた (Fig. 3-3-7a)。そこで、Et-*Zm* 株と Et-*Go* 株の培養評価を再度実施したところ (Fig. 3-3-7bc)。エタノールの最大濃度は Et-*Zm* 株の 79.3 mg/L/5days から Et-*Go* 株の 88.9 mg/L/5days にまで有意に向上していることが明らかとなった (p 値 = 0.024)。また、非天然配列を発現する株については Et-*Zm* 株よりも高いエタノール生産性を達成することはできなかったものの、Et-ASR2 株は評価した配列の中で 3 番目に高い生産性を示しており (Fig. 3-3-

7a)、ASR は特定の配列に類似する周辺配列を設計・評価するうえで有用な方法であることが示唆された。

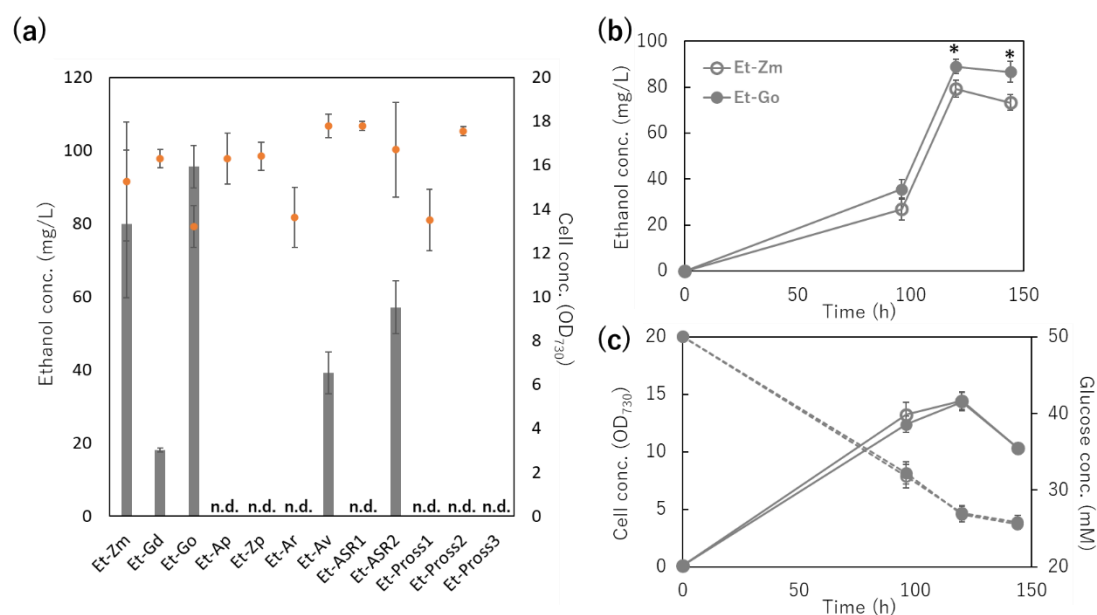


Fig. 3-3-7 Ethanol titer of bacterial PDCs and non-native sequences

すべてのデータは平均値±SDで示した (n=3)。(a) 120 時間の培養終了時点における培養上清中のエタノール濃度 (左軸、棒グラフ) と菌体濃度 (右軸、散布図)。上清中にエタノールが検出されなかった株は n.d. で示した。(b) Et-Zm 株と Et-Go 株の培養上清中エタノール濃度の経時変化。2 株のエタノール濃度に有意な差がある ($p < 0.05$ (片側 t 検定)) サンプルング点に * を付した。(c) Et-Zm 株と Et-Go 株の菌体濃度 (実線、左軸) と培養上清中のグルコース濃度 (破線、右軸)。

3.3.8 細胞内 PDC 活性の測定

エタノール生産株及び非生産株において細胞内の PDC 活性変化を確認すべく、培養開始 96 時間目の細胞を回収し、細胞内 PDC 活性の測定を行った (Fig. 3-3-8)。その結果、Et-Go 株は Et-Zm 株よりも細胞内 PDC 活性が低い傾向にあることが確認された (Fig. 3-3-8a, p 値 = 0.065)。しかしながら、Et-Go 株と同程度の細胞内 PDC 活性を有する Et-Gd 株のエタノール生産性は低く (Fig. 3-3-8a)、Go PDC による高い生産性は単純な細胞内 PDC 活性の強弱のみに寄らないことが確認された。また、Et-Zm、Et-Ar、Et-ASR2 株の細胞内 PDC 活性についても同程度であるにも関わらず、Et-Ar または Et-ASR2 株のエタノール生産性は Et-Zm 株よりも低くなっており、生産性と細胞内 PDC 活性が一致していなかった。その一方で、エタノール非生産株については細胞内の PDC 活性が PDC を導入していない PCC6803 野生株と同程度であり、生産性と細胞内 PDC 活性が一

致していた。PCC6803 及びその他のエタノール非生産株で検出されるわずかな PDC 活性はクルードの活性測定系において何らかの要因で NADH の酸化が進行するバックグラウンドに相当すると考えている。

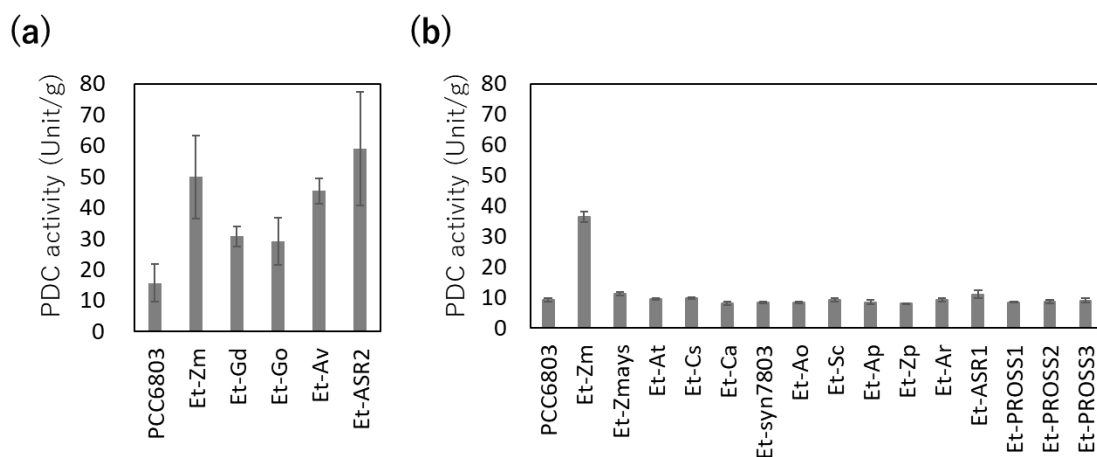


Fig. 3-3-8 PDC activity of Ethanol-producing strains

培養開始 96 時間目の細胞内 PDC 活性。すべてのデータは平均値±SD で示した (n = 3)。(a) エタノール生産株、(b) エタノール非生産株の細胞内 PDC 活性を示しており、培養回及び活性測定回が異なるため、図を分けて示している。

3.3.9 SDS-PAGE による PDC の発現確認

本研究ではシアノバクテリアでのエタノール生産性向上を目指し、細胞内での発現性と活性を包括した上で有用な配列を探索した。その一方で、各配列のエタノール生産性と細胞内での発現性及び活性を切り分けて考えることができれば、配列探索結果のより精度の高い理解が可能になると考えられる。そこで培養開始 96 時間目の細胞を回収し、破碎して得られたクルードの SDS-PAGE を行うことで細胞内での PDC 発現を確認することを試みた (Fig. 3-3-9)。

PDC を発現している株は 60 kDa 付近に新規のタンパク質バンドが確認できるはずであるが、いずれの株においても PCC6803 野生株のレーン 2 と比較して明瞭な新規のバンドを確認することはできなかった。これは細胞内タンパク質に占める PDC の割合が低いことによると考えられる。本研究では細胞内での発現性と活性を包括した上で有用な配列を探索したため、発現量を定量するための系を導入していなかった。正確な発現量の定量のためには細胞内で発現する PDC に His-tag を融合しておき、ウェスタンブロットによる定量を行うことが有効であったと考えられる。しかしながら、His-tag の融合はタンパク質発現に影響を及ぼすことが知られているため (Booth *et al.*, 2018)、

本研究のように発現量の議論が必須ではない場合には融合しない方が好ましいといえる。

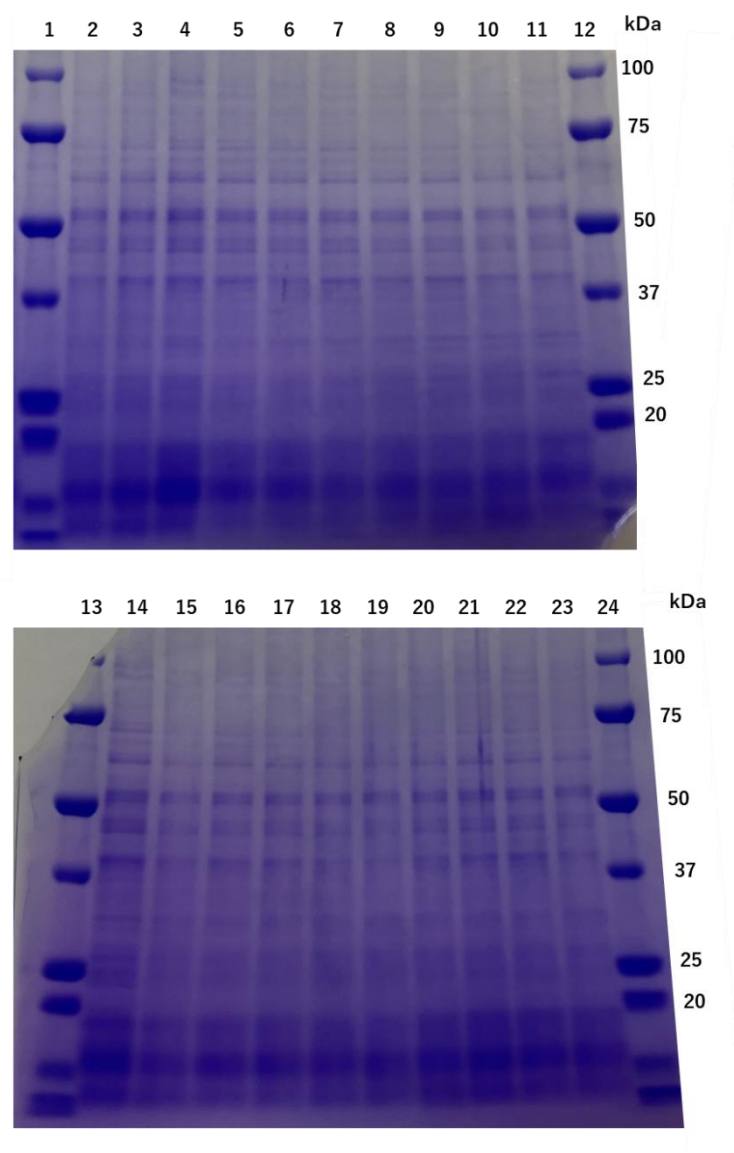


Fig. 3-3-9 様々な PDC を発現する株のクルード SDS-PAGE

培養開始 96 時間目の細胞を回収し、クルードサンプルを調製した。全て同一培養回かつ同時に泳動と染色を行った。各レーンに流したクルードサンプルと推定分子量を下記に示した。 1: protein marker, 2: PCC6803 (non), 3: Et-Zm (60.93 kDa), 4: Et-Gd (59.14 kDa), 5: Et-Zmays (65.44 kDa), 6: Et-At (66.22 kDa), 7: Et-Cs (60.16 kDa), 8: Et-Ca (62.57 kDa), 9: Et-syn7803 (60.19 kDa), 10: Et-Ao (62.93 kDa), 11: Et-Sc (61.5 kDa), 12: protein marker, 13: protein marker, 14: Et-Go (60.81 kDa), 15: Et-Ap (59.87 kDa), 16: Et-Zp (60.12 kDa), 17: Et-Ar (59.19 kDa), 18: Et-Av (60.45 kDa), 19: Et-ASR1 (60.02 kDa), 20: Et-ASR2 (60.14 kDa), 21: Et-Pross1 (61.24 kDa), 22: Et-Pross2 (61.19 kDa), 23: Et-Pross3 (61.39 kDa), 24: protein marker.

3.4 考察

本研究では PDC を題材として多様な配列の中から目的に合った酵素配列を探索することを試みた。最初の発現候補選抜では、BRENDA に登録されている PDC 活性を有することが実験的に確認されている配列や、Cyanobase を元に PCC6803 株の近縁種が有する PDC 様の配列を中心に、*Zm* PDC を含む配列多様性に富んだ 9 配列を選抜した。しかしながら、ADHII と共にシアノバクテリアで発現した際のエタノール生産性を評価したところ、エタノールの生産が確認されたのはバクテリア由来 PDC である *Zm* PDC と *Gd* PDC を発現する 2 株のみであった。先行研究において、PCC6803 での *S. cerevisiae* 由来 PDC の発現により、エタノールを生産した例が報告されていたにもかかわらず (Velmurugan *et al.*, 2019)、本研究で検討した *S. cerevisiae* 由来 PDC を発現する株はエタノールを生産しなかった。これは、先行研究との発現・培養条件の違いや、発現した *S. cerevisiae* 由来 PDC の配列がわずかに異なることによると考えられる (配列相同性 98.2 %)。また、異種シアノバクテリアの PDC 様の配列の発現によってもエタノールが生産できなかったことは、これらの遺伝子が PDC の機能を欠失していることによる可能性がある。

2 回目の発現候補選抜は、シアノバクテリアにおけるエタノール生産に適していることが示唆された *Zm* PDC 周辺のバクテリア由来 PDC から行った。バクテリア由来 PDC クレードの系統解析に基づいて新たに 5 つの発現候補が選抜された。しかしながら、これらの PDC 配列はいずれも最初の検討で高い生産性を示した *Zm* PDC との配列相同性が高くなく、PROSS または ASR によって *Zm* PDC と配列相同性の高い非天然配列の設計を行った。PROSS を用いた設計によって *Zm* PDC との配列相同性が 96.4%、94.8%、87.4% の PDC 配列を、ASR を用いた設計によって相同性が 71.5% と 76.2% の PDC 配列を獲得し、シアノバクテリアでの発現候補に加えた。2 つの方法で設計した配列間の相同性は最大でも 75.9% であり、2 つの方法の活用によって多様な非天然配列を発現候補に加えることができていたことが分かる。

新たに選抜した 10 配列をそれぞれ発現するシアノバクテリア株を構築し、エタノールの生産性を確認したところ、*Go* PDC を発現する Et-*Go* 株で最も高いエタノール生産性が確認された (Fig. 3-3-7a)。継時の生産挙動を確認したところ (Fig. 3-3-7b)、この株のエタノール生産性は 88.9 mg/L/5days であり、Et-*Zm* 株の 79.3 mg/L/5days から有意に向上していることが明らかとなった。両株でグルコースの消費速度には差がないことから、Et-*Go* 株では CO₂ やグルコースに由来する炭素を効率よく物質生産に向けることができていると言える。生産性が向上したにもかかわらず、Et-*Go* 株は Et-*Zm* 株と比較して細胞内 PDC 活性が低下傾向にあった (p 値 = 0.065)。このことはエタノールの生産性が必ずしも細胞内 PDC 活性の高さのみによって決まるわけではないことを示している。PCC6803 の細胞内 pH は 6.8 ~ 7.0 という中性付近であることが報告されており (Jiang *et al.*, 2013)、pH 6.5 における K_M 値は *Zm* PDC よりも *Go* PDC の方がわずかに小さい (*Zm*

PDC : $K_M = 1.3$ mM, *Go* PDC : 1.2 mM)(Van *et al.*, 2014)。エタノールの前駆体であるピルビン酸の PCC6803 細胞内濃度は PDC の K_M 値と比較して十分に小さいため (0.29 mM)(Nishiguchi *et al.*, 2020)、わずかな K_M 値の違いがエタノールの生産性に影響したことが考えられる。加えて、ピルビン酸の供給能がエタノールの生産性に影響を与えることが報告されており (Luan *et al.*, 2015)、*Zm* PDC と *Go* PDC それぞれの発現による毒性などの因子 (Kaur *et al.*, 2018) が代謝に影響を与え、ピルビン酸供給能を変化させたことによってエタノールの生産性に差が表れた可能性も考えられる。

一方で、本研究で選抜または設計された PDC の多くは、エタノールを生産することができなかった。人工的に設計された配列を除くと、これらの配列は、本研究で同定された高いエタノール生産性を示す *Zm* PDC および *Go* PDC との配列相同性が 65% 以下と低く、配列比較によるエタノール生産に重要な残基の特定は困難であった。非生産株の細胞内 PDC 活性の測定結果によると (Fig. 3-3-8b)、いずれの株も PDC 活性が低く、PDC を発現していない PCC6803 野生株と同等のレベルであった。非生産株のうち、Et-Zmays、Et-At、Et-Ca、Et-Ao、および Et-Sc 株に導入された PDC 配列は、BRENDA に PDC 活性を有する配列として登録されている。したがって、これらの株においてエタノールが生産できなかった要因は、細胞内で PDC を適切に発現することができなかったことによると考えられる。残りの非生産株がエタノールを生産できなかった要因が PDC を適切に発現できていないことによるか、細胞内での酵素活性が低いことによるかを知るためにはさらなる研究が必要である

また、本研究の中で、PROSS または ASR で設計した非天然配列をシアノバクテリアの代謝工学に応用することを試みた。その結果、Et-ASR2 株は比較的高い生産性 (Fig. 3-3-7a、テストした配列の中で Et-Zm 株に次いで 3 番目に高い生産性) と、Et-Zm 株と同程度の細胞内 PDC 活性を有していることが確認された。加えて、生産性は高いものの、これまでに異種発現によるエタノール生産への利用がされてこなかった *Ancylobacter rudongensis* (*Ar*) 由来 PDC が異種細胞内で活性発現することも明らかとなった。*Go* PDC を含め、本研究でシアノバクテリアにおけるエタノール生産に利用可能であることが明らかとなった配列は PCC6803 のエタノール生産で良く用いられる *Zm* PDC との相同性が 50.8 ~ 76.2% と高くなく、同定は容易でなかったといえる。*Zm* PDC のようなバクテリア由来の PDC は比較的まれであることから (Van *et al.*, 2014)(Fig. 3-3-3)、異種発現によるエタノール生産に利用可能な配列情報を拡張できたことは、他の微生物におけるエタノール生産での配列の利用や、タンパク質配列比較に基づいた変異導入による PDC の高機能化に応用できるという点で価値がある。

人工的に設計した非天然配列にはエタノールを生産できない配列も多く存在した。ASR による設計では、ASR1 を発現する株はエタノールを生産することが出来なかった。これは、系統樹の根に近い、より祖先型の配列であるほど配列推定の精度が落ち、機能的な配列を得ることが出来なかったことによると考えている。ASR によって非天然配列を設計する際には、現存配列に近い分岐を選択した方が良い可能性がある。一方、

PROSS での設計では、エタノールを生産可能な配列が得られなかった。この要因としては入力とした *Zm* PDC と類似した配列がデータベース上に存在せず、MSA を元にした許容される変異の計算がうまくいかなかったことが挙げられる。PROSS を有効に活用するためには、変異体を設計したいクエリ配列の周辺に、十分な近縁配列が存在することが適用基準になると考えられる。近縁配列が少ないケースにおいても PROSS を活用する必要がある場合には、適切なアライメントが取れていると予測されている領域のみで変異を許容するなど、PROSS の適用方法にさらなる試行錯誤が必要かもしれない。また、PROSS によって性能が向上した Polyethylene terephthalate (PET) 分解酵素を獲得した例では、PROSS で設計された変異体を既存の高活性変異体と配列比較し、高活性変異体が有する変異のみを実験的検証に進めていた (Rennison *et al.*, 2021)。この文献と同様に PDC 活性を有することが実験的に確認されていた *Go* PDC などが有する変異のみを許容して検証することで活性を有する PDC を設計出来た可能性も考えられる。本研究で実施した、第一段階としての広い配列空間の探索・評価と、第二段階における絞り込んだ配列空間での非天然配列を含めた探索・評価というステップワイズな戦略は、様々な有用酵素の選抜や利用可能な酵素配列の拡張に有用であると考えられる。

3.5 結言

本研究では多様なタンパク質配列の中から目的の配列を獲得することを目指し、配列的に多様な酵素の評価（配列空間を広く探索）と有望な酵素の周辺を追加で評価する（配列空間を密に探索）という 2 段階の探索を実施した。シアノバクテリアのエタノール生産に利用可能な PDC に関して初めて系統的な探索を行ったところ、*Zm* PDC より高いエタノール生産性を達成可能な *Go* PDC を同定することに成功した。これまでのエタノール生産性向上に関する知見とは作用点が異なり、既存知見との組み合わせによるさらなる生産性向上が期待できる。

1 段階目では、配列的に多様でかつ、BRENDA などのデータベースに PDC 活性を有すると記載されている配列を中心に発現検討を行った。このステップでは、配列空間のどの領域（どの配列の周辺）がシアノバクテリアにおけるエタノール生産に有望であるかを当たり付けすることを目指した。このような広い探索においては、各クレードから少数の配列しか発現検討へと進めることが出来ないため、できる限り発現可能性の高い配列を検討に進めることが重要となる。そのために、BRENDA などの実験に基づいたデータベースは極めて有用である。従来良く用いられている *Zm* PDC 以外の領域に最適な配列空間がある可能性を考慮して検討を進めたものの、本研究の例では *Zm* PDC 周辺のバクテリア由来配列が適しているという結果になった。これは、候補として選抜した酵素をシアノバクテリア細胞内で正しく発現できていないことによると考えており（3.4 項で議論した）、シアノバクテリアは細胞内で発現可能な異種タンパク質が限られていることが示唆された。このことは、シアノバクテリアで発現する異種遺伝子を選抜するうえで有用な知見になり得る。

2 段階目では、シアノバクテリアにおけるエタノール生産に有用であることが確認されたバクテリア由来 PDC の周辺配列について発現検討を行った。このステップでは、バクテリア由来 PDC（特に *Zm* PDC 周辺）の配列空間を密に探索すべく、天然配列から新たな候補を選抜するとともに非天然の配列を設計し、評価した。その結果、*Zm* PDC よりも生産性が高い配列として *Go* PDC を同定することに成功し、*ASR2* や *Ar* PDC といったシアノバクテリアでのエタノール生産に利用可能な PDC 配列を獲得することができた。バクテリア由来の PDC、特にシアノバクテリアでのエタノール生産に利用可能な PDC の情報は乏しいため、利用可能な配列を拡張できたことは有益であるといえる。エタノールを生産可能な配列群は配列間の相同性が 76%以下と乏しく、配列間の比較による重要なアミノ酸残基の特定は困難であった。しかしながら、配列相同性が 55%以上の配列であればタンパク質のキメラを作製できる可能性がある (Smith *et al.*, 2014)。エタノールを生産可能な配列間でキメラを作製することで、好ましい領域を組み合わせた変異体を獲得することができれば、さらなる生産性向上に寄与できると考えられる。

上記の様な2段階の探索手法は、非常に広い配列空間(PDCのアミノ酸を500 aaと考えると 20^{500})に対する網羅性としては不十分であるものの、既存の酵素よりも適したものを探索するうえでは有用な方法であると考えている。まず、1段階目の多様な配列群の評価によって、既存の酵素性能を大きく上回る配列群が存在する場合、それらを同定し、その周辺のさらなる探索を行うことができる。1段階目で既存の酵素よりも性能の高い配列を得られなかった場合も、既存の酵素の周辺を非天然配列も含めて密に探索することで既存の酵素よりも高い性能を有する酵素を獲得できる可能性がある。本研究はこのケースに合致する。本研究で提案する探索手法を適用するためには、機能がアノテーションされた配列的に多様な酵素配列の情報が必要となる。機能と配列の関係づけとしては、InterProに3000を超えるタンパク質ファミリーとそこに含まれるタンパク質の配列情報が収録されており (Paysan-Lafosse *et al.*, 2023)、ベンチマークとなる既存酵素が存在すれば機能アノテーションを活用することで本研究と同様の探索が可能であると考えられる。また、本研究のようにバクテリア由来の配列が少なく、発現性に課題がある酵素種ではなく、原核生物でも発現が容易なバクテリア由来の配列群が多く存在するケースでは配列探索の難易度も低くなると考えられる。本研究で活用した探索手法は、タンパク質配列情報が増加していく中で、様々な有用物質生産に利用可能な酵素を同定できるものであると考えている。

また、本研究ではPROSSやASRのような非天然配列の設計手法についても検討し、異種酵素探索に活用する際の適用可能性に関する知見も得られた。PROSSは、変異体を設計したい酵素の周辺に十分な数の近縁配列が存在する酵素ファミリーに対して有効である可能性が高い。ASRは、設計の際に現存配列に近い分岐を選択することで機能的な配列を設計できる可能性が高くなると考えられる。これらの知見は、今後の非天然配列設計における指針となりうる。

第4章 結論

4.1 結果のまとめ

シアノバクテリアの代謝を理解し、有用物質生産に適した状態へと改変することは代謝工学の目的のひとつである。改変の方向性として、酸化還元制御などによる酵素機能低下の緩和や有用物質の生合成経路を構成する高活性な代謝酵素の獲得が考えられる。本研究では、タンパク質配列情報をアミノ酸残基の保存度と有用配列情報のリソースとして活用し、シアノバクテリアの代謝工学へと応用する方法論を示すことを目的とした。

第2章では、タンパク質配列情報から得られるアミノ酸残基の保存度を指標に細胞内での機能性を推察し、変異体構築と細胞内での発現を行った。カルビンベンソン回路の重要酵素である Phosphoriblokinase (PRK) の活性制御の緩和を対象として検討を進めたところ、最終的に、PCC6803 由来 PRK の C 末端ループ上に Cys230Ser と Cys236Ser の変異を有する変異体を発現するシアノバクテリア株において Photomixotroph 条件における細胞内 PRK 活性が増加し、有機酸の蓄積量が増加するという物質生産に適した形質が確認された。変異の対象とした Cys230 及び Cys236 はシアノバクテリア内で保存性が高く、活性中心から離れた露出残基であった。保存度が高いアミノ酸残基には細胞内で機能的な制御に関連する残基が含まれている可能性があり、変異点として有効であると考えられる。しかしながら、保存度が高いアミノ酸残基には酵素の活性中心や酵素の折りたたみに関連する残基も含まれるため、変異点の慎重な選択が必要になる。その際、活性中心離れた位置に存在するアミノ酸残基や埋没していない残基が変異対象として有用であると考えられる。本章に示した①系統的な保存度の解析と機能に重要なアミノ酸残基の絞り込み、②維持したい機能に重要な領域 (活性中心など) から離れた残基を選択し変異を導入する、という方法は効果的に機能を改変した酵素を獲得できる方法になり得ると考えている。特に、シアノバクテリアや植物に由来する酵素は細胞内で様々な酸化還元制御を受けていることが知られているため、保存度が高いシステイン残基への変異が特に有効な可能性がある。

第3章では、改変の対象となる内在性の遺伝子を持たないケースにおいて、配列データベース上の多様なタンパク質配列の中から有用な異種配列を探索した。探索の対象はこれまでに系統的な配列探索と評価が行われてこなかったシアノバクテリアにおけるエタノール生産に適した pyruvate decarboxylase (PDC) とした。ここではタンパク質配列情報を有用配列情報のリソースとして活用することに加え、PDC の系統解析や人工配列設計に活用した。PDC の探索は2段階で実施した。1段階目では、PDC 活性を有する可能性が高く (BRENDA などに活性情報が記載)、かつ配列的に多様な PDC を評価することで有望な配列空間の当たり付けを行った (広い探索)。そこから得られたバクテリア由来 PDC がシアノバクテリアに適しているという知見を元に、2段階目はバクテリア由来 PDC 配列の候補を追加してさらに評価した (密な評価)。その際、祖先型

配列設計や PROSS で人工的に設計した非天然配列を候補に加えることで、*Zymomonas mobilis* (Zm) PDC の周辺の配列空間をより密に評価した。最終的に、Zm PDC よりも高い生産性を示す配列として、天然配列から *Gluconobacter oxydans* (Go) PDC を同定することに成功した。加えて、ASR2 や *Ancylobacter rudongensis* がシアノバクテリアにおけるエタノール生産に利用可能であることも明らかとなった。利用可能な配列の拡張は、それらの配列のキメラ化などによって、好ましい領域を組み合わせた変異体を獲得することにつながる可能性がある。本研究の 2 段階の探索手法は、非常に広い配列空間 (PDC のアミノ酸を 500 aa と考えると 20^{500}) に対する網羅性としては不十分であるものの、既存の酵素よりも適したものを探索するうえでは有用な方法であると考えている。まず、1 段階目の多様な配列群の評価によって、既存の酵素性能を大きく上回る配列群が存在する場合、それらを同定し、その周辺のさらなる探索を行うことができる。1 段階目で既存の酵素よりも性能の高い配列を得られなかった場合も、既存の酵素の周辺を非天然配列も含めて密に探索することで既存の酵素よりも高い性能を有する酵素を獲得できる可能性がある。本研究で活用した探索手法は、タンパク質配列情報が増加していく中で、様々な有用物質生産に利用可能な酵素を同定できるものであると考えている。

これらの結果から、本研究の保存度を指標とした変異導入と 2 段階での配列探索は、タンパク質配列情報をシアノバクテリアの代謝工学に応用するための非常に有用な方法論であるといえる。本研究で得られた有用物質生産性向上のための新たな知見は、他の知見との組み合わせによってさらなる生産性向上に寄与できると期待している。タンパク質配列情報は、単なる配列の羅列ではなく、酵素の機能や特性を理解し、代謝工学の目標達成に貢献する上で極めて貴重な情報源であるといえる。

4.2 本研究の成果と方法論の適用範囲

本研究では、微細藻類、特にシアノバクテリアの有用物質生産性を向上するための方法論を示すことを目的とした。シアノバクテリアによる物質生産性向上の例は複数報告されているものの、実用化のためにはさらなる生産性向上が必要である。シアノバクテリアにおける 2,3-ブタンジオールの生産性向上に成功した先行研究 (Kanno *et al.*, 2017) においては、生産性向上後の値を元にコスト試算を行っても、生産にかかるコストと 2,3-ブタンジオールの市場価格は同程度で利益がでない構造となっていた。さらにこの試算は、培養のスケールアップによっても生産性が低下しないという仮定に基づいていた。これらの因子を考慮すると、利益を得ながら大規模培養で生産した 2,3-ブタンジオールを上市するためには 5-10 倍程度の生産性向上が必要と推定される。また、2,3-ブタンジオールはシアノバクテリアでの有用物質生産において非常に生産性が高い物質であり (Kanno *et al.*, 2017)、その他の物質を上市するためにはより大きな生産性向上幅を達成することが必要である (10-50 倍程度と推察)。

こうした目標設定から考えると、本研究で提案する方法論の適用例として挙げた変異 PRK 発現による有機酸蓄積の増加 (ピルビン酸濃度 1.5 倍) や *Go* PDC 発現によるエタノール生産性の増加 (1.12 倍) は小さいものであるといえる。しかしながら、目標となるような大幅な生産性向上は一つの因子の改変で到達できるようなものではなく、複合的な宿主・培養条件などの改変によって初めて達成されうるものであると考えている。本研究で提案する 2 つの方法論は多様な代謝酵素の改変・選抜に応用できるものであり、様々な代謝酵素の最適化によって物質生産性を向上していくために非常に有用であると考えている。Table 4.2.1 に本研究で提案した方法論の概要と適用対象をまとめた。近年、配列情報や酵素に関する特徴づけが増加してきており、適用条件を満たす代謝酵素は増加してきている。今後、これらの代謝酵素の最適化によりさらなる物質生産性の向上につながることを期待される。

Table 4-4-1 本研究で提案した方法論の適用範囲

方法論1：保存度を活用した 内在性酵素 への変異導入
適応対象：細胞内での活性制御が報告されている代謝酵素（例：酸化還元制御、アロステリック制御）
好ましい対象の条件：
<ul style="list-style-type: none"> ・同一ファミリー内に十分な配列数が存在し、保存度の解析が可能 ・活性中心や構造維持に重要な残基（金属リガンドが配位する残基など）から離れた位置に保存性の高い残基を有する ・上記を満たす残基が好ましくはシステイン残基
候補となる代謝酵素の例：
<ul style="list-style-type: none"> ・ 3-Deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase <i>Chlamydomonas reinhardtii</i>においてフィードバック制御が起きていることが示唆 (Niraula <i>et al.</i>, 2025) ・ FB Pase/SBPase <i>Synechocystis</i> sp. PCC 6803において、TrxAによる酸化還元制御を受けていることを確認 (Mallén <i>et al.</i>, 2024)
方法論2：段階的な 異種酵素 探索
適応対象：宿主に内在しない、あるいは内在酵素の活性が低い代謝酵素（有用物質生産における生合成経路関連遺伝子など）
好ましい対象の条件：
<ul style="list-style-type: none"> ・ベンチマークとなる酵素が存在 ・InterProやPfam等のデータベース上に対象の酵素ファミリーが定義され、多数の配列が登録 ・（非天然配列設計を行うためには）有望と同定されたクレードの中にある程度の近縁配列を含む
候補となる代謝酵素の例：
<ul style="list-style-type: none"> ・ L-Lactate Dehydrogenase <i>Synechocystis</i> sp. PCC 6803に<i>Bacillus subtilis</i>由来酵素を導入することでL-乳酸の生産に成功 (Knoot <i>et al.</i>, 2018) ・ Ethylene-Forming Enzyme <i>Synechocystis</i> sp. PCC 6803に<i>Pseudomonas syringae</i>由来酵素を導入することでエチレンの生産に成功 (Knoot <i>et al.</i>, 2018)

4.3 タンパク質配列情報の代謝工学への活用の展望

本研究における保存度と有用配列情報のリソースとしてのタンパク質配列情報の活用という手法は、シアノバクテリアだけでなく他の微生物や酵素の研究においても利用できる可能性がある。第2章で示した PRK の C 末端ループにおける酸化制御の緩和は、同様の酸化還元制御が他の藻類や植物の様々な代謝酵素で働いていることが知られており (Michelet *et al.*, 2013)、これらの生物種における代謝酵素の機能改変や活性制御に応用できる可能性がある。また、第3章の段階的な配列探索や非天然配列の設計

といった戦略は特定の微生物に限らず、様々な代謝経路に関わる酵素の探索や機能改善に応用可能であると考えられる。有用物質生産においては、中央代謝による基質供給と、目的物質の生合成経路の双方が律速点になるケースも多い (Liu *et al.*, 2020; Ren *et al.*, 2023)。こうした課題を中央代謝酵素の保存度を指標にした改変と、目的物質を生産するために最適な異種遺伝子を探索することで解決し、生産性を向上することが期待できる。

さらに、本研究の手法は最新のタンパク質の機能予測や構造予測のさらなる発展により、精度を向上できる可能性がある。例えば、近年目覚ましい発展を遂げている AlphaFold といったタンパク質構造予測ツール (Yang *et al.*, 2023) を用いることで、実験的に構造が決定されていないタンパク質に対しても高精度な三次元構造モデルを構築することが可能になってきた。このような構造情報は、保存度の高いアミノ酸残基がタンパク質の機能や安定性にどのように寄与しているかをより深く理解する上で非常に有用であり、第 2 章で示したような保存度の高い残基への変異導入の精度を向上することが期待される。また、機械学習に基づいたタンパク質機能予測技術 (Lawson *et al.*, 2020) と組み合わせることで、配列データベースから候補となる酵素配列を絞り込む際の精度を向上させ、第 3 章で実施したようなデータベース探索において、目的とする酵素活性を持つ配列を効率的に見つけ出すことが可能になると考えられる。さらに、タンパク質の機能予測や構造予測を活用することで、非天然配列の設計においても、機能的な酵素が得られる可能性が高くなると考えられる (Yang *et al.*, 2024)。

本研究では、タンパク質配列情報をシアノバクテリアの代謝工学に応用し、代謝改変と有用物質の生産性を向上する方法論を示した。タンパク質配列の保存度に基づく合理的な変異体設計と、配列データベースからの系統的な異種酵素探索という 2 つのアプローチは、それぞれ異なるメカニズムでシアノバクテリアの代謝を効果的に改変し、有用物質生産へと繋がる知見が得られた。これらの手法は、他の微生物や酵素の研究においても応用可能であり、今後の代謝工学研究の発展に貢献できると考えている。さらに、最新のタンパク質機能予測や構造予測技術の発展により、変異設計や酵素探索の精度はさらに向上すると期待している。本研究の成果は微生物の物質生産性を向上し、持続可能な社会を実現することに貢献すると考えている。

参考文献

Ana Carolina de Lima Barizão, Luiz Eduardo de Oliveira Gomes, Larissa Lamburghini Brandão, Igor Carvalho Fontes Sampaio, Isabela Viana Lopes de Moura, Ricardo Franci Gonçalves, Jairo Pinto de Oliveira, Sérgio Túlio Cassini, Microalgae as tertiary wastewater treatment: Energy production, carbon neutrality, and high-value products, *Algal Research*, 2023, Volume 72, 103113, ISSN 2211-9264, <https://doi.org/10.1016/j.algal.2023.103113>.

Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* 2004 Jan 1;32(Database issue):D115-9. doi: 10.1093/nar/gkh131. PMID: 14681372; PMCID: PMC308865.

Aresta M, Dibenedetto A. Utilisation of CO₂ as a chemical feedstock: opportunities and challenges. *Dalton Trans.* 2007 Jul 28;(28):2975-92. doi: 10.1039/b700658f. Epub 2007 Jun 26. PMID: 17622414.

Atsumi S, Higashide W, Liao JC. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat Biotechnol.* 2009 Dec;27(12):1177-80. doi: 10.1038/nbt.1586. PMID: 19915552.

Bajaj K, Chakrabarti P, Varadarajan R. Mutagenesis-based definitions and probes of residue burial in proteins. *Proc Natl Acad Sci U S A.* 2005 Nov 8;102(45):16221-6. doi: 10.1073/pnas.0505089102. Epub 2005 Oct 26. PMID: 16251276; PMCID: PMC1283427.

Bernardo Llamas, María C. Suárez-Rodríguez, Cynthia V. González-López, Pedro Mora, F. Gabriel Acien., Techno-economic analysis of microalgae related processes for CO₂ bio-fixation, *Algal Research*, 2021, Volume 57, 102339, ISSN 2211-9264, <https://doi.org/10.1016/j.algal.2021.102339>.

Biovia, D.S. (2019) Discovery Studio Visualizer. San Diego.

Blanc-Garin V, Veaudor T, Sétif P, Gontero B, Lemaire SD, Chauvat F, Cassier-Chauvat C. First in vivo analysis of the regulatory protein CP12 of the model cyanobacterium *Synechocystis* PCC 6803: Biotechnological implications. *Front Plant Sci.* 2022 Sep 13;13:999672. doi: 10.3389/fpls.2022.999672. PMID: 36176677; PMCID: PMC9514657.

Booth WT, Schlachter CR, Pote S, Ussin N, Mank NJ, Klapper V, Offermann LR, Tang C, Hurlburt BK, Chruszcz M. Impact of an N-terminal Polyhistidine Tag on Protein Thermal Stability. *ACS Omega.* 2018 Jan 31;3(1):760-768. doi: 10.1021/acsomega.7b01598. Epub 2018 Jan 22. PMID: 29399652; PMCID: PMC5793033.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009 Aug 1;25(15):1972-3. doi: 10.1093/bioinformatics/btp348. Epub 2009 Jun 8. PMID: 19505945; PMCID: PMC2712344.

Castaño-Cerezo S, Chamas A, Kulyk H, Treitz C, Bellvert F, Tholey A, Galéote V, Camarasa C, Heux S, Garcia-Alles LF, Millard P, Truan G. Combining systems and synthetic biology for in vivo enzymology. *EMBO J.* 2024 Nov;43(21):5169-5185. doi: 10.1038/s44318-024-00251-w. Epub 2024 Sep 25. PMID: 39322757; PMCID: PMC11535393.

Chikunova A, Ubbink M. The roles of highly conserved, non-catalytic residues in class A β -lactamases. *Protein Sci.* 2022 Jun;31(6):e4328. doi: 10.1002/pro.4328. PMID: 35634774; PMCID: PMC9112487.

Choi YN, Park JM. Enhancing biomass and ethanol production by increasing NADPH production in *Synechocystis* sp. PCC 6803. *Bioresour Technol.* 2016 Aug;213:54-57. doi: 10.1016/j.biortech.2016.02.056. Epub 2016 Mar 3. PMID: 26951740.

Comino N, Cifuentes JO, Marina A, Orrantia A, Eguskiza A, Guerin ME. Mechanistic insights into the allosteric regulation of bacterial ADP-glucose pyrophosphorylases. *J Biol Chem.* 2017 Apr 14;292(15):6255-6268. doi: 10.1074/jbc.M116.773408. Epub 2017 Feb 21. PMID: 28223362; PMCID: PMC5391755.

Davidi D, Noor E, Liebermeister W, Bar-Even A, Flamholz A, Tumbler K, Barenholz U, Goldenfeld M, Shlomi T, Milo R. Global characterization of in vivo enzyme catalytic rates and their correspondence to in vitro *k_{cat}* measurements. *Proc Natl Acad Sci U S A.* 2016 Mar 22;113(12):3401-6. doi: 10.1073/pnas.1514240113. Epub 2016 Mar 7. PMID: 26951675; PMCID: PMC4812741.

Deng MD, Coleman JR. Ethanol synthesis by genetic engineering in cyanobacteria. *Appl Environ Microbiol.* 1999 Feb;65(2):523-8. doi: 10.1128/AEM.65.2.523-528.1999. PMID: 9925577; PMCID: PMC91056.

Dörte Gocke, Thorsten Graf, Helen Brosi, Ilona Frindi-Wosch, Lydia Walter, Michael Müller, Martina Pohl, Comparative characterisation of thiamin diphosphate-dependent decarboxylases, *Journal of Molecular Catalysis B: Enzymatic*, Volume 61, Issues 1–2, 2009, Pages 30-35, ISSN 1381-1177, <https://doi.org/10.1016/j.molcatb.2009.03.019>.

Ethan I. Lan, Soo Y. Ro and James C. Liao. Oxygen-tolerant coenzyme A-acylating aldehyde dehydrogenase facilitates efficient photosynthetic n-butanol biosynthesis in cyanobacteria. *Energy & Environmental Science.* 2013. Vol. 6(9):2672. DOI: 10.1039/c3ee41405a

Fernandes R, Campos J, Serra M, Fidalgo J, Almeida H, Casas A, Toubarro D, Barros AIRNA. Exploring the Benefits of Phycocyanin: From *Spirulina* Cultivation to Its Widespread Applications. *Pharmaceuticals (Basel).* 2023 Apr 14;16(4):592. doi: 10.3390/ph16040592. PMID: 37111349; PMCID: PMC10144176.

Foley G, Mora A, Ross CM, Bottoms S, Sützl L, Lamprecht ML, Zaugg J, Essebier A, Balderson B, Newell R, Thomson RES, Kobe B, Barnard RT, Guddat L, Schenk G, Carsten J, Gumulya Y, Rost B, Haltrich D, Sieber V, Gillam EMJ, Bodén M. Engineering indel and substitution variants of diverse and ancient enzymes using Graphical Representation of Ancestral Sequence Predictions (GRASP). *PLoS Comput Biol.* 2022 Oct 24;18(10):e1010633. doi: 10.1371/journal.pcbi.1010633. PMID: 36279274; PMCID: PMC9632902.

Fukui K, Yoshida K, Yokochi Y, Sekiguchi T, Wakabayashi KI, Hisabori T, Mihara S. The Importance of the C-Terminal Cys Pair of Phosphoribulokinase in Phototrophs in Thioredoxin-Dependent Regulation. *Plant Cell Physiol.* 2022 Jun 15;63(6):855-868. doi: 10.1093/pcp/pcac050. PMID: 35413120; PMCID: PMC9199185.

Gaucher EA, Thomson JM, Burgan MF, Benner SA. Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature*. 2003 Sep 18;425(6955):285-8. doi: 10.1038/nature01977. PMID: 13679914.

Gerth ML, Lutz S. Mutagenesis of non-conserved active site residues improves the activity and narrows the specificity of human thymidine kinase 2. *Biochem Biophys Res Commun*. 2007 Mar 16;354(3):802-7. doi: 10.1016/j.bbrc.2007.01.070. Epub 2007 Jan 23. PMID: 17266931; PMCID: PMC1853344.

Goldenzweig A, Goldsmith M, Hill SE, Gertman O, Laurino P, Ashani Y, Dym O, Unger T, Albeck S, Prilusky J, Lieberman RL, Aharoni A, Silman I, Sussman JL, Tawfik DS, Fleishman SJ. Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Mol Cell*. 2016 Jul 21;63(2):337-346. doi: 10.1016/j.molcel.2016.06.012. Epub 2016 Jul 14. Erratum in: *Mol Cell*. 2018 Apr 19;70(2):380. doi: 10.1016/j.molcel.2018.03.035. PMID: 27425410; PMCID: PMC4961223.

Gomez-Fernandez BJ, Garcia-Ruiz E, Martin-Diaz J, Gomez de Santos P, Santos-Moriano P, Plou FJ, Ballesteros A, Garcia M, Rodriguez M, Risso VA, Sanchez-Ruiz JM, Whitney SM, Alcalde M. Directed -in vitro- evolution of Precambrian and extant Rubiscos. *Sci Rep*. 2018 Apr 3;8(1):5532. doi: 10.1038/s41598-018-23869-3. PMID: 29615759; PMCID: PMC5883036.

Heidorn T., Camsund, D., Huang, H.-H., Lindberg, P., Oliveira, P., Stensjö, K., Lindberg, P. "Synthetic biology in cyanobacteria: engineering and analyzing novel functions." *Methods in enzymology*. Vol. 497. Academic Press, 2011. 539-579.

Hitchcock A, Hunter CN, Canniffe DP. Progress and challenges in engineering cyanobacteria as chassis for light-driven biotechnology. *Microb Biotechnol*. 2020 Mar;13(2):363-367. doi: 10.1111/1751-7915.13526. Epub 2019 Dec 27. PMID: 31880868; PMCID: PMC7017823.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol*. 2018 Feb 1;35(2):518-522. doi: 10.1093/molbev/msx281. PMID: 29077904; PMCID: PMC5850222.

IHI Corporation. Supplying bio-jet fuel produced from microalgae to regular domestic flights-
| June 18, 2021 | Press Release |
https://www.ihl.co.jp/en/all_news/2021/other/1197475_3370.html (accessed May 5, 2025).

Jeske L, Placzek S, Schomburg I, Chang A, Schomburg D. BRENDA in 2019: a European ELIXIR core data resource. *Nucleic Acids Res*. 2019 Jan 8;47(D1):D542-D549. doi: 10.1093/nar/gky1048. PMID: 30395242; PMCID: PMC6323942.

Jiang HB, Cheng HM, Gao KS, Qiu BS. Inactivation of Ca(2+)/H(+) exchanger in *Synechocystis* sp. strain PCC 6803 promotes cyanobacterial calcification by upregulating CO(2)-concentrating mechanisms. *Appl Environ Microbiol*. 2013 Jul;79(13):4048-55. doi: 10.1128/AEM.00681-13. Epub 2013 Apr 26. PMID: 23624472; PMCID: PMC3697565.

Jiao X, Yang L, An M, Chen W. A modified amino acid network model contains similar and dissimilar weight. *Comput Math Methods Med*. 2013;2013:197892. doi: 10.1155/2013/197892. Epub 2013 Jan 2. PMID: 23365624; PMCID: PMC3549380.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017 Jun;14(6):587-589. doi: 10.1038/nmeth.4285. Epub 2017 May 8. PMID: 28481363; PMCID: PMC5453245.

Kanno M, Carroll AL, Atsumi S. Global metabolic rewiring for improved CO₂ fixation and chemical production in cyanobacteria. *Nat Commun*. 2017 Mar 13;8:14724. doi: 10.1038/ncomms14724. PMID: 28287087; PMCID: PMC5355792.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013 Apr;30(4):772-80. doi: 10.1093/molbev/mst010. Epub 2013 Jan 16. PMID: 23329690; PMCID: PMC3603318.

Kaur J, Kumar A, Kaur J. Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. *Int J Biol Macromol*. 2018 Jan;106:803-822. doi: 10.1016/j.ijbiomac.2017.08.080. Epub 2017 Aug 19. PMID: 28830778.

Klomsiri C, Karplus PA, Poole LB. Cysteine-based redox switches in enzymes. *Antioxid Redox Signal*. 2011 Mar 15;14(6):1065-77. doi: 10.1089/ars.2010.3376. Epub 2010 Sep 17. PMID: 20799881; PMCID: PMC3064533.

Knoot CJ, Ungerer J, Wangikar PP, Pakrasi HB. Cyanobacteria: Promising biocatalysts for sustainable chemical production. *J Biol Chem*. 2018 Apr 6;293(14):5044-5052. doi: 10.1074/jbc.R117.815886. Epub 2017 Oct 2. PMID: 28972147; PMCID: PMC5892586.

König S. Subunit structure, function and organisation of pyruvate decarboxylases from various organisms. *Biochim Biophys Acta*. 1998 Jun 29;1385(2):271-86. doi: 10.1016/s0167-4838(98)00074-0. PMID: 9655918.

Kono T, Mehrotra S, Endo C, Kizu N, Matusda M, Kimura H, Mizohata E, Inoue T, Hasunuma T, Yokota A, Matsumura H, Ashida H. A RuBisCO-mediated carbon metabolic pathway in methanogenic archaea. *Nat Commun*. 2017 Jan 13;8:14007. doi: 10.1038/ncomms14007. PMID: 28082747; PMCID: PMC5241800.

Krigbaum, W. R., & Komoriya, A. (1979). Local interactions as a structure determinant for protein molecules: II. *Biochim Biophys Acta*, 576(1), 204-248. [https://doi.org/10.1016/0005-2795\(79\)90498-7](https://doi.org/10.1016/0005-2795(79)90498-7)

Lai MC, Lan EI. Advances in Metabolic Engineering of Cyanobacteria for Photosynthetic Biochemical Production. *Metabolites*. 2015 Oct 27;5(4):636-58. doi: 10.3390/metabo5040636. PMID: 26516923; PMCID: PMC4693188.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007 Nov 1;23(21):2947-8. doi: 10.1093/bioinformatics/btm404. Epub 2007 Sep 10. PMID: 17846036.

Lawson CE, Martí JM, Radivojevic T, Jonnalagadda SVR, Gentz R, Hillson NJ, Peisert S, Kim J, Simmons BA, Petzold CJ, Singer SW, Mukhopadhyay A, Tanjore D, Dunn JG, Garcia Martin H. Machine learning for metabolic engineering: A review. *Metab Eng*. 2021 Jan;63:34-60. doi: 10.1016/j.ymben.2020.10.005. Epub 2020 Nov 19. PMID: 33221420.

Li G, Yao J. A Review of Algae-Based Carbon Capture, Utilization, and Storage (Algae-Based CCUS). *Gases*. 2024; 4(4):468-503. <https://doi.org/10.3390/gases4040024>

Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006 Jul 1;22(13):1658-9. doi: 10.1093/bioinformatics/btl158. Epub 2006 May 26. PMID: 16731699.

Lim JM, Kim G, Levine RL. Methionine in Proteins: It's Not Just for Protein Initiation Anymore. *Neurochem Res*. 2019 Jan;44(1):247-257. doi: 10.1007/s11064-017-2460-0. Epub 2018 Jan 11. PMID: 29327308; PMCID: PMC6446232.

Lindberg P, Park S, Melis A. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metab Eng*. 2010 Jan;12(1):70-9. doi: 10.1016/j.ymben.2009.10.001. Epub 2009 Oct 13. PMID: 19833224.

Liu P, Zhang B, Yao ZH, Liu ZQ, Zheng YG. Multiplex Design of the Metabolic Network for Production of l-Homoserine in *Escherichia coli*. *Appl Environ Microbiol*. 2020 Oct 1;86(20):e01477-20. doi: 10.1128/AEM.01477-20. PMID: 32801175; PMCID: PMC7531971.

Luan G, Qi Y, Wang M, Li Z, Duan Y, Tan X, Lu X. Combinatory strategy for characterizing and understanding the ethanol synthesis pathway in cyanobacteria cell factories. *Biotechnol Biofuels*. 2015 Nov 21;8:184. doi: 10.1186/s13068-015-0367-z. PMID: 26594240; PMCID: PMC4654843.

Lucius S, Theune M, Arrivault S, Hildebrandt S, Mullineaux CW, Gutekunst K, Hagemann M. CP12 fine-tunes the Calvin-Benson cycle and carbohydrate metabolism in cyanobacteria. *Front Plant Sci*. 2022 Oct 11;13:1028794. doi: 10.3389/fpls.2022.1028794. PMID: 36330266; PMCID: PMC9623430.

M. Prussi, W. Weindorf, M. Buffi, J. Sánchez López, N. Scarlat, Are algae ready to take off? GHG emission savings of algae-to-kerosene production, *Applied Energy*, 2021, Volume 304, 117817, ISSN 0306-2619, <https://doi.org/10.1016/j.apenergy.2021.117817>.

Man Kee Lam, Keat Teong Lee, Abdul Rahman Mohamed, Current status and challenges on microalgae-based carbon capture, *International Journal of Greenhouse Gas Control*, 2012, Volume 10, Pages 456-469, ISSN 1750-5836, <https://doi.org/10.1016/j.ijggc.2012.07.010>.

Matson MM, Atsumi S. Photomixotrophic chemical production in cyanobacteria. *Curr Opin Biotechnol*. 2018 Apr;50:65-71. doi: 10.1016/j.copbio.2017.11.008. Epub 2017 Nov 24. PMID: 29179151.

Matsuda F, Toya Y, Shimizu H. Learning from quantitative data to understand central carbon metabolism. *Biotechnol Adv*. 2017 Dec;35(8):971-980. doi: 10.1016/j.biotechadv.2017.09.006. Epub 2017 Sep 18. PMID: 28928003.

McFarlane CR, Shah NR, Kabasakal BV, Echeverria B, Cotton CAR, Bubeck D, Murray JW. Structural basis of light-induced redox regulation in the Calvin-Benson cycle in cyanobacteria. *Proc Natl Acad Sci U S A*. 2019 Oct 15;116(42):20984-20990. doi: 10.1073/pnas.1906722116. Epub 2019 Sep 30. PMID: 31570616; PMCID: PMC6800369.

Mengesha I, Roy D. Carbon pricing drives critical transition to green growth. *Nat Commun.* 2025 Feb 3;16(1):1321. doi: 10.1038/s41467-025-56540-3. PMID: 39900895; PMCID: PMC11791185.

Miao R, Liu X, Englund E, Lindberg P, Lindblad P. Isobutanol production in *Synechocystis* PCC 6803 using heterologous and endogenous alcohol dehydrogenases. *Metab Eng Commun.* 2017 Jul 29;5:45-53. doi: 10.1016/j.meteno.2017.07.003. PMID: 29188183; PMCID: PMC5699533.

Michelet L, Zaffagnini M, Morisse S, Sparla F, Pérez-Pérez ME, Francia F, Danon A, Marchand CH, Fermari S, Trost P, Lemaire SD. Redox regulation of the Calvin-Benson cycle: something old, something new. *Front Plant Sci.* 2013 Nov 25;4:470. doi: 10.3389/fpls.2013.00470. PMID: 24324475; PMCID: PMC3838966.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol.* 2020 May 1;37(5):1530-1534. doi: 10.1093/molbev/msaa015. Erratum in: *Mol Biol Evol.* 2020 Aug 1;37(8):2461. doi: 10.1093/molbev/msaa131. PMID: 32011700; PMCID: PMC7182206.

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods.* 2022 Jun;19(6):679-682. doi: 10.1038/s41592-022-01488-1. Epub 2022 May 30. PMID: 35637307; PMCID: PMC9184281.

Moore V, Vermaas W. Functional consequences of modification of the photosystem I/photosystem II ratio in the cyanobacterium *Synechocystis* sp. PCC 6803. *J Bacteriol.* 2024 May 23;206(5):e0045423. doi: 10.1128/jb.00454-23. Epub 2024 May 2. PMID: 38695523; PMCID: PMC11112997.

Nakajima T, Kajihata S, Yoshikawa K, Matsuda F, Furusawa C, Hirasawa T, Shimizu H. Integrated metabolic flux and omics analysis of *Synechocystis* sp. PCC 6803 under mixotrophic and photoheterotrophic conditions. *Plant Cell Physiol.* 2014 Sep;55(9):1605-12. doi: 10.1093/pcp/pcu091. Epub 2014 Jun 26. PMID: 24969233.

Nakajima T, Yoshikawa K, Toya Y, Matsuda F, Shimizu H. Metabolic Flux Analysis of the *Synechocystis* sp. PCC 6803 Δ nrtABCD Mutant Reveals a Mechanism for Metabolic Adaptation to Nitrogen-Limited Conditions. *Plant Cell Physiol.* 2017 Mar 1;58(3):537-545. doi: 10.1093/pcp/pcw233. PMID: 28130420.

Nakamura Y, Kaneko T, Tabata S. CyanoBase, the genome database for *Synechocystis* sp. strain PCC6803: status for the year 2000. *Nucleic Acids Res.* 2000 Jan 1;28(1):72. doi: 10.1093/nar/28.1.72. PMID: 10592184; PMCID: PMC102459.

Namakoshi K, Nakajima T, Yoshikawa K, Toya Y, Shimizu H. Combinatorial deletions of *glgC* and *phaCE* enhance ethanol production in *Synechocystis* sp. PCC 6803. *J Biotechnol.* 2016 Dec 10;239:13-19. doi: 10.1016/j.jbiotec.2016.09.016. Epub 2016 Sep 28. PMID: 27693092.

Nevoigt E. Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev.* 2008 Sep;72(3):379-412. doi: 10.1128/MMBR.00025-07. PMID: 18772282; PMCID: PMC2546860.

Nishiguchi H, Hiasa N, Uebayashi K, Liao J, Shimizu H, Matsuda F. Transomics data-driven, ensemble kinetic modeling for system-level understanding and engineering of the cyanobacteria central metabolism. *Metab Eng*. 2019 Mar;52:273-283. doi: 10.1016/j.ymben.2019.01.004. Epub 2019 Jan 8. PMID: 30633975.

Nishiguchi H, Liao J, Shimizu H, Matsuda F. Novel allosteric inhibition of phosphoribulokinase identified by ensemble kinetic modeling of *Synechocystis* sp. PCC 6803 metabolism. *Metab Eng Commun*. 2020 Nov 27;11:e00153. doi: 10.1016/j.mec.2020.e00153. PMID: 33312875; PMCID: PMC7721636.

Omage FB, Salim JA, Mazoni I, Yano IH, Borro L, Gonzalez JEH, de Moraes FR, Giachetto PF, Tasic L, Arni RK, Neshich G. Protein allosteric site identification using machine learning and per amino acid residue reported internal protein nanoenvironment descriptors. *Comput Struct Biotechnol J*. 2024 Oct 23;23:3907-3919. doi: 10.1016/j.csbj.2024.10.036. PMID: 39559776; PMCID: PMC11570862.

Paysan-Lafosse T, Blum M, Chuguransky S, et al. InterPro in 2022. *Nucleic Acids Research*. 2023 Jan;51(D1):D418-D427. DOI: 10.1093/nar/gkac993. PMID: 36350672; PMCID: PMC9825450.

Pei XY, Erixon KM, Luisi BF, Leeper FJ. Structural insights into the prereaction state of pyruvate decarboxylase from *Zymomonas mobilis*. *Biochemistry*. 2010 Mar 2;49(8):1727-36. doi: 10.1021/bi901864j. PMID: 20099870; PMCID: PMC2855724.

Peleg Y, Vincentelli R, Collins BM, Chen KE, Livingstone EK, Weeratunga S, Leneva N, Guo Q, Remans K, Perez K, Bjerga GEK, Larsen Ø, Vaněk O, Skořepa O, Jacquemin S, Poterszman A, Kjær S, Christodoulou E, Albeck S, Dym O, Ainbinder E, Unger T, Schuetz A, Matthes S, Bader M, de Marco A, Storici P, Semrau MS, Stolt-Bergner P, Aigner C, Suppmann S, Goldenzweig A, Fleishman SJ. Community-Wide Experimental Evaluation of the PROSS Stability-Design Method. *J Mol Biol*. 2021 Jun 25;433(13):166964. doi: 10.1016/j.jmb.2021.166964. Epub 2021 Mar 27. PMID: 33781758; PMCID: PMC7610701.

Perozich J, Hempel J, Morris SM Jr. Roles of conserved residues in the arginase family. *Biochim Biophys Acta*. 1998 Jan 15;1382(1):23-37. doi: 10.1016/s0167-4838(97)00131-3. PMID: 9507056.

Pleiss J. Protein design in metabolic engineering and synthetic biology. *Curr Opin Biotechnol*. 2011 Oct;22(5):611-7. doi: 10.1016/j.copbio.2011.03.004. Epub 2011 Apr 20. PMID: 21514140.

Pohl M. Protein design on pyruvate decarboxylase (PDC) by site-directed mutagenesis. Application to mechanistical investigations, and tailoring PDC for the use in organic synthesis. *Adv Biochem Eng Biotechnol*. 1997;58:15-43. PMID: 9103910.

Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2007 Jan;35(Database issue):D61-5. doi: 10.1093/nar/gkl842. Epub 2006 Nov 27. PMID: 17130148; PMCID: PMC1716718.

Quinn L, Armshaw P, Soulimane T, Sheehan C, Ryan MP, Pembroke JT. *Zymobacter palmae* Pyruvate Decarboxylase is Less Effective Than That of *Zymomonas mobilis* for Ethanol

Production in Metabolically Engineered *Synechocystis* sp. PCC6803. *Microorganisms*. 2019 Oct 27;7(11):494. doi: 10.3390/microorganisms7110494. PMID: 31717863; PMCID: PMC6920748.

Ren X, Wei Y, Zhao H, Shao J, Zeng F, Wang Z, Li L. A comprehensive review and comparison of L-tryptophan biosynthesis in *Saccharomyces cerevisiae* and *Escherichia coli*. *Front Bioeng Biotechnol*. 2023 Dec 4;11:1261832. doi: 10.3389/fbioe.2023.1261832. PMID: 38116200; PMCID: PMC10729320.

Rennison A, Winther JR, Varrone C. Rational Protein Engineering to Increase the Activity and Stability of IsPETase Using the PROSS Algorithm. *Polymers (Basel)*. 2021 Nov 10;13(22):3884. doi: 10.3390/polym13223884. PMID: 34833182; PMCID: PMC8621346.

Roussou S, Albergati A, Liang F, Lindblad P. Engineered cyanobacteria with additional overexpression of selected Calvin-Benson-Bassham enzymes show further increased ethanol production. *Metab Eng Commun*. 2021 Jan 11;12:e00161. doi: 10.1016/j.mec.2021.e00161. PMID: 33520653; PMCID: PMC7820548.

Schütz A, Golbik R, Tittmann K, Svergun DI, Koch MH, Hübner G, König S. Studies on structure-function relationships of indolepyruvate decarboxylase from *Enterobacter cloacae*, a key enzyme of the indole acetic acid pathway. *Eur J Biochem*. 2003 May;270(10):2322-31. doi: 10.1046/j.1432-1033.2003.03602.x. PMID: 12752452.

Sedivy JM, Babul J, Fraenkel DG. AMP-insensitive fructose biphosphatase in *Escherichia coli* and its consequences. *Proc Natl Acad Sci U S A*. 1986 Mar;83(6):1656-9. doi: 10.1073/pnas.83.6.1656. PMID: 3006063; PMCID: PMC323142.

Sengupta A, Bandyopadhyay A, Schubert MG, Church GM, Pakrasi HB. Antenna Modification in a Fast-Growing Cyanobacterium *Synechococcus elongatus* UTEX 2973 Leads to Improved Efficiency and Carbon-Neutral Productivity. *Microbiol Spectr*. 2023 Aug 17;11(4):e0050023. doi: 10.1128/spectrum.00500-23. Epub 2023 Jun 15. PMID: 37318337; PMCID: PMC10433846.

Sharir-Ivry A, Xia Y. Quantifying evolutionary importance of protein sites: A Tale of two measures. *PLoS Genet*. 2021 Apr 7;17(4):e1009476. doi: 10.1371/journal.pgen.1009476. PMID: 33826605; PMCID: PMC8026052.

Shivange AV, Hoeffken HW, Haefner S, Schwaneberg U. Protein consensus-based surface engineering (ProCoS): a computer-assisted method for directed protein evolution. *Biotechniques*. 2016 Dec 1;61(6):305-314. doi: 10.2144/000114483. PMID: 27938322.

Smith, M.A., Arnold, F.H. (2014). Designing Libraries of Chimeric Proteins Using SCHEMA Recombination and RASPP. In: Gillam, E., Copp, J., Ackerley, D. (eds) *Directed Evolution Library Creation. Methods in Molecular Biology*, vol 1179. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-1053-3_22

Spence MA, Kaczmarek JA, Saunders JW, Jackson CJ. Ancestral sequence reconstruction for protein engineers. *Curr Opin Struct Biol*. 2021 Aug;69:131-141. doi: 10.1016/j.sbi.2021.04.001. Epub 2021 May 21. PMID: 34023793.

Sumida KH, Núñez-Franco R, Kalvet I, Pellock SJ, Wicky BIM, Milles LF, Dauparas J, Wang J, Kipnis Y, Jameson N, Kang A, De La Cruz J, Sankaran B, Bera AK, Jiménez-Osés G, Baker D. Improving Protein Expression, Stability, and Function with ProteinMPNN. *J Am Chem Soc*. 2024 Jan 24;146(3):2054-2061. doi: 10.1021/jacs.3c10941. Epub 2024 Jan 9. PMID: 38194293; PMCID: PMC10811672.

Sunden F, Peck A, Salzman J, Ressler S, Herschlag D. Extensive site-directed mutagenesis reveals interconnected functional units in the alkaline phosphatase active site. *Elife*. 2015 Apr 22;4:e06181. doi: 10.7554/eLife.06181. PMID: 25902402; PMCID: PMC4438272.

The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res*. 2017 Jan 4;45(D1):D158-D169. doi: 10.1093/nar/gkw1099. Epub 2016 Nov 29. Erratum in: *Nucleic Acids Res*. 2018 Mar 16;46(5):2699. doi: 10.1093/nar/gky092. PMID: 27899622; PMCID: PMC5210571.

Tokuyama K, Toya Y, Horinouchi T, Furusawa C, Matsuda F, Shimizu H. Application of adaptive laboratory evolution to overcome a flux limitation in an *Escherichia coli* production strain. *Biotechnol Bioeng*. 2018 Jun;115(6):1542-1551. doi: 10.1002/bit.26568. Epub 2018 Mar 8. PMID: 29457640.

Tungtur S, Meinhardt S, Swint-Kruse L. Comparing the functional roles of nonconserved sequence positions in homologous transcription repressors: implications for sequence/function analyses. *J Mol Biol*. 2010 Jan 29;395(4):785-802. doi: 10.1016/j.jmb.2009.10.001. Epub 2009 Oct 8. PMID: 19818797; PMCID: PMC2813367.

UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic Acids Res*. 2025 Jan 6;53(D1):D609-D617. doi: 10.1093/nar/gkae1010. PMID: 39552041; PMCID: PMC11701636.

van Zyl LJ, Schubert WD, Tuffin MI, Cowan DA. Structure and functional characterization of pyruvate decarboxylase from *Gluconacetobacter diazotrophicus*. *BMC Struct Biol*. 2014 Nov 5;14:21. doi: 10.1186/s12900-014-0021-1. PMID: 25369873; PMCID: PMC4428508.

Van Zyl LJ, Taylor MP, Eley K, Tuffin M, Cowan DA. Engineering pyruvate decarboxylase-mediated ethanol production in the thermophilic host *Geobacillus thermoglucosidasius*. *Appl Microbiol Biotechnol*. 2014 Feb;98(3):1247-59. doi: 10.1007/s00253-013-5380-1. Epub 2013 Nov 26. PMID: 24276622.

Velmurugan R, Incharoensakdi A. Metal Oxide Mediated Extracellular NADPH Regeneration Improves Ethanol Production by Engineered *Synechocystis* sp. PCC 6803. *Front Bioeng Biotechnol*. 2019 Jun 19;7:148. doi: 10.3389/fbioe.2019.00148. PMID: 31275934; PMCID: PMC6593046.

Vioque A. Transformation of cyanobacteria. *Adv Exp Med Biol*. 2007;616:12-22. doi: 10.1007/978-0-387-75532-8_2. PMID: 18161487.

Wadano, A., Nishikawa, K., Hirahashi, T. et al. Reaction mechanism of phosphoribulokinase from a cyanobacterium, *Synechococcus* PCC7942. *Photosynthesis Research* 56, 27–33 (1998). <https://doi.org/10.1023/A:1005979801741>

Weinstein JJ, Goldenzweig A, Hoch S, Fleishman SJ. PROSS 2: a new server for the design of stable and highly expressed protein variants. *Bioinformatics*. 2021 Apr 9;37(1):123-125. doi: 10.1093/bioinformatics/btaa1071. PMID: 33367682; PMCID: PMC7611707.

Williams, J. G. K.: Construction of specific mutations in photosystem II photosynthetic reaction center by genetic engineering methods in *Synechocystis* 6803., *Method Enzymol*, 167, 766-778 (1988).

Wilson RH, Hayer-Hartl M, Bracher A. Crystal structure of phosphoribulokinase from *Synechococcus* sp. strain PCC 6301. *Acta Crystallogr F Struct Biol Commun*. 2019 Apr 1;75(Pt 4):278-289. doi: 10.1107/S2053230X19002693. Epub 2019 Apr 2. PMID: 30950829; PMCID: PMC6450518.

Wu W, Wang Z, Cong P, Li T. Accurate prediction of protein relative solvent accessibility using a balanced model. *BioData Min*. 2017 Jan 24;10:1. doi: 10.1186/s13040-016-0121-5. PMID: 28127402; PMCID: PMC5259893.

Xu Y, Verma D, Sheridan RP, Liaw A, Ma J, Marshall NM, McIntosh J, Sherer EC, Svetnik V, Johnston JM. Deep Dive into Machine Learning Models for Protein Engineering. *J Chem Inf Model*. 2020 Jun 22;60(6):2773-2790. doi: 10.1021/acs.jcim.0c00073. Epub 2020 May 5. PMID: 32250622.

Yang J, Li FZ, Arnold FH. Opportunities and Challenges for Machine Learning-Assisted Enzyme Engineering. *ACS Cent Sci*. 2024 Feb 5;10(2):226-241. doi: 10.1021/acscentsci.3c01275. PMID: 38435522; PMCID: PMC10906252.

Yang Z, Zeng X, Zhao Y, Chen R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct Target Ther*. 2023 Mar 14;8(1):115. doi: 10.1038/s41392-023-01381-z. PMID: 36918529; PMCID: PMC10011802.

Yep A, Kenyon GL, McLeish MJ. Determinants of substrate specificity in KdcA, a thiamin diphosphate-dependent decarboxylase. *Bioorg Chem*. 2006 Dec;34(6):325-36. doi: 10.1016/j.bioorg.2006.08.005. Epub 2006 Oct 9. PMID: 17028071.

Yoshikawa K, Hirasawa T, Shimizu H. Effect of malic enzyme on ethanol production by *Synechocystis* sp. PCC 6803. *J Biosci Bioeng*. 2015 Jan;119(1):82-4. doi: 10.1016/j.jbiosc.2014.06.001. Epub 2014 Jul 9. PMID: 25022874.

Yu A, Xie Y, Pan X, Zhang H, Cao P, Su X, Chang W, Li M. Photosynthetic Phosphoribulokinase Structures: Enzymatic Mechanisms and the Redox Regulation of the Calvin-Benson-Bassham Cycle. *Plant Cell*. 2020 May;32(5):1556-1573. doi: 10.1105/tpc.19.00642. Epub 2020 Feb 25. PMID: 32102842; PMCID: PMC7203937.

Zhang H, Stephanopoulos G. Engineering *E. coli* for caffeic acid biosynthesis from renewable sugars. *Appl Microbiol Biotechnol*. 2013 Apr;97(8):3333-41. doi: 10.1007/s00253-012-4544-8. Epub 2012 Nov 21. PMID: 23179615.

Zheng J, Liu HQ, Qin X, Yang K, Tian J, Wang XL, Wang YR, Wang Y, Yao B, Luo HY, Huang HQ. Identification and Mutation Analysis of Nonconserved Residues on the TIM-Barrel Surface of GH5_5 Cellulases for Catalytic Efficiency and Stability Improvement. *Appl Environ*

Microbiol. 2022 Sep 13;88(17):e0104622. doi: 10.1128/aem.01046-22. Epub 2022 Aug 24. PMID: 36000858; PMCID: PMC9469711.

Zhengxu Gao, Hui Zhao, Zhimin Li, Xiaoming Tana and Xuefeng Lu : Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria, Energy Environ. Sci., 2012,5, 9857-9865

Zhou J, Zhang H, Meng H, Zhu Y, Bao G, Zhang Y, Li Y, Ma Y. Discovery of a super-strong promoter enables efficient production of heterologous proteins in cyanobacteria. Sci Rep. 2014 Mar 28;4:4500. doi: 10.1038/srep04500. PMID: 24675756; PMCID: PMC3968457.

Zhou HX, Pang X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. Chem Rev. 2018 Feb 28;118(4):1691-1741. doi: 10.1021/acs.chemrev.7b00305. Epub 2018 Jan 10. PMID: 29319301; PMCID: PMC5831536.

付録

1. 略語

2PG,2-phosphoglycerate; 3PG,3-phosphoglycerate; 6PG,6-phosphogluconate; 6PGDH,6-phosphogluconate dehydrogenase; 6PGL,6-phosphogluconolactonase; AcCoA,acetyl coenzyme A; ACO,cis-aconitate; aKG, α -ketoglutarate; Cit,citrate; CoA,coenzyme A; CS,citrate synthase; DHAP,dihydroxyacetone phosphate; E4P,erythrose 4-phosphate; ENO,enolase; F6P,fructose 6-phosphate; FBA,fructose-bisphosphate aldolase; FBP,fructose 1,6-bisphosphate; FBPase,fructose 1,6-bisphosphatase; FUM,fumaric acid or fumarase; G6P,glucose-6-phosphate; G6PDH,glucose-6-phosphate dehydrogenase; GAP,glyceraldehyde 3-phosphate; GAPDH,glyceraldehyde 3-phosphate dehydrogenase; GPM,phosphoglycerate mutase; ICD,isocitrate dehydrogenase; IsoCit,isocitrate; ME,malic enzyme; MAL,malate; MDH,malate dehydrogenase; NADH,nicotinamide adenine dinucleotide; NADPH,nicotinamide adenine dinucleotide phosphate; OGDC,oxoglutarate dehydrogenase complex; oxPPP,oxidative pentose phosphate pathway; PDH,pyruvate dehydrogenase; PEP,phosphoenolpyruvate; PFK,phosphofructokinase; PGI,phosphoglucose isomerase; PGK,phosphoglycerate kinase; PPC,phosphoenolpyruvate carboxylase; PRK,phosphoribulokinase; PS,photosystem; PTS,phosphotransferase system; PYK,pyruvate kinase; PYR,pyruvate; R5P,ribose 5-phosphate; RPE,ribulose-phosphate 3-epimerase; RPI,ribose-5-phosphate isomerase; Ru5P,ribulose 5-phosphate; RuBP,ribulose 1,5-bisphosphate; RuBisCO,ribulose 1,5-bisphosphate carboxylase/oxygenase; S7P,sedoheptulose 7-phosphate; SBP,sedoheptulose 1,7-bisphosphate; SDH,succinate dehydrogenase; S7P,sedoheptulose 1,7-bisphosphatase; SUC,succinate; TAL,transaldolase; TCA cycle,tricarboxylic acid cycle; TKL,transketolase; TPI,triase isomerase; Xu5P,xylulose 5-phosphate

2. Phosphoenolpyruvate carboxylase (PPC) の保存度解析

大腸菌由来 PPC 配列 (NCBI アクセッション No. WP_001005582.1) をクエリとして、RefSeq Select proteins database (Pruitt et al., 2006) に対する Blast 検索によって E-value 上位 2000 配列を取得した。得られた配列について大腸菌由来 PPC 配列の番号付けで各位置における各アミノ酸残基の出現回数をカウントし、配列数で割ることで保存度とした。

3. 人工合成した PDC 遺伝子の DNA 配列

全ての配列は PCC6803 にコドン最適化されている。小文字は終止コドンと Shine-Dalgarno sequence を含む共通配列である。

>*Zymomonas mobilis*

ATGAGTTATACAGTTGGGACGTACCTGGCTGAGCGGTTAGTGCAAATCGGGCTGAAA
CATCATTTTGC GGTTGCAGGCGATTATAATTTAGTTCTACTGGACA ACTTGCTACTAAA

TAAGAATATGGAGCAGGTGTACTGTTGCAACGAATTAAATTGCGGATTCAGCGCAGA
GGGCTATGCGCGCGCGAAAGGAGCAGCGGCAGCCGTAGTCACGTATTCTGTGGGCG
CCCTCAGTGCCTTCGATGCGATTGGAGGTGCTTACGCAGAAAATCTGCCGGTGATCT
TGATCAGCGGCGCCCCCAATAACAACGATCACGCAGCAGGACACGTTTTACATCACG
CGTTGGGGAAGACTGATTACCATTACCAGCTAGAAATGGCCAAGAATATCACGGCGG
CTGCCGAGGCGATTTACACCCCTGAAGAGGCTCCCGCAAAAATTGACCACGTTATTA
AAACGGCTCTACGCGAGAAGAAACCCGTATACCTAGAGATCGCGTGCAACATCGCG
AGCATGCCCTGCGCCGCTCCAGGGCCGGCCAGCGCTTTGTTCAATGATGAGGCATCC
GATGAGGCTAGTTTGAACGCGGCAGTTGAGGAACTTTGAAGTTTATCGCCAATCGA
GATAAGGTCGCTGTATTGGTCGGTTCTAAACTGCGCGCAGCAGGAGCAGAGGAAGC
GGCCGTTAAGTTCGCTGATGCACTGGGTGGAGCGGTAGCTACCATGGCGGCGGCGA
AAAGTTTTTTTCCCGAGGAGAACCCCCATTATATTGGGACTAGTTGGGGTGAAGTCT
CTTATCCTGGCGTCGAAAAGACTATGAAAGAAGCGGATGCTGTTATCGCATTAGCGC
CTGTGTTCAATGACTATTCCACCACCGGCTGGACGGACATCCCCGACCCCAAGAAAC
TCGTGCTGGCTGAACCGCGAAGCGTGGTAGTTAACGGGATTCGTTTTCTTCTGTGC
ATCTGAAGGATTATTTGACACGCCTGGCCCCAAAAGGTGAGCAAGAAAACCTGGGGCA
CTCGATTTCTTCAAGTCCCTCAATGCAGGTGAACTGAAGAAGGCCGCTCCGGCCGAC
CCCTCCGCCCCCTTTAGTTAACGCTGAAATTGCACGACAAGTTGAGGCCCTGCTGACG
CCAAATACGACGGTGATTGCTGAGACGGGCGATAGTTGGTTTAAACGCGCAGCGTATG
AAGCTGCCCAACGGTGCCCGAGTTGAGTACGAAATGCAATGGGGCCACATCGGTTG
GTCCGTTCCGGCGGCCTTTGGCTACGCCGTCGGGGCACCGGAGCGTCGAAATATCTT
AATGGTTCGGTGACGGGTCTTTCCAGCTCACTGCGCAGGAGGTGGCCCAAATGGTAC
GTCTAAAGTTACCTGTCATCATTTTTTTGATCAACAATTACGGATATACCATCGAGGTC
ATGATCCATGATGGGCCTTACAACAACATCAAGAATTGGGATTACGCGGGGTTGATG
GAGGTGTTTAAACGGGAATGGGGGTACGACAGTGGCGCCGGGAAAGGGCTCAAAG
CTAAGACAGGCGGAGAATTAGCTGAAGCCATTAAAGTTGCATTGGCAAATACTGACG
GTCCGACTCTAATTGAGTGTTTCATCGGGCGTGAGGATTGCACTGAAGAATTAGTTA
AATGGGGTAAACGTGTCGCTGCTGCGAACTCTCGCAAACCAGTAAACAAGCTACTat
aatactagagtagtggaggttactag

>*Gluconacetobacter diazotrophicus*

ATGACCTACACGGTTGGTCGGTACTTGGCAGACCGACTGGCTCAGATTGGACTGAAA
CACCACCTTTGCCGTTGCTGGGGATTACAATCTGGTGTTACTCGACCAGTTGCTACTGA
ATACAGACATGCAACAAATTTATTGCAGCAATGAATTGAACTGTGGATTCTCTGCAGA
AGGCTACGCTCGCGCGAACGGAGCGGCTGCGGCGATCGTGACTTTTTCCGTGGGAG
CTTTGAGTGCAATTAATGCGTTGGGTGGCGCATACGCTGAGAATCTCCCTGTCATCCT
GATTAGCGGTGCGCCCAATGCAAATGACCACGGGACTGGGCACATCCTACACCATAC
ACTAGGGACGACCGACTACGGGTATCAGCTGGAAATGGCGCGCCACATCACATGCGC

AGCAGAAAGTATTGTTGCGGCTGAGGATGCGCCTGCCAAGATTGACCACGTGATCCG
CACGGCACTGCGAGAAAAGAAGCCGGCCTACTTGGAATCGCCTGTAACGTGCTG
GTGCCCCATGTGTACGCCCTGGGGGTATTGACGCACTCCTGTCTCCACCCGCACCTG
ATGAAGCTTCTCTCAAAGCAGCCGTCGACGCAGCACTGGCCTTTATCGAACAGCGCG
GCTCCGTTACGATGTTGGTAGGCTCCCGCATTGCGGCCGCCGGCGCGCAAGCACAAG
CAGTAGCCTTGGCAGATGCTCTGGGTGCGCTGTAACGACAATGGCGGCCGCGAAG
TCTTTCTTCCCCGAGGACCACCCTGGATACCGCGGTCATTACTGGGGAGAAGTATCC
AGCCCAGGCGCGCAACAAGCAGTCGAGGGTGCTGATGGCGTAATCTGTCTCGCCCC
GGTTTTTAATGACTACGCAACCGTTGGTTGGAGCGCCTGGCCGAAAGGAGATAACGT
GATGCTAGTAGAGCGACATGCGGTGACCGTCGGGGGAGTGGCGTACGCGGGGATTG
ACATGCGAGATTTCTTAACACGGTTAGCAGCGCATAAGTTCGCCGTGACGCAACGG
CTCGCGGGGGAGCGTATGTCACACCCCAAACCTCCGGCCGCTGCGCCTACTGCCCCC
TAAACAATGCAGAGATGGCGCGGCAGATTGGGGCGCTACTGACCCCCCGGACTACG
TTAACCGCCGAAACAGGGGACTCTTGGTTCAATGCCGTGCGCATGAACTACCTCAT
GGAGCCCGGGTGGAACCTCGAAATGCAATGGGGCCATATTGGATGGTCTGTACCAGCC
GCTTTTGGGAACGCATTAGCTGCCCCAGAACGCCAACATGTACTCATGGTTGGTGAT
GGCAGCTTTCAACTGACGGCTCAAGAGGTGGCTCAAATGATTGACATGATTTACCC
GTGATCATCTTCCTAATCAATAACCATGGTTACACGATTGAGGTCATGATCCACGATG
GCCCTTACAACAATGTAAAAAACTGGGACTATGCAGGTTTGATGGAAGTCTTCAACG
CTGGTGAAGGGAATGGCTTGGGTCTACGAGCCCGGACTGGCGGCGAACTCGCTGCA
GCCATCGAGCAGGCGCGAGCAAATCGCAACGGCCCTACCCTAATCGAGTGCACACT
AGATCGTGATGACTGTACACAAGAGTTGGTAACCTGGGGTAAGCGTGTTGCCGCAGC
GAACGCGCGTCCGCCACGAGCAGGT

>Zea mays

ATGGAAACACTGCTCGCGGGTAATCCAGCAAATGGAGTGGCTAAACCGACATGCAA
CGGTGTAGGTGCTTTGCCAGTCGCTAATAGCCACGCTATTATCGCCACGCCTGCCGCC
GCGGCTGCGACGTTGGCACCAGCTGGAGCTACATTGGGGCGTCACTTAGCTCGACG
ACTCGTGCAAATTGGTGCAAGCGATGTATTTGCCGTCCCGGGAGACTTTAATCTCACT
TTGCTGGACTATCTCATCGCCGAGCCTGGATTAACTCTCGTAGGTTGTTGTAACGAAT
TAAACGCGGGGTACGCAGCGGACGGCTATGCACGTAGTCGCGGAGTCGGTGCCTGC
GCTGTTACCTTTACCGTCGGGGGCTTATCTGTTTTAAACGCGATTGCGGGGGCGTACT
CTGAAAATTTACCCGTGGTCTGTATCGTAGGAGGCCCGAATAGTAATGACTATGGTAC
TAATCGGATCTTACACCACACCATTGGATTGCCGGATTTTAGTCAAGAATTACGATGT
TTCAAACCTATCACCTGTTACCAGGCCATTATTAACAACCTAGATGACGCGCATGAGC
AGATTGATACGGCTATCGCAACAGCCCTGCGAGAGTCTAAGCCCGTTTATATCAGCGT
CTCTTGTAATCTAGCAGGTCTGAGTCACCCACGTTCTCTCGCGATCCAGTACCGATG
TTCATTAGCCACGGTTAAGTAACAAGGCTAATTTGGAATATGCAGTGGAAGCTGCC

GCGGACTTTCTGAATAAGGCAGTAAAACCAGTCATGGTTGGCGGGCCAAAAATTCGT
GTAGCGAAGGCACGAGAGGGCGTTTGCTGCAGTGGCTGACGCCTCTGGCTATCCATTC
GCTGTCATGCCTGCTGCCAAAGGCCTCGTACCTGAACATCACCCGCGTTTCATTGGC
ACTTACTGGGGAGCTGTGTCTACTACGTTTTTGCGCCGAGATTGTCGAGTCTGCGGAT
GCTTACCTCTTCGCAGGCCCCGATTTTCAATGACTACTCCTCCGTGGGCTACTCTCTCC
TATTAACGAGAAAAGGCTGTAATCGTGCAGCCGGATCGAATGGTAGTTGGGGACG
GGCCGGCGTTTCGGCTGCATCCTGATGCCGGAGTTCCTCCGAGCCCTAGCCAAGCGAT
TACGACGTAATACCACTGCATACGACAACCTACCGGCGTATCTTTGTTCCGGATCGTGA
ACCGCCAAATGGAAAGCCTAATGAACCGCTACGGGTAAATGTTCTGTTTAAACATATC
AAAGGCATGCTAAGCGGGGACTCTGCTGTCGTCGCCGAGACTGGTGATAGCTGGTTC
AATTGCCAAAAGCTCCGACTACCAGAAGGTTGTGGCTATGAGTTCCAAATGCAGTAC
GGGAGTATTGGGTGGAGCGTTGGAGCTACACTCGGTTACGCTCAAGCAGCAAAAAGA
TAAGCGGGTAATCGCGTGCATTGGGGATGGGTCTTTTCAAGTTACGGCCCAAGATGT
GAGTACTATGCTACGCTGCGGTCAAAAAAGCATCATCTTCCTCATCAATAACGGGGG
TTACACCATTGAAGTAGAAATCCACGACGGACCCTATAATGTCATCAAAAACTGGGA
TTACTGTCCTGGTGAACGCTATCCACAACCTCCGAAGGCAATTGCTGGACCATGAA
AGTTCGTACAGAAGAGCAATTGAAAGAGGCAATTGCGACAGTCACGGGCGCCAAAA
AGGATTGTCTGTGTTTCATCGAAGTCATTGTACATAAGGACGATACGTCTAAGGAGTT
ATTGGAATGGGGTAGCCGGGTATCTGCCGCGAACAGTCGACCTCCAAACCCCCAAtaa
tactagagtagtgagggttactag

>*Arabidopsis thaliana*

ATGGACACGAAAATTGGGTCCATCGACGATTGCAAACCGACGAATGGAGACGTATGT
AGTCCAACCAACGGTACTGTAGCAACAATTCACAACCTCCGTTCCATCCTCTGCCATC
ACCATCAATTATTGTGATGCAACCCTGGGCCGCCACTTAGCCCCGGCGATTAGTACAAG
CCGGAGTCACAGATGTGTTTACGCGTTCCCGGCGATTTTAACTAACATTATTGGATCA
CCTGATGGCCGAACCAGATCTCAACCTAATCGGTTGTTGTAACGAACTAAACGCAGG
ATACGCCGCAGACGGCTATGCACGGTCCCGCGGTGTAGGAGCATGTGTCGTCACGTT
TACTGTGGGGGGTCTCAGTGTCTCAATGCTATCGCAGGTGCTTATAGCGAGAATCTA
CCCTTGATCTGTATTGTGGGTGGCCCAAACCTCCAATGATTATGGTACTAACCGCATCT
TACATCATACGATCGGGCTCCCAGATTTTTCTCAAGAATTGCGTTGCTTTCAGACCGT
GACCTGTTATCAAGCCGTGGTGAATAACCTAGATGATGCCACGAACAGATCGATAA
AGCTATCAGCACTGCGCTCAAAGAGTCCAAGCCTGTTTACATCAGTGTATCCTGCAAT
CTGGCGGGCGATTCCACATCATACGTTTAGTCGAGACCCGGTCCCATTCTCCTTAGCGC
CACGCCTATCCAATAAAATGGGTTTGGAAGCAGCCGTGGAAGCTACATTAGAGTTTC
TAAATAAAGCAGTGAAGCCCGTAATGGTCGGCGGCCCAAATTACGGGTCGCGAAG
GCTTGCGATGCATTTGTTGAATTAGCGGACGCTTCTGGTTACGCACTGGCGATGATGC
CGAGCGCAAAGGGATTTCGTGCCCGAGCACCACCCCACTTTATCGGTACGTACTGGG

GCGCGGTCTCCACGCCCTTTTGCTCTGAAATTGTGGAGAGCGCTGACGCCTACATTT
 TTGCAGGCCCCATCTTCAACGACTACAGCAGCGTTGGTTACTCCTTACTCTTAAAAA
 AAGAAAAGGCCATTGTAGTCCAGCCCGATCGCATCACAGTCGCCAATGGGCCGACG
 TTTGGTTGCATCTTAATGAGTGACTTTTTTCGAGAACTGAGTAAGCGCGTAAAGCGA
 AACGAAACCGCCTACGAAAATTACCACCGAATTTTCGTTCTGAAGGTAAACCACTC
 AAGTGCGAGTCCCGTGAACCTCTACGAGTTAATACTATGTTTCAGCACATCCAGAAA
 ATGTTATCTAGTGAGACAGCAGTCATTGCCGAAACCGGGGACTCCTGGTTTAATTGC
 CAGAAGCTGAAATTACCAAAGGGGTGCGGCTACGAATTCCAAATGCAGTATGGGAG
 CATTGGCTGGAGCGTGGGAGCGACCCTCGGTTATGCTCAGGCAAGCCCTGAGAAAC
 GGGTCCTCGCATTTATTGGGGATGGGTCTTTCAGGTGACGGTCCAGGACATCTCCA
 CCATGCTGCGAAATGGACAAAAGACGATTATCTTCCTAATTAATAATGGAGGATATAC
 GATTGAAGTCGAAATCCATGACGGCCCTTACAACGTAATCAAAAACCTGGAACCTACAC
 TGGCTTAGTCGACGCCATTCACAACGGAGAAGGTAACCTGCTGGACCGCGAAGGTAC
 GGTATGAGGAAGAGTTGGTAGAAGCGATTACGACGGCCACGACCGAGAAAAAAGA
 CTGTCTGTGTTTCATCGAAGTGATTTTACATAAAGACGACACGTCCAAAGAGTTACTA
 GAATGGGGCTCTCGCGTAAGTGCTGCAAACAGCCGGCCTCCAAACCCACAAAtaatactag
 agtagtggaggttactag

>*Crocospaera subtropica*

ATGGAAAATAATAAGAGTATTGGCAACTATCTAATTGAACGTCTCCTCCAACCTAGGGG
 TGAACCATGTATTTGGGGTTCCAGGGGACTTTGTGTTGGGGTTCAATAAGTTATTGGA
 GAAAAGTGAGTTGGAATTTATTAACACTTGCGACGAGCAAGGGGCCGTTTCGCCG
 CTGACGCTTACGCGCGCCTCCGCGGACTGGGGGTTGTATGTGTGACCTATTGCGTAG
 GCGGGCTAAAAATCGCCAACACAACCTGCGCAGTGCTTTGCGGAGAAAAGTCCTGTT
 GTAGTCATTTCCGGGTCTCCCGGCGTGAATGAGCGCACCAAAAATCCACTGTTACAT
 CACAAGGTAAAGGAATTCGATACACAGTACAAAGTGTTCCAGGAGATCACTGTTCGC
 GAGTACGGTCTTGGACAATCCGATACTGCTTACTCCGAAATCGAGCGTGTCTAAC
 AGCCGCCTTACGATATAAACGACCCGTATACATTGAAATTCCGCGAGATATGGTGAAT
 GTGACTTTAGGGGCAGATAACCAGCCGAGTTCTAATAATTCTAGCAGTAATCCGGAC
 GCTCTCCAGGAAGCCCTGGAAGAGGCGGTGAATCTCATCAACCAAGCACAAACCCC
 AGTCATTCTGGCAGGAGTCGAGATCCATCGTTTTTAAGCTCCAAGAATCTCTCTTAAA
 ATTAGTGGAGAAGACTAATCTCCCGGTAGCGGAGACCTTACTGGGGAAATCCGTAAT
 CAATGAAATGCACCCCAATAATCTGGGGATCTACGAGGGGGCGATGGGTAAAGAGTT
 CACCCGCAAATACGTAGAGGAAAGCGATTGTGTAATCGCCTTGGGAACCTTCTTGAG
 TGACGTGAATCTAGGAATTTTACAGCTAAATTGAATCCCCAGAGTTTCATTGATGTT
 AATTCTGAAAAGACCAGCATCCACTTTCACAACTACGAAGATATTAGTCTCGTGGAC
 TTTCTAAATGGCCTCTTGAATGCAGACTTGAAGCATCGCCAGGTTGACTTACCAAGT
 CGGTTTTTATTGCCACCAACTTCTCTGTCAAACCTGGGGAGAAGATCACGGTACAG

CGCCTCTTCGAGCGGCTAAACTTATTTATCTCTAACGACATGATTGTGATTGCAGACG
TTGGCGACGCTTTATTTGCGGGAGCCGACCTGGTGGTACACCAGAAAAGCCGGTTCT
TGTCCCCAGCATATTATGCCTCCTTAGGGTTTGCGGTCCCGGCAAGCATCGGGGCTCA
GATGGCTAATAGTACGTTGCGTCCGCTCGTGCTGGTCGGCGACGGAGCTTTCCAAAT
GACCGGTATGGAATTAAGCACTATCGTGCGATATGGACTCAATCCTATTATCATTGTGT
TAAATAACTTGGGGTATGGCACAGAGCGACCCATGCAGGACGGAAAATTTAATGACA
TCCTCTTATGGAATTACTCCCAACTACCTACGATTTTCAATGCGGGAAAGGGGTTTGA
CATTTCGGACAGAGGATGAGTTAGAGTTAGGACTAGAAAAATCTCAAACCTTATACCGA
AGGCTTTTGCATTTTAGACGTCCATTTAGACCCACAAGACTCCAGCCTCGCGCTCAA
ACGGCTGACGAAGGCATTGCGTGAGAAAGTCTaatactagagtagtgagggttagtag

>*Clostridium acetobutylicum*

ATGAAGAGTGAGTATACCATTGGACGTTATCTCCTGGACCGTCTCTCCGAGCTGGGA
ATTCGGCACATTTTTGGTGTACCTGGAGACTATAACCTGTCTTTCCTAGATTATATCAT
GGAGTATAAAGGTATCGACTGGGTAGGGAATTGTAATGAACTGAATGCAGGCTACGC
TGCTGATGGGTACGCCCCGTATTAATGGGATCGGAGCCATTCTCACTACATTTGGAGTT
GGTGAACCTCTCTGCAATCAATGCAATTGCAGGAGCATAACGCAGAGCAAGTACCTGTCT
GTAAAGATTACGGGAATTCCGACAGCCAAAGTACGTGACAATGGATTGTATGTCCAT
CACACGCTCGGAGATGGCCGTTTCGATCATTTTTTTGAAATGTTCCGAGAAGTTACG
GTCGCAGAGGCCTTATTAAGCGAAGAAAACGCGGCACAGGAGATCGATCGAGTGCT
AATTAGTTGTTGGCGTCAGAAGCGACCAGTGCTAATTAATCTCCCAATCGATGTATAT
GACAAACCAATCAATAAACCTCTGAAGCCGCTGTTGGACTACACTATTAGTAGTAAC
AAAGAGGCCGCTTGCGAATTTGTTACTGAAATTGTCCTTATCATTAACCGTGCGAAA
AAGCCCGTGATTTTAGCCGATTATGGCGTCTACCGCTATCAAGTACAACACGTTTTGA
AGAACTTAGCGGAAAAAACGGGCTTTTCTGTTGCCACACTCTCTATGGGGAAAGGG
GTATTTAACGAAGCCCACCCTCAGTTCATTGGTGTGTACAATGGGGACGTAAGCAGT
CCGTACCTCCGGCAGCGTGTAGACGAGGCGGATTGCATTATTAGTGTTGGAGTAAAA
TTGACAGACTCTACAACAGGAGGCTTCTCTCATGGATTCTCCAAACGCAACGTGATT
CATATCGATCCTTTTTCTATTAAAGCCAAAGGCAAGAAGTATGCTCCTATTACTATGAA
GGATGCCCTGACAGAGCTGACAAGTAAGATTGAGCATCGCAACTTCGAGGACCTGG
ATATTAAACCTTACAAAAGCGATAACCAGAAGTACTTCGCCAAAGAAAAGCCGATTA
CTCAGAAGCGATTTTTCGAGCGCATCGCTCACTTTATCAAGGAGAAAGACGTCCTAT
TAGCTGAACAGGGAACCTTGCTTTTTTGGAGCTTCCACCATCCAACCTGCCGAAAGATG
CTACGTTTCATCGGACAGCCGTTATGGGGCTCTATCGGGTACACGCTCCCGGCACTATT
GGGGTCCCAGCTCGCTGATCAAAAGCGTCGCAATATTCTCCTGATTGGTGTATGGGGC
TTTTCAGATGACCGCACAGGAGATCAGTACTATGCTCCGCCTACAAATCAAACCGATT
ATTTTCTTGATTAAACAACGATGGGTATACCATTGAACGCGCTATCCACGGCCGTGAGC
AGGTGTACAACAACATTCAGATGTGGCGGTACCACAACGTTTCCTAAAGTCTTGGGGC

CTAAGGAGTGCAGCCTCACCTTTAAGGTTTCAGTCCGAGACCGAATTGGAGAAGGCC
TTGTTGGTGGCAGATAAAGACTGTGAACACCTAATTTTCATCGAGGTGGTTATGGATC
GGTACGACAAGCCGGAACCGTTGGAGCGACTCAGTAAACGTTTCGCAAACCAAAAC
AATtaactagagtagtgagggttactag

>*Synechococcus* sp. (strain WH7803)

ATGCCCCCTAGCGTAGTTACATACGTACTAGACCGCTTGGCGGACCTGGGCATCGGTC
ATGTCTTCGGGGTTCCTGGTGACTATTCTTTCCCGCTCAATGACGCGGTTCGAGGTACA
TCCGCGATTGCAGTGGGTTCCTAAGTGCCAACGAACGCGAGCGTACGCGGCTG
ACGGCTACGCGCGTTCGACGCGGAGCCGGCATTGTATGTACAACCTTATGGTGTGGTG
AGCTCTCCGCACTAAACGGAGTTATGGGTAGCATGGCCGAACGGCTCCAGTATTTC
ATTTGGTCGGCACTCCATCCGTGCGGATTGTCCGGCAGGGTTTGATCTGCCACCATAAC
GTTGGGAGATACTCGCTATGATCGGTTTCGAAGCTATCAGTGCCGCAGCTGGCTGCGT
CTCCGCACGCCTGACACCCGAAAATGCTGTAGTCGAACTCGAACGCGTGATCGACA
AAGCGTTGGAAGACAGTCGGCCCGCATACCTGACTGTACCTATGGATTTGGCACTCA
TGCCTATTACCGGCACACCTATTCAGGGGACCCCCATCGGTAGCATTGACCAGCACG
CATCTGTACAGTTGAACTCGATGCCGTGTTGGATCTGGTTATGGAACGCCTAGCGA
AGGCGACGCGGCCTCTGGTAATGCCAACAGTTACTCTGAAGCGGTTTCGGACTCGTC
GAGACATTCGCGACCTTTCTCGAAGTCAGCGGTTTGGCCTATGCTACAACCTCCTATGG
ATAAAGCTCTGTTGTCCGAGGGACACCCAGCTTTTCTCGGGATGTACAACGGCGTAC
GCTCCACGCCCGCCGCCTTGCAGTCCGTGGTGGAGGGGGCGGATGTTTTACTGGATG
TCGGTGGTTTAGTAATGGAAGATCTCAATACTGGTCTCTGGAGTGGATACCTCGATAG
TCGCCGAGTTATCTCTCTGCACGCGGACTGGGTTCAAGCTGGCGATCAAGTATTCAC
GAGTGTCTCTCTGAGTGAGGTCCTGGCAGGTCTCATCAAGCGGTTTCAAGCAGCCG
ACGCGAAACCATCTCAATGGGGCGAACAGCGCCCTGTGCAACCGGAACCACTCATT
CCTCTAGCAGGCGAAGGAGATCAACCCACGGCTAGTGCCAACGTTTATCCTCGGCTC
CAGCGTTTTCTGCGTCCAACAGACTTACTCATGAGCGATACTGGGACTTCCTTGCTG
AAACTCAACGCCATGCGCCTCCCAGATGGGGTCGCCATTGAGACACAGACACTATGG
GGCAGTATTGGTTGGGCGACTCCGGCCGCACTGGGCTGTGCACTCGCCGATGCCGA
ACGCCGGGTAGTACTGGTGACGGGGGATGGAGCTCATCAGCTGACAGTACAAGAAA
TTGGGGTCATGGGTTTCATGAAGGTTAAGCCCGTCGTAATCGTACTGAACAATGGTCT
GTACGGTGTGGAAGCCTTGCTCAGCGAGACAGGTCATGCGTACAACGACCTCCCAC
CATGGCGGTACGCGCAACTGCCAGAAGCGTTTCGGTTGCCAGGGGTGGTGGTGTGGC
AAGGCGAGTACAGTTGCTGAGTTAGAGCAGGCATTGGCTGCAATCAACGCACATAAT
GGGGCTGCTTACCTAGAAGTAAGTATCCCGCCAGAAGAGAGCCTGCCCTTGCTGAA
GCAATGATTGAGACGCTGCATCAAACCTGCTACACCCtaactagagtagtgagggttactag

>*Aspergillus oryzae*

ATGGCGACAGATATCGCTACTCGTGACTTACGGAAGCCTATTGACGTGGCGGAATAC
CTGTTCCGTCGCTTACGTGAAGTTGGTGTACGAGCCGTCCACGGCGTGCCAGGCGAC
TATAATCTCGTAGCGTTAGATTATCTCCCCAAGTGCGACTTAACTGGGTTCGGAATT
GCAACGAGCTAAATGCGGGATATGCTGCCGATGGCTATGCCCCGTATTAACGGAATGTC
CGCCTTAGTTACTACTTTCGGTGTGGAGAGCTGAGTGCGCTCAACGCCATTGCTGG
GGCTTACTCCGAGTTTGTTCCTTGTCCATATTGTGGGTGAGCCTCATACCAAAAGT
CAGAAGGATGGGATGCTGCTGCACCACACGTTAGGGAACGGCGATTTTAACGTATTC
ACTCGAATGTCCGCGGATATTTCTTGTACATTGGGGTGTCTAAATTCTACACATGAGG
TCGCCACGCTCATCGACAATGCGATCCGAGAATGTTGGATTCTAGCCGCCCTGTTTA
CATTTCCCTCCCTACAGACATGGTAACGAAGAAAATTGAAGGGGAACGCCTAGATAC
GCCCTTAGACCTGTCTTTACCACCCAACGACCCCGAGAAAGAGGATTACGTCGTCGA
TGTCGTAATAAAATACTTACACGCGGCAAAAAAGCCTGTAATTCTCGTTGATGCGTGT
GCAATTCGCCATCGCGTGTAGACGAGGTACACGAGTTCGTAGAGAAGAGCGGTCTA
CCGACTTTTGTGCGGCCAATGGGTAAAGGGGCGGTTGATGAGACCCACAAAAATTAC
GGTGGTGTCTATGCAGGAACAGGGTCCAATCCCGGTGTGCGAGAACAGGTGGAGAG
CAGCGACCTCATCCTAAGTATTGGGGCAATCAAATCCGACTTTAATACAACGGGCTTC
AGTTACCGTATTGGGCAGTTAAATACGATTGACTTTCCTCTACGTACGTCCGAGTCC
GTTACTCTGAATATCCAGACATTAACATGAAAGGCGTACTCCAAAAAATCGTGCAGC
GAATGGGGAATCTGAATGTGGGACCAGTTTCTCCGCCTTCCAACCTGCTACCTGATA
ACGAAAAAGCGTCTACAGAACAAGCGATTACGCATGCATGGCTCTGGCCGACCGTC
GGTCAATGGCTAAAAGAGAAGGATGTCTGATCACTGAAACCGGGACTGCTAACTT
TGGGATTTGGGATACTCGCTTCCCGGCCGCGCTCACCGCAATTTCCCAAGTGCTCTG
GGGCAGCATTGGGTACTCCGTAGGCGCCTGTCAAGGCGCTGCCCTAGCTGCTAAAGA
ACAAGGGCGGCGGACTGTGCTATTCGTGGGCGATGGAAGTTTCCAGCTCACTCTGCA
GGAAGTAAGCACCATGATTCGAAATAATTTGAATCCGATCATTTTTGTTATCTGCAAC
GAGGGGTACACCATCGAACGATACATCCATGGGTGGGAGGCCGTCTACAATGACATC
CAACCATGGGACTTCCTAAATATTCCTGTAGCCTTTGGAGCAAAGGATAAATATAAGG
GCTACAAGGTAACCACGCGTGACGAACTCCGGGAATTGTTGCTAACGAGGAGTTC
GCTAGTGCGCCGTGCCTACAGTTGGTCGAGCTCCACATGCCACGGGACGACTGTCCG
GCTAGTCTAAAGCTGACGGCCGAGTCCGCGGCTGAACGAAACAAATCTCTGtaatactag
agtagtggaggttactag

>*Saccharomyces cerevisiae*

ATGAGTGAGATTAACTGGGTAAGTATTTGTTTGAGCGCTTGAAGCAGGTAAACGTG
AATACAGTCTTCGGGTTGCCAGGCGACTTCAATTTATCCCTCCTGGACAAGATTTACG
AAGTTGAGGGAATGCGTTGGGCTGGCAATGCAAATGAATTGAACGCGGCTTACGCG
GCAGACGGCTATGCGCGGATTAAGGGAATGTCTTGTATTATCACAACCTTTCGGGGTG
GGGGAATTGTCTGCGCTCAACGGGATTGCCGGTTCCTACGCCGAGCATGTCGGAGTG

CTACACGTAGTAGGGGTGCCCAGTATCAGTGCTCAGGCAAAACAATTGCTGCTACAT
CACACGTTGGGGAATGGGGACTTTACAGTATTCCATCGCATGAGCGCGAATATCAGT
GAAACGACTGCTATGATTACGGATATTGCCACGGCCCCTGCGGAAATTGACCGATGT
ATCCGCACCACCTATGTTACACAACGACCCGTTTATTTGGGCTTGCCAGCTAATTTGG
TAGATTTAAATGTGCCCCGCTAAGTTACTCCAGACTCCGATCGATATGTCCTTAAAGCC
TAATGATGCCGAGAGTGAAAAGGAAGTAATTGATACCATCCTGGCACTCGTTAAAGA
TGCCAAGAACCCAGTTATCCTCGCTGATGCATGTTGTAGTCGACATGACGTCAAAGC
GGAGACAAAAAAGCTCATTGACCTAACGCAATTTCTGCGTTCGTGACGCCAATGGG
TAAGGGAAGCATCGATGAGCAGCATCCACGCTACGGGGGCGTCTACGTTGGCACGCT
CTCTAAGCCAGAAGTTAAAGAAGCTGTGCAATCCGCTGATCTCATTTTAAGCGTTGG
TGCCTTATTGTCTGATTTCAACACGGGATCTTTCAGTTATTCTTACAAAACGAAAAAC
ATTGTGCAATTTATAGCGACCACATGAAGATCCGCAATGCCACATTCCAGGAGTAC
AAATGAAGTTTGTTTTACAGAAGTTGCTGACGACTATTGCCGACGCCGCAAAAGGAT
ACAAACCTGTGCTGTACCCGCGCGTACTCCCGCCAATGCTGCAGTACCTGCGTCCA
CGCCCCTGAAGCAAGAATGGATGTGGAATCAATTGGGCAACTTCTTGCAGGAAGGA
GATGTCGTTATCGCCGAGACAGGTACTTCTGCGTTTGGTATTAACCAAACGACTTTTC
CCAATAATACCTACGGTATTTCCCAAGTTCTGTGGGGTTCTATCGGTTTCACTACTGGT
GCCACTTTGGGTGCGGCATTTCGCCGCGGAAGAAATTGACCCCAAGAAGCGAGTTAT
CTTATTCATTGGAGACGGAAGCTTGCAGCTAACCGTGCAGGAAATCTCCACGATGAT
TCGTTGGGGGTTAAAACCATATTTATTTGTGCTAAACAATGATGGGTACACGATCGAG
AAACTGATCCACGGTCCCAAGGCTCAGTACAACGAAATTCAGGGATGGGATCACCTC
TCTTTGCTACCTACTTTTGGGGCAAAGGATTATGAAACTCATCGGGTAGCTACGACAG
GTGAATGGGATAAACTCACTCAAGACAAAAGCTTTAACGATAACTCTAAAATTCGGA
TGATCGAGATTATGTTACCTGTGTTTCGACGCCCCGCAAAATTTAGTTGAGCAAGCGA
AACTCACAGCCGCAACGAATGCAAAGCAAaataactagagtagtggaggttag

>*Gluconobacter oxydans*

ATGACTTATACCGTAGGCCACTATCTAGCAGAGCGCTTGACGCAAATTGGTTTAAAGC
ACCACTTTGCCGTAGCAGGCGATTATAACCTCGTGTTGTTAGATCAACTGATTGAACA
GGGTGGGACTAAGCAAATCTATGATTGTAATGAATTGAACTGTAGTTTCGCAGCTGA
AGGCTACGCACGCGCAAATGGAGCGGCCGCGGCTGTAATTACATTCAGTGTTGGGGC
GATCTCCGCAATGAATGGATTGGGGGGCGCATAACGCGGAAAATCTCCCCATTCTCGT
GATTAGCGGGGCCCCGAACCTCTAACGATCATGGTTCTGGACATGTTCTACATCATACA
ATCGGTACAACCTGACTACTCCTACCAGATGGAGATGGCTAAGCATGTTACATGCGCA
GCAGAATCCATTACAAGTGCAGAGACAGCGCCCCGCCAAAATTGATCACGTAATCCGC
ACAATGCTCCGTGAGAAGAAGCCTGCATTTCTCGAGATCGCATGTAATATCAGCGCC
GCACCCTGCGTACGGCCTGGACCAGTGAGCAGTCTCCATGCACACCCCCGTCCCGAT
GAAGCGAGTCTAAAGGCGGCCCTAGACGAATCTCTCTCCTTTCTCAACAAAACAAAT

AAGGTTGCCATCCTGGTAGGGACTAAACTACGGGCAGCAGAAGCTCTGAAAGAGAC
CGTGGAAGTAGCTGATAAATTGGGATGTCCTGTAACCGTTATGGCCGCGGCCAAGAG
CTACTTTCCTGAAACACACCCTGGGTTTCGTGGAGTATATTGGGGCGACGTGAGCTC
TCCGGGGGCTCAGGAAATTATTGAGGGGGCGGACGCGGTGATCTGCCTAGCACCAG
TATGGAACGACTACTCCAGTGGCGGGTGGAAATCCGTTGTGCGAGGTGAAAAAGTG
TTGGAGGTGGATCCTAACCGGGTTACAGTCAATGGGAAAACCTTTTGAAGGGTTTCGA
CTCAAAGAGTTTGTTAAGGCGCTCACCGAGAAGGCACCTAAGAAGAGCGCAGCGCT
GACCGGCGAGTACAAACAGTCATGCTACCGAAAGCAGATCCCAGTAAACCGCTGA
GCAACGACGAGATGACACGTCAAATCAACGAACTGGTCGATGGTAATACCACGCTAT
TCGCTGAGACAGGAGACTCTTGGTTCAATGCAGTGCGGATGCATTTGCCGGAAGGG
GCCAAAGTGGAACGGAGATGCAATGGGGGCACATTGGATGGAGCGTGCCCTCCAT
GTTTCGGCAATGCAACGGCATCTCCAGAGCGTAAACATGTCCTAATGGTCGGAGACGG
TTCTTTTCAATTAAGTCTCAAGAAGTAGCTCAGATGGTACGGTACGAGCTGCCGGT
TATTATTTTCCTAGTAAACAATCATGGCTATGTTATTGAGATGGCAATTCACGACGGAC
CGTATAATTACATTCAAAATTGGGATTATGCGGCTCTAATGCAGTGTTTTAATCAAGGC
GTTCTTGGAGAAGAGAGTGGCAAGTATGGACTAGGACTACACGCCACAACAGGCGC
GGAGCTGGCAGAGGCCATTGCAAAGGCTAAAAAAACACACGAGGTCCTACTCTGA
TTGAATGCAAACCTGGACCGCACTGACTGCACCAAGACGTTGGTGGAGTGGGGGAAG
GCAGTAGCAGCAGCAAACCTCTCGTAAACCGCAAAGCGTTtaatactagagtagtgagggttactag

>*Acetobacter pasteurianus*

ATGACGTACACAGTTGGCATGTACCTGGCTGAGCGCTTGGTGCAAATTGGGCTCAAA
CACCACCTTTGCCGTCGGTGGGAGACTACAACCTGGTGTACTGGACCAACTGCTCCTC
AATAAGGATATGAAGCAAATTTATTGTTGTAATGAGTTAAACTGTGGGTTTCAGCGCGG
AAGGTTACGCTCGATCCAATGGCGCGGCGGCCGAGTAGTTACATTTCCGTCGGAG
CGATTAGCGCCATGAACGCGCTAGGCGGTGCATACGCTGAGAACCTCCCTGTAATCT
TGATCAGCGGAGCGCCTAACAGTAATGATCAGGGGACCGGACACATTCTGCACCATA
CAATCGGTAAAACCGACTACAGCTACCAATTGGAGATGGCTCGTCAGGTGACCTGTG
CAGCAGAGAGTATTACTGACGCACACTCCGCACCAGCTAAAATTGACCACGTAATCC
GCACAGCCCTCCGCGAACGGAAGCCCGCCTATTTAGACATCGCTTGCAATATTGCTTC
TGAGCCTTGTGTCCGTCCGGGGCCGGTCAGTAGTCTCCTATCCGAGCCTGAGATTGA
TCATACTAGCCTGAAGGCGGCGGTGGACGCCACGGTCGCGCTCCTAAAAAATCGAC
CTGCCCCTGTTATGTTACTGGGGTCTAAGCTGCGCGCAGCGAATGCGTTGGCGGCAA
CAGAAACCCTAGCAGATAAGCTACAGTGCGCAGTAACCATCATGGCGGCCGCAAG
GGCTTTTTCCCGGAGGATCACGCTGGCTTCCGGGGACTGTATTGGGGCGAGGTTTCC
AACCCTGGTGTTTCAGGAGCTCGTCGAAACGAGCGATGCTCTCTTATGTATCGCACCT
GTTTTCAATGACTACTCCACAGTAGGTTGGTCTGGGATGCCCAAGGGGCCAAACGTG
ATTTTGGCCGAGCCTGACCGTGTAACGGTTGATGGACGTGCCTATGATGGTTTTACTC

TCCGGGCATTCTTACAGGCTTTGGCGGAGAAAGCCCCGGCACGCCCAGCTAGTGCC
CAGAAATCCTCCGTGCCAACTTGCTCCTTAACAGCCACGTCCGATGAGGCAGGGTTG
ACGAACGACGAGATCGTACGGCATATTAACGCTCTCCTCACGTCCAACACTACGTTG
GTCGCTGAGACAGGCGACTCTTGTTCAATGCAATGCGCATGACTCTGGCTGGAGCT
CGCGTGGAGTTAGAGATGCAGTGGGGCCATATCGGTTGGAGTGTACCGAGCGCATTC
GGGAACGCAATGGGTTCCCAGGATCGACAGCATGTGGTCATGGTAGGCGACGGATC
CTTCCAGCTGACCGCACAAGAAGTCGCCCCAAATGGTGCGTTACGAGTTACCAGTTAT
CATTTTTCTGATCAACAACCGTGGTTATGTCATCGAAATCGCCATCCACGACGGCCCA
TATAATTATATTA AAAACTGGGACTACGCTGGACTGATGGAGGTATTTAACGCGGGCG
AGGGTCATGGACTAGGCCTGAAAGCGACAACCCCGAAGGAATACTGAGGCAATC
GCACGGGCTAAAGCCAACACTCGGGGCCCTACACTCATCGAATGCCAGATTGACCG
GACGGATTGTACTGACATGCTGGTTCAATGGGGGCGCAAAGTTGCCAGCACAAATGC
TCGGAAGACGACCTTGGCAtaatactagagtagtgagggttag

>*Zymobacter palmae*

ATGTATACGGTCGGCATGTACCTAGCCGAACGCTTGGCTCAAATCGGGCTCAAGCAC
CATTTTGCTGTCGCAGGTGACTACAACCTTAGTACTCCTAGATCAGTTACTATTGAATA
AGGACATGGAACAGGTCTATTGCTGTAATGAGCTAAACTGTGGATTACAGCGCAGAAG
GGTATGCTCGAGCTCGTGGCGCTGCTGCAGCGATCGTCACTTTCAGCGTGGGGGCCA
TCTCCGCTATGAACGCTATCGGTGGAGCTTACGCGGAAAATCTCCCTGTAATCCTCAT
TAGCGGGTCCCCGAATACGAATGATTATGGCACGGGACATATTTTACATCATACGATC
GGTACGACCGATTACAACCTATCAGCTCGAGATGGTAAAGCATGTGACGTGTGCACGC
GAATCCATTGTTTCTGCGGAGGAAGCGCCAGCCAAAATTGATCACGTTATCCGCACT
GCCTTACGAGAGCGTAAACCTGCTTACTTAGAGATCGCCTGTAACGTTGCGGGCGCT
GAATGTGTGCGTCCAGGCCCAATTAATAGCTTGTTGCGCGAACTCGAGGTGGACCAA
ACATCTGTCACTGCTGCAGTCGACGCAGCCGTGGAGTGGTTACAAGATCGACAAAA
CGTGGTAATGCTAGTGGGGAGCAAACCTGCGTGCTGCGGCGGCGGAGAAACAGGCG
GTGGCGCTCGCCGATCGTTTAGGATGCGCAGTAACCATTATGGCTGCTGAGAAAGGA
TTCTTTCCGGAGGACCATCCTAATTTTCGCGGCCTGTACTGGGGTGAGGTTTCCTCCG
AGGGTGCGCAAGAGCTCGTGGA AAAATGCGGATGCTATTTTATGCCTGGCTCCCGTATT
TAACGACTATGCAACTGTCGGTTGGAATAGCTGGCCCCAAAGGTGACAATGTTATGGT
GATGGATACAGACCGAGTGACCTTCGCCGGTCAATCCTTCGAGGGCTTAAGCTTGAG
CACATTCGCTGCTGCCCTAGCGGAAAAAGCGCCGTCCCGGCCTGCCACAACGCAGG
GCACTCAGGCTCCGGTGCTGGGAATCGAAGCGGCCGAACCGAATGCGCCCTTGACG
AATGACGAGATGACCCGTCAAATTCAATCCCTCATTACGAGTGATACTACACTAACTG
CCGAAACGGGCGATAGTTGGTTTAATGCAAGCCGTATGCCCATTCCTGGCGGTGCCC
GTGTGCAACTCGAAATGCAGTGGGGACACATCGGGTGGAGCGTCCCGTCCGCTTTC
GGGAACGCTGTTGGAAGCCCAGAGCGCCGCCATATCATGATGGTCGGTGACGGGTCT

TTTCAATTAACGGCACAGGAAGTGGCCCAGATGATTTCGGTATGAAATCCCCGTTATTA
TCTTCCTCATTAATAATCGCGGGTACGTTATCGAAATCGCCATTACGATGGTCCTTAC
AACTATATTAAAACTGGAATTATGCAGGACTCATTGATGTCTTTAACGACGAGGACG
GCCATGGCCTCGGATTGAAGGCGTCCACGGGCGCAGAACTAGAGGGTGCCATCAAA
AAGGCGTTAGATAACCGACGAGGGCCTACTTTGATCGAATGCAATATTGCCCAGGAC
GACTGTACCGAAACCTCATCGCATGGGGTAAACGGGTAGCTGCGACCAATTCTCGG
AAACCACAAGCGtaatactagagtagtggaggttactag

>*Ancylobacter rudongensis*

ATGACATATACAGTGGGTAGCTATCTAGCCGCCC GCCTCGCGCAGATCGGCCTCAAGC
ACCACTTCGCTGTCGCTGGGGACTTTAACCTAGCCCTCCTAGACCAGTTGTTGACTA
ATGAGGAAACAGAGCAGGTGTATTGCTGTAACGAATTAACTGTGGTTACGCTGCGG
AGGGTTACGCGCGCGCTAATGGCGCCGCTGCCGCCGTCGTAACATTACGCGTCGGAG
GACTCAGTATTATTAACGCAATTGCGGGCGCGAATGCCGAAAATTTACCGGTCATTCT
AATTAGCGGTGCACCAAATACCAACGACCGCGCTACTGACCACTTACTACATCATAC
GATGGCGACAAGCGATCTGTCTTATCAGTTGGATATTGCAGCTCGTATTACCTGTTGC
GCAGTCGCTATCACAAGTCCAGAACAGGCTCCGAGCCAAATCGATACCGCTATCCGG
ACTGCTCTCCGCGAATCTAAACCCGCTTATATTGAAATTGCGTGCAATATCGCGGGAG
CACCTGTGCCGCTCCGGGACCAAGTTAGCGGATTGTTAGCGGGCGCACCTTCCGATG
CCGCAACATTAGAAGCCGCTGTAGCAGCAGCGGCTGATTTTCATCAAGGGACGAAAA
AAACCTGTCATTCTCATCGGGTGTAAGCTCCGGGCGGCCAAAGCGGAGCGAGCCGC
ATTAGCCCTAGCACAGGCGCTCGGCTGCCCTGTAGCCGTGATGGCCGCAGCAAAGA
GCTTTTTTCCCAGAAACGCATCCGCAATTTGCCGGGATTTACTGGGGCGAGGTGTCCG
CGCCAGGGACCCGCGAGATCGTAGACTGGAGCGACGGCCTGATTTGTATTGGCACG
CTATTTAACGATTACAGTACCGTGGGTTGGACGTCCAATCCCGCGGGTGCAAATGTGT
TAACACTCGACATCGACCACGTTACATTGGCCGGGCATGACTTCTCCGGACTACATTT
GGCGGAAGTTTTGGACGGACTAGCCGCACGTGTGGATGCCAACCCGACGGCTCTCA
ATGAATTCCACCGGTTGCGCGCGGCCCCCGCTACAGTACCAGAAGCCCCTCCGGAG
GCTAAGCTCTCTCGTACGGAAATGATGCGCCAGATTCAAGGATTACTACCGGCGAC
AGTACTGTATTGGTAGAGACGGGAGACAGCTGGTTTAACGGCATGCGACTCGCTCTG
CCTGAAGGGGCTCGCTTTGAGATTGAAATGCAGTGGGGTTCTATCGGGTGGGCGGTA
CCGGCCACATTCGGATATGCTGTTGGAGCACCTGGCCGTCGCGTCATCGCTATGATTG
GTGATGGCTCTTTTCAATTGACGGCCCAAGAAGTCGCACAAATGATTCGCCGCAAGC
TACCGATTATCATCTTTCTCATCAATAACCGAGGCTATACAATTGAAGTAGAGATTCAT
GACGGACCGTATAATAACATTAAGAAGTGGGACTATGCTGGGCTAATTAATGTTTTCA
ACGCGGAAGACGGCGAAGGCTTGGGGCTGCGGGCCGGAAGTGGCGGTGAATTGGC
AGAAGCGATCACTCGAGCGCTCGGAAACACTAAGGGCCCAACTTTGATTGAGTGTG

CACTCGACCGCGACGACTGCAGCCCTGATCTCATTAGTTGGGGCCGCCGAGTAGCCA
CTGCTAATGCTCGGCCCCCACAGCCTAAGtaatactagagtagtggaggttactag

>*Acetobacter vaccinii*

ATGACTTACACCGTCGGCACTTACTTGGCGGAACGATTAAGCCAAATCGGGCTGAAA
CACCATTTCGCCGTAGCGGGGGACTATAACCTAGTACTACTGGACCAACTGCTCCTA
AACAAAGACTGTGAGCAAGTATACTGTTGCAATGAGCTAAACTGCGGTTTCAGTGCT
GAGGGGTACGCTCGGGCAAATGGAGCGGCTGCAGCCGTCGTGACGTTCAATGTGCG
CGCATTCTCCGCATTAAATGCGATCGGAGGAGCTTATGCGGAAAATCTACCTGTGATC
CTCATTAGTGGTGCTCCTAACAGTAACGACTATGGTAGTGGGCACATCCTGCACCAC
ACGATTGGTGTACCTGACTATGGCTATCAGCTGGAAATTGCGAAAAAGCTCACTTGC
GCGGCGGTATCTATCACCAGCGCGGAAGACGCGCCAGAGAAAATCGATCATGCTATT
CGGACGGCTTTGCGCGAAAAAAGCCGGCGTACATCGAAATTTCTTGTAATGTAAGC
GCGCAGCCATGTGCCGCGCCAGGTCCGATTTCTGCCGTGCTGGATATTCGGCCTAGC
GATCAAGAACTCTAGCAGCCGCAGTTGCTGCAGTGGCGGAGTTCATCGAGAAACA
TGAAAAACCGGCCATCCTGATTGGGTCTCGCCTACGAGCGGCGGGGGCTGAAGCCG
CTGCAGTGAAGCTAGCCGACGCCTTGGGCTGTGTGGTCGCGACAATGGCCGCCGCG
AAATCTTTCTTCCCCGAAGATCACGCTGGCTATGTAGGTACGTACTGGGGGTCCGCTA
GCACACCTGGTGTTAACGAAATCTTCAATTGGACAGATAGTATCTTGGCTCTCGCCCC
GATCTTTAACGATTATAGCACCGAGGGGTGGACTGCTTGGCCGAAAGGGCCGAATGT
TGCAGTATTAGACAAAAATATTTCTAAGGTCGAAGGCCACAGTTTTGATAATATTCAT
TTAACGGAGCTGTTGGAGGCCCTAGCTGAGCACTTTAAAGATAAACCCAAAAAAGA
CGCTACATTGACACAGTATAAACGAGTCCACACTACTACTATTACGGAAACGACTGG
TAGTGGAGACGACAAGTTAAGCCGCGCAGATATTCAAAAAACCTTACAGTCCTTAGT
AACACCCGAGACAACCTATCCTAGGTGAGACAGGGGATTCTTGGTTCAATGTAGTCCG
CACTAAGTTACCCCGCGGGGCTCGTGTTGAATTTGAAATGCAGTGGGGTCATATTGG
CTGGAGCGTACCGGCAGCGTTTGGCTATGCTGTTGGAGCGCCTGAGCGGCAAACGA
TCCTGATGGTGGGTGACGGAAGCTTTCAGTTAACC GCGCAAGAGGTGCGCCAGATG
ACGCGCCGCAAGCTGCCGGTTATCATTTTCCTAATTAACAATGCGGGTTACACAATTG
AAGTTGAGATTCATGATGGGCCTTACAATAATATCAAAAACTGGGACTATGCGGCGGT
CGTGAACGCGTTCAACGCTGAGGACGGGAAGGGGAAGGGACTCCGTGCCACGACA
GCTGGGGAGCTCGCGGAGGCCGTCAAGGTCGCCCTAGATAACAAAGAAGGTCCCAC
GTTAATTGAGTGCGTGATCCCGCGTGATGACTGCACACCTGAGCTGATTTCTTGGGG
TCGCCACGTTGCTACAGCTAACGCACGTCCACCAGCAAAAAtaatactagagtagtggaggttactag

>ASR1

ATGACCTATACTGTGCGGACTTATCTCGCCGAGCGTTTGGCGCAAATCGGTCTCAAAC
ACCATTTCGCTGTGGCCGGCGACTACAATCTCGTATTGTTAGATCAGCTACTGAAAAA
TAAGGATATGGAACAGGTTTATTGTTGCAACGAGCTGAATTGCGGATTCTCCGCGGA

AGGCTACGCACGTGCGAACGGTGCTGCTGCCGCAGTGGTTACCTTTAGTGTCGGTGC
GCTGAGCGCATTCAATGCAATCGGTGGTGCCTATGCTGAGAATTTACCCGTAATCCTC
ATCAGTGGTGCACCCAATTCTAACGACCATGGTAGCGGTCATATTTTGCACCATACAA
TCGGAACCCCAGATTATTCTTACCAGCTAGAAATGGCTAAACATATCACCTGTGCGGC
AGTTAGTATTACCTCCGCCGAGGAAGCCCCTGCCAGATCGACCATGCAATCCGTAC
TGCATTGCGTGAGAAGAAACCTGCATATATTGAGATTGCATGTAACGTTGCTGGGGCC
CCATGCGCGGCACCAGGTCCAATCTCCGCTGTGTTGTCTGAAGAGCCGAGCGACGA
AGCTAGCCTCGAAGCGGCGGTTGATGCCGCGTTGGAATTCATTGAGAAACGCGAAA
AACCCGTCTTGCTCGTGGGTTCTAAACTACGAGCCGCCGAGCGGAAGAAGCTGCA
GTAAAGTTAGCCGATGCGTTGGGCTGCGCGGTAGCCACTATGGCGGCTGCGAAGAGT
TTTTTCCCAGAGGACCATCCTCATTATGTGGGAACGTATTGGGGCGAGGTTTCTTCCC
CAGGGGTCCAAGAGATTGTTGAGAGTAGCGACGGGGTAATTTGCCTGGGTCCCGTG
TTCAACGACTATAGCACAGTGGGATGGACCGCCTGGCCCAAAGGGGAAAACGTTCT
GTTGGCTGATCCACATCATGTAACGGTAGATGGGCACGCATTCTCCGGTATTCACTTG
CGAGATTTTTTAGCGGCGTTAGCAGAACGGGTTTCCAAGAAAGACGCAACACTGGA
GGAGTTCCGACGGATTACACACCTACTGCCGAAGTGGCTGCTGCAGCGCCCGACG
CACCTTTGACACGTGCTGAGATGGCTCGACAGATCCAGGGTCTACTGACTCCCAATA
CTACCGTCATCGCTGAGACGGGTGACTCTTGGTTCAATGCAATGCGAATGAAGCTCC
CTAATGGTGCGCGCGTGGAGTTCGAAATGCAATGGGGACACATTGGATGGAGTGTCC
CGGCAGCATTGGGATACGCGGTGGGCGCACCGAAGACGACGTGTCATCGCGATGGTC
GGTGACGGGAGTTTTTCACTGACCGCACAAGAAGTGGCGCAGATGATCCGACACAA
ACTGCCCCGTAATCATCTTCCTCATTAACAATCATGGGTACACAATCGAGGTAGAAATC
CACGATGGTCCTTACAACAATATCAAAAACCTGGGACTACGCTGGCCTGATCGAAGTA
TTTAACGCGGAAGACGGCCATGGGAAAGGTCTCCGTGCTAAGACTGGAGGCGAACT
CGCCGAAGCTATCAAGACTGCATTAGCGAATAAGGAAGGTCCTACCTTGATCGAGTG
TGCTATTGATCGGGATGACTGTACTAGTGAGTTAGTTAAATGGGGAAAACGTGTGCG
AGCAGCCAACAGCCGACCCCCTCGGGCTAAAtaatactagagtagtgagggtactag

>ASR2

ATGACTTATACGGTAGGCACGTATCTGGCCGAACGGCTGGCACAGATCGGTTTGAAG
CACCATTTCGCTGTAGCCGGTGACTACAATTTAGTACTACTCGATCAATTATTGCTCAA
TAAGGACATGGAACAAGTTTATTGTTGTAATGAATTAAATTGTGGCTTCTCTGCCGAA
GGGTACGCGCGAGCAAATGGCGCCGCTGCCGCGGTCGTAACGTTCTCCGTGGGAGC
TCTAAGTGCCTTTAATGCCATTGGAGGAGCATAACGCTGAAAATCTACCCGTCATTCTC
ATCAGTGGCGCTCCAAATAGCAACGATCACGGCTCTGGCCATATTTTACACCATACGA
TTGGGACTACTGATTATTCTTACCAACTGGAAATGGCCAAACATATTACCTGCGCCGC
TGAGAGCATTACATCCGCGGAGGAGGCCCCCGCAAAAATCGATCACGTGATTTCGGAC
AGCCCTCCGCGAAAAAAAACCTGCCTACCTGGAGATTGCGTGCAATGTCGCGGGTG

CTCCCTGCGCCGCTCCAGGCCCAATTAGCGCATTATTATCCGAAGAGGCCAGCGATG
AAGCAAGCCTAAACGCTGCAGTGGACGCTGCTCTCGAGTTTATCGAGAAACGCGAG
AAAGTAGTGCTATTAGTTCGGCAGTAACTACGCGCAGCGGGTGCCGAAGAGGCGGC
CGTGAAGCTGGCTGATGCCCTCGGATGTGCAGTCGCGACCATGGCCGCTGCGAAATC
TTTCTTTCCTGAGGACCATCCTCACTACGTTGGTACCTATTGGGGAGAAGTATCTAGT
CCGGGAGTCCAGGAGATTGTCGAATCCGCGGATGGAGTCATTTGCTTAGCACCGGTG
TTCAACGACTACTCTACGGTGGGATGGACGGCCTGGCCTAAGGGCGAAAATGTACTC
CTGGCTGAACCCCATCACGTTACTGTAGATGGGCACGCCTTTAGCGGCATTCACTTAC
GTGACTTCCTCACTGCGTTGGCTGAACGCGTTAGCAAGAAAGATGCTACACTCGAG
GGGTTCCGTCGGATTACACACCCGACCCTCGAGGTTGCGGCTGCAGCTCCAGACGC
GCCACTGACCAACGCCGAAATGGCTCGGCAAATTCAAGCATTACTACCCCCAAACA
CAACGGTGATCGCAGAAACGGGTGACAGCTGGTTCAACGCTATGCGGATGAAACTG
CCCAATGGCGCGCGGGTCGAATTTGAGATGCAATGGGGTCATATTGGATGGAGCGTG
CCAGCCGCATTCGGCTACGCCGTTGGGGCGCCTGAGCGCCGCCACATTCTGATGGTA
GGAGACGGTTCCTTTCAGCTAACGGCACAAGAAGTTGCGCAAATGATCCGCCATAA
GTTGCCGGTAATTATCTTTTAAATCAACAACCACGGATACACTATTGAGGTTATGATTC
ACGATGGCCCTTATAACAATATTAAGAATTGGGACTATGCCGGACTGATTGAGGTTTT
CAACGCAGAGGATGGACATGGGAAAGGACTGCGTGCTAAAACTGGGGGTGAGCTG
GCGGAGGCTATCAAACTGCATTGGCTAACAAGGACGGCCCGACCCTCATTGAGTGT
GCCATTGACCGCGATGATTGTACAGAAGAACTCGTCAAGTGGGGGAAACGAGTTGC
AGCTGCTAATAGCCGACCCCTCGGGCTAAGTaatactagagtagtgaggttactag

>PROSS1

ATGTCTTATACCGTGGGAACGTATCTCGCTGAGCGCCTCGTTCAGATTGGATTGAAGC
ACCACTTCGCCGTTGCGGGTGATTACAACCTAGTACTACTCGATCAGTTACTGTAAAA
TAAAAACATGGAACAAGTATATTGTTGCAACGAGCTAAATTGTGCGTTTGCCGCAGA
GGGGTACGCGCGTGCTAAAGGTGTAGCGGCGGCTGTCGTCACGTATAGCGTGGGCG
CACTAAGCGCTTTTGATGCTATTGGTGGGGCGTATGCAGAGAATCTACCAGTGATTTT
GATTAGCGGAGCACCAAACAATAACGATCATGCTGCAGGTCATGTTTTGCACCATGC
GCTGGGCAAAACCGACTATCATTACCAATTGGAGATGGCAAAAAATATTACCGCCGC
GGCGGAAGCAATTTATACACCTGAGGAGGCACCAGCTAAGATTGATCACGCAATCAA
GACCGCGTTACGGGAGAAGAAACCTGTTTACTTAGAAATTGCCTGTAACATTGCGAG
CATGCCCTGCGCGGCTCCAGGACCGGCGAGCGCTCTGTTTAAATGACGAAGCCTCCGA
TGAGGCGTCTCTCAACGCCGCGGTAGAAGAAACACTCAAATTTATTGCAAACCGTG
ATAAGGTTGTACTACTCGTCGGACCCAAACTACGTGCCGCCGGTGCAGAGGAAGCCT
TTGTCAAACCTAGCAGATGCTCTAGGATGTGCCGTAGCTACAATGCCGGCTGCAAAAG
GGTTTTTTCCCGAAGAGAACCCGCATTACATTGGTACATATTGGGGCGAGGTATCCAC
GCCAGGGGTAGAGAAGACCATGAAGGAGGCGGATGCCGTAATCGCTATCGCTCCAG

TTTTTAACGACTATTCCACGGGAGGTTGGACTCTAATCCCCGATCCTAAAAAATTGAT
TATCGCCGAACCTCGGTCTGTTGTGGTTAATGGAATCCGATTTCATCTGTACACCTG
AAGGACTACCTCACTCGGCTCGCCCAGAAGGTCTCCAAAAAGACAGCCGCCCTAGA
TTTTTTTAAGAGCCTGAATGCAGGTGAATTGAAGAAAGCTGCGCCAGCTGATCCGAG
TGCACCCCTGGTAAATGCCGAGATTTTCCGGCAGATTGAAGCATTACTGACCCCCAA
TACCACGGTCATCGTCGAGACTGGTGATAGTTGGTTTAATGCCCAGCGGATGAAGCT
GCCTAATGGGGCACGAGTGGAATACGAAATGCAATGGGGTCATATTGGATGGAGTGT
TCCGGCAGCCTTCGATACGCTCTCGGTGCTCCTGAGCGGCGTGTAATCCTCATGGT
CGGGGACGGTTCCTTCCAGCTAACTGCACAAGAAGTGGCCCAGATGGTGCGTTATAA
GCTCCCTGTTATTATCTTTCTAATCAATAATTACGGGTATACTATTGAGGTCATGATCCA
TGACGGACCTTACAACAATATTA AAAAATTGGGATTATGCGGGCCTCATGGAAGTTTTC
AACGGCAATGGAGGATATGATTCTGGGGCCGGAAAAGGATTGAAAGCAAAGACCGC
AGGGGAGTTAGCGGAGGCCATTAAGGTGGCTCTCGCAAATACTGACGGACCCACATT
GATTGAATGTTTTATTGGCCGGGAGGACTGTAGTGAAGAGTTAGTTAAGTGGGGTAA
ACGCGTAGCAGCGGCTAATTCTCGCAAACCCGTGAATAAGTTGCTCtaatactagagtagtgga
ggttactag

>PROSS2

ATGTCCTATACGGTGGGGACGTATCTGGCGGAACGCCTCGTTCAAATCGGACTAAAA
CACTACTTCGCTGTAGCGGGGGATTATAATCTCGTTTTGTAGACCAACTACTCGCTA
ACAAAAACATGAATCAAGTCTATTGCTGCAATGAGCTAAACTGTGTCTTCGCTGCGG
ATGGATATGCGCGGGCCCGAGGCGTTGCCGCTGCGTTGTGACGTTTTCTGTTGGCG
CGCTAAGTGCGTTTCGATGCTATTGGAGGAGCATACGCCGAAAATCTGCCTGTGATCCT
GATTAGCGGCGCCCCGAACAACAACGACCATGCTGCGGGTCACGTCCTACATCATGC
TTTGGGAAAGACAGATTACCACTACCAGTTGGAGATGGCCAAAAACATCACCTGCG
CCGCGGTGGCGATCTACACGCCTGAAGAAGCGCCTGCTAAAATCGACCACGCGATC
AAGACAGCGTTACGTGAAAAGAAACCCGTTTATATTGAGATCGCCTGCAACATTGCG
TCCATGCCTTGTGCTGCCCCCTGGTCCGGCCTCCGCTTTATTTAATGACGAGGCTTCCG
ATCCAGCAAGCTTAGAAGCAGCAGTCGAGGAACTTTGAAGTTTATTGCGAATCGA
GATAAAGTTGTCATCCTCGTAGGCCCCAACTCCGAGCTGCGGGGGCAGAGGAAGC
TTTTGTAAAGTTAGTAGACGCGCTCGGCTGCGCCGTAGCTGTTATGCCCGCCGCCAA
GGGCCTATTTCCGGAGGACAACCCGCACTTCATTGGGACCTATTGGGGCGAGATCTC
TACTCCCGGCGTGGAGGAAGTAATGAAAAGTGCTGATGCGGTCATTTGTATTGCGCC
CGTATTCAACGACTACTCTACGGCCGGCTGGACCCTGATTCCTGATCCGAAGAAGGT
AATCATCGTAGAACCACGCAGTGTCGTAGTGAACGGAATCCGCTTCCCGAGTGTCCA
CTTAAAAGATTTTCTCACAGCCTTGGCCCAGAAAGTACCAAAAAATACAGCTGCCTT
GGATTTTTTTTAAATCTCTAAATGCGGGCGAGTTAAAGAAGGCCGCCCCAGCGGACCC
TTCCGCGCCATTAGTTAACGCTGAAATTTCCGTCAGATTCAGGCCATGTTGACACCG

AATACTACGGTGATCGTGGAGACTGGGGACAGCTGGTTCAATGCACAACGCATGAA
ATTGCCAGAAGGGGCGCGGGTTCGAATTCGAGATGCAGTGGGGACACATTGGCTGGA
GCGTCGGCGCGGCTTTTCGGGTACGCACTAGCTGCGCCGGACCGTCGGGTCAATTCTAA
TGGTAGGAGATGGGTCTTTTCAGCTCACGGCACAGGAGGTCGCGCAGATGGTACGC
TATAAACTACCGGTTATCATCTTTTGGATCAATAACTATGGTTACACTATTGAGGTGAT
GATCCACGATGGACCCTATAACAACATCAAAAACCTGGGATTATGCTGGTCTCATGGA
GGTCTTCAATGGTAATGGGGGATACGACTCCGGTGCTGGTAAGGGTCTCAAAGCTAA
GACGGCCGAGGAACTCGCAGAGGCGATTAAAGTAGCCCTCGCTAACACTGACGGGC
CTACACTGATTGAGTGTTTTATCGGACGCGACGACTGCAGTGAAGAGCTGATCAAGT
GGGGGAAGCGGGTGGCGGCTGCAAATAGCCGGAACCTGTAAATAAGCTGTTAataac
tagagtagtggaggttag

>PROSS3

ATGTCCTACACTGTAGGAATGTATTTGGCGGAGCGATTAGTTCAGATTGGGCTCAAGC
ATCATTTTCGCGGTGGCCGGAGACTATAACTTGGTTCTATTAGATCAACTCTTACTCAA
CAAGAACATGGAACAAGTGTACTGCTGCAATGAGTTGAATTGCGGGTTTTCTGCGGA
AGGTTATGCCCCGTGCGAAGGGTGCAGCTGCAGCAGTAGTAACATATTCCGTTGGCGC
GTTGTCCGCTTTCGACGCGATCGGCGGAGCCTACGCCGAGAATCTACCTGTTATTTTG
ATTAGTGGCGCCCCCAACAACAACGACCATGCCGCGGGTCACGTGTTACATCACGCC
TTAGGAAAGACAGACTACCACTACCAACTCGAAATGGCGAAGAATATTACTTGTGCA
GCTGAGGCAATTTACACACCGGAAGAAGCCCCTGCAAAGATCGATCATGTGATTA
ACGGCTTTACGGGAGAAGAAGCCAGTTTATTTGGAGATCGCTTGCAATATTGCTTCTA
TGCCGTGCGCGGCTCCGGGACCGGCTTCTGCTTTATTTAACGACGAGGCTTCCGACG
AAGCTTCCTTAAATGCCGCTGTGGAGGAGACATTGAAATTTATTGCTAACCGAGACA
AGGTTGTATTGTTAGTCGGCTCTAAGTTACGAGCAGCCGGTGCTGAGGAAGCAGCGG
TTAAATTTGCGGATCGGCTGGGAGGTGCTGTAGCTGTTATGGCTGCCGCGAAAGGTT
TCTTCCCTGAAGAGAACCCACACTACATCGGGATTACTGGGGAGAAATTAGCTACC
CGGGAGTAGAGAAGCTGATGAAAGAGGCCGACGCTGTAATCTGTTTGGCACCCGTG
TTCAACGATTATAGCACCACGGGTTGGACGGATATTCCCGATCCGAAGAACTCATCC
TCGCAGAGCCGCGGAGCGTGGTTGTAAATGGGATTCGGTTTCCTTCCGTACATCTGA
AAGACTACTTAACGCGCTTAGCACAGAAGGTGCCCAAGAAAACCGCTGCTCTCGAC
TTTTTCAAATCTCTAAACGCTGGTGAATAAGAAAGCAGCCCCCGCGGATCCAAGC
GCCCCACTCGTAAACGCGGAGATTGCACGACAAGTAGAGGCTCTATTGACGCCGAA
CACCACCTTTGATCGCCGAGACAGGGGATTCCTGGTTCAACGCGATGCGTATGAAGCT
ACCGAATGGCGCGCGCGTTGAATACGAGATGCAATGGGGTCACATTGGGTGGTCTGT
CCCAGCAGCGTTCGGGTATGCTGTGGGTGCTCCCGAGCGCCGCAACATCCTCATGGT
AGGGGACGGATCCTTTTCAGTTGACAGCTCAAGAGGTGGCGCAGATGGTACGGTATA
AACTGCCTGTGATTATCTTCCTCATCAACAACCTACGGATATACGATTGAAGTAATGATC

CATGATGGGCCTTACAACAACATTA AAAATTGGGATTACGCGGGGCTGATGGAGGTT
TTTAATGGGAACGGTGGTTACGATTCCGGGGCCGGTAAGGGCTTAAAGGCCAAGAC
ACCTGGTGA ACTGGCCGAGGCCATTAAGGTTGCACTAGCAAACACGGACGGTCCGA
CTTTGATTGAATGCTTCATCGGTCGAGAGGATTGCACGGAGGAGCTCGTGAAATGGG
GAAAACGGGTTGCCGCCGCAAATAGCCGCAAGCCTGTTAACAAACTCCTAtaatactagag
tagtggaggttactag

謝辞

本研究を遂行するにあたり、社会人としての博士取得に向けて多大なるご配慮を賜り、研究と博士論文執筆の機会と場を与えてくださり、また研究内容について議論させていただく貴重な機会も頂戴し、研究室での活発な議論を通して学びと刺激を得ることができました、大阪大学大学院情報科学研究科バイオ情報工学専攻代謝情報工学講座、清水浩教授に、心より感謝申し上げます。また、研究に関する議論の中で多くのご助言をいただくとともに、研究室に伺った際にはいつも気さくにお声がけくださり、論文執筆においては細部にわたるご指摘と丁寧なご指導を賜りました、大阪大学大学院情報科学研究科バイオ情報工学専攻代謝情報工学講座、戸谷吉博准教授に、深く感謝申し上げます。研究室生活におけるさまざまな面でのご支援にも、心より御礼申し上げます。あわせて、私の知識が十分でないタンパク質研究に関して、多くの示唆に富んだご意見を賜り、研究を深めることができましたことに加え、日々の研究室生活においても温かいご支援をいただきました、大阪大学大学院情報科学研究科バイオ情報工学専攻代謝情報工学講座、二井手哲平助教に、心より感謝申し上げます。さらに、シアノバクテリアの培養に関する技術的なご指導や、研究設備の使用方法について丁寧に教えていただき、研究活動を円滑に進めることができました、大阪大学大学院情報科学研究科バイオ情報工学専攻代謝情報工学講座、今田辰海先生に、厚く御礼申し上げます。また、博士後期課程のバイオ情報工学専攻における発表の機会に折々にふれて、有益なご助言、ご指導をいただきました大阪大学大学院情報科学研究科バイオ情報工学専攻バイオ情報計測学講座、松田史生教授に心より感謝申し上げます。博士論文審査委員として、また、博士後期課程のバイオ情報工学専攻における発表の機会に折々にふれて、有益なご助言、ご指導をいただきました大阪大学大学院情報科学研究科バイオ情報工学専攻ゲノム情報工学講座、瀬尾茂人准教授に心より感謝申し上げます。また、日々の研究活動において、研究設備の使用方法をご丁寧にご教示いただき、研究室での会話も楽しく、大きな励みとなりました、大阪大学大学院情報科学研究科バイオ情報工学専攻代謝情報工学講座、一色衣香技術補佐員に心より感謝申し上げます。事務的な面においても、通常と異なる対応が必要な場面も多々ある中で、常に温かく、柔軟にご対応いただき、研究生活を円滑に進める上で多大なご支援を頂きました、大阪大学大学院情報科学研究科バイオ情報工学専攻代謝情報工学講座の事務員である石原かおりさん、嘉藤田裕美さんに、心よりお礼申し上げます。

最後に、博士後期課程の進学を後押しし、悔いのない研究生活を送らせてくださった妻に心から感謝いたします。