The University of Osaka
Institutional Knowledge Archive

Title	メタゲノム法により枝葉コンポストから単離した新規セルラーゼとエステラーゼの構造と機能に関する研究
Author(s)	岡野, 啓志
Citation	大阪大学, 2015, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/52152
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

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

The University of Osaka

博士学位論文

メタゲノム法により枝葉コンポストから単離した新規セルラーゼとエステラーゼの構造と機能に関する研究

Studies on the structures and functions of novel cellulases and esterase isolated from leaf-branch compost by a metagenomic approach

岡野　啓志

２０１５年１月

大阪大学大学院工学研究科

生命先端工学専攻　物質生命工学講座

極限生命工学領域　金谷研究室

目次

序論
1. 新規酵素取得におけるメタゲノム解析法の有用性
2. 好熱性新規酵素取得元としての枝葉コンポスト
3. セルラーゼ
4. 油脂分解酵素
5. 本研究の目的

第 1 章　*Rhodothermus marinus* 由来 Cel12A と高い相同性を示す新規 GHF12 セルラーゼ LC-CelA の構造と安定性の解析
1-1. はじめに
1-2. 実験材料および方法
1-2-1.　菌体、プラスミド
1-2-2.　DNA ライブラリーの構築とスクリーニング
1-2-3.　プラスミド構築
1-2-4.　大量発現と精製
1-2-5.　配列分析
1-2-6.　酵素活性
1-2-7.　CD (Circular dichroism) スペクトル
1-2-8.　熱変性測定
1-2-9.　結晶化
1-2-10.　X 線回折データ収集と構造決定
1-2-11.　Protein Data Bank accession number
1-3. 実験結果
1-3-1.　メタゲノム DNA ライブラリー由来セルラーゼ遺伝子のクローニング
1-3-2.　Pre-LC-CelA~J のアミノ酸配列
1-3-3.　Pre-LC-CelA と Pre-*Rm*Cel12A のアミノ酸配列比較
1-3-4.　LC-CelA と LC-CelA-His の大量発現と精製
1-3-5.　LC-CelA-His の活性
1-3-6.　LC-CelA-His の耐熱性
1-3-7.　LC-CelA-His の熱変性

1-3-8.　LC-CelA の活性と安定性
1-3-9.　LC-CelA の結晶構造
1-3-10.　ΔFL-LC-CelA-His と E34A-LC-CelA-His の安定性
1-3-11.　DTT 存在下における LC-CelA-His の安定性
1-3-12.　ΔFL-LC-CelA-His と E34A-LC-CelA-His の活性
1-4. 考察

第 2 章　N 末端に Ig-like ドメインを有する GHF9 セルラーゼ LC-CelG の構造、活性および安定性の解析
2-1. はじめに
2-2. 実験材料および方法
2-2-1.　プラスミド構築
2-2-2.　大量発現と精製
2-2-3.　酵素活性
2-2-4.　CD (Circular dichroism) スペクトル
2-2-5.　熱変性測定
2-2-6.　結晶化
2-2-7.　X 線回折データ収集と構造決定
2-2-8.　Protein Data Bank accession number
2-3. 実験結果
2-3-1.　LC-CelG のアミノ酸配列
2-3-2.　LC-CelG と His-LC-CelG の大量発現と精製
2-3-3.　His-LC-CelG の活性
2-3-4.　His-LC-CelG の耐熱性
2-3-5.　LC-CelG の結晶構造
2-3-6.　基質結合ポケット
2-3-7.　金属結合部位
2-3-8.　Ig-like ドメインと触媒ドメインの相互作用
2-3-9.　LC-CelG の活性や耐熱性における Ig-like ドメインの重要性
2-4. 考察

第 3 章　N 末端に長い伸長領域を有する新規エステラーゼ LC-Est1 の構造と諸特性解析
3-1. はじめに
3-2. 実験材料および方法
3-2-1. 菌体、プラスミド
3-2-2. プラスミド構築
3-2-3. 大量発現と精製
3-2-4. 酵素活性
3-2-5. CD (Circular dichroism) スペクトル
3-2-6. 熱変性測定
3-2-7. 結晶化
3-2-8. X 線回折データ収集と構造決定
3-2-9. Protein Data Bank accession number
3-3. 実験結果
3-3-1. メタゲノムから単離された新規エステル/脂肪分解酵素の同定
3-3-2. LC-Est1 – 6 のアミノ酸配列
3-3-3. LC-Est1 と CSu-Est、Tm-EstA のアミノ酸配列比較
3-3-4. LC-Est1、LC-Est1C、LC-Est1C*の大量発現と精製
3-3-5. LC-Est1C*の結晶構造
3-3-6. LC-Est1 と LC-Est1C の酵素活性
3-3-7. LC-Est1 と LC-Est1C の耐熱性
3-3-8. LC-Est1C*と Tm-EstAC の構造的特徴の比較
3-4. 考察

総括

参考文献

本研究に関する論文

謝辞

【序論】

1. 新規酵素取得におけるメタゲノム解析法の有用性

地球上に生息する多種多様な微生物が生産する酵素には、多様な基質特異性を有するものの他、高温、低温、高 pH、低 pH、高塩濃度環境に対応したものや、界面活性剤や有機溶媒などに対して高い安定性を有するものも多く、様々な用途に応用されている。産業等に利用される様々な酵素は、通常、その酵素を生産する微生物を土壌、湖沼、河川、海洋など自然界からスクリーニングすることにより獲得される。しかし、地球上に生息する微生物のうち、人工的な環境で単離・培養が可能なものは僅か 1%未満にすぎず、大半は難培養微生物である。ゆえに自然界には未知の新規有用酵素が数多く存在していると考えられる。

これら難培養微生物由来の遺伝子資源を単離する方法として、近年メタゲノム解析法が注目されている [1-6]。この方法は、環境中から微生物の単離・培養を介することなく、様々な微生物由来のゲノム DNA を環境サンプルから直接抽出し、遺伝子資源を解析する方法である。まず、環境中から抽出した DNA をシェアリングにより約 25 kb のサイズに断片化し、フォスミドベクターとライゲーションを行い、λファージに封入し、大腸菌を感染させることにより、多数のクローンからなるメタゲノムライブラリーを構築する。次に、得られたフォスミドライブラリーを用いて、目的の活性を有する酵素遺伝子のスクリーニングが行われる。この方法を用いると、これまで解析されたことのない難培養微生物や未発見の微生物が持つ酵素遺伝子を解析できるようになるので、従来よりも新規酵素遺伝子を効率的に取得できる。

2. 好熱性新規酵素取得元としての枝葉コンポスト

大阪府吹田市万博記念公園では、園内で発生する植物残材を園内で有効活用する「植物残材ゼロミッション」という取り組みが行われており、その一環としてコンポスト (堆肥) 化事業が行われている。本事業では、園内で剪定や伐採により発生した幹や枝葉を粉砕したもの (ウッドチップ) に尿素を添加して、屋外の堆肥化ヤードに積み上げ、水分調整や切り返しなどの処理を行いながら自然に発生する微生物による分解や発酵を促進させて堆肥化させている。この堆肥化過程で、積み上げられたウッドチップ内部の温度は発酵熱で最高 80°C に達する。そのため、本コンポストは多種多様な好熱性微生物の宝庫であると考えら

れる。またセルロースやエステルを含む細胞壁を分解する微生物が多く存在すると考えられる。

本研究室ではこれまでに、4 か月間発酵させて温度が 60°C に上昇した枝葉コンポスト（pH 6.0）からメタゲノム法により新規クチナーゼや新規リボヌクレアーゼ H (RNase H) [7] を単離している。このクチナーゼは細胞壁の外側にあるクチクラと呼ばれる透明で水を通さない層を形成するクチンを分解し、かつポリエチレンテレフタレート(PET)を分解するため、産業酵素としての利用が期待されている [8]。またこのクチナーゼは極めて変性速度の遅い耐熱性の高い酵素であることが分かった [12]。一方で、新規 RNase H のうち、LC9-RNase H1 はユニークな活性部位のアミノ酸モチーフを有する RNase H1 であること [9]、LC11-RNase H1 はユニークな基質認識機構を有する RNase H1 であることが発見された [10-11]。このように、これらの RNase H の研究は、RNase H の分子多様性や基質認識機構に関して新たな知見を与える。しかしこの枝葉コンポストからセルラーゼやエステラーゼはまだ単離されていない。

3. セルラーゼ

セルロースはグルコースがβ-1,4 結合により連なった直鎖状のポリマーであり、植物細胞壁の主成分である。これは植物の光合成により作られる再生可能な物質であるため、地球上に豊富に存在し、枯渇することのない資源である。近年「バイオリファイナリー」という概念が広まっており、エネルギーや化学品の原料としてセルロースのような草木系・木質系バイオマスに注目が集まっている。特にスイッチグラスなどのリグノセルロース系バイオマスは、食糧生産と競合しないために安定供給が可能である。この豊富な資源を工業的な規模でグルコースに転換する技術開発ができれば、微生物を用いた発酵によりアルコール、有機酸やアミノ酸といった化学製品を、化石資源に頼ることなく獲得することが可能となる。特に化石燃料由来のエネルギー資源は、環境汚染やその枯渇が大きな問題となっており、セルロースの有効活用はそれら問題の解決につながると期待されている。セルロースをグルコースに転換する方法としては、これまで酸糖化法の研究が進んでいるが、厳しい反応条件を必要とし、工業的規模で生産するにはコスト面が大きな問題となる。そこで比較的温和な条件で副産物を伴わずにグルコースまで分解する方法として、酵素による糖化が注目されている [13-17]。

細胞壁のセルロースは主要成分であり、水素結合を介してセルロース間を架橋するヘミセルロースにより、細胞壁に強度と柔軟性を与えている。またペクチンは隙間の拡充やネットワーク形成を行っている。このように細胞壁は複雑で強固な構造を取っている。セルロース糖化過程は、まず高温高圧条件下で物理化学的前処理を施し、セルロース分解酵素による糖化を行う。様々な酵素が細胞壁の分解に関わっているが、特にセルロースの加水分解に関わっている酵素を総称してセルラーゼと呼ぶ。これは基質の切断様式により次の 3 つに分類される [18, 19]。非結晶セルロース鎖の内部をランダムに切断するエンドグルカナーゼ (EC 3.2.1.4)、セルロースの還元・非還元末端からそれぞれ二糖単位で分解するセロビオヒドロラーゼ (EC 3.2.1.91, 3.2.1.176)、これら酵素により生じたセロビオースやセロオリゴ糖をグルコース単糖まで分解するβ-グルコシダーゼ (EC 3.2.1.21) である。糖質のグリコシド結合を加水分解する糖質加水分解酵素は、アミノ酸配列の類似性に基づき分類され、それらは CAZy database (http://www.cazy.org/) にまとめられている。現在までに Glycoside hydrolase family (GHF) 1 – 133 まで分類されており、エンドグルカナーゼ、セロビオヒドロラーゼ、β-グルコシダーゼはそれぞれ合計 12、5、6 種類のファミリーに分類される。。

セルラーゼは前述のバイオエタノール生産だけではなく、洗剤、食品加工、繊維加工、医薬品生産などその用途は多岐にわたる。酵素糖化に用いられる酵素として、現在は糸状菌 *Tricoderma reesei* 由来酵素が注目されている。本菌は細胞壁の分解に関わる酵素を全て有し、かつその分泌量が非常に多いことから、本菌の培養液を用いるだけで効率良く糖化が行えるという利点がある。一方で好熱嫌気性細菌 *Clostridium thermocellum* にも注目が集まっている。これはセルラーゼ、ヘミセルラーゼ等が規則的に配列し、巨大な酵素複合体として結晶性セルロースにも効率的に作用することができるセルロソームを生産する。また耐熱性に優れた酵素を生産することができるため、高温条件下での使用にも応用が可能である。糖化にかかるコスト削減は工業化に向けた最大の課題であるが、高温にするほど反応速度や基質の溶解度が上昇し、グルコース消費に伴う収率の低下を引き起こす雑菌の繁殖を防ぐことができるため、活性の高い新規耐熱性セルラーゼの単離が望まれている。またこの新規セルラーゼの構造や機能を解析することにより、セルラーゼの安定性や活性を高める方法について新たな知見を得ることが望まれている。

4. エステラーゼ/リパーゼ

エステラーゼやリパーゼは、脂肪酸エステルやトリグリセリドなどのカルボン酸エステルを分解する酵素であり、微生物からヒトまであらゆる生物が有している。エステラーゼとリパーゼは基質特異性の違いに基づき、短鎖カルボン酸エステルを優先的に分解するものがエステラーゼ (EC3.1.1.1)、高級脂肪酸エステルを優先的に分解するものがリパーゼ (EC3.1.1.3) と分類されている [20]。医薬品、食品の油脂加工、廃水中の油脂処理をはじめ、近年ではバイオディーゼル燃料の合成や生分解性プラスチックの加水分解といったバイオマス産業においても、エステラーゼ/リパーゼの用途は拡大している。油脂の分解やエステル合成を行う際に、水酸化ナトリウムの代わりにこれらの酵素を利用することにより、環境負荷が軽減されると期待されている [21]。このような様々な用途に適したエステラーゼ/リパーゼを効率よく見つけるためには、既存のエステラーゼ/リパーゼのバリエーションを増やす必要がある。従って、新規エステラーゼ/リパーゼの単離が望まれている。またこの新規エステラーゼ/リパーゼの構造や機能を解析することにより、エステラーゼ/リパーゼの安定性や活性を高める方法について新たな知見を得ることが望まれている。

エステラーゼはカルボン酸とアルコール間のエステル結合の加水分解を水溶媒中で触媒する [22]。まず触媒セリン残基のヒドロキシル基が基質によりアシル化され、アシル-酵素中間体を形成し、エステル結合を加水分解する。エステラーゼはα/β hydrolase 構造をとり、セリン、ヒスチジン、アスパラギン酸 (あるいはグルタミン酸) が触媒トライアドを形成している。このセリン残基は保存された 5 アミノ酸残基モチーフ Gly-X-Ser-X-Gly に位置する。エステラーゼの構造はリパーゼとよく似ている [22-23]。しかしリパーゼは水に不溶性の基質を分解するため、触媒部位を覆う lid 構造を有している。この lid 構造が開くことにより、基質を認識することが可能になる [24]。一方でエステラーゼの基質は水溶性であるため、このような lid 構造は有していない。これらエステラーゼ/リパーゼは、アミノ酸配列や生化学的な特徴に基づき、8 つのファミリー (Family I – VIII) に分類される [79]。エステラーゼは様々なエステル結合の加水分解を触媒し、高い構造安定性を有し、立体選択性を示す。また、有機溶媒中では合成反応を促進するため、産業用生体触媒として非常に注目されている [20, 25, 26]。

5. 本研究の目的

一般的に高温環境で安定な酵素は、生産コストが抑えられ、長期間の保存が可能となる傾向にあるため、熱安定な酵素は産業利用上有用である [27-29]。また反応温度が高くなると、反応速度が増加し、基質や生産物の可溶性が増加し、反応溶液の粘度が低下し、コンタミネーションを防ぐことが可能である。特に、上記にも示した通り、セルラーゼによる分解産物であるグルコースの収率を上げるためには、コンタミネーションを防ぐことは不可欠である。そこで本研究では、枝葉コンポストからメタゲノム法により耐熱性のある新規セルラーゼと新規エステラーゼ/リパーゼを単離することを目的とした。そのために、まず枝葉コンポストから抽出したメタゲノムを用いて遺伝子ライブラリーを構築し、CM セルロースやトリブチリンを含むプレート上でハローを形成するコロニーをスクリーニングした。その結果、10 種類の新規セルラーゼ遺伝子と 6 種類の新規エステラーゼ/リパーゼ遺伝子をクローニングすることに成功した。これらの遺伝子がコードする酵素のうち、２つのセルラーゼ LC-CelA と LC-CelG、および１つのエステラーゼ LC-Est1 について、その構造ならびに諸特性を解析した。LC-CelA は N 末端にフレキシブルリンカー（FL）を持つが、非常に耐熱性の高い *Rhodothermus marinus* 由来セルラーゼと高い相同性を示す。LC-CelG は N 末端に Ig-like ドメインを持つが、耐熱性に優れ強い細胞壁分解活性を有することから産業利用に向けた研究が進んでいる好熱嫌気性細菌 *Clostridium thermocellum* 由来エンドグルカナーゼと比較的高い相同性を示す。LC-Est1 は N 末端に機能未知の長い伸長領域を有するが、そのホモログも含めて構造や機能の解析はまだされていない。一般的なエステラーゼ/リパーゼ取得源である土壌細菌から単離されたエステラーゼ/リパーゼから実用化された例はまだ少なく、LC-Est1 の機能未知ドメインのように新規性の高い構造を持つエステラーゼの取得は大変興味深い。また、これまでにセルラーゼやエステラーゼの触媒機構や基質認識機構についての研究は進んでいるが、その末端伸長領域やドメインの役割についてはまだ分かっていない点も多い。そこで本研究では LC-CelA の FL、LC-CelG の Ig-like ドメイン、LC-Est1 の N 末端伸長領域に注目し、結晶構造解析や諸特性解析を通して、これらが活性や安定性にどのような影響を及ぼすのか解析した。

第 1 章　*Rhodothermus marinus* 由来 Cel12A と高い相同性を示す新規 GHF12 セルラーゼ LC-CelA の構造と安定性の解析

1-1.　はじめに

序論で述べたとおりセルラーゼは、バイオエタノール生産のためのセルロース糖化酵素、洗剤添加物、食品加工用酵素、医薬品合成用酵素、繊維加工用酵素などとして産業的に広く利用されている。セルラーゼの工業化における最大の課題は酵素製造にかかるコスト削減である。高温では、酵素の反応速度や基質の溶解度が上昇し、雑菌の繁殖に伴う収率の低下などが防げるため、高温でも安定な耐熱性セルラーゼの単離が望まれている。万博記念公園の枝葉コンポストはセルラーゼの基質となるセルロースを多く含み、発酵により内部温度が 60°C 以上の高温となるため、新規耐熱性セルラーゼの取得源として期待される。またメタゲノム法は、微生物の単離・培養を介することなく、様々な微生物由来のゲノム DNA を環境サンプルから直接抽出するため、新規酵素の取得を効率良く行うことが可能である。そこで本章では、このメタゲノム法を用いて、枝葉コンポストから新規耐熱性セルラーゼを取得することを目的とした。

さらに、今回枝葉コンポストから単離された 10 種類の新規セルラーゼのうち、LC-CelA は耐熱性に極めて優れた *Rhodothermus marinus* 由来のセルラーゼと高い相同性を示したことから、LC-CelA も同様に耐熱性に優れたセルラーゼであると期待される。そこで LC-CelA の構造と諸特性を調べることを次の目的とした。また構造解析により N 末端付近に特徴的な水素結合やジスルフィド結合が確認されたことから、それらの安定性における役割を調べた。

1-2. 実験材料および方法

1-2-1. 菌体、プラスミド

大腸菌 BL21-CodonPlus(DE3)-RP は Stratagene 社から、プラスミド pET25b は Novagen 社からそれぞれ購入した。BL21-CodonPlus(DE3)-RP 形質転換体は LB 培地 (トリプトン 10 g L^{-1}、Yeast extract 5 g L^{-1}、NaCl 10 g L^{-1}) に 50 mg L^{-1} アンピシリンを添加して培養した。

1-2-2. DNA ライブラリーの構築とスクリーニング

尿素を加え 4 ヶ月間発酵させた万博記念公園の枝葉コンポスト(温度 67°C、pH 7.5) から、メタゲノム DNA の抽出を行った。Fosmid Library Production kit (EPICENTRE Biotechnologies, Madison, WI, USA) を用い、メタゲノム DNA ライブラリーの構築を行った [8]。12.5 μg mL^{-1} クロラムフェニコール、0.01% L-アラビノース、0.5% CM セルロース、0.1 mg mL^{-1} トリパンブルーを含む LB 寒天培地を用いて、この DNA ライブラリーを植菌し、CM セルロース分解活性を有するクローンを探索した [33]。2、3 日間 37°C で培養後、50、80°C でインキュベートし、CM セルロースの加水分解によるハローの形成を確認した。CM セルロース分解酵素遺伝子の ORF 配列は EZ-Tn5TM<T7/KAN-2> Promoter Insertion kit (EPICENTRE Biotechnologies)を用いて *in vitro* トランスポゾン変異法により決定した。DNA の塩基配列は ABI Prism 310 DNA sequencer (Applied Biosystems) を用いて決定した。

1-2-3. プラスミド構築

Pre-LC-CelA (Met1-Arg261) が導入されたフォスミドベクターを鋳型として、LC-CelA (Leu20-Arg261)、C 末端に His-tag を導入した LC-CelA-His、さらに N 末端フレキシブルループ (FL) を欠損させたΔFL-LC-CelA-His (Thr36-Arg261) の発現用ベクターを構築した。また LC-CelA-His 発現ベクターpET-LC-CelA-His を鋳型とし、Glu34 のコドン GAG を GCT (Ala) になるようにプライマーを設計して、E34A-LC-CelA-His 発現ベクターを構築した。プライマーの合成は北海道システムサイエンスに依頼した。

1-2-4.　大量発現と精製

LC-CelA、LC-CelA-His、ΔFL-LC-CelA-His および E34A-LC-CelA-His は、1-2-3 で構築した発現ベクターを用い、それぞれ大腸菌 BL21(DE3) Codon Plus の中で封入体として以下の方法で発現させた。組換え大腸菌を OD_{600} の値が 0.5 程度になるまで 37°C で振盪培養し、終濃度が 0.5 mM になるように Isopropyl β-D-thio galactopyranoside (IPTG) を加えて更に 4 時間培養した。菌体を 6,000 *g*、10 分の遠心分離により回収し、TE バッファー (10 mM Tris-HCl (pH 8.0)、1 mM EDTA) に再懸濁した。ソニケーションによる菌体破砕後、30,000 *g* で 30 分遠心分離し、上清を回収した。TE バッファーに透析した後、70°C で 30 分間熱処理し、30,000 *g*、30 分の遠心分離により沈殿を除去した。LC-CelA-His、ΔFL-LC-CelA-His、E34A-LC-CelA-His の精製は、熱処理後の上清をバッファーA (20 mM Tris-HCl (pH 7.0)、10 mM イミダゾール、0.3 M NaCl) に透析し、バッファーA で平衡化した HiTrap Chelating HP カラム (GE Healthcare) に供し、イミダゾール濃度を 10 mM から 300 mM まで直線的に上げて目的タンパク質をカラムから溶出させることにより行った。回収したタンパク質は 10 mM Tris-HCl (pH 7.0) に透析した後保存した。His-tag を持たない LC-CelA の精製は、熱処理後上清を 1 mM DTT を含む TE バッファーにより平衡化した HiTrap Q HP カラム (GE Healthcare) に供し、NaCl 濃度を 0 M から 1 M まで直線的に上げて目的タンパク質をカラムから溶出させることにより行った。目的タンパク質を含むフラクションを回収し、50 mM NaCl を含む TE バッファーで平衡化した Hi-Load 16/60 Superdex 200 pg カラム (GE Healthcare) に供し、目的タンパク質を含むフラクションを回収した。精製したタンパク質の純度は SDS-PAGE により確認した [34]。N 末端アミノ酸配列は、Procise automated sequencer model 491 (Applied Biosystems) により決定した。タンパク質の濃度は、280 nm における Tyr と Trp の分子吸光係数 1,576、5,225 M^{-1} cm^{-1} [35] を用いて計算した。1 mg mL^{-1} のタンパク質溶液の吸光係数 ($A_{280}^{0.1\%}$) に基づき決定した。この値は LC-CelA、LC-CelA-His、ΔFL-LC-CelA-His、E34A-LC-CelA-His に対し、それぞれ 3.37、2.96、3.15、2.97 であった。

1-2-5.　配列分析

メタゲノムから単離された新規セルラーゼのアミノ酸配列のホモロジーサーチは DDBJ blastp search tool (http://blast.ddbj.nig.ac.jp/blastn?lang=en/) により行った。PROTSCALE tool (http://web.expasy.org/protscale/) によりフレキシブル領域、

親水性領域の予測を行った。Compute p*I*/Mw tool (http://web.expasy.org/compute_pi/) により等電点 (p*I*) を算出した。SMART tool (http://smart.embl.de/) によりドメインサーチを行った。

1-2-6. 酵素活性

酵素活性は CM セルロースを基質として dinitrosalicylic acid (DNS) stopped method により測定した [36]。反応溶液 (100 mM sodium phosphate (pH 7.0)、1% (w/v) CM セルロース (low viscosity grade, Sigma-Aldrich)) 90 μL に 0.1 mg mL^{-1} 酵素溶液 10 μL を加えて反応を開始させ、10 分後に 10% SDS を 10 μL 加え、3 分ボイルすることで反応を停止させた。この溶液に DNS 試薬 [36] を 300 μL 加え、5 分ボイルし、氷冷した。17,000 *g* で 5 分遠心し、上清 100 μL を MilliQ 100 μL と混合した。このサンプルの波長 500 nm の吸光度 (A_{500}) を測定した。グルコースを標準サンプルとして、その吸光度から描いた標準直線から、酵素反応により生じた還元糖の量を推定した。1 分間に 1 μmol の還元糖を生成する酵素量を 1 unit と定義した。温度依存性解析では pH 7.0、40 – 100°C の各温度条件で測定した。pH 依存性解析は 90°C、pH 4.0 – 10.5 の各 pH 条件で測定した。この測定において、pH 4.0 – 6.0 の範囲では 100 mM sodium citrate、pH 6.0 – 8.0 の範囲では 100 mM sodium phosphate、pH 8.0 – 10.5 の範囲では 100 mM glycine-NaOH のバッファーを用いた。

1-2-7. CD (circular dichroism) スペクトル

Far-UV (200 – 260 nm) CD スペクトルは J-725 spectropolarimeter (Japan Spectroscopic) を用いて 25°C で測定した。測定用バッファーは 10 mM Tris-HCl (pH 7.0) を用いた。測定には 0.1 mg mL^{-1} のサンプル、2 mm 光路長のセルを用いた。平均残基分子楕円率[θ] (deg cm^2 $dmol^{-1}$) はアミノ酸の平均分子量 110 を用いて計算した。

1-2-8. 熱変性測定

熱変性曲線は温度変化に伴う波長 222 nm の CD 値をモニターすることで得られた。サンプルは 3.0 M グアニジン塩酸を含む 10 mM Tris-HCl (pH7.0) に溶かした。測定はタンパク質濃度 0.1 mg mL^{-1}、光路長 2 mm、昇温速度 3.0°C min^{-1}

の条件で行った。変性中点温度 (T_m) は最小二乗法を基にしたカーブフィッティングにより計算した。

1-2-9.　結晶化

結晶化に用いる LC-CelA は 10 mM Tris-HCl (pH8.0) に透析し、Centricon (Millipore) で 10 mg mL^{-1} に濃縮した。まず、Hampton Research 社の結晶化キット (Crystal Screen I, II) と Emerald BioStructures 社の結晶化キット (Wizard I - IV) を用いて結晶化条件のスクリーニングを行った。条件探索は 4°C、20°C の温度条件の下、96-well Corning CrystalEX Microplates (Hampton Research) を用いて、シッティングドロップ蒸気拡散法により行った。結晶化ドロップは 1 μL の LC-CelA 溶液と 1 μL のリザーバー溶液を混合して調製し、100 μl のリザーバー溶液に対して蒸気拡散平衡化させた。数週間後、4°C の温度条件で、Wizard I No.28 (0.1 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.5)、0.2 M NaCl、20% (w/v) polyethylene glycol (PEG) 3000) をリザーバー溶液として用いた条件で結晶が得られた。結晶の質を向上させるために、リザーバー溶液の最適化を行った。その結果、11.7 mg mL^{-1} のタンパク質溶液 1 μL と最適化されたリザーバー溶液 (0.1 M HEPES (pH 7.5)、0.2 M NaCl、23% (w/v) PEG 3350) 1 μL を混合して調製したドロップをリザーバー溶液 100 μL に対して数週間かけて蒸気拡散平衡化させることにより、X 線回折データ収集に適した結晶が得られた。

1-2-10.　X 線回折データ収集と構造決定

X 線回折データセットの収集は、SPring-8 のビームライン BL44XU のシンクロトロン放射光で行った。データ収集時には、クライオプロテクタント溶液は使用せず、結晶を CryoLoop (Hampton Research) にセットし、-173°C の窒素ガスを吹き付けた状態で、0.9 Å の波長でデータ収集を行った。HKL2000 [37] を用いて、得られた回折データを処理した。計算された電子密度図を元に、*Rm*Cel12A (PDB code: 1H0B) の構造を鋳型とし、CCP4 [38] 内の MOLREP [39] による分子置換法により、LC-CelA の初期分子モデルを構築した。ArpWarp [40] を用いてモデルを自動的に作成した。モデルの精密化は REFMAC [41] を用いて行い、COOT [42] によりモデルをさらに修正した。LC-CelA 構造の図は PyMol (http://www.pymol.org) を用いて作成した。

1-2-11. Protein Data Bank accession number

解析した LC-CelA の構造座標および構造因子は PDB コード番号 3WX5 として、Protein Data Bank に登録した。

1-3. 実験結果

1-3-1. メタゲノム DNA ライブラリー由来セルラーゼ遺伝子のクローニング

枝葉コンポストからメタゲノム DNA を抽出し、メタゲノムライブラリーを構築した。本ライブラリーは、約 35 kb のメタゲノム DNA 断片が挿入された 21,000 個のクローンから成っている。ついで、このうち 6,000 個のクローンから CM セルロースとトリパンブルーを含むプレート上で 50°C においてハローを形成するクローンをスクリーニングした。その結果、37°C では 24 個のクローンがハローを形成し、10 個は 50°C でもハローを形成した。これらのクローンの遺伝子配列を決定することにより、10 種類の新規セルラーゼ遺伝子を同定した。これらの遺伝子がコードするセルラーゼは N 末端にシグナルペプチド (SP) を有するため、それぞれ Pre-LC-CelA~J と名付けた。このうち、Pre-LC-CelA、Pre-LC-CelD、Pre-LC-CelF、Pre-LC-CelG、Pre-LC-CelI は 80°C でもハローを形成したので、高い熱安定性を有することが示唆された。

1-3-2. Pre-LC-CelA~J のアミノ酸配列

Pre-LC-CelA~J は 261 - 782 残基のアミノ酸から成る。BlastX を用いてこれらセルラーゼのアミノ酸配列のホモロジーサーチを行ったところ、いずれも既存のセルラーゼと 42 - 76%の相同性を示した (Table 1-1)。つまり、いずれも既存のセルラーゼとは異なる配列を持っており、メタゲノム法が新規セルラーゼの探索に有用であることが確認された。Pre-LC-CelA~J のアミノ酸配列は accession number KF626648-KF626657 として GenBank に登録した。

Pre-LC-CelA~J の一次構造を模式的に Figure 1-1 に示す。ドメイン解析ツール SMART [43] を用いて、SP、セルラーゼドメイン、セルロース結合ドメインを予測した。いずれも N 末端に 16 – 36 残基から成る SP を有しており、分泌タンパク質であることが示唆された。またいずれもセルラーゼドメインを有しており、Pre-LC-CelA、Pre-LC-CelD、Pre-LC-CelE は GHF12 セルラーゼ、Pre-LC-CelF、Pre-LC-CelG は GHF9 セルラーゼ、Pre-LC-CelB、Pre-LC-CelJ は GHF6 セルラーゼ、Pre-LC-CelI は GHF3 セルラーゼ、Pre-LC-CelH は GHF44 セルラーゼ、Pre-LC-CelC は GHF51 セルラーゼの触媒ドメインを有している。加えて、Pre-LC-CelD と Pre-LC-CelF は cellulose binding domain II (CBDII) 、Pre-LC-CelF は cellulose binding module 3 (CBM3) も有している。PROTSCALE によると、Pre-LC-

CelG 以外は、いずれも SP と触媒ドメインの間に 14 - 29 残基から成る親水性の flexible loop (FL) を有している。これらの FL は酸性で、その p*I* 値は 3.28 – 5.52 である。

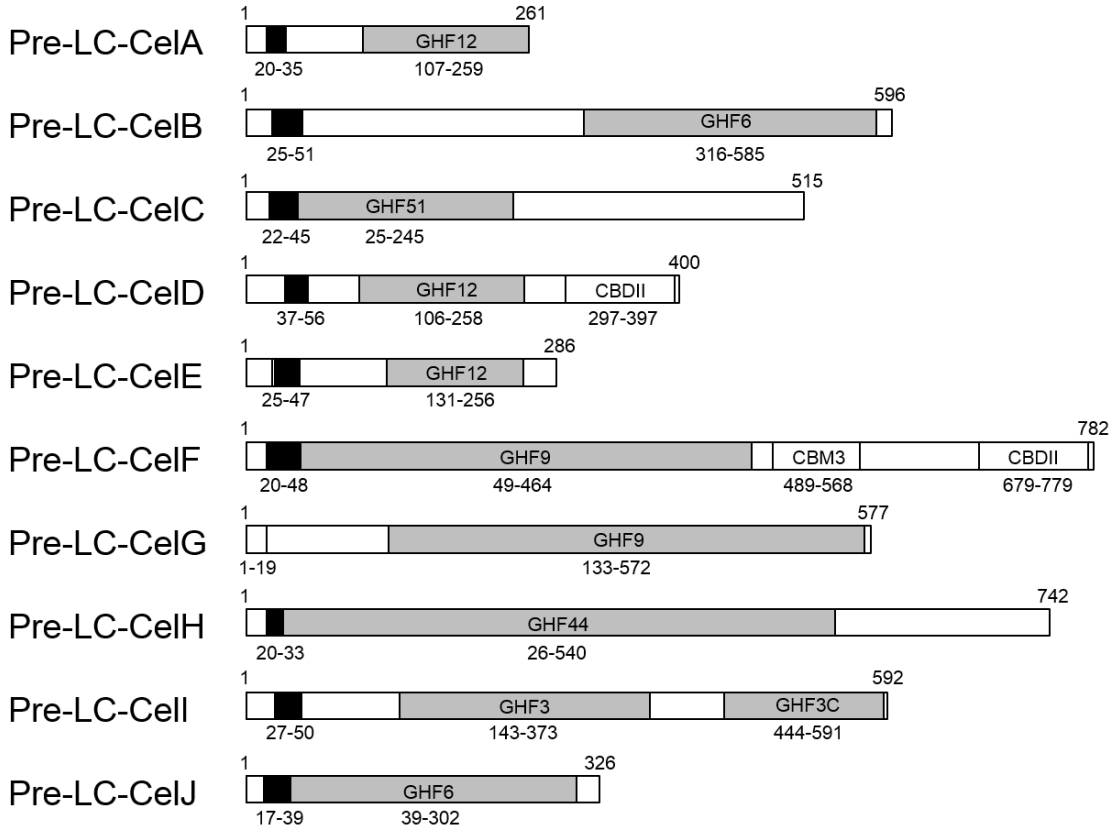


Fig. 1-1. Schematic representation of the primary structures of metagenome-derived cellulases from leaf-branch compost. Putative signal peptides (N-terminal open boxes), putative flexible linkers (black boxes), glycoside hydrolase family (GHF) domains (grey boxes), and cellulose binding domains/modules (C-terminal open boxes) are shown. The numbers above the sequence represent the positions of the N- and C-terminal residues of each protein. The numbers below the sequence represent the positions of the N- and C-terminal residues of each region or domain.

Table 1-1. List of cellulases isolated from leaf-branch compost and proteins with the highest amino acid sequence identities

Cellulases	No. of residues	Protein with the highest sequence identity			
		Protein	Source organism[a]	Accession No.	Identity (%)
Pre-LC-CelA	261	Glycoside hydrolase family 12	*Rhodothermus marinus* (65°C)	G2SJ29	76
Pre-LC-CelB	596	Cellulase	*Plesiocystis pacifica* SIR-1 (18°C)	A6G3Z7	50
Pre-LC-CelC	515	α-L-Arabinofuranosidase-like protein	*Opitutus terrae* (30°C)	B1ZZJ5	45
Pre-LC-CelD	400	Glycosyl hydrolase family 12	*Thermobispora bispora* (50°C)	D6Y3Q3	45
Pre-LC-CelE	286	Cellulase 12A	*Streptomyces sp.* 11AG8 (35°C)	Q9KIH1	43
Pre-LC-CelF	782	Uncharacterized protein	*Streptosporangium roseum* (26°C)	D2B808	62
Pre-LC-CelG	577	Glycosyl hydrolase family 9	*Microcoleus sp.* PCC 7113 (25°C)	K9WM66	42
Pre-LC-CelH	742	Cellulase	*Candidatus Methylomirabilis oxyfera* (25°C)	D5MF13	64
Pre-LC-CelI	592	Glucosidase-like glycosyl hydrolase	*Xenococcus sp.* PCC 7305 (25°C)	L8M232	57
Pre-LC-CelJ	326	Glycoside hydrolase family 6	*Frankia sp.* EUN1f (30°C)	D3D8M1	53

[a] The optimum growth temperature of each organism is shown in parenthesis.

SP はタンパク質が分泌されると除去されるので、これらのタンパク質は LC-CelA~J として機能すると考えられる。このうち、LC-CelA (Leu20-Arg261) は極めて高い耐熱性を示すことが知られている *Rm*Cel12A (Cys18-Gln261) と 76% のアミノ酸配列相同性を示すため、特に熱安定性の高いセルラーゼであることが予測された (Table 1-1)。

1-3-3. Pre-LC-CelA と Pre-*Rm*Cel12A のアミノ酸配列比較

Pre-LC-CelA と Pre-*Rm*Cel12A のアミノ酸配列を比較した (Fig. 1-2)。いずれも 261 残基から成る。しかし Pre-LC-CelA の SP (Met1 - Val19)、FL (Leu20 - Pro35) と Pre-*Rm*Cel12A の SP (Met1 - Gly17)、FL (Cys18 - Pro37) の領域は僅かに相違がある [44]。ΔFL-*Rm*Cel12A (Thr38 - Lys261) の基質複合体構造 [32] によると、Trp9、Trp26、Trp69、Trp159、Trp161、Tyr163、Trp209 (Pre-*Rm*Cel12A ではそれぞれ Trp45、Trp62、Trp104、Trp195、Trp197、Tyr199、Trp245 に相当) が活性部位クレフト付近で広域にわたり芳香族ネットワークを形成しており、基質結合に関わると考えられる。Asn24、His67、Arg100、Met136、Pro137、Gly138 (Pre-*Rm*Cel12A ではそれぞれ Asn60、His103、Arg136、Met172、Pro173、Gly174 に相当) もこのクレフトの形成に関わっている。2 つの触媒残基 Glu124、Glu207 (Pre-*Rm*Cel12A ではそれぞれ Glu160、Glu243 に相当) はこのクレフトに位置している。これらの残基は全て Pre-LC-CelA でも保存されている。

*Rm*Cel12A の FL 欠損変異体の T_m 値は野生型よりも 8.4°C 低く、90°C における半減期が 5 時間から 2 時間に低下することから [44] 、FL が *Rm*Cel12A の耐熱化に寄与することが示唆されている。しかし、FL を含んだ *Rm*Cel12A の結晶構造は解かれていないため [31, 32]、この耐熱化機構はまだ分かっていない。そこで、LC-CelA が *Rm*Cel12A 同様に高い熱安定性を有するのかどうか、また、FL がどのような構造を形成するのかを調べるために、LC-CelA および His-tag を C 末端に導入した LC-CelA-His をそれぞれ大腸菌で大量生産し、精製後、諸特性解析と結晶構造解析を行った。

```
                                                   βA1       βB1       βB2
                                                    o
Pre-LC-CelA    MKHFYWTLFLIGFLAGCSVLFPE-KNESNPQAEDEPTATVCDRWGSRDVAGGRYRVINN  58
Pre-RmCel12A MNVMRAVLVLSLLLLFGCDW***DGD*GKE*EP*P***VEL*G**DA*************  60(24)
                     βA2       βA3                βB3       βC1
                                 o
Pre-LC-CelA   VWGAETAQCIEIGLETGNFILTRAEHSNGDNVAAYPAIYLGCHWGACTSQSGLPLRVEAI 118
Pre-RmCel12A  **********V*******TI***D*D**N*********F*****N****R**QEL 120(84)
                 βA5          βB5                βB6          βB9
                                                   ●
Pre-LC-CelA   SRLQSNWRLTPISSGRWNAAYDLWFSPSLTSTNGYSGGAELMIWLNWRGNVMPGGNRVAT 178
Pre-RmCel12A  *DVRTS*T****TT*******I****VTN*G***************N*G****S**** 180(144)
                      βB8       βB7       βA6       α          βB4
Pre-LC-CelA   VSLAGATWEVWFADWDWNYIAYRRTTPVTEVTQLDLKAFIDDAVARGYISPTWYLHAVET 238
Pre-RmCel12A  *E*********Y*****************T*S*SE****************R*E******** 240(204)
                                βA4
                 ●
Pre-LC-CelA   GFELWEGGRGLKSSDFSVTLTAR                                      261
Pre-RmCel12A  ********A**R*A*****VQ                                        261(225)
```

Fig. 1-2. Alignment of the amino acid sequences of pre-LC-CelA and pre-Cel12A from *Rhodothermus marinus* (pre-*Rm*Cel12A). The accession numbers are KF626648 for pre-LC-CelA and G2SJ29 for pre-*Rm*Cel12A. For the pre-*Rm*Cel12A sequence, the identical residues with those in the pre-LC-CelA sequence are indicated by asterisks (*). Gaps are shown by dashes. A putative signal peptide (SP) is underlined. The region predicted as a flexible linker (FL) between a putative SP and a catalytic domain is boxed. The amino acid residues responsible for dimerization and catalytic activity of *Rm*Cel12A are denoted above the sequences by open and closed circles respectively. The aromatic and non-aromatic amino acid residues that form the active site cleft of *Rm*Cel12A and corresponding residues of LC-CelA are shown in boldface. Numbers represent the positions of the amino acid residues that start from the initiator methionine residue for each protein. Numbers in parentheses represent the positions of the amino acid residues that start from the initiator methionine for the *Rm*Cel12A derivative without SP and FL used to determine the crystal structure [Met-(Thr38-Gln261)]. The ranges of the secondary structures of LC-CelA, which are identical to those of *Rm*Cel12A, are shown above the sequences.

1-3-4. LC-CelA と LC-CelA-His の大量発現と精製

Pre-LC-CelA の SP は、Pre-LC-CelA がその宿主細胞から分泌されるときに除去される。しかし、大腸菌を用いて発現させても Pre-LC-CelA が分泌され、その SP が除去されるとは限らないので、本研究では SP を含まない LC-CelA (Leu20 – Arg261) と LC-CelA-His をそれぞれ大腸菌で大量発現させた。LC-CelA は N 末端に Met-Asp が導入されており、LC-CelA-His は N 末端に Met、C 末端に His-tag

が導入されている。発現誘導後、ソニケーションにより菌体を破砕し、遠心分離により可溶性画分と不溶性画分に分離したところ、LC-CelA、LC-CelA-His いずれも大腸菌菌体内に蓄積し、そのうち約 20 – 30%が可溶性であることが分かった。可溶性タンパク質を回収し、それぞれ 1-2-4 の手順に従って SDS-PAGE で単一バンドが得られるまで精製を行った。いずれも 1 L の培養液から約 3 mg の精製タンパク質が得られた。LC-CelA の N 末端アミノ酸配列は Met-Asp-Leu-Phe-であることから、FL 領域の欠損は見られなかった。

ゲルろ過クロマトグラフィーによると、LC-CelA の分子量はおおよそ 29 kDa であると推定された。これはアミノ酸配列から計算された値 (26.9 kDa) とほぼ一致するため、LC-CelA は溶液中において単量体で存在することが示唆された。一方、ΔFL-*Rm*Cel12A は二量体で存在することが報告されている [32]。ΔFL-*Rm*Cel12A の二量体構造は Glu40 と Arg83 の間に形成される塩橋により安定化される。しかし、LC-CelA においては、Arg83 は Arg81 として保存されているものの Glu40 は Thr38 に置換されている (Fig. 1-2)。従って、LC-CelA は塩橋構造を形成できず、二量体構造を形成しないと考えられる。

1-3-5. LC-CelA-His の活性

LC-CelA-His の酵素活性の温度依存性を、CM セルロースを基質として 40 - 100°C の範囲で測定した。Figure 1-3A に示すように、LC-CelA-His は 90°C で最大活性を示す。至適温度が 100°C 以上の *Rm*Cel12A と比べるとやや低い温度であるが、ほぼ匹敵している [44]。一方、60°C での LC-CelA-His の比活性 4.2 units mg^{-1} は、RmCel12A の 65°C における比活性 3.1 units mg^{-1} よりわずかに高い [44]。つまり LC-CelA-His は *Rm*Cel12A 同様非常に耐熱性の高い酵素であり、僅かに *Rm*Cel12A よりも活性が強いと言える。

LC-CelA-His の酵素活性の pH 依存性を、pH 4.0 – 10.5 の範囲で測定した。Figure 1-3B に示すように、LC-CelA-His は pH 5.0 – 9.0 の広い範囲で最大活性に近い活性を示した。*Rm*Cel12A も pH 5.0 – 8.0 の範囲で最大活性に近い活性を示したが、pH 9.0 では最大活性の半分程度の活性であった [30]。よって LC-CelA-His の方が *Rm*Cel12A よりも弱アルカリ環境において安定であることが示唆される。

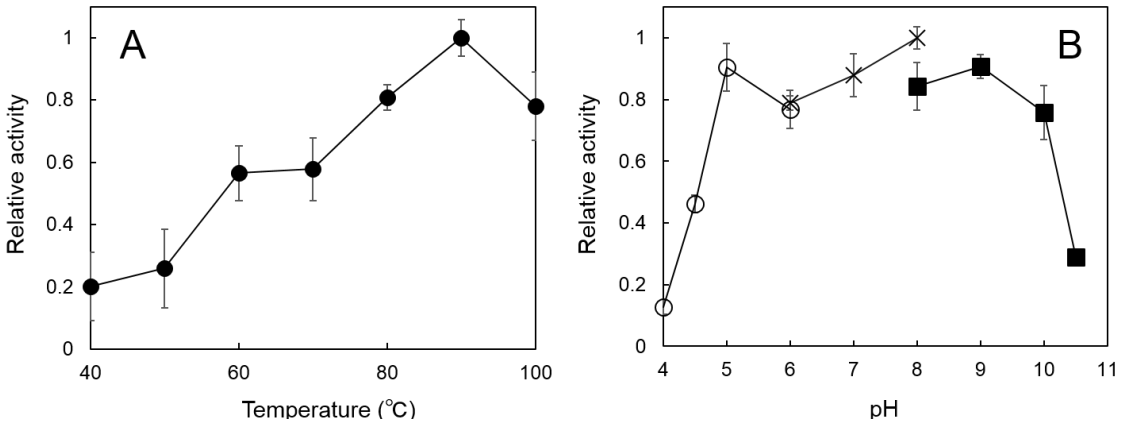


Fig. 1-3. Optimum temperature and pH for activity of LC-CelA-His. The temperature (A) and pH (B) dependencies of the enzymatic activity of LC-CelA-His are shown. The activity was determined at pH 7.0 and the temperatures indicated (A) or at 90°C and the pHs indicated (B) using 1% (w/v) carboxymethyl-cellulose (CM-cellulose) as a substrate, as described in Experimental procedures. The buffers used to analyze the pH dependence of the activity were 100 mM sodium citrate (pH 4.0-6.0), 100 mM sodium phosphate (pH 6.0-8.0) and 100 mM Glycine-NaOH (pH 8.0-10.5). The experiment was carried out at least twice, and errors from the average values are indicated by *vertical lines*.

1-3-6. LC-CelA-His の熱失活

LC-CelA-His の耐熱性を調べるため、100 mM sodium phosphate (pH 7.0) に透析した酵素 (0.05 mg mL^{-1}) を 60 - 100°C の各温度で 30 分間インキュベートし、残存活性を pH 7.0、60°C の条件で求めた。Figure 1-4 に示すように、LC-CelA-His は 90°C、30 分間のインキュベート後では活性をほぼ保持していたが、95°C、30 分間のインキュベート後はほとんど失活していた。この結果は至適温度が 90°C であることとほぼ一致する (Fig. 1-3A)。一方で 100°C、30 分間のインキュベート後の残存活性はほぼ失われていたが (Fig. 1-4)、100°C における活性は最大活性の 80%を保持していた (Fig. 1-3A)。これは、この測定条件では酵素が完全に変性する前に、基質を加水分解するからであると考えられる。

1-3-7. LC-CelA-His の熱変性

より定量的に LC-CelA-His の耐熱性を調べるために、3 M グアニジン塩酸存在下 pH 7.0 で 222 nm の CD 値をモニターすることにより熱変性曲線を測定した。熱変性はこの条件では可逆的に起こる。また≤2 M グアニジン塩酸存在下では、100°C でも完全には変性しなかった。Figure 1-5 に 3 M グアニジン塩酸存在

下での変性曲線を示す。LC-CelA-His の変性中点温度 (T_m) は 86.8°C であった。グアニジン塩酸非存在下における *Rm*Cel12A の T_m 値は 102.9°C であるが[44]、LC-CelA-His はグアニジン塩酸非存在下では、100°C では十分に変性をしないため、*Rm*Cel12A 同様極めて耐熱性の高い酵素であることが分かった。

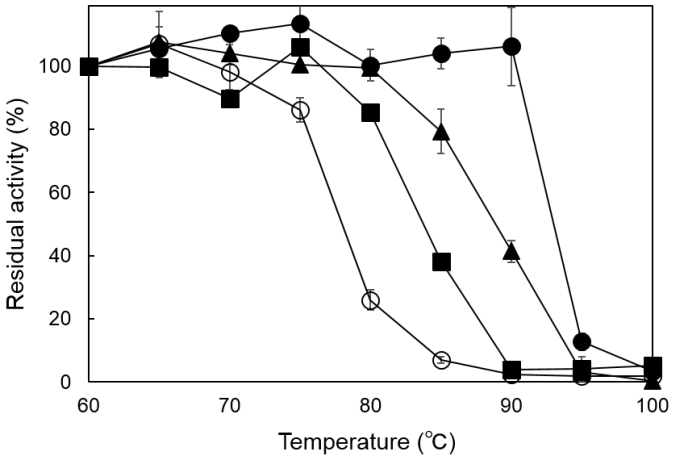


Fig. 1-4. Stability of LC-CelA-His and its derivatives against heat inactivation. LC-CelA-His was incubated at pH 7.0 and the temperatures indicated for 30 min in the absence of DTT (filled circle) or in the presence of 1mM DTT (open circle). ΔFL-LC-CelA-His (filled square) and E34A-LC-CelA-His (filled triangle) were also incubated at the same condition in the absence of DTT. The residual activities were determined at pH 7.0 and 60°C using CM-cellulose as a substrate, as described in Experimental procedures. The experiment was carried out at least twice, and errors from the average values are indicated by *vertical lines*.

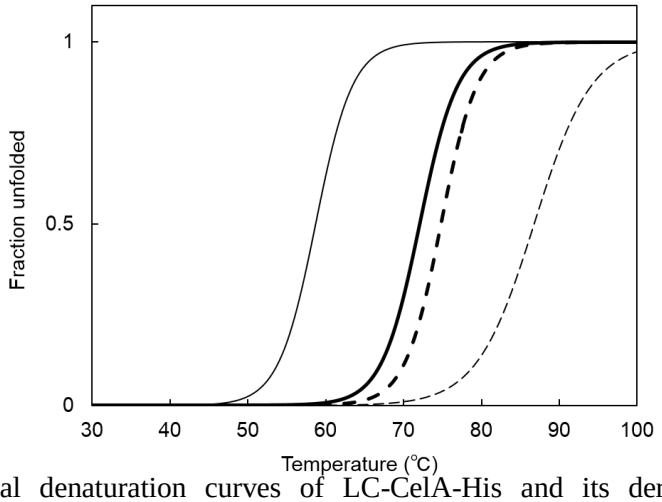


Fig. 1-5. Thermal denaturation curves of LC-CelA-His and its derivatives. Thermal denaturation curves of LC-CelA-His (thin dashed line), ΔFL-LC-CelA-His (thick solid line) and E34A-LC-CelA-His (thick dashed line) obtained in the absence of DTT and that of LC-CelA-His obtained in the presence of 1 mM DTT (thin solid line) are shown. These curves were obtained at pH 7.0 in the presence of 3M GdnHCl by monitoring the change in CD values at 222 nm as described in Experimental procedures.

1-3-8. LC-CelA の活性と安定性

C 末端に導入した His-tag の影響を調べるために、LC-CelA の活性と安定性を決定した。CM セルロースを基質として用いたときの 60°C での LC-CelA の活性は 4.3 units mg^{-1} であり、これは LC-CelA-His (4.2 units mg^{-1}) とほぼ等しい。また LC-CelA の T_m 値は 86.0°C であり、これも LC-CelA-His (86.8°C) とほぼ等しい。これらのことから、His-tag の存在は LC-CelA の活性や安定性にほとんど影響しないと言える。

1-3-9. LC-CelA の結晶構造

LC-CelA の FL がどのように折り畳まれているかを調べるために、LC-CelA の結晶構造を 1.85 Å の分解能で決定した。非対称ユニット中に 2 分子 (A、B) 存在した。いずれの分子も N 末端領域には不規則な揺らぎがあり、N 末端の Met-Asp と FL 領域の大部分 (Leu20 – Glu32) の電子密度を観測することはできなかった。分子 A と B の構造は極めて類似しており、その平均二乗偏差 (RMSD) 値は 0.08 Å であった。データ収集および精密化の統計値を Table 1-2 に示す。

LC-CelA (分子 A) の全体構造を Figure 1-6A に示す。LC-CelA は典型的なβ-ジェリーロール構造を持ち、2 つのβシートと 1 つのαヘリックスから成る。FL 領域を除いた LC-CelA の構造はΔFL-*Rm*Cel12A (PDB code: 1H0B) の構造と極めて類似しており、その RMSD 値は 0.33 Å であった。LC-CelA の活性部位周辺の構造をΔFL-*Rm*Cel12A-セロペントース複合体構造 (PDB code: 2BWC) [32] と重ね合わせたところ、2 つの触媒残基 (Glu158、Glu241) を含め、基質結合ポケットを形成すると予測されるアミノ酸残基の配置は ΔFL-*Rm*Cel12A とほぼ一致した (Fig. 1-6B)。LC-CelA の基質結合ポケットは、*Rm*Cel12A のように芳香族残基 (主に Trp) が一連のグルコース結合サブサイトを形成し、グルコースユニットとスタッキング相互作用を形成するか、極性残基が糖分子と特異的な水素結合を形成すると考えられる [32]。またΔFL-*Rm*Cel12A の結晶中では HEPES 分子が基質結合部位に結合していた [31]。しかし LC-CelA の結晶構造中では、HEPES 分子を表す明確な電子密度は見られなかったことから、LC-CelA の活性部位には HEPES 分子は強くは結合していないことが示唆された。これは活性部位のアミノ酸配置の僅かな違いによるものかもしれない。

Table 1-2. Data collection and refinement statistics of LC-CelA

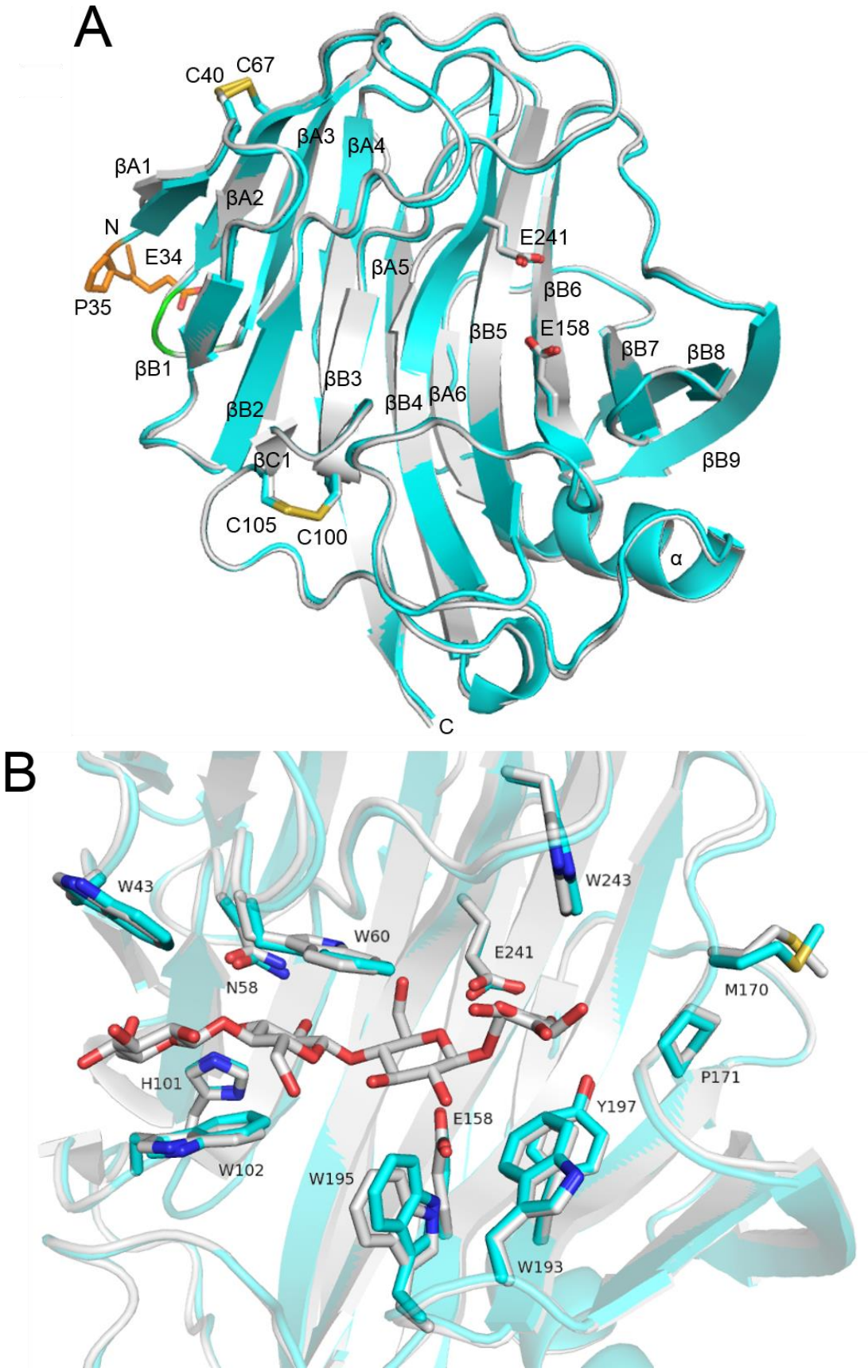
Crystal	LC-CelA
Wavelength (Å)	0.900
Space group	C121
Cell parameters	
a, b, c (Å)	130.78, 59.58, 74.95
α= γ, β (°)	90.00, 122.89
Molecules/asymmetric unit	2
Resolution range (Å)	50.00-1.85 (1.88-1.85)[a]
Reflections measured	297,834
Unique reflections	41,281
Redundancy	7.2 (7.4)
Completeness (%)	99.4 (99.6)[a]
R_{merge} (%)[b]	12.0 (37.1)[a]
Average $I/\sigma(I)$	21.8 (5.6)[a]
Refinement statistics	
Resolution limits (Å)	62.9-1.9
No. of atoms	
Protein/water	1798/507
R_{work} (%) / R_{free} (%)[c]	16.2/20.5
Rms deviations from ideal values	
Bond lengths (Å)	0.012
Bond angles (°)	1.200
Average B factors (Å^2)	
Protein	15.2
Water	27.2
Ramachandran plot statistics	
Preferred regions (%)	96.9
Allowed regions (%)	3.1
Molprobity score[d]	1.35

[a] Values in parentheses are for the highest-resolution shell.

[b] $R_{merge} = \sum|I_{hkl} - \langle I_{hkl}\rangle|/\sum I_{hkl}$, where I_{hkl} is an intensity measurement for reflection with indices *hkl* and $\langle I_{hkl}\rangle$ is the mean intensity for multiply recorded reflections.

[c] Free *R*-value was calculated using 5% of the total reflections chosen randomly and omitted from refinement.

[d] MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution [53].


A
C40
C67
βA3
βA4
βA1
βA2
N
E34
P35
βB1
βA5
E241
βB6
βB5
E158
βB7
βB8
βB3
βB4
βA6
βB2
βC1
βB9
C105
C100
α
C
B
W43
W243
W60
E241
N58
M170
H101
P171
Y197
E158
W102
W195
W193

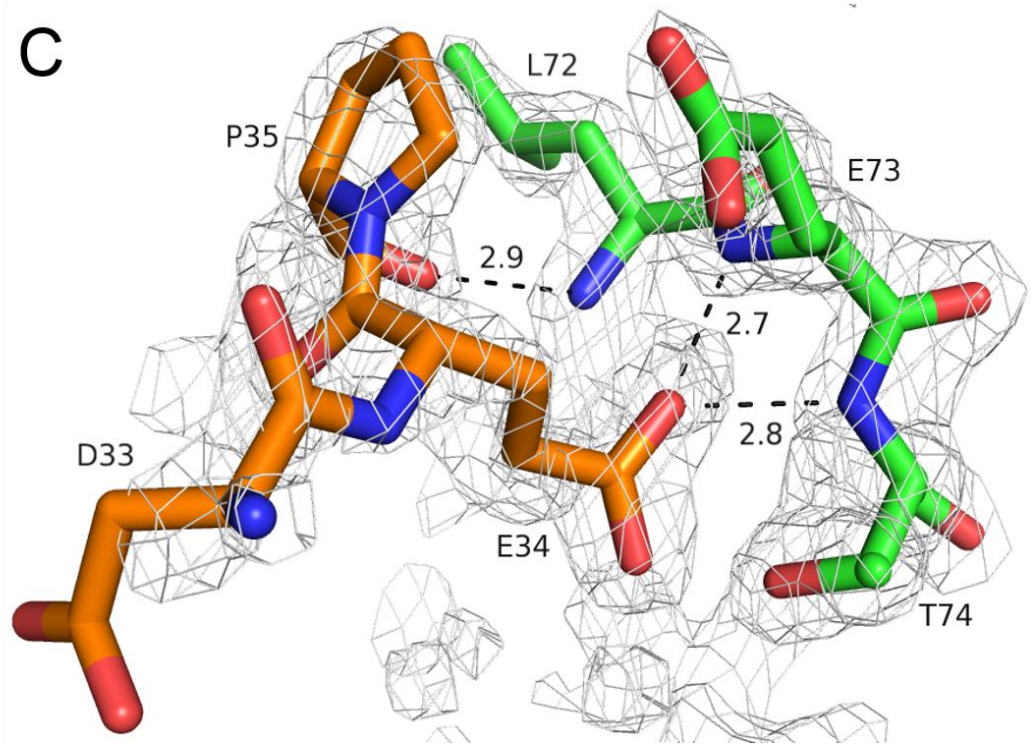


Fig. 1-6. Crystal structure of LC-CelA. (A and B) The entire structure (A) and the structure around the active site (B) of LC-CelA are superimposed on those of *Rm*Cel12A in complex with the substrate (D(+)-cellopentose) (PDB: 2BWC). The structures of LC-CelA and *Rm*Cel12A are colored cyan and gray respectively. The region predicted as a flexible linker (FL) of LC-CelA is colored orange. The central region of LC-CelA that contacts FL is colored green. Two acidic active site residues (Glu158 and Glu241), two disulfide bonds (Cys40–Cys67 and Cys100–Cys105), twelve residues forming the substrate binding site of LC-CelA, and C-terminal two residues of FL (Glu34 and Pro35) are shown as stick models with labels. The corresponding residues of *Rm*Cel12A, except for those corresponding to Glu34 and Pro35, and the substrate bound to *Rm*Cel12A are also shown as stick models. In these stick models, the oxygen, nitrogen and sulfur atoms are colored red, blue and yellow respectively. The substrate and the residues forming the substrate binding site are not shown in panel A for simplicity. Six β-strands forming sheet A (βA1–βA6), nine β-strands forming sheet B (βB1–βB9), βC1-strand, and one α-helix are shown in panel A. N and C represent the N- and C-termini. (C) Electron density around the region, where hydrogen bonds are formed between the FL and central regions of LC-CelA. The 2Fo–Fc map contoured at the 1.5ρ level is shown. The side chain and main chain atoms are shown as stick models, where the oxygen and nitrogen atoms are colored red and blue respectively, and the carbon atoms are colored orange and green for the FL and central regions respectively as in panel A. Three hydrogen bonds formed between Glu34 $O^{\delta 1}$ and Thr74 N, between Glu34 $O^{\delta 1}$ and Glu73 N, and between Pro35 O and Leu72 N are shown by broken lines together with their distances.

LC-CelA の構造中に、Cys40 – Cys67、Cys100 – Cys105 の 2 つのジスルフィド結合が見られた。Cys40 – Cys67 は触媒ドメインの N 末端付近のβA1 鎖と βB1 鎖の間のループとβA2 鎖を繋ぐ。また、Cys100 – Cys105 は短いループ (His101 – Ala104) を形成する。Cys40 – Cys67 のジスルフィド結合は GHF12 セルラーゼに広く保存されているが、Cys100 – Cys105 のジスルフィド結合は *Rm*Cel12A と *Streptomyces* 属由来 GHF12 セルラーゼでのみ保存されている。

FL 領域の大部分は不規則な揺らぎがあり、その電子密度を観測することはできなかったが、C 末端 3 残基 (Asp33-Pro35) のみはっきりとした電子密度が観測された。興味深いことに、このうち 2 残基は、3 本の水素結合を介して酵素本体の Leu72-Thr74 の主鎖と相互作用していた (Fig. 1-6C)。興味深いことに、この領域は 3 つの水素結合 (Glu34 $O^{\delta 1}$ と Thr74 N 間、Glu34 $O^{\delta 1}$ と Glu73 N 間、Pro35 O と Leu72 N 間) を介して中心領域 (βA2 鎖とβA3 鎖の間のループ) と相互作用している。これらの残基は *Rm*Cel12A でも保存されているため、これらの水素結合は RmCel12A でも同様に形成されると考えられる。

1-3-10. ΔFL-LC-CelA-His と E34A-LC-CelA-His の安定性

上述の N 末端付近の水素結合が LC-CelA の安定性に寄与するかどうかを調べるために、FL 欠損変異体 ΔFL-LC-CelA-His と Glu34→Ala 変異体 E34A-LC-CelA-His を構築し、熱処理、ニッケルアフィニティクロマトグラフィーにより精製した。収量はいずれも LC-CelA-His と同程度であった(培養液 1 L あたり 3 mg)。いずれも Far-UV CD スペクトルは LC-CelA-His と類似することから(Fig. 1-7)、FL 領域の欠損や変異は全体構造にはほとんど影響しないことが示唆された。

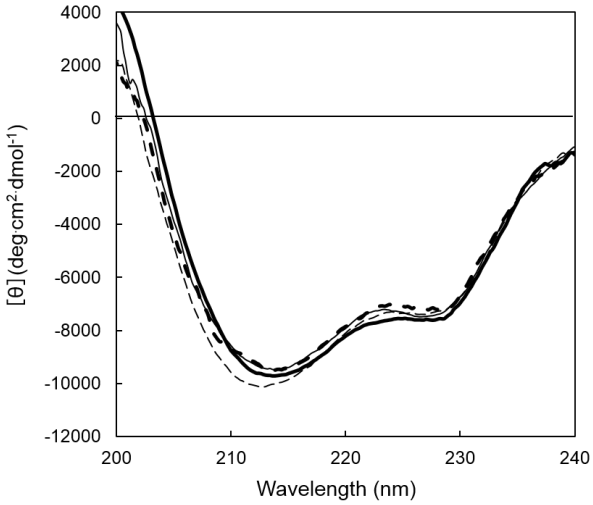


Fig. 1-7. CD spectra of LC-CelA-His and its derivatives. The far-UV CD spectrum of LC-CelA-His was measured at pH 7.0 and 25°C either in the absence of DTT (thin dashed line) or in the presence of 1 mM DTT (thin solid line), as described in Experimental procedures. The far-UV CD spectra of ΔFL-LC-CelA-His (thick solid line) and E34A-LC-CelA-His (thick dashed line) were also measured at pH 7.0 and 25°C in the absence of DTT.

LC-CelA-His と同様に、ΔFL-LC-CelA-His と E34A-LC-CelA-His を 60 - 100°C の各温度で 30 分間インキュベートし、残存活性を pH 7.0、60°C の条件で求めた (Fig. 1-4)。LC-CelA-His は 90°C、30 分間のインキュベート後もほぼ最大活性を保持していたのに対し、ΔFL-LC-CelA-His は 85°C でインキュベートした後は 60%、90°C でインキュベートした後はほぼ 100%活性が失われていた。同様に E34A-LC-CelA-His も 85°C でインキュベートした後は 20%、90°C でインキュベートした後は 60%程度活性が低下していた。つまり耐熱性は LC-CelA-His> E34A-LC-CelA-His> ΔFL-LC-CelA-His の順に低下した。

続いて、ΔFL-LC-CelA-His と E34A-LC-CelA-His について、3 M グアニジン塩酸存在下 pH 7.0 で熱変性曲線を測定した (Fig. 1-5)。それぞれの T_m 値を Table 1-3 にまとめる。ΔFL-LC-CelA-His と E34A-LC-CelA-His の T_m 値 は、LC-CelA-His よりそれぞれ 14.7°C、12.0°C 低下していた。従って、LC-CelA の FL は上記の水素結合によって、*Rm*Cel12A の FL 同様、LC-CelA の耐熱化に寄与することが示唆された。また T_m 値の差より、Pro35 よりも Glu34 が形成する水素結合の方が、FL による LC-CelA の安定化に強く寄与することが示唆された。水素結合は一般的にタンパク質の安定化に大きく寄与することが報告されている [45]。一方で不規則に揺らいでいる FL 領域 (Leu20 – Glu32) は LC-CelA の安定化にほとんど寄与しないと考えられた。

Table 1-3. Activities and stabilities of LC-CelA-His and its derivatives

Protein	[DTT] (mM)	Specific activity[a] (U/mg)	Relative activity (%)	T_m[b] (°C)	ΔT_m[c] (°C)
LC-CelA-His	0	4.2	100	86.8	–
	1	4.2	100	58.8	-28.0
ΔFL-LC-CelA-His	0	3.9	93	72.1	-14.7
E34A-LC-CelA-His	0	4.0	96	74.8	-12.0

[a] The specific activity was determined at 60°C either in the presence or absence of DTT using CM-cellulose as a substrate, as described in Experimental procedures. Each experiment was carried out at least twice and the average value is shown. Errors are within 10% of the values reported.
[b] The melting temperature (T_m), which is the temperature of the midpoint of the transition, was determined from the thermal denaturation curves shown in Fig. 5.
[c] ΔT_m = Tm determined - 86.8°C.

1-3-11.　DTT 存在下における LC-CelA-His の安定性

ジスルフィド結合がタンパク質の耐熱化に寄与する例は数多く報告されている [12, 46-49]。そこで LC-CelA に存在する 2 つのジスルフィド結合が耐熱化に寄与するかどうかを調べるため、3 M グアニジン塩酸存在下 pH 7.0 で熱変性曲線を測定した。この熱変性曲線を Figure 1-5 に示す。1 mM DTT 存在下における LC-CelA-His の T_m 値は DTT 非存在下のときと比べ 28.0°C 低下した (Table 1-3)。よってジスルフィド結合も LC-CelA の耐熱化に大きく寄与することが示唆された。なお、10 mM DTT 存在下での T_m 値は 59.8°C で、1 mM DTT 存在下での T_m 値 (58.8°C) とほぼ同じであるため、LC-CelA-His のジスルフィド結合は 1 mM DTT で十分還元されていると考えられる。

いずれのジスルフィド結合が LC-CelA の耐熱化に強く寄与するのかは分かっていない。ただし、一般にジスルフィド結合間の残基数が多いほど安定性に強く寄与することが報告されている [46]。Cys40 – Cys67 間と Cys100 – Cys105 間のアミノ酸残基数はそれぞれ 26 残基、4 残基であるため、恐らく前者のジスルフィド結合が主に耐熱化に寄与していると考えられた。

ジスルフィド結合は LC-CelA あるいは *Rm*Cel12A と *Streptomyces* 種由来 GHF12 セルラーゼの耐熱性の違いの主な要因ではないと考えられる。なぜなら、LC-CelA の 2 つのジスルフィド結合は、これらのタンパク質に保存されているためである。*Rm*Cel12A と *Streptomyces* sp. 11AG8 由来 GHF12 セルラーゼ (*St*Cel12A) の変性剤非存在下での変性中点温度 T_m 値はそれぞれ 102.9°C [44]、65.7°C [50] であり、*Rm*Cel12A の方が *St*Cel12A よりも極めて耐熱性が高い。LC-CelA の耐熱性も *Rm*Cel12A に匹敵するため、*St*Cel12A よりも耐熱性は極めて高いと考えられる。タンパク質は一般に、イオン結合や水素結合、ジスルフィド結合、ループ領域中のプロリン残基の数が多いほど、またタンパク質内部の疎水性度が高いほど耐熱性が高いとされている [45]。そこで、LC-CelA、*Rm*Cel12A、*St*Cel12A (PDB: 1OA4) についてこれらの比較を行った。LC-CelA の内部残基全体に対する内部の非極性残基の割合は 66.7%であり、これは *Rm*Cel12A (65.2%) と同程度である一方、*St*Cel12A (60.8%) よりも大幅に高い。また LC-CelA のイオン結合の数は 10 で、*Rm*Cel12A (8) と同程度であるが、*St*Cel12A (5) よりも多い。一方、水素結合やループ中のプロリン残基の数にはこのような傾向は見られなかった。よって主に酵素内部の疎水性度とイオン対の数が、LC-CelA あるいは *Rm*Cel12A と *St*Cel12A の安定性の差を生んでいると考えられる。

1-3-12　ΔFL-LC-CelA-His と E34A-LC-CelA-His の活性

ΔFL-LC-CelA-His と E34A-LC-CelA-His は LC-CelA の至適温度である 90°C では失活する恐れがあるため (Fig. 1-4)、これらの酵素の活性は CM セルロースを基質として 60°C で測定した。その結果、いずれの活性もほぼ LC-CelA-His と同じであった (Table 1-3)。よって、FL の欠損や変異は LC-CelA の酵素活性にはほぼ影響を及ぼさないことが示唆された。同様に DTT 存在、非存在条件で LC-CelA-His の活性に差は見られなかったため、ジスルフィド結合の還元も LC-CelA の活性にはほどんど影響しないと考えられた。

1.4　考察

枝葉コンポストからメタゲノム法により 10 種類の耐熱性の高い新規セルラーゼを単離した。その中で、耐熱性に優れた *R. marinus* 由来の GHF12 セルラーゼ (*Rm*Cel12A) のホモログである LC-CelA の結晶構造を決定した。これは N 末端のフレキシブルリンカー (FL) を含む GHF12 セルラーゼとしては初めての構造である。しかし、FL 領域の大半は明確な電子密度を与えず、フレキシビリティの高いことが分かった。この結果は、酵素と細胞膜を繋ぐ疎水性度の高いシグナルペプチドと酵素の触媒コアを分離するために FL が必要であるという仮説に一致する [44]。しかし FL の C 末端 2 残基 (Glu34、Pro35) は明確な電子密度を与え、中心領域 (Leu72 – Thr74) と 3 つの水素結合を介して相互作用していることが明らかになった。また Cys40 – Cys67、Cys100 – Cys105 間に 2 つのジスルフィド結合が形成されていた。

LC-CelA は至適温度が 90°C の極めて耐熱性の高い酵素である。LC-CelA の変性温度は FL の欠損により 14.9°C 低下した。これは FL と中心領域間の水素結合を失うことが主な原因である。*Rm*Cel12A においても同様に FL を欠損させると変性温度は 8.4°C 低下した [44]。これも LC-CelA の Glu34 と Pro35 に相当する残基を失うことに起因すると考えられる。データベースサーチによるとこれらの残基は LC-CelA と *Rm*Cel12A のみに保存され、FL と中心領域間に形成される水素結合による耐熱化はこれら酵素に特有であることが示唆される。一方 DTT によるジスルフィド結合の還元により、LC-CelA の変性温度は 28 °C 低下した。N 末端付近に位置する Cys40 – Cys67 間のジスルフィド結合の方が、Cys100 – Cys105 間のジスルフィド結合よりもループ間のアミノ酸残基数が多いため、耐熱化に強く寄与すると推測された。

GHF11 キシナラーゼは GHF12 セルラーゼとアミノ酸配列の相同性は低いが、その結晶構造は極めて類似している。N 末端にある 11 残基のユニークな伸長配列と中心領域の間に形成される水素結合やジスルフィド結合により耐熱性を向上させている GHF11 キシラナーゼが最近報告された [51]。また GHF11 キシラナーゼの耐熱性を、N 末端への変異導入 [52] や N 末端のジスルフィド結合の導入 [53-56] により向上させるという研究が複数報告されている。また、数は少ないが、GHF12 セルラーゼの活性や耐熱性を向上させるタンパク質工学的研究も報告されている[57, 58]。本研究の成果は、GHF12 セルラーゼや GHF11 キシラナーゼの耐熱性を向上させるための新たな手法としての応用が期待できる。

第 2 章　N 末端に Ig-like ドメインを有する GHF9 セルラーゼ LC-CelG の構造、活性および安定性の解析

2-1.　はじめに

第 1 章で述べたとおり、メタゲノム法により枝葉コンポストから新規セルラーゼを 10 種類単離した。このうち、新規 GHF9 セルラーゼである LC-CelG は 577 アミノ酸残基から成り、N 末端にシグナルペプチド (Met1 – Ala19) を持つ。シグナルペプチドを除いた LC-CelG は N 末端 Ig-like ドメイン (Leu20 – Pro132) と C 末端触媒ドメイン (Val133 – Leu577) から成る。LC-CelG は *Microcoleus* sp. PCC 7113 由来 GHF9 酵素 (accession No. K9WM66) と 42%のアミノ酸配列相同性がある。また結晶構造が決定されているものでは、*Alicyclobacillus acidocaldarius* 由来 Cel9A (AaCel9A) (3EZ8) [63]、*C. thermocellum* 由来 CelD (CtCelD) (1CLC) [64]、*C. thermocellum* 由来セロビオヒドロラーゼ CbhA (CtCbhA) (1UT9) [65] とそれぞれ 31%、31%、29%の比較的低いアミノ酸配列相同性を示す。これらのことより、LC-CelG も高い耐熱性を示すのか、またこれらの GHF9 セルラーゼと類似の構造を有するのかを調べることは大変興味深い。

GHF9 セルラーゼはよく研究が進んでいるセルラーゼファミリーの 1 つであり、大半はエンドグルカナーゼであるが、セロビオヒドロラーゼや他の糖分解酵素も含まれている。これまでに 100 種類以上の GHF9 酵素の諸特性解析が進んでいる。そのうち 12 種類は結晶構造が解かれており、ドメインの数や様式に基づき、3 つのグループに分けることができる。1 つ目は *C. cellulolyticum* 由来 Cel9M (PDB ID 1IA6) [59]、*C. thermocellum* 由来 CelT (2YIK) [60]、ミミズ (*Eisenia fetida*) 由来 EF-EG2 (3WC3) [61]、シロアリ (*Nasutitermes takasagoensis*) 由来 NtEgl (1KSC) [62] が含まれ、いずれも触媒ドメインのみから成っている。2 つ目は AaCel9A、CtCelD、CtCbhA が含まれ、いずれも N 末端 Immunoglobulin-like (Ig-like) ドメインと触媒ドメインから成る。3 つ目には *C. cellulolyticum* 由来 Cel9G (1G87) [66]、*Thermomonospora fusca* 由来 endo/exocellulase E4 (1TF4) [67] が含まれ、C 末端ファミリー3 carbohydrate-binding module (CBM3) と触媒ドメインから成る。GHF9 酵素の触媒ドメインは$(\alpha/\alpha)_6$-バレル構造を有し、触媒残基は 2 つのアスパラギン酸と 1 つのグルタミン酸である。これら 2 つのアスパラギン酸残基は、脱プロトン化により求核試薬として働く水分子を活性化させ、グルタミン酸は一般酸 (プロトンドナー) として働く [68]。2 つのアスパラギン酸残基はこ

の水分子と結合している。なお、CtCbhA は Ig-like ドメイン、触媒ドメインの他に、N 末端領域に CBM4、X11、X12 モジュールを持つ。しかし、CtCbhA は Ig-like ドメインと触媒ドメインのみで活性を有し、N 末端領域のモジュールを含まない変異体の結晶構造が解かれている [65]。また CtCbhA は Ig-like ドメインの欠損により、不活性化することが報告されている [69]。しかし GHF9 酵素における Ig-like ドメインの役割はまだ明確には分かっていない。

そこで本章では、LC-CelG を大腸菌で大量生産し、精製後、諸特性解析と結晶構造解析を行った。その結果、LC-CelG が熱安定なエンドグルカナーゼであること、また結晶構造から他の Ig-like ドメインを有する GHF9 酵素と類似の構造を有することを明らかにした。また Ig-like ドメインの欠損変異体と、Ig-like ドメインと触媒ドメイン間の相互作用に関わる残基の変異体を構築し、活性や耐熱性を測定した。そしてこれらの結果より、Ig-like ドメインの役割を議論した。

2-2.　実験材料および方法

2-2-1.　プラスミド構築

Pre-LC-CelG (Met1 – Leu577) が導入されたフォスミドベクターを鋳型として、N 末端に Met を導入した LC-CelG (Leu20 – Leu577)、N 末端に His-tag を導入した His-LC-CelG の発現用ベクター (それぞれ pET-LC-CelG、pET-His-LC-CelG)を構築した。さらに Ig-like ドメイン欠損変異体 His-ΔIg-CelG の発現用ベクターpET-His-ΔIg-CelG を構築した。また pET-His-LC-CelG を鋳型とし、Gln40 のコドン CAG、Asp99 のコドン GAC を GCC (Ala) になるようにプライマーを設計して、His-Q40A-CelG、His-D99A-CelG、His-Q40A/D99A-CelG の発現ベクターをそれぞれ構築した。PCR は KOD DNA ポリメラーゼ (Toyobo) を用いて、Gene Amp PCR system 2400 (Applied Biosystems) により行った。プライマーの合成は北海道システムサイエンスに依頼した。構築したプラスミドの DNA 塩基配列は Prism 310 DNA sequencer を用いて決定した。

2-2-2.　大量発現と精製

LC-CelG やその変異体は、2-2-1 で構築した発現ベクターを用い、それぞれ大腸菌 BL21(DE3) の中で封入体として以下の方法で発現させた。組換え大腸菌を OD_{600} の値が 0.5 程度になるまで NZCYM 培地を用いて 37°C で振盪培養し、終濃度が 0.5 mM になるように Isopropyl β-D-thio galactopyranoside (IPTG) を加えて更に 4 時間培養した。8,000 *g*、10 分間の遠心分離により菌体を回収し、1 mM DTT を含む 10 mM Tris-HCl (pH 8.0) に再懸濁した。ソニケーションによる菌体破砕後、30,000 *g* で 30 分間遠心分離し、上清を回収した。50%硫酸アンモニウムを用いてタンパク質を沈殿させ、1 mM DTT、1 mM $CaCl_2$、1 mM $ZnCl_2$ を含む 10 mM Tris-HCl (pH 8.5) に溶解させた。そして 10 mM 2-メルカプトエタノール、1 mM $CaCl_2$ を含む 10 mM Tris-HCl (pH 8.5) に一晩透析した。そして 60°C で 60 分間熱処理をして、30,000 *g*、30 分の遠心分離により沈殿を除去した。続いて LC-CelG の精製においては、遠心分離後の上清を 10 mM Tris-HCl (pH 8.5) により平衡化した HiTrap Q HP カラム (GE Healthcare) に供し、NaCl 濃度を 0 M から 1 M まで直線的に上昇させることにより目的タンパク質をカラムから溶出させた。目的タンパク質を含むフラクションを回収し、50 mM NaCl を含む 10 mM Tris-HCl (pH 8.0) で平衡化した Hi-Load 16/60 Superdex 200 pg カラム (GE Healthcare)

に供し、目的タンパク質を含むフラクションを回収した。次に His-LC-CelG、His-ΔIg-CelG、His-Q40A-CelG、His-D99A-CelG、His-Q40A/D99A-CelG の精製においては、遠心分離後の上清を 1 mM DTT、10 mM イミダゾール、0.3 M NaCl を含む 10 mM Tris-HCl (pH 8.0) (バッファーA) に透析し、バッファーA で平衡化した HiTrap Chelating HP カラム (GE Healthcare) に供し、イミダゾール濃度を 10 mM から 300 mM まで直線的に上昇させることにより目的タンパク質をカラムから溶出させた。回収したフラクションは 10 mM Tris-HCl (pH 8.0) に透析した。

2-2-3. 酵素活性

1-2-6 と同様の手法で行った。ただし酵素濃度は 0.01 mg mL^{-1} とした。

2-2-4. CD (circular dichroism) スペクトル

1-2-7 と同様の手法で行った。

2-2-5. 熱変性測定

熱変性曲線は温度変化に伴う波長 222 nm の CD 値をモニターすることで得られた。サンプルは 5.0 mM $CaCl_2$ を含む 10 mM Tris-HCl (pH7.0) に溶かした。測定はタンパク質濃度 0.1 mg mL^{-1}、光路長 2 mm、昇温速度 1.0°C/min の条件で行った。変性中点温度 ($T_{1/2}$) は最小二乗法を基にしたカーブフィッティングにより計算した。

2-2-6. 結晶化

1-2-9 と同様に結晶化条件を検討した。ただし結晶化に用いた LC-CelG の濃度は 10.3 mg mL^{-1} とした。数か月後、20°C の温度条件の下、Crystal Screen I No.6 (0.1 M Tris-HCl (pH 8.5)、0.2 M magnesium chloride hexahydrate、30% (w/v) polyethylene glycol (PEG) 4000) をリザーバー溶液として用いた条件で良質な直方体の単結晶が得られた。

2-2-7. X 線回折データ収集と構造決定

1-2-10 と同様の手法で行った。ただし CtCelD (PDB code: 1CLC) の構造を鋳型とした。

2-2-8. Protein Data Bank accession number

解析した LC-CelG の構造座標および構造因子は PDB コード番号 3X17 として、Protein Data Bank に登録した。

2-3. 実験結果

2-3-1. LC-CelG のアミノ酸配列

シグナルペプチドを除いた LC-CelG (Leu20 – Leu577) のアミノ酸配列を CtCelD、AaCel9A、CtCbhA と比較し、Figure 2-1 に示す。CtCelD、AaCel9A、CtCbhA のアミノ酸残基数はそれぞれ 649、537、813 である。GHF9 セルラーゼの 3 つの活性部位残基は、LC-CelG において Asp194、Asp197、Glu558 として保存されている。これらの残基は CtCelD では Asp198、Asp201、Glu555、AaCel9A では Asp143、Asp146、Glu515、CtCbhA では Asp383、Asp386、Glu795 として保存されている。AaCel9A とセロテトラオースとの共結晶構造によると [70]、Asp143、Asp146、Tyr150、Phe221、Gly298、Tyr300、Trp343、Ile400、Trp401、His461、Arg463、His485、Gln487、Tyr511、Glu515、Tyr519 が基質結合ポケットを形成している。LC-CelG において、Phe221、Ile400、Tyr519 はそれぞれ Trp270、His450、Asp562 に置換されているが、その他の残基は全て保存されている。よって LC-CelG の基質結合メカニズムや触媒メカニズムは AaCel9A と類似していることが示唆された。

金属イオン結合部位の数はこれら 4 つの GHF9 セルラーゼの間で異なっている。CtCelD は 1 つの Zn^{2+}結合部位 (Zn site) と 3 つの Ca^{2+}結合部位 (Ca1 – Ca3 site) を持つ。これらのうち Ca2 site のみが AaCel9A と CtCbhA でも保存されている。Zn site と Ca3 site は AaCel9A のみ保存されており、Ca1 site は CtCbhA のみ保存されている。つまり、AaCel9A は Zn、Ca2、Ca3 site を持ち、CtCbhA は Ca1、Ca2 site を持つ。Zn site は CtCelD では Cys155、Cys173、His174、His197 により形成され、AaCel9A では Cys104、Cys121、His122、His142 により形成される。LC-CelG において、対応する残基は Cys150、Asp166、His167、His193 であることから、Zn site は LC-CelG でも保存されていると考えられた。Ca1 site は CtCelD では水分子、Glu236、Asn239、Ile241、Asp243、Asp246 により形成され、CtCbhA では Asp421、Glu428、Asn431、Tyr433、Asp435、Asp438 により形成される。LC-CelG において、対応する残基は Asp225、Glu232、Asn235、Val237、Asp239、Asp242 であり、Ca1 site も LC-CelG で保存されていると考えられた。Ca2 site は CtCelD では Thr356、Ser358、Asp361、Asp362、Asp401 により形成され、AaCel9A では Asp302、Glu304、Asp307、Glu308、Ala344 により形成され、CtCbhA では Asp557、Tyr559、Asp562、Asp563、Gly617 により形成される。LC-CelG において、対応する残基は Asp351、Asp353、Asp356、Asp357、Val395 であ

り、Ca2 site も LC-CelG で保存されていると考えられた。Ca3 site は CtCelD では Ser465、Asp468、Val470 により形成され、AaCel9A では Ser520、Asp523、Ile525 により形成される。LC-CelG において、対応する残基は Val514、Ser517、Val519 であるため、Ca3 site は LC-CelG では保存されていないと考えられた。

2-3-2. LC-CelG と His-LC-CelG の大量発現と精製

構造解析と諸特性解析のために、シグナルペプチドを除き、N 末端に Met を導入した LC-CelG (Leu20 – Leu577)、N 末端に His-tag を導入した His-LC-CelG を構築した。大腸菌内で大量発現を行い、菌体破砕後、遠心分離により可溶型として得られた。それぞれ 2-2-2 の手順に従って SDS-PAGE で単一バンドが得られるまで精製を行った。1 L の培養液から LC-CelG は約 3 mg、His-LC-CelG は約 4 mg の精製タンパク質が得られた。LC-CelG の N 末端アミノ酸配列は Met-Leu-Ala-Gly-であり、シグナルペプチド領域以外の欠損は見られなかった。ゲルろ過クロマトグラフィーによると LC-CelG の分子量は 60 kDa と見積もられた。これはアミノ酸配列から計算した値 (62.6 kDa) とほぼ一致したので、LC-CelG は溶液中においてモノマーで存在することが示唆された。

2-3-3. His-LC-CelG の活性

LC-CelG が CtCelD や AaCel9A のようにエンドグルカナーゼであるか、CtCbhA のようにセロビオヒドロラーゼであるかを確かめるために、CM セルロースと p-ニトロフェニルセロビオシドを基質として、100 mM sodium phosphate (pH 7.0)、50°C の条件下で His-LC-CelG の加水分解活性を調べた。酵素、CMC、p-ニトロフェニルセロビオシドの濃度はそれぞれ 0.01 mg mL^{-1}、1% (w/v)、1 mM とした。His-LC-CelG は CM セルロースには活性を示したが、p-ニトロフェニルセロビオシドには不活性であったため、LC-CelG はエンドグルカナーゼであることが示唆された。

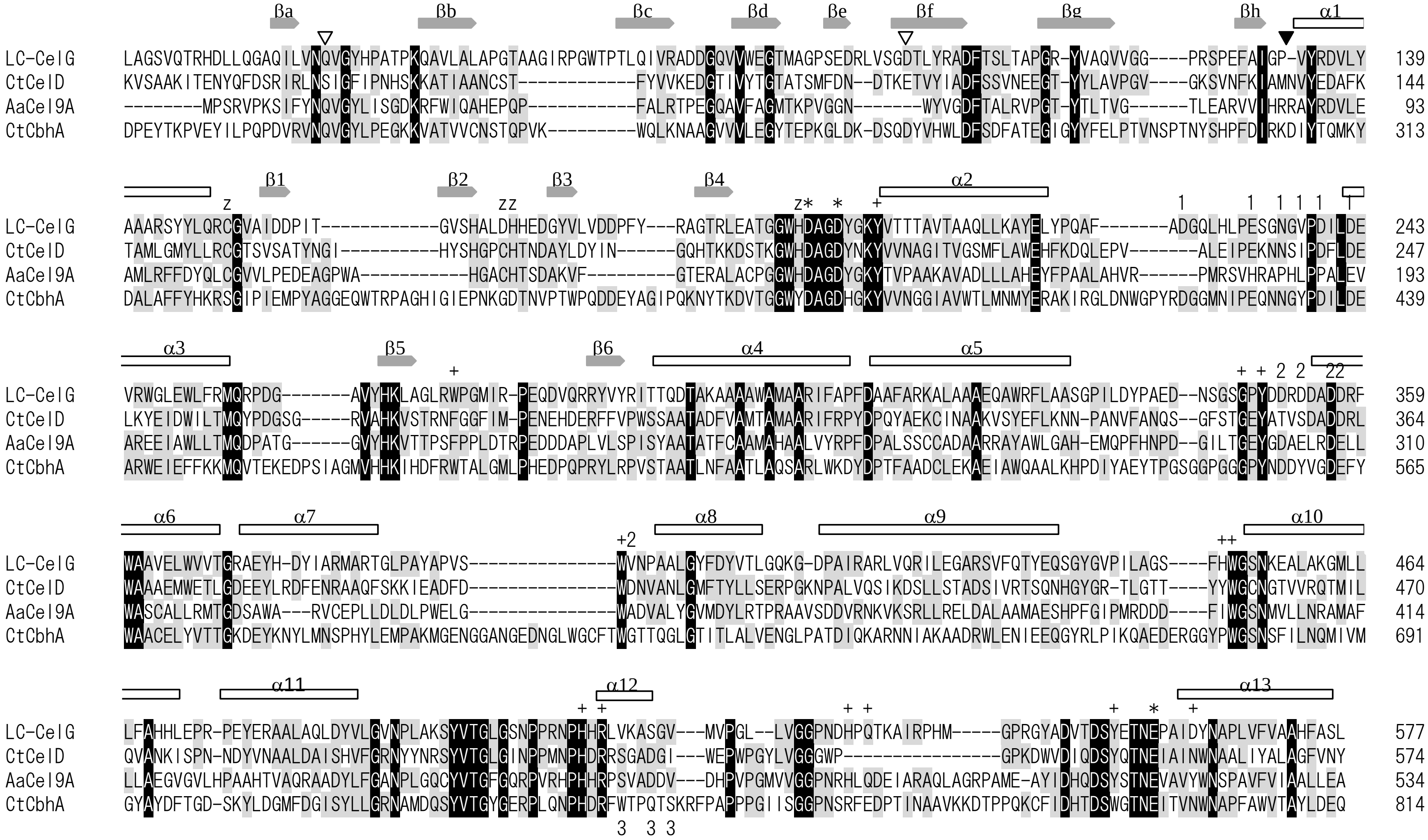

βa βb βc βd βe βf βg βh α1
LC-CelG LAGSVQTRHDLLQGAQILVNQVGYHPATPKQAVLALAPGTAAGIRPGWTPTLQIVRADDGQVVWEGTMAGPSEDRLVSGDTLYRADFTSLTAPGR-YVAQVVGG----PRSPEFAIGP-VYRDVLY 139
CtCelD KVSAAKITENYQFDSRIRLNSIGFIPNHSKKATIAANCST------------FYVVKEDGTIVYTGTATSMFDN--DTKETVYIADFSSVNEEGT-YYLAVPGV----GKSVNFKIAMNVYEDAFK 144
AaCel9A --------MPSRVPKSIFYNQVGYLISGDKRFWIQAHEPQP-----------FALRTPEGQAVFAGMTKPVGGN-------WYVGDFTALRVPGT-YTLTVG------TLEARVVIHRRAYRDVLE 93
CtCbhA DPEYTKPVEYILPQPDVRVNQVGYLPEGKKVATVVCNSTQPVK---------WQLKNAAGVVVLEGYTEPKGLDK-DSQDYVHWLDFSDFATEGIGYYFELPTVNSPTNYSHPFDIRKDIYTQMKY 313
β1 β2 β3 β4 α2
z zz z* * + 1 1 1 1 1 1
LC-CelG AAARSYYLQRCGVAIDDPIT------------GVSHALDHHEDGYVLVDDPFY---RAGTRLEATGGWHDAGDYGKYVTTTAVTAAQLLKAYELYPQAF-------ADGQLHLPESGNGVPDILDE 243
CtCelD TAMLGMYLLRCGTSVSATYNGI----------HYSHGPCHTNDAYLDYIN------GQHTKKDSTKGWHDAGDYNKYVVNAGITVGSMFLAWEHFKDQLEPV-------ALEIPEKNNSIPDFLDE 247
AaCel9A AMLRFFDYQLCGVVLPEDEAGPWA-----------HGACHTSDAKVF---------GTERALACPGGWHDAGDYGKYTVPAAKAVADLLLAHEYFPAALAHVR------PMRSVHRAPHLPPALEV 193
CtCbhA DALAFFYHKRSGIPIEMPYAGGEQWTRPAGHIGIEPNKGDTNVPTWPQDDEYAGIPQKNYTKDVTGGWYDAGDHGKYVVNGGIAVWTLMNMYERAKIRGLDNWGPYRDGGMNIPEQNNGYPDILDE 439
α3 β5 β6 α4 α5
+ + + 2 2 22
LC-CelG VRWGLEWLFRMQRPDG-------AVYHKLAGLRWPGMIR-PEQDVQRRYVYRITTQDTAKAAAAWAMAARIFAPFDAAFARKALAAAEQAWRFLAASGPILDYPAED--NSGSGPYDDRDDADDRF 359
CtCelD LKYEIDWILTMQYPDGSG-----RVAHKVSTRNFGGFIM-PENEHDERFFVPWSSAATADFVAMTAMAARIFRPYDPQYAEKCINAAKVSYEFLKNN-PANVFANQS--GFSTGEYATVSDADDRL 364
AaCel9A AREEIAWLLTMQDPATG------GVYHKVTTPSFPPLDTRPEDDDAPLVLSPISYAATATFCAAMAHAALVYRPFDPALSSCCADAARRAYAWLGAH-EMQPFHNPD--GILTGEYGDAELRDELL 310
CtCbhA ARWEIEFFKKMQVTEKEDPSIAGMVHHKIHDFRWTALGMLPHEDPQPRYLRPVSTAATLNFAATLAQSARLWKDYDPTFAADCLEKAEIAWQAALKHPDIYAEYTPGSGGPGGGPYNDDYVGDEFY 565
α6 α7 α8 α9 α10
+2 ++
LC-CelG WAAVELWVVTGRAEYH-DYIARMARTGLPAYAPVS---------------WVNPAALGYFDYVTLGQKG-DPAIRARLVQRILEGARSVFQTYEQSGYGVPILAGS----FHWGSNKEALAKGMLL 464
CtCelD WAAAEMWETLGDEEYLRDFENRAAQFSKKIEADFD---------------WDNVANLGMFTYLLSERPGKNPALVQSIKDSLLSTADSIVRTSQNHGYGR-TLGTT----YYWGCNGTVVRQTMIL 470
AaCel9A WASCALLRMTGDSAWA---RVCEPLLDLDLPWELG---------------WADVALYGVMDYLRTPRAAVSDDVRNKVKSRLLRELDALAAMAESHPFGIPMRDDD----FIWGSNMVLLNRAMAF 414
CtCbhA WAACELYVTTGKDEYKNYLMNSPHYLEMPAKMGENGGANGEDNGLWGCFTWGTTQGLGTITLALVENGLPATDIQKARNNIAKAADRWLENIEEQGYRLPIKQAEDERGGYPWGSNSFILNQMIVM 691
α11 α12 α13
+ + + + + * +
LC-CelG LFAHHLEPR-PEYERAALAQLDYVLGVNPLAKSYVTGLGSNPPRNPHHRLVKASGV---MVPGL--LVGGPNDHPQTKAIRPHM-----GPRGYADVTDSYETNEPAIDYNAPLVFVAAHFASL 577
CtCelD QVANKISPN-NDYVNAALDAISHVFGRNYYNRSYVTGLGINPPMNPHDRRSGADGI---WEPWPGYLVGGGWP----------------GPKDWVDIQDSYQTNEIAINWNAALIYALAGFVNY 574
AaCel9A LLAEGVGVLHPAAHTVAQRAADYLFGANPLGQCYVTGFGQRPVRHPHHRPSVADDV---DHPVPGMVVGGPNRHLQDEIARAQLAGRPAME-AYIDHQDSYSTNEVAVYWNSPAVFVIAALLEA 534
CtCbhA GYAYDFTGD-SKYLDGMFDGISYLLGRNAMDQSYVTGYGERPLQNPHDRFWTPQTSKRFPAPPPGIISGGPNSRFEDPTINAAVKKDTPPQKCFIDHTDSWGTNEITVNWNAPFAWVTAYLDEQ 814
3 3 3

Figure 2-1. Alignment of amino acid sequences of LC-CelG, CtCelD, AaCel9A and CtCbhA. The accession numbers are KF626654 for LC-CelG, CAA28255 for CtCelD, ACV59481 for AaCel9A, and ABN51651 for CtCbhA. The amino acid sequence of LC-CelG without a putative signal peptide (residues 20-577) and the corresponding regions of other three proteins are shown. The amino acid residues, which are conserved in all four proteins, are denoted with white letters and highlighted in black. The amino acid residues, which are conserved in two or three different proteins, are highlighted in gray. The ranges of the secondary structures of LC-CelG (βa-βh strands for the Ig-like domain, and β1-β6 strands and α1-α13 helices for the cellulase domain) are shown above its sequence based on its crystal structure. The residues that form the catalytic site (Asp194, Asp197 and Glu558), substrate binding site, Zn site (Cys150, Asp166, His167 and His193), Ca1 site (Asp225, Glu232, Asn235, Val237, Asp239 and Asp242) and Ca2 site (Asp351, Asp353, Asp356, Asp357 and Val395) are indicated by "∗", "+", "z", "1" and "2" respectively, above the LC-CelG sequence. The residues that form the Ca3 site of CtCelD and AaCel9A are indicated by "3" below the sequences. The position, at which N-terminal Ig-like domain of LC-CelG is truncated to construct His-ΔIg-CelG, is shown by solid arrow head above the LC-CelG sequence. Likewise, Gln40 and Asp99 of LC-CelG, which are mutated in this study, are indicated by open arrow heads.

His-LC-CelG の酵素活性の pH 依存性を、60°C 、pH 4.0 – 10.0 の範囲で CM セルロースを基質として測定した。Figure 2-2A に示すように、His-LC-CelG は pH 5.0 – 9.0 の範囲で最大活性の 80%以上の活性を示した。同様に His-LC-CelG の酵素活性の温度依存性を、CM セルロースを基質として pH 7.0、40 - 90°C の範囲で測定した。Figure 2-2B に示すように、His-LC-CelG は 70°C で最大活性を示した。70°C、pH 7.0 における His-LC-CelG の比活性は 50.1 ± 0.3 units mg^{-1} であった。70°C、pH 5.5 における AaCel9A の比活性は 80 units mg^{-1} であることが報告されており [71]、His-LC-CelG は AaCel9A よりも活性が若干弱いが、同程度の耐熱性を示すことが示唆された。一方で、AaCel9A は最大活性の 50%以上を示す範囲が pH 4.5 – 7.0 であるため、His-LC-CelG の方が AaCel9A よりも pH に対する適合範囲が広いことが示唆された。なお、LC-CelG の pH や温度依存性は His-LC-CelG と類似していたので、N 末端への His-tag の付加は LC-CelG の活性にはほとんど影響を及ぼさない。

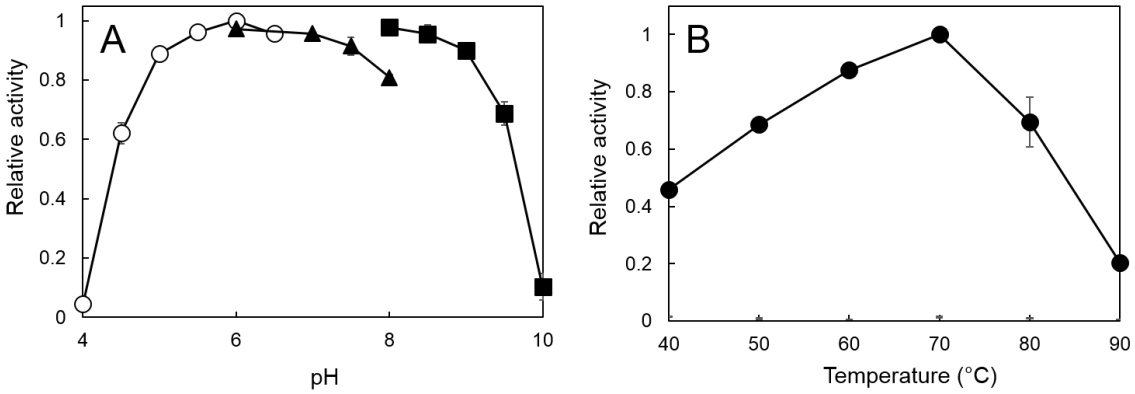


Figure 2-2. Optimum pH and temperature for activity of His-LC-CelG. The pH (A) and temperature (B) dependencies of the enzymatic activity of His-LC-CelG are shown. The activity was determined at 60°C at the pH indicated (A) or at pH 7.0 and the temperatures indicated (B) using 1% (w/v) CM-cellulose as a substrate, as described in Materials and Methods. The buffers used to analyze the pH dependence of the activity were 100 mM sodium citrate (pH 4.0-6.5), 100 mM sodium phosphate (pH 6.0-8.0) and 100 mM Glycine-NaOH (pH 8.0-10.0). The experiment was carried out at least twice, and errors from the average values are indicated by *vertical lines*.

2-3-4. His-LC-CelG の耐熱性

His-LC-CelG の耐熱性を調べるため、5 mM $CaCl_2$ 存在下、pH 7.0 の条件で熱変性曲線を測定した。なおこの条件での熱変性は不可逆であった。Figure 2-3 に His-LC-CelG の熱変性曲線を示す。変性中点温度 ($T_{1/2}$) は 81.4°C であった。AaCel9A の $T_{1/2}$ 値は 77.9°C であるため [69]、His-LC-CelG の耐熱性が僅かに高い。また His-LC-CelG の $T_{1/2}$ 値は活性の至適温度 (70°C) よりも高く、これは活性部位付近が温度変化に伴い局所的に先に変性を起こすためであると考えられる。なお、LC-CelG の $T_{1/2}$ 値は His-LC-CelG と類似していたので、N 末端への His-tag の付加は LC-CelG の安定性にはほとんど影響を及ぼさない。

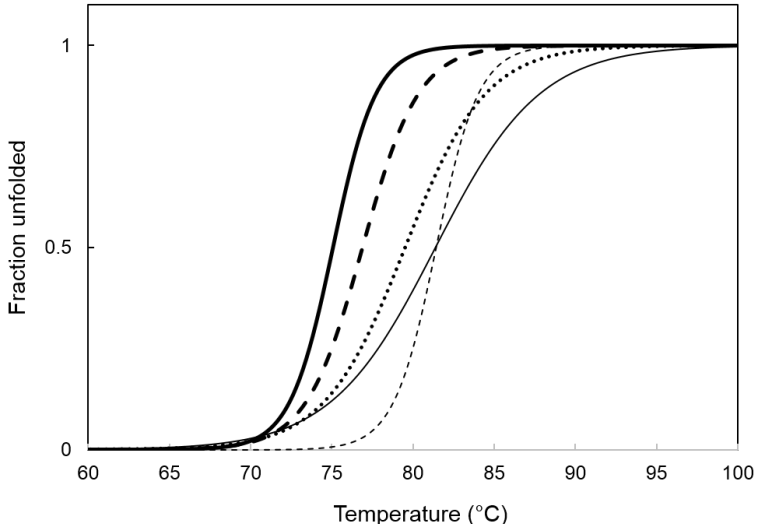


Figure 2-3. Thermal denaturation curves of His-LC-CelG and its derivatives. The thermal denaturation curves of His-LC-CelG (thin solid line), His-ΔIg-CelG (thick solid line), His-Q40A-CelG (thin dashed line), His-D99A-CelG (thick dotted line) and His-Q40A/D99A-CelG (thick dashed line) are shown. These curves were obtained in the presence of 5 mM $CaCl_2$ at pH 7.0 by monitoring the change in CD values at 222 nm as described in Materials and Methods.

2-3-5.　LC-CelG の結晶構造

LC-CelG と他の GHF9 セルラーゼの構造が類似しているかどうかを調べるために、LC-CelG の結晶構造を 2.15 Å の分解能で決定した。非対称ユニット中に 2 分子 (A、B) 存在した。LC-CelG は溶液中ではモノマーで存在するため、結晶中に見られる分子間相互作用は結晶のパッキングによるものと考えられた。いずれも N 末端の Met 残基と Ig-like ドメインの一部 (Leu20 – Asp29) は不規則な揺らぎのため、その電子密度を観測することはできなかった。分子 A と B の構造は極めて類似しており、その平均二乗偏差 (RMSD) 値は 0.30 Å であった。データ収集および精密化の統計値を Table 2-1 に示す。

Table 2-1. Data collection and refinement statistics of LC-CelG

Crystal	LC-CelG
Wavelength (Å)	0.900
Space group	$P2_12_12_1$
Cell parameters	
a, b, c (Å)	84.631, 89.913, 151.157
α= β=γ (°)	90.00
Molecules/asymmetric unit	2
Resolution range (Å)	50.00-2.15 (2.19-2.15)[a]
Reflections measured	941,321
Unique reflections	63,355
Redundancy	14.9 (15.0)
Completeness (%)	99.9 (100.0)[a]
R_{merge} (%)[b]	13.3 (68.7)[a]
Average $I/\sigma(I)$	28.0 (2.6)[a]
Refinement statistics	
Resolution limits (Å)	77.28-2.15
No. of atoms	
Protein/water	4239/410
R_{work} (%) / R_{free} (%)[c]	18.7/24.6
Rms deviations from ideal values	
Bond lengths (Å)	0.011
Bond angles (°)	1.190
Average B factors (Å^2)	
MonomerA / B	47.6 / 39.6
Zn^{2+}/Ca^{2+}1/Ca^{2+}2 in monomer A	40.6 / 79.3 / 46.0
Zn^{2+}/Ca^{2+}1/Ca^{2+}2 in monomer B	34.3 / 56.8 / 33.6
Water	44.3
Ramachandran plot statistics	
Preferred regions (%)	95.3
Allowed regions (%)	4.5

[a] Values in parentheses are for the highest-resolution shell.

[b] $R_{merge} = \sum|I_{hkl} - \langle I_{hkl}\rangle|/\sum I_{hkl}$, where I_{hkl} is an intensity measurement for reflection with indices *hkl* and $\langle I_{hkl}\rangle$ is the mean intensity for multiply recorded reflections.

[c] Free *R*-value was calculated using 5% of the total reflections chosen randomly and omitted from refinement.

LC-CelG と CtCelD (PDB code: 1CLC) の結晶構造の重ね合わせ図を Fig. 2-4A に示す。これらの構造は比較的よく類似しており、RMSD 値は 1.58 Å であった。同様に LC-CelG の構造は AaCelA や CtCbhA とも類似しており、RMSD 値はそれぞれ 1.49 Å、1.92 Å であった。LC-CelG の触媒ドメインの構造は典型的な $(\alpha/\alpha)_6$-バレル構造であり、中心の$(\alpha/\alpha)_6$-バレルを形成する 12 本の長いα–ヘリックスと 2 本の短いα–ヘリックス、及び 6 本ずつの逆平行β鎖で形成される 2 つのβ-シートから成る。Ig-like ドメインは 4 本ずつの逆平行β鎖で形成される 2 つのβ-シートから成る。3 つの触媒残基 (Asp194、Asp197、Glu558) の立体配置は CtCelD (Fig. 2-4A)、AaCel9A (Fig. 2-4B)、CtCbhA とほぼ一致していた。Asp194 と Asp197 は求核試薬として働くと考えられる水分子と水素結合を形成していた。Asp194、Asp197 とその水分子の距離はそれぞれ 2.6 Å、2.7 Å である。

2-3-6. 基質結合ポケット

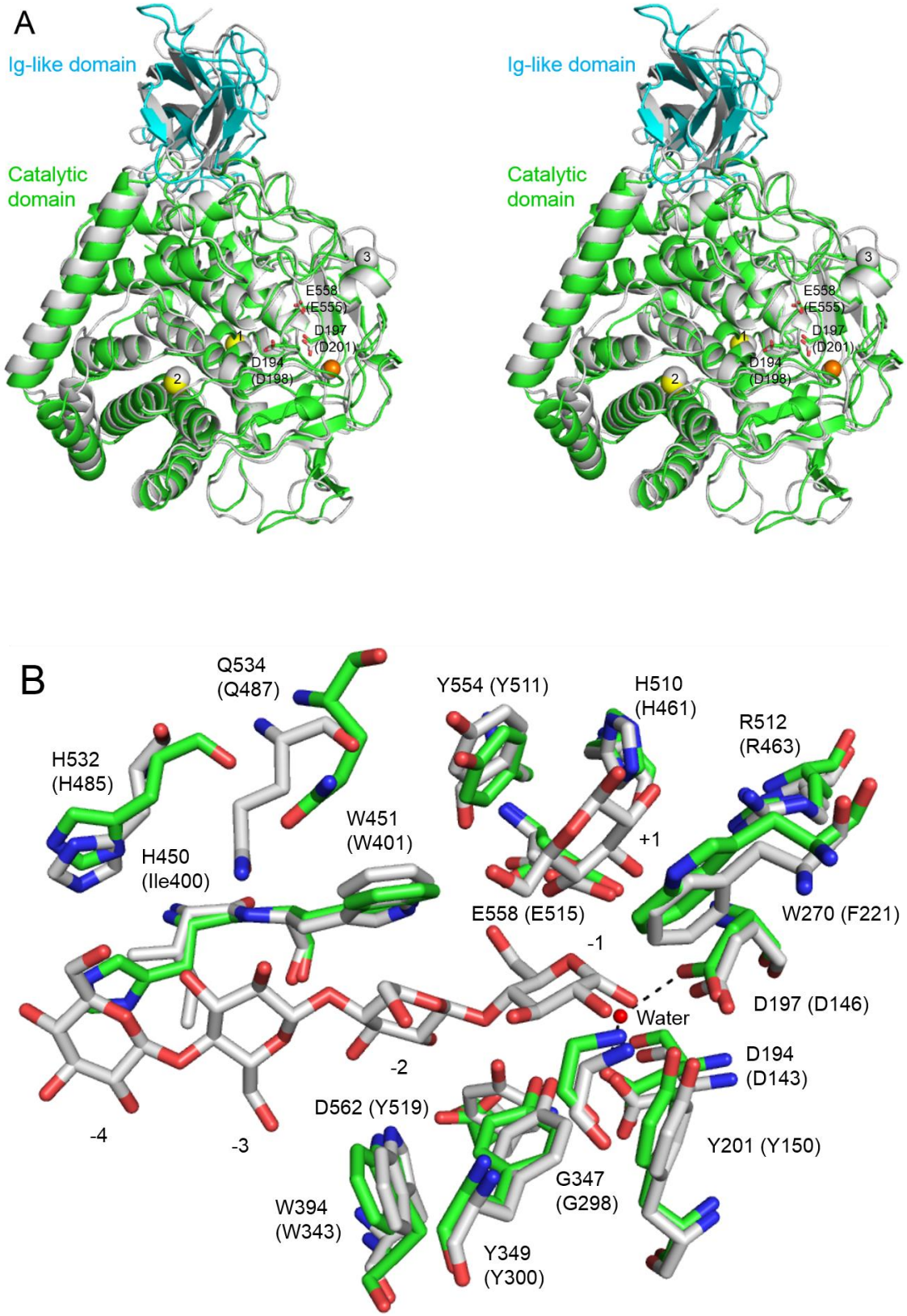
AaCel9A のセロテトラオース複合体構造 (PDB code: 3H3K) を基に、LC-CelG の基質結合ポケット構造を比較した (Fig. 2-4B)。LC-CelG の基質結合ポケットを形成するアミノ酸の立体配置は AaCel9A と非常に類似しており、これらの残基が AaCel9A と同様に基質と相互作用することが示唆された。Tyr201 OH、Gly347 N、Tyr349 OH、Glu558 $O^{\varepsilon 1}$、Asp562 $O^{\delta 1}$ はサブサイト-1 のグリコシルユニット (Glc(-1))と水素結合を形成すると考えられた。Trp394 $N^{\varepsilon 1}$ は Glc(-2)と水素結合を形成し、Trp451 はスタッキング相互作用によりその結合を促進させると考えられた。His450 と His532 はスタッキング相互作用により Glc(-4)の結合を促進すると考えられた。His510 と Arg512 は Glc(+1)と水素結合を形成し、Trp270 と Tyr554 はスタッキング相互作用によりその結合を促進すると考えられた。一方、AaCel9A において Gln487 N^{ε}が Glc(-3)と水素結合を形成するが、LC-CeG では対応する残基である Gln534 N^{ε}の位置が異なるため、同様に水素結合を形成するかどうかは不明である。またサブサイト-1 と+1 の間で基質が大きく捻じ曲がらないとサブサイト+1 には結合できないため、AaCel9A-セロテトラオース複合体構造のサブサイト+1 が基質結合に関わるかどうかはまだ不明である。セロテトラオースは AaCel9A により加水分解されるため、AaCel9A-基質複合体構造の決定に利用された [70]。しかし、AaCel9A の結晶へのソーキングを行った際、セロテトラオースはサブサイト-4 から-1 に生成物のような形で結合した [70]。

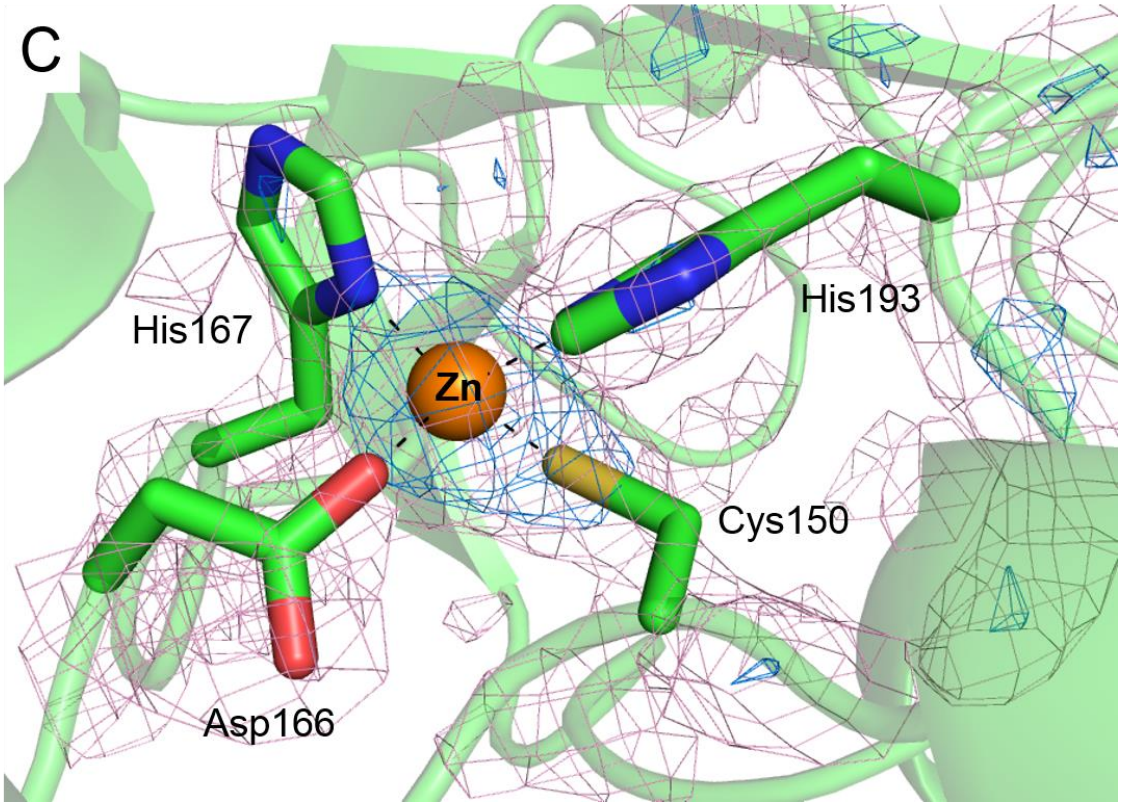
AaCel9A の基質複合体構造によると、Asp143 と Asp146 (LC-CelG において

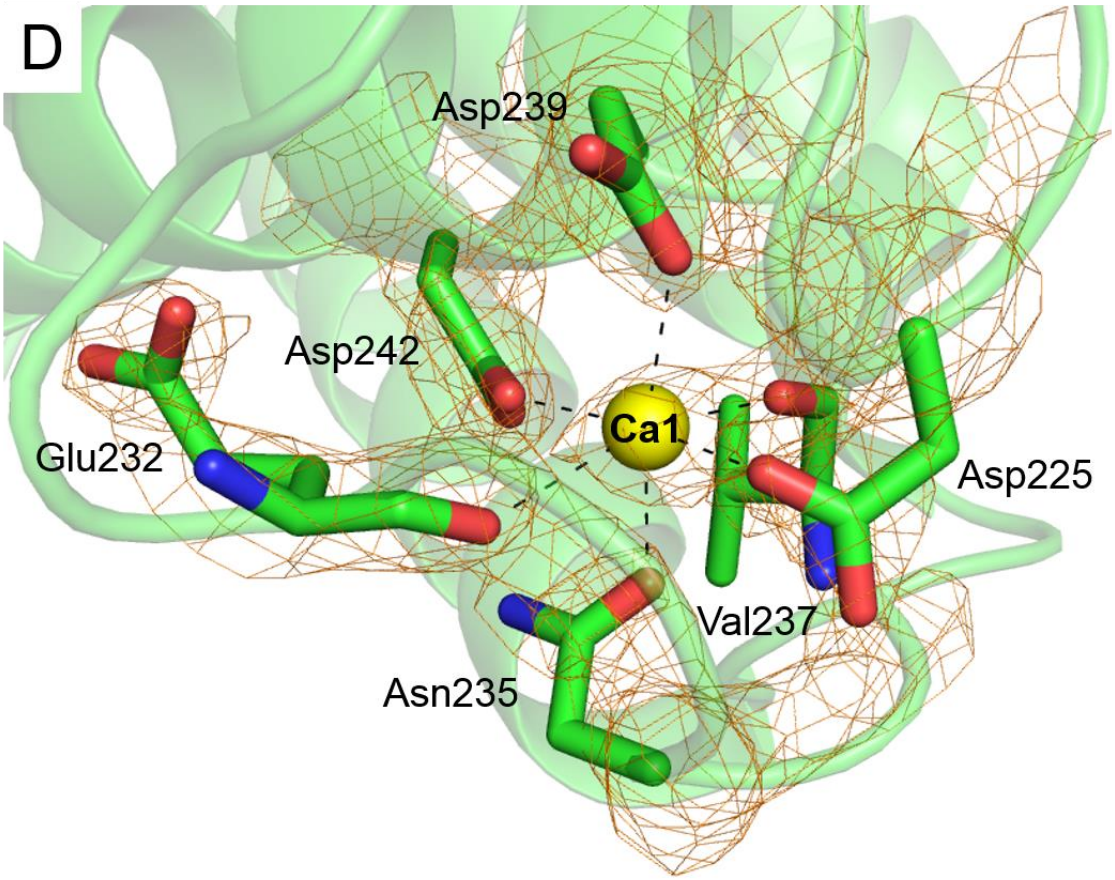
Asp194 と Asp197) は Glc(-1)の O1 原子と水素結合を形成する。しかしセロテトラオースが結合していない構造によると、求核試薬として機能すると考えられる水分子が O1 原子の位置に結合することが示唆された [70]。この O1 原子は Glc(-1)と Glc(+1)がグリコシド結合を形成しているときは存在しないが、この結合が切断されたときは存在した。この結果より、Glc(-1)と Glc(+1)の間の C1-O4 結合が開裂した後の O1 原子のように、水分子は LC-CelG の Asp194 と Asp197 に結合することが示唆された。

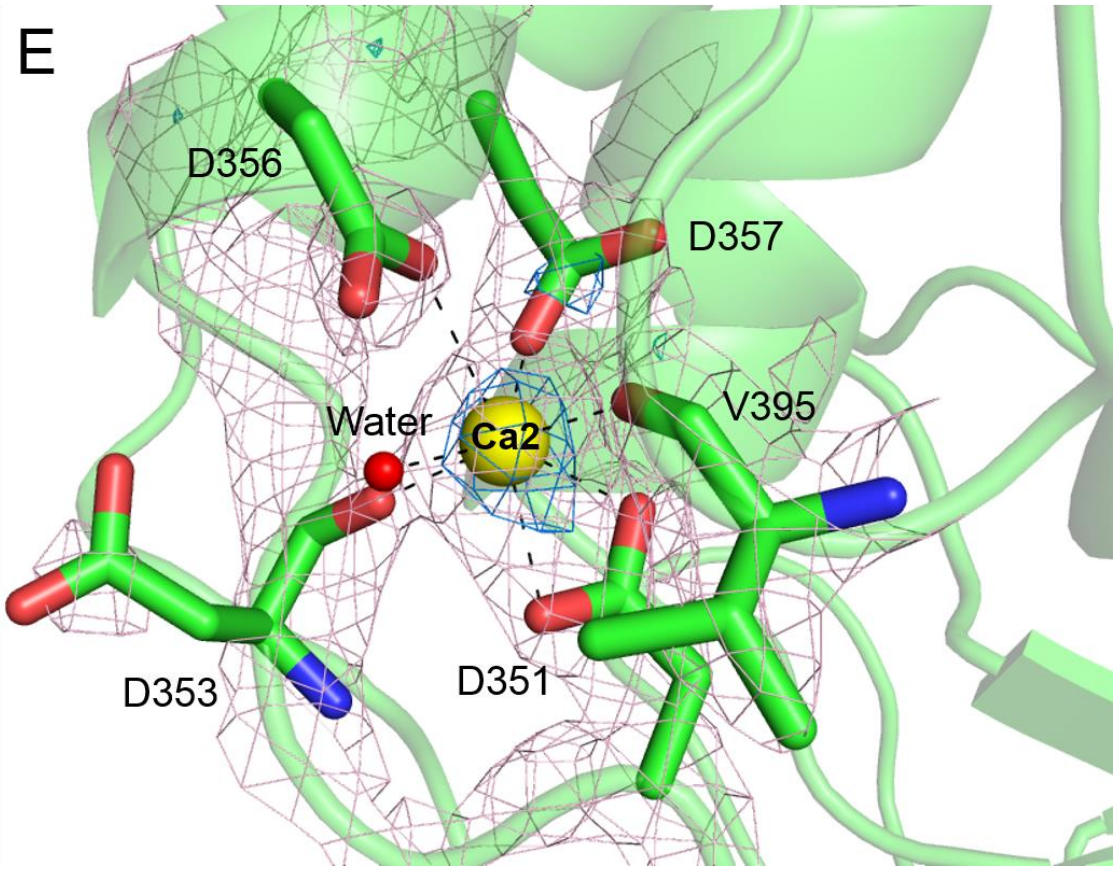
2-3-7. 金属結合部位

CtCelD や AaCel9A、CtCbhA と LC-CelG のアミノ酸配列を比較すると、LC-CelG は 1 つの Zn^{2+}結合部位 (Zn site) と 2 つの Ca^{2+}結合部位 (Ca1 site、Ca2 site) を持つことが予測された (Fig. 2-1)。CtCelD の Zn site、Ca1 site、Ca2 site に対応する LC-CelG の各部位の 2*F*o-*F*c マップをそれぞれ Figure 2-4C、D、E に示す。結晶化試薬に金属イオンは存在せず、Zn^{2+}イオンの配位構造は Ca^{2+}イオンとは異なるため、それぞれの電子密度ピークを Zn^{2+}イオン、Ca^{2+}イオン (Ca1、Ca2) と同定した。CtCelD の Ca3 site に対応する領域には明確な電子密度ピークは観測されなかったため、LC-CelG は Ca3 site を持たないと考えられた。Zn^{2+}イオンは Cys150 SH、Asp166 $O^{\delta2}$、His167 $N^{\delta1}$、His193 $N^{\varepsilon2}$ と 4 配位している (Fig. 2-4C)。Ca1 site の Ca^{2+}イオンは Asp225 $O^{\delta1}$、Glu232 O、Asn235 O^{δ}、Val237 O、Asp239 $O^{\delta1}$、Asp242 $O^{\delta2}$ と 6 配位している (Fig. 2-4D)。Ca2 site の Ca^{2+}イオンは Asp351 $O^{\delta1}$、Asp351 $O^{\delta2}$、Asp353 O、Asp356 $O^{\delta1}$、Asp357 $O^{\delta2}$、Val395 O、水分子と 7 配位している (Fig. 2-4E)。ただし Ca1 site の Ca^{2+}イオンの電子密度ピークは他の 2 つの金属イオンよりも弱い。これは恐らくこの部位への Ca^{2+}イオンの結合が弱く、結晶化や特性解析を行う際に Ca^{2+}イオンを含まない緩衝液に透析することにより、この部位から Ca^{2+}イオンが部分的に解離したためであると考えられる。しかし 10 mM $CaCl_2$ 存在下での酵素活性は、非存在下とほぼ同程度であったことから、この部位への Ca^{2+}イオンの結合は活性には特に影響しないと考えられた。一方、他の 2 つの金属イオンは、金属を含まないバッファーに透析した後も LC-CelG に強く結合していた。


A
Ig-like domain
Catalytic domain
3
E558
(E555)
D197
(D201)
1
D194
(D198)
2
Ig-like domain
Catalytic domain
3
E558
(E555)
D197
(D201)
1
D194
(D198)
2
B
Q534
(Q487)
Y554 (Y511)
H510
(H461)
R512
(R463)
H532
(H485)
W451
(W401)
+1
H450
(Ile400)
E558 (E515)
W270 (F221)
-1
D197 (D146)
Water
D194
(D143)
-2
D562 (Y519)
-4
-3
Y201 (Y150)
G347
(G298)
W394
(W343)
Y349
(Y300)


C
His167
Zn
His193
Cys150
Asp166


D
Asp239
Asp242
Ca1
Glu232
Asp225
Val237
Asn235


E
D356
D357
Water
Ca2
V395
D353
D351

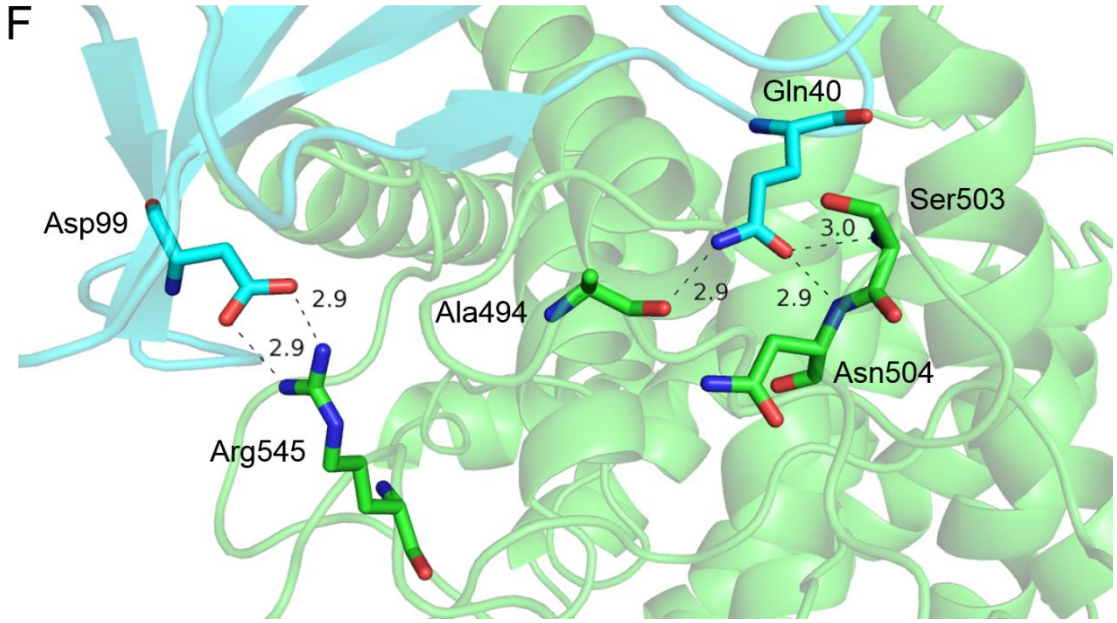

F
Gln40
Ser503
Asp99
3.0
2.9
2.9
2.9
2.9
Ala494
Asn504
Arg545

Figure 2-4. Crystal structure of LC-CelG. (A) A stereo view of the overall structure of LC-CelG. The structure of LC-CelG is superimposed on that of CtCelD (PDB code 1CLC). For the LC-CelG structure, the Ig-like and catalytic domains are colored cyan and green respectively. One zinc and two calcium ions (Ca1 and Ca2) are shown as orange and yellow spheres, with the numbers of the calcium ions indicated. Three active site residues (Asp194, Asp197 and Glu558) are indicated by green stick models, in which the oxygen and nitrogen atoms are colored red and blue respectively. The entire CtCelD structure, including one zinc and three calcium ions (Ca1-Ca3), is colored gray. Three active site residues (Asp198, Asp201 and Glu555) are indicated by gray stick models. They are labeled in parentheses. (B) The structure around the substrate binding pocket. The structure around the substrate binding pocket of LC-CelG is superimposed on that of AaCel9A in complex with cellotetraose (PDB code 3H3K). The residues forming the substrate binding pocket of LC-CelG are indicated by green stick models, and the corresponding residues and cellotetraose in the AaCel9A structure are indicated by gray stick models. In these stick models, the oxygen and nitrogen atoms are colored red and blue respectively. The labels for the AaCel9A structure are shown in parentheses. The five subsites are labeled from −4 to +1. The water molecule is shown as red sphere. Dashed lines represent hydrogen bonds. (C-E) Electron density around the binding sites of zinc ion (C), and calcium ions Ca1 (D) and Ca2 (E). The zinc and calcium ions are shown as orange and yellow spheres respectively. The water molecule is shown as red sphere. The residues coordinated with these metal ions are indicated as shown in (B). In (C) and (E), the $2F_o$–F_c maps contoured at the 2.0σ and 4.0σ levels are shown in magenta and blue respectively. In (D), the $2F_o$–F_c map contoured at the 1.5σ level is shown in orange. (F) The structure around Gln40 and Asp99. Gln40 and Asp99 located in the Ig-like domain are shown as cyan stick models. The residues that are located in the catalytic domain and form hydrogen bonds with Gln40 (Ala494, Ser503 and Asn504) or salt bridge with Asp99 (Arg545) are shown by green stick models. In these stick models, the oxygen and nitrogen atoms are colored red and blue respectively

2-3-8. Ig-like ドメインと触媒ドメインの相互作用

CtCelD や AaCel9A、CtCbhA と同様に、LC-CelG の Ig-like ドメインと触媒ドメインの界面には多数の親水性、疎水性相互作用が形成される。しかし、LC-CelG の Ig-like ドメインに存在する Gln40 と、触媒ドメインの α11-α12 ヘリックス間のループの間に形成される水素結合は、4 つの GHF9 セルラーゼの間で保存されている唯一のドメイン間相互作用である。Gln40 は AaCel9A では Gln13、CtCbhA では Gln218 として保存されている。CtCelD では Ser58 となっているが、他と同様に α11-α12 ヘリックス間のループと水素結合を形成している。Gln40 の側鎖が形成する水素結合を Figure 2-4F に示す。Gln40 N^{ε}と Ala494 O 間、Gln40 N^{ε}と Ser503 N 間、Gln40 O^{ε}と Asn504 N 間の距離はそれぞれ 2.9、3.0、2.9 Å である。また、LC-CelG の Asp99 と Arg545 の間に形成される塩橋は、Ig-like ドメインと触媒ドメインの間に形成される唯一の塩橋であり (Fig. 2-4F)、CtCelD (Glu103 と Lys542) と CtCbhA (Asp267 と Lys782) でも保存されている。AaCel9A においては、この塩橋を形成する酸性残基が位置するループを欠いていることから (Fig. 2-1)、この塩橋は形成されていない。

2-3-9. LC-CelG の活性や耐熱性における Ig-like ドメインの役割

CtCbhA は Ig-like ドメインの欠損により不活性化することが報告されているが [69]、GHF9 セルラーゼにおける Ig-like ドメインの役割はまだそれほど良く分かっていない。CtCbhA の Ig-like ドメインと触媒ドメイン間に形成される 1 つ (Asp264 – Tyr676)、あるいは 2 つ (Thr230 – Gly661、Asp262 – Gly661) の水素結合を取り除くことにより、耐熱性はそれぞれ 8.7°C、6.0°C 低下するが、活性は保持している [69]。しかし Ig-like ドメインの欠損により、他の GHF9 酵素も同様に失活するかどうかはまだ分かっていない。また GHF9 セルラーゼに保存されている Ig-like ドメインと触媒ドメイン間の相互作用が安定性や活性に影響するかどうかはまだ調べられていない。なお、CtCbhA における Asp264 – Tyr676、Thr230 – Gly661、Asp262 – Gly661 間の相互作用は LC-CelG では保存されていない。

そこで LC-CelG が Ig-like ドメインの欠損により失活するのかどうかを調べるために、Ig-like ドメイン欠損変異体 His-ΔIg-CelG を構築した。また、Gln40 が形成する水素結合と、Asp99-Arg545 間の塩橋が His-LC-CelG の活性や安定性に影響するかどうかを調べるために、Gln40、Asp99 のいずれか、あるいは両方を

Ala に変異した His-Q40A-CelG、His-D99A-CelG、His-Q40A/D99A-CelG をそれぞれ構築した。いずれも 2-2-2 の手順に従い精製を行った。いずれの変異体も大腸菌における発現レベルは His-LC-CelG と同程度であった。しかし His-ΔIg-CelG はほとんどが不溶化してしまい、可溶性タンパク質として菌体内に蓄積した量はごく僅かであったため、精製後の収量が他と比べて極めて少なくなった。他の変異体はいずれも大部分が可溶性タンパク質として菌体内に蓄積した (Fig. 2-5)。結果として、1 L の培養液から精製された量は、His-ΔIg-CelG の場合は 0.3 mg、他の変異体の場合は His-LC-CelG と同程度の約 4 mg であった。

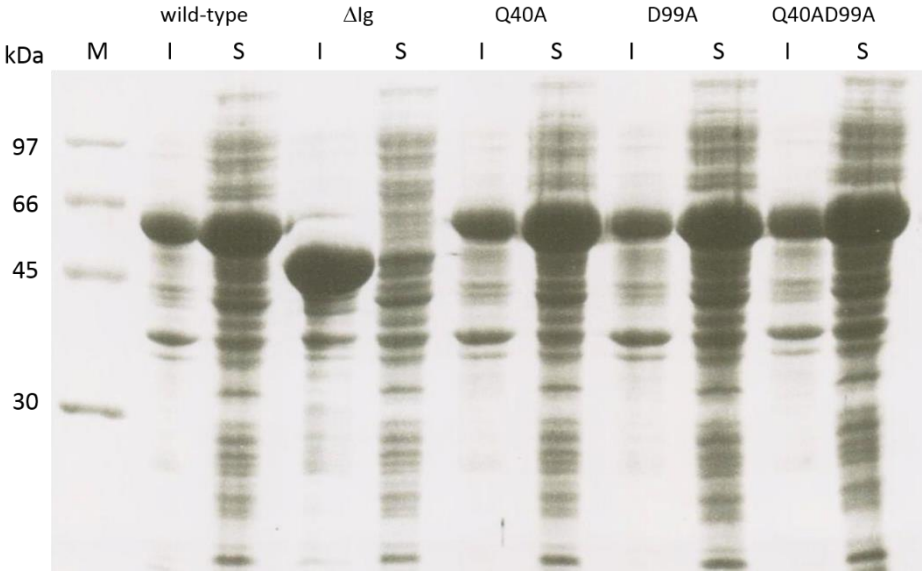


Figure 2-5. SDS-PAGE analysis of His-LC-CelG and its derivatives. “I” and “S” indicate insoluble form and soluble form after sonication, respectively.

これらの変異タンパク質の Far-UV CD スペクトルを pH 7.0、25°C の条件で測定し、His-LC-CelG のスペクトルと共に Figure 2-6 にまとめた。His-ΔIg-CelG 以外のスペクトルは類似しており、Gln40 や Asp99 の変異が His-LC-CelG の二次構造にほとんど影響しないことが示唆された。His-ΔIg-CelG のスペクトルの形は His-LC-CelG に類似しているが、210 – 220 nm の範囲の谷が深いため、His-ΔIg-CelG のヘリックス含量は His-LC-CelG よりも多いことが示唆された。結晶構造より、His-LC-CelG、His-ΔIg-CelG のヘリックス含量はそれぞれ 36.6%、45.8%であり、これは Far-UV CD スペクトルの結果と一致している。これにより Ig-like ドメインの欠損による触媒ドメインの大きな構造変化は生じていないことが示唆された。

pH 7.0、50°C の条件で CM セルロースを基質として酵素活性を測定した。His-LC-CelG の酵素活性は Ig-like ドメインの欠損により 1/100 にまで低下したのに対して、Gln40 や Asp99 の変異によりほとんど低下しなかった (Table 2-2)。His-ΔIg-CelG は 37°C でもほとんど活性を示さなかった。この結果より、Ig-like ドメインの欠損は His-LC-CelG の酵素活性を大幅に低下させるが、Gln40 や Asp99 の変異は酵素活性には大きな影響を及ぼさないと考えられた。

これらの変異体の熱変性を、5 mM $CaCl_2$ 存在下、pH 7.0 の条件で、222 nm の CD 値をモニターすることにより解析した。なおこの条件での熱変性は不可逆的であった。Figure 2-3、Table 2-2 に熱変性曲線と $T_{1/2}$ 値をそれぞれまとめる。His-ΔIg-CelG と His-Q40A/D99A-CelG の $T_{1/2}$ 値は His-LC-CelG よりもそれぞれ 6.3°C、5.0°C 低下した。この結果より、Ig-like ドメインは His-LC-CelG の耐熱化に寄与するが、その 80%程度は Gln40 と Asp99 が介するドメイン間相互作用によるものであることが示唆された。しかし His-Q40A-CelG と His-D99A-CelG の $T_{1/2}$ 値は、それぞれ His-LC-CelG と同程度か 1.9°C のみ低いことから、Gln40 と Asp99 が介するドメイン間相互作用は協調的に耐熱化に寄与することが示唆された。

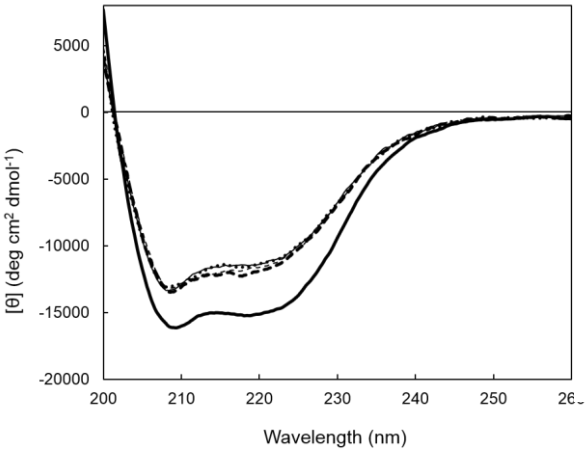


Figure 2-6. CD spectra of His-LC-CelG and its derivatives. The far-UV CD spectra of His-LC-CelG (thin solid line), His-ΔIg-CelG (thick solid line), His-Q40A-CelG (thin dashed line), His-D99A-CelG (thick dotted line) and His-Q40A/D99A-CelG (thick dashed line) were measured at pH 7.0 and 25°C, as described in Materials and Methods.

Table 2-2. Activities and stabilities of His-LC-CelG and its mutants

Protein	Specific activity[a] (Units mg^{-1})	Relative activity[b] (%)	$T_{1/2}$[c] (°C)	$\Delta T_{1/2}$[d] (°C)
His-LC-CelG	34.4±0.4	100	81.4±0.2	–
His-Q40A-CelG	32.2±0.9	94	81.4±0.2	0
His-D99A-CelG	36.3±1.2	106	79.5±0.1	-1.9
His-Q40A/D99A-CelG	33.2±0.1	97	76.4±0.2	-5.0
His-ΔIg-CelG	0.4±0.2	1.1	75.1±0.1	-6.3

[a] The specific activity was determined at pH 7.0 and 50°C using CM-cellulose as a substrate, as described in Materials and Methods. Each experiment was carried out at least twice and the average value is shown together with the error.

[b] The relative activity was calculated by dividing the specific activity of the protein by that of His-LC-CelG.

[c] The temperature of the midpoint of the thermal denaturation transition, $T_{1/2}$, was determined from the thermal denaturation curves shown in Figure 2-3.

[d] $\Delta T_{1/2} = T_{1/2}$ determined – 81.4°C .

2-4. 考察

LC-CelG は N 末端に Ig-like ドメインを有する GHF9 セルラーゼである。今回 LC-CelG と 29-31%のアミノ酸配列相同性を示す 3 つの GHF9 セルラーゼの結晶構造を基に、Ig-like ドメインの役割について活性や耐熱性の面から考察した。His-LC-CelG の Ig-like ドメイン欠損変異体 His-ΔIg-CelG はほとんど失活し、耐熱性は 6.3°C 低下した (Table 2-1)。CtCbhA においても Ig-like ドメインの欠損変異体は失活し、熱変性温度は 10.3°C 低下した [69]。CtCbhA における考察を基にすると [69]、His-LC-CelG の Ig-like ドメインを欠損することにより、Ig-like ドメインと触媒ドメイン間の複数の相互作用が失われ、活性部位のコンフォメーションが僅かに変化するという仮説が考えられた。触媒残基 Glu558 や基質結合ポケットを形成する残基が複数位置しているα11 とα13 ヘリックスの間にある長いループ領域は、Ig-like ドメインと複数の相互作用を形成している。ゆえに Ig-like ドメインの欠損によりこのループのコンフォメーションが変化すると考えられる。つまり Ig-like ドメインは活性部位のコンフォメーションの形成と基質結合ポケットの機能を持たせるために必要であるのかもしれない。AaCel9A の分子動力学シミュレーションによると、Ig-like ドメインは基質結合ポケットや触媒部位の残基と動力学的に相関があると示唆された [72]。また CtCbhA は p-ニトロフェニルセロビオシドの加水分解活性を有するが、Ig-like ドメインの欠損変異体はその活性を失うため、Ig-like ドメインが基質認識に直接関与する可能性は低いと考えられた [69]。この基質のセロビオシド部分は、Ig-like ドメインから離れたサブサイト-1 と-2 に結合する。つまり Ig-like ドメインは間接的に活性に強く関与していると考えられる。今回の研究では、His-LC-CelG の Gln40 と Asp99 の二重変異体は熱変性温度が 5.0°C 低下したが、活性は保持していた (Table 2-1)。つまり Gln40、Asp99 の変異のみでは、この長いループ領域のコンフォメーションを変化させるのに十分ではないと考えられる。

Ig-like ドメインは His-LC-CelG や CtCbhA の安定化に寄与していることが分かった。このことにより、これらの GHF9 セルラーゼが高温環境に適応するのにこのドメインを必要とすること示唆された。しかし耐熱性の高い GHF9 酵素全てがこのドメインを有しているわけではない。様々な生物から単離された Ig-like ドメインを有さない GHF9 酵素も、様々な温度環境に適応している [73-75]。このことから Ig-like ドメインは GHF9 セルラーゼの耐熱化に必ずしも必要ではないことが示唆された。

Ig-like ドメインを欠損させると、大腸菌内での AaCel9A の発現レベルが大幅に低下することが報告されている [72]。これは、Ig-like ドメインの欠損により、触媒ドメインが正しくフォールディングされないことによるものであると推測されている。LC-CelG においても同様の結果が得られた。His-ΔIg-CelG は大腸菌内で大半が不溶型として蓄積した。これは Ig-like ドメインの欠損が His-LC-CelG のフォールディングに影響するためと考えられた。しかし、大腸菌内で可溶型として蓄積した His-ΔIg-CelG は、CD スペクトルによると His-LC-CelG の触媒ドメインと類似した二次構造を有していることが予測された。一方でその活性は極めて弱く、His-LC-CelG の 1%程度であった。これらの結果より、Ig-like ドメインは GHF9 酵素の触媒ドメインのフォールディングを助ける分子内シャペロンのような機能を有すると考えられる。このドメインを有していない場合、触媒ドメインはほとんど正しくフォールディングできないか、あるいはフォールディングが非常に遅くなり、フォールディングプロセスが完了する前に封入体を形成してしまうのかもしれない。

第 3 章　N 末端に長い伸長領域を有する新規エステラーゼ LC-Est1 の構造と諸特性解析

3-1.　はじめに

エステラーゼ/リパーゼはあらゆる生物が保有している酵素であり、その触媒機構や基質認識機構などについてはこれまでよく研究されている。エステラーゼ/リパーゼは脂肪酸エステルやトリグリセリドのカルボン酸エステルを加水分解するが、カルボン酸エステル結合を有する化合物は多岐にわたり、その基質特異性に従ってエステラーゼ/リパーゼは多数に分類されている。実際、基質特異性の違いに基づき、医薬品、食品の油脂加工、廃液中の油脂処理をはじめ、近年ではバイオディーゼル燃料の合成や生分解性プラスチックの加水分解といったバイオマス産業など、エステラーゼ/リパーゼは多種多様な分野で利用されている。このような様々な用途に適したエステラーゼ/リパーゼを効率よく見つけるためには、既存のエステラーゼ/リパーゼのバリエーションを増やす必要がある。従って、新規エステラーゼ/リパーゼの単離は現在でも望まれている。

今回新規酵素の取得源として注目している枝葉コンポストは、第 1 章で示したセルラーゼ基質であるセルロースだけではなく、植物性油脂や天然樹脂ポリマーも豊富であると考えられる。ゆえにこれらに特異的に働くエステラーゼ/リパーゼを生産する好熱性微生物群が枝葉コンポストには集積していると予想される。また、土壌細菌から単離されたエステラーゼ/リパーゼで、バイオマス産業の分野で実用化されている例はまだ少ないため、土壌細菌が多数存在すると考えられる枝葉コンポストは新規エステラーゼ/リパーゼの取得源として有望である。そこで本章では、メタゲノム法を用いて枝葉コンポストから新規のエステラーゼ/リパーゼを取得することを目的とした。

今回枝葉コンポストから単離された 6 種類の新規エステラーゼ/リパーゼのうち、LC-Est1 は N 末端に長い伸長領域を有する新しいタイプのエステラーゼである。この伸長領域と相同性の高い酵素は 1 つしか存在せず、その構造や機能は明らかになっていない。そこでこの伸長領域が活性や安定性に必要なのかどうかを調べることを次の目的とし、結晶構造解析や諸特性解析を行った。

3-2.　実験材料および方法

3-2-1.　菌体、プラスミド

大腸菌、プラスミド、培地はそれぞれ 1-2-1 と同様のものを使用した。

1-2-2.　スクリーニング

12.5 μg mL^{-1} クロラムフェニコール、0.2% Tween 80、2% トリブチリンを含む LB 寒天培地を用いて、メタゲノム DNA ライブラリーを植菌し、トリブチリン分解活性を有するクローンを探索した [33]。2、3 日間 37°C で培養後、50°C でインキュベートし、トリブチリンの加水分解によるハローの形成を確認した。トリブチリン分解酵素遺伝子の ORF 配列は 1-2-2 と同様に決定した。

3-2-3.　プラスミド構築

LC-Est1 遺伝子が導入されたフォスミドベクターを鋳型として、いずれも N 末端に His-tag を導入した LC-Est1 (Gln26 – Lys510)、LC-Est1C (Glu284 – Lys510)、LC-Est1C* (Pro304 – Lys510) の発現用ベクターを構築した。プライマーの合成は北海道システムサイエンスに依頼した。DNA の塩基配列は ABI Prism 310 DNA sequencer を用いて決定した。

3-2-4.　大量発現と精製

LC-Est1、LC-Est1C、LC-Est1C*は、3-2-3 で構築した発現ベクターを用い、それぞれ大腸菌 BL21-CodonPlus(DE3)-RP の中で封入体として以下の方法で発現させた。組換え大腸菌を OD_{600} の値が 0.5 程度になるまで 37°C で振盪培養し、終濃度が 0.5 mM になるように Isopropyl β-D-thio galactopyranoside (IPTG) を加えて更に 4 時間培養した。菌体を 8,000 *g*、10 分の遠心分離により回収し、1 mM EDTA を含む 20 mM phosphate buffer (pH 7.0) に再懸濁した。ソニケーションによる菌体破砕後、30,000 *g* で 30 分遠心分離し、上清を回収した。20 mM phosphate buffer (pH 6.0) により透析した後、同バッファーにより平衡化した HiTrap SP HP カラム (GE Healthcare) に供し、NaCl 濃度を 0 M から 1 M まで直線的に上昇させることにより目的タンパク質をカラムから溶出させた。目的タンパク質を含むフラクションを回収し、10 mM イミダゾール、0.3 M NaCl を含む 20 mM phosphate buffer (pH 7.0) に透析した。同バッファーで平衡化した HiTrap Chelating

HP カラムに供し、イミダゾール濃度を 10 mM から 300 mM まで直線的に上昇させることにより目的タンパク質をカラムから溶出させた。回収したタンパク質は 10 mM Tris-HCl (pH 7.5) に透析した。

LC-Est1C*のセレノメチオニン置換体はメチオニン要求大腸菌 B834(DE3) pLysS を用いて、セレノメチオニンを含む最小培地で大量生産した [76]。精製は上記と同様に行った。

大腸菌における発現レベル、精製純度の確認は SDS-PAGE により行った [34]。タンパク質の濃度は、280 nm における Tyr と Trp の分子吸光係数 1,576、5,225 M^{-1} cm^{-1} [35] を用いて計算した。1 mg mL^{-1} のタンパク質溶液の吸光係数 ($A_{280}^{0.1\%}$) に基づき決定した。この値は LC-Est1、LC-Est1C、LC-Est1C*に対し、それぞれ 0.75、0.98、1.01 であった。

3-2-5. 酵素活性

基質としては *p*-nitrophenyl (NP) acetate (C2)、*p*-NP butyrate (C4)、*p*-NP caproate (C6)、*p*-NP caprylate (C8)、*p*-NP decanoate (C10)、*p*-NP laurate (C12)、*p*-NP myristate (C14)を使用した。活性測定は 10% アセトニトリル、1 mM 基質を含む 20 mM Tris-HCl (pH7.5) を用いて所定の温度で行った。基質から遊離した *p*-ニトロフェノールの量は U-2810 形分光光度計 (HITACHI) を用い、412 nm の吸収を測定することにより決定した。この際、*p*-ニトロフェノールの 412 nm における分子吸光係数 14,200 M^{-1} cm^{-1} を用いた。酵素活性の 1 Unit は 1 分間に 1 μmol の *p*-ニトロフェノールを生産する酵素量と定義した。比活性はタンパク質 1 μmol あたりの酵素活性と定義した。

温度依存性解析では 20 mM Tris-HCl (pH 7.5) を用い、20 – 80°C の各温度条件で測定した。pH 依存性解析は 30°C の温度で、pH 4.0 – 9.0 の各 pH 条件で測定した。この測定において、pH 4.0 – 6.0 の範囲は 20 mM sodium citrate、pH 6.0 – 8.0 の範囲では 20 mM sodium phosphate、pH 7.0 – 9.0 の範囲では 20 mM Tris-HCl を用いた。

カイネティックパラメーターの決定には、*p*-NP butyrate (C4) を基質として用いた。基質濃度は 0.25 – 6.0 mM の範囲とした。この加水分解はミカエリス・メンテン式に従っており、Lineweaver-Burk plot によりカイネティックパラメーターを決定した。

3-2-6.　CD (circular dichroism) スペクトル

1-2-7 と同様の手法で行った。

3-2-7.　熱変性測定

熱変性曲線は温度変化に伴う波長 222 nm の CD 値をモニターすることで得られた。サンプルは 10 mM Tris-HCl (pH7.0) に溶かした。測定はタンパク質濃度 0.1 mg mL^{-1}、光路長 2 mm、昇温速度 1.0°C の条件で行った。変性剤を含まない条件では、熱変性は不可逆であった。変性中点温度 ($T_{1/2}$) は最小二乗法を基にしたカーブフィッティングにより計算した。

3-2-8.　結晶化

結晶化に用いる LC-Est1*の天然体とセレノメチオニン置換体は 10 mM Tris-HCl (pH8.0) に透析し、Centricon で 19 mg mL^{-1} に濃縮した。Hampton Research 社の結晶化キット(Crystal Screen I, II) と Emerald BioStructures 社の結晶化キット(Wizard I - IV) を用いて結晶化条件のスクリーニングを行った。条件探索には 4°C、20°C の温度条件の下、96-well Corning CrystalEX Microplates (Hampton Research) を用いて、シッティングドロップ蒸気拡散法により行った。結晶化ドロップはタンパク質溶液 1 μL とリザーバー溶液 1 μL を混合して調製し、100 μl のリザーバー溶液に対して蒸気拡散平衡化させた。天然体の結晶は 20°C で、Index No.71 (0.1 M Bis-Tris (pH 6.5)、0.2 M sodium chloride、25% (w/v) polyethylene glycol 3,350) をリザーバー溶液として用いた条件で 1 週間後に得られた。セレノメチオニン置換体は結晶の質を向上させるためにさらにリザーバー溶液の最適化を行った。その結果、最適化されたリザーバー溶液 (0.1 M Bis-Tris (pH 6.5)、0.18 M sodium chloride、21% (w/v) polyethylene glycol 3,350) を用いて、4°C で 2 週間蒸気拡散平衡化させることにより、良質な結晶が得られた。

3-2-9.　X 線回折データ収集と構造決定

天然体とセレノメチオニン置換体の LC-Est1C*の X 線回折データセットの収集は、SPring-8 のビームライン BL44XU のシンクロトロン放射光を用い、-173°C の窒素ガスを吹き付けた状態で、0.9 Å の波長でデータ収集を行った。HKL2000 を用いて、得られた回折データを処理した [37]。LC-Est1C*の構造は HKL2MAP [77] 中の SHELX [78] を用いて、多波長異常分散 (MAD) 法により

決定した。そして ArpWarp [40] を用いてモデルを自動的に決定した。モデルの精密化は REFMAC [41] を用いて行い、COOT [42] によりモデルをさらに修正した。LC-CelA 構造の図は PyMol (http://www.pymol.org) を用いて作成した。

3-2-10. Protein Data Bank accession number

解析した LC-Est1C*の構造座標および構造因子は PDB コード番号 3WYD として、Protein Data Bank に登録した。

3-3.　実験結果

3-3-1.　メタゲノムから単離された新規エステラーゼ/リパーゼの同定

枝葉コンポストからメタゲノム法により新規エステラーゼ/リパーゼを単離できるかどうかを調べるために、第 1 章で構築したメタゲノムライブラリーを用いて、トリブチリンを含むプレート上で 50°C においてハローを形成するクローンをスクリーニングした。その結果得られたクローンのうち 6 つは 50°C で比較的大きなハローを形成する。これらのクローンから得られるエステラーゼ/リパーゼは、活性や安定性、分泌レベルが比較的高いと考えられる。そこで、これらのクローンが新規のエステラーゼ/リパーゼ遺伝子を持つかどうかを調べるために、フォスミドベクターをこれらのクローンから抽出し、DNA 挿入塩基配列を決定した。挿入配列の平均サイズは 35 kb である。BlastX を用いたホモロジーサーチによりオープンリーディングフレームを解析することにより、5 種類の新規エステラーゼと 1 種類の新規リパーゼ遺伝子を同定した。これらの遺伝子がコードするタンパク質を LC-Est1 – 6 と名付けた。

3-3-2.　LC-Est1 – 6 のアミノ酸配列

LC-Est1 – 6 のアミノ酸残基数と、これらのタンパク質に対して最も高いアミノ酸配列相同性を示すタンパク質をそれぞれ Table 3-1 にまとめる。LC-Est1 – 6 のアミノ酸残基数はいずれも 265 - 587 の間である。データベース上に存在する LC-Est1 – 6 と最も相同性の高いタンパク質は、それぞれ LC-Est1 – 6 と 44 – 73%の相同性を示す。LC-Est1 – 6 のアミノ酸配列は accession number KM40609 – KM406411、KM406413 – KM406415 として GenBank に登録した。

Table 3-1. List of esterases/lipases isolated from leaf-branch compost and proteins with the highest amino acid sequence identities

Cellulases	No. of residues	Protein with the highest sequence identity			
		Protein	Source organism	Accession No.	Identity (%)
LC-Est1	510	Esterase	*Candidatus Solibacter usitatus*	ABJ82142	46
LC-Est2	383	Carboxylesterase typeB	*Planctomyces brasiliensis*	ADY57734	51
LC-Est3	265	Lipase	(bacterium enrichment culture)	ADR10200	73
LC-Est4	534	Putative carboxylesterase	*Bradyrhizobium* sp. STM 3843	WP_008970463	52
LC-Est5	303	Esterase/lipase	(uncultured bacterium)	AGF91880	53
LC-Est6	587	Putative esterase	*Geobacillus thermoglucosidasius*	GAJ43165	44

LC-Est1 – 6 の一次構造を模式的に Figure 3-1 に示す。SignalIP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP-3.0/) を用いた N 末端配列分析と SMART (http://smart.embl.de) を用いたドメインサーチにより、シグナルペプチドと esterase/lipase/peptidase (Pfam domain) の触媒ドメインを予測した。LC-Est1、LC-Est2、LC-Est4 は N 末端にシグナルペプチドを有するため、分泌タンパク質であることが示唆された。その他はシグナルペプチドを持たないため、細胞質タンパク質であるか、Sec を介さない経路により分泌されることが示唆された。LC-Est1 は“esterase” domain、LC-Est2 と LC-Est4 は“COesterase”(carboxylesterase) domain、LC-Est5 と LC-Est6 は “Abhydrolase” (α/β hydrolase) domain を持つため、これらは恐らくエステラーゼである。LC-Est6 は “PepX C” (X-Pro dipeptidyl-peptidase C-terminal domain) domain も有することから、エステラーゼ活性だけではなく、ペプチダーゼ活性も持つことが示唆された。LC-Est3 は “Lipase 2” domain を持つため、リパーゼと考えられる。Arpigny と Jaeger による分類法 [79] に基づき、触媒ドメインのアミノ酸配列の違いにより LC-Est1 – 6 を分類すると、LC-Est1 は Family V、LC-Est2 は Family I-6、LC-Est3 は Family III、LC-Est4 は Family VII、LC-Est5 は Family IV、LC-Est6 は Family VI esterases/lipases に属する。

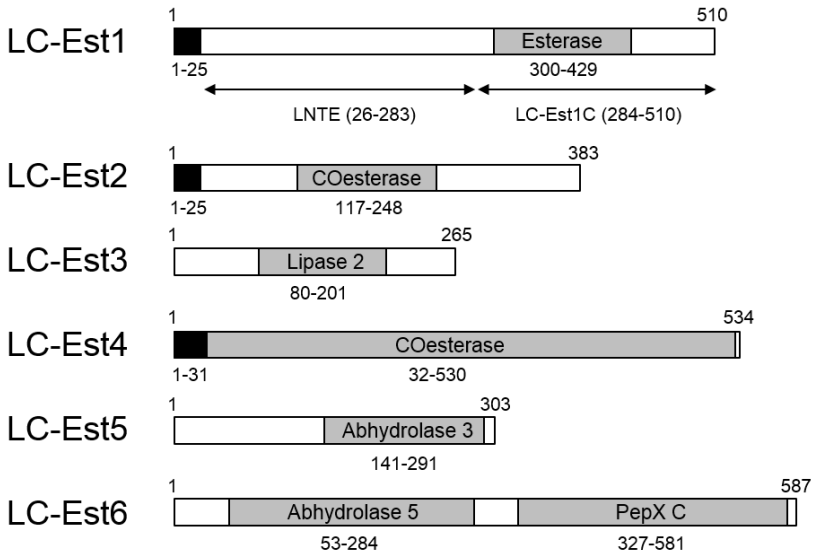


Figure 3-1. Schematic representation of the primary structures of metagenome-derived esterolytic/lipolytic enzymes from leaf-branch compost. A putative signal peptide and a putative esterase/lipase/peptidase domain are shown by black and grey boxes respectively. “COesterase”, “Abhydrolase”, and “PepX C” represent “carboxylesterase”, “α/β-hydrolase”, and “X-Pro dipeptidyl-peptidase C-terminal domain” respectively. The numbers above the sequence represent the positions of the N- and C-terminal residues of each protein. The numbers below the sequence represent the positions of the N- and C-terminal residues of each domain. The regions that are defined as long N-terminal extension (LNTE) and C-terminal esterase domain (LC-Est1C) of LC-Est1 in this study are also shown.

LC-Est1 は 510 アミノ酸から成り、分子量 55,620 kDa、等電点 (p*I*) 8.67 で、シグナルペプチドを除いた形 (Gln26 – Lys510) で分泌されると考えられる。LC-Est1 は機能未知の長い N 末端ドメイン (Long N-terminal extension、LNTE) を有する (Fig. 3-1)。ここでは、Gln26 – Phe283 を LNTE、Glu284 – Lys510 を C 末端エステラーゼドメインと定義した。SMART によると LNTE は既存の機能や構造の分かっているいかなるドメインとも相同性を示さなかった。ホモロジーサーチによると *Candidatus Solibacter usitatus* Ellin6076 由来 putative esterase (CSu-Est、ABJ82142) のみが LC-Est1 全体のアミノ酸配列と有意な類似性を示した。*C. S. usitatus* Ellin6076 は *Acidobacteria* 門の亜門 3 に属し、9.9 Mbp の巨大なゲノムを持っている [80]。*Acidobacteria* 門は環境中に豊富に存在する門の一つであり、土壌や堆積物中に見つかっており、26 の亜門に分類される。CSu-Est のみが LC-Est1 同様 LNTE を持ち、両者のアミノ酸配列の相同性は 46%である。しかし、CSu-Est の構造や機能についての研究はされておらず、LNTE の役割も分かっていない。一方、LC-Est1 の C 末端エステラーゼドメイン (LC-Est1C、Glu284-Lys510) は他のエステラーゼやリパーゼと最大 30%程度のアミノ酸配列相同性を示す。そのうち、結晶構造が決定されている *Thermotoga maritima* MSB8 由来耐熱性エステラーゼ EstA (Tm-EstA、NP_227849) [81] は、LC-Est1 の C 末端エステラーゼドメインと 29%のアミノ酸配列相同性を示す。

3-3-3. LC-Est1 と CSu-Est、Tm-EstA のアミノ酸配列比較

LC-Est1、CSu-Est、Tm-EstA のアミノ酸配列のアライメントを Figure 3-2 に示す。CSu-Est と Tm-EstA は N 末端に 15 残基のシグナルペプチドを有するため、LC-Est1 同様分泌タンパク質であることが示唆された。エステラーゼ/リパーゼの触媒トライアドを構成するアミノ酸残基 (LC-Est1 においては Ser399、Asp447、His479) は全て保存されている。エステラーゼ/リパーゼでよく保存されている触媒セリン残基を含む GXSXG モチーフも、LC-Est1 において GHSMG (Gly397 – Gly401) として保存されている。Tm-EstA の結晶構造 (PDB code: 3DOH) によると、Tm-EstA は N 末端 Ig-like ドメイン (Gln16-Asp157) と C 末端エステラーゼドメイン (Tm-EstAC、Asp158-Arg395) から成る。N 末端 Ig-like ドメインは Tm-EstA のオリゴマー化に寄与しており、また Tm-EstA の活性や安定性にも重要であると報告されている [81]。しかし、LC-Est1 の LNTE は Tm-EstA の Ig-like ドメインと相同性が低いため、LC-Est1 や CSu-Est の LNTE は Tm-EstA の Ig-like ドメインとは異なる構造を形成することが示唆される。

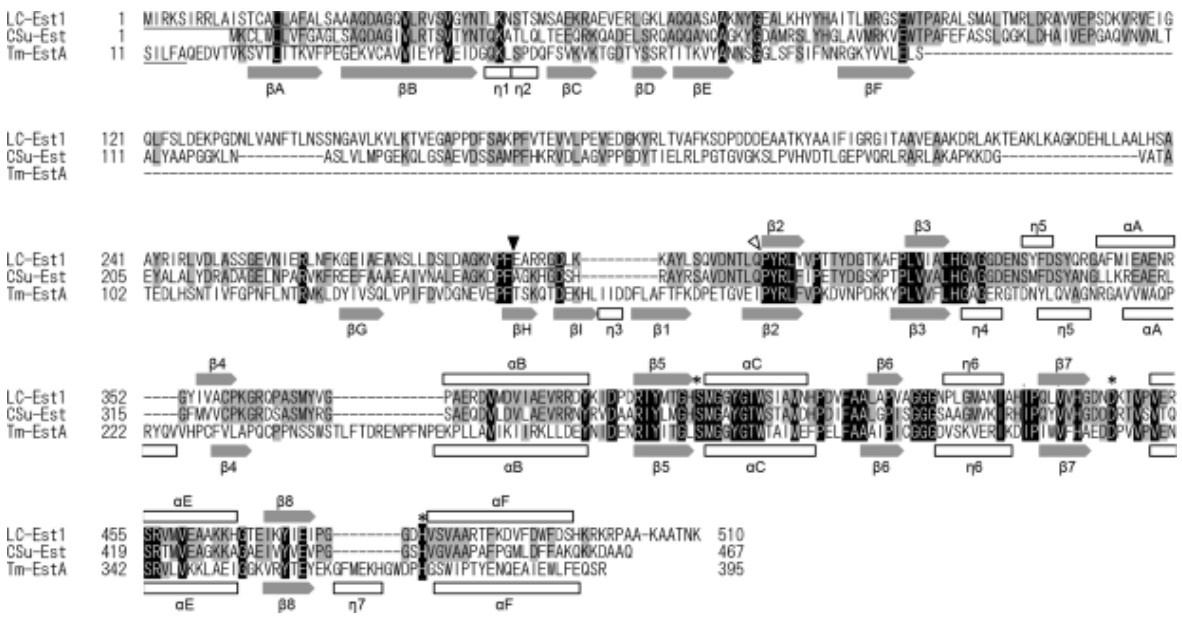


Figure 3-2. Alignment of amino acid sequences of LC-Est1, CSu-Est and Tm-EstA. The amino acid residues, which are conserved in all three proteins, are denoted with white letters and highlighted in black. The amino acid residues, which are conserved in two different proteins, are highlighted in gray. The amino acid residues that form a catalytic triad (Ser399, Asp447 and His479 for LC-Est1) are denoted by asterisks. A signal peptide of each protein predicted by the SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP-3.0/) is underlined. Solid and open triangles above the sequences represent the positions, at which N-terminal regions of LC-Est1 are truncated to construct LC-Est1C and LC-Est1C* respectively. The ranges of the secondary structures of LC-Est1C* and Tm-EstA are shown above and below the sequences respectively. “α”, “β” and “η” represent α helix, β strand, and 3_{10} helix respectively. Gaps are denoted by dashes. The numbers represent the positions of the amino acid residues starting from the N terminus of the protein. The accession numbers of these sequences are KM406409 for LC-Est1, ABJ82142 for CSu-Est, and NP_227849 for Tm-EstA.

3-3-4. LC-Est1、LC-Est1C、LC-Est1C*の大量発現と精製

LNTE がどのような構造を形成するのか、また LNTE は LC-Est1 の活性や安定性に必要なのかを調べるために、LC-Est1 (Gln26 – Lys510)、LC-Est1C (Glu284 – Lys510)、LC-Est1C* (Pro304 – Lys510) を構築し、大腸菌で大量発現した。LC-Est1 の Glu284 に対応する Tm-EstA の Thr145 は、N 末端 Ig-like ドメインのβH 鎖に位置し、これはη3 3_{10} ヘリックスを介して C 末端エステラーゼドメインに繋がっている (Fig. 3-2)。同様に、LC-Est1 の Pro304 に対応する Tm-EstA の Pro173

は、エステラーゼドメインのβ2 鎖に位置する。ゆえに、LC-Est1C はエステラーゼドメイン全体を含むのに対し、LC-Est1C*はβ1 鎖を欠いていることが予測される。

大腸菌内で発現を行うと、いずれも可溶型で目的タンパク質が得られた。それぞれ 3-2-3 の手順に従って SDS-PAGE で単一バンドが得られるまで精製を行った。1 L の培養液から LC-Est1 は 6 mg、LC-Est1C は 2 mg、LC-Est1C*は 3 mg の精製タンパク質がそれぞれ得られた。ゲルろ過クロマトグラフィーによると、LC-Est1、LC-Est1C、LC-Est1C*の分子量はそれぞれ 58 kDa、28 kDa、26 kDa と推定され、これはアミノ酸配列から計算した値 (それぞれ 55,620 Da、27,382 Da、25,094 Da) とほぼ一致することから、いずれも溶液中においてモノマーで存在することが示唆された。

LC-Est1 と LC-Est1C の Far-UV CD スペクトルを Figure 3-3 に示す。いずれも 220 nm 付近で[θ]値が最も小さな曲線を描く。LC-Est1C の Far-UV CD スペクトルは LC-Est1 よりも 220 nm 付近の谷の深さが浅かったので、LC-Est1C のヘリックス含量は LC-Est1 よりも低いことが示唆された。なお LC-Est1C*のスペクトルは LC-Est1C と類似していた。

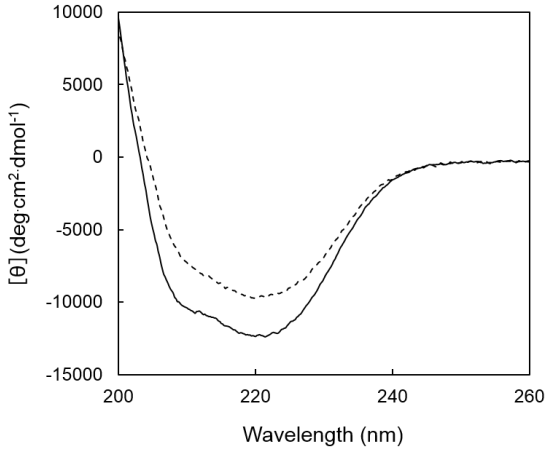


Figure 3-3. Far-UV CD spectra. The far-UV CD spectra of LC-Est1 (solid line) and LC-Est1C (broken line) were measured at 25°C in 10 mM Tris-HCl (pH 7.0) as described in Materials and Methods.

3-3-5. LC-Est1C*の結晶構造

LNTE の欠損が LC-Est1 の C 末端エステラーゼドメインの構造に影響するかどうかを調べるには、LC-Est1 と LC-Est1C の結晶構造を決定する必要があった。結晶化条件検討の末、X 線回折実験に適する結晶は LC-Est1C*のみ得られた。LC-Est1 と LC-Est1C の結晶化も試みたが、構造解析に適した良質の結晶を得ることはできなかった。これは恐らく LNTE や LC-Est1C の N 末端領域がフレキシブルで、良質な結晶の形成を妨げているからと考えられる。

LC-Est1C*の結晶構造を MAD 法により 1.53 Å の分解能で決定した。データ収集および精密化の統計値を Table 3-2 に示す。非対称ユニット中に 2 分子 (A、B) 存在した。分子 A の Arg361 – Met366 と Ala503 – Lys510、分子 B の Arg361 – Met366 と Arg501 – Lys510 は不規則な揺らぎのために、その電子密度を観測することはできなかった。分子 A と B の構造は極めて類似しており、その RMSD 値は 0.24 Å であった。本研究では分子 A の構造を使用した。

LC-Est1C*の結晶構造によると、LC-Est1C*は典型的な α/β hydrolase fold を有していた (Fig. 3-4)。LC-Est1C*の構造は Tm-EstAC (PDB code: 3DOH) の構造と良く類似しており、RMSD 値は 1.25 Å であった。この結果、LC-Est1 のエステラーゼドメインの構造は LNTE の欠損によりそれほど変化しないことが示唆された。LC-Est1C*の触媒トライアドを形成する活性部位残基 (Ser399、Asp447、His479) の立体構造は Tm-EstA の触媒残基 (Ser286、Asp334、His374) とほぼ一致した。LC-Est1C*の構造は、β4 鎖とαB ヘリックスの間のループ、β8 鎖とαF ヘリックスの間のループを除いて、Tm-EstA のエステラーゼドメインと類似していた。LC-Est1C*のβ4 鎖とαB ヘリックスの間のループ (Lys359 – Gly369) は短く、全体の電子密度ははっきりとしていない。一方 Tm-EstA の対応するループ (Pro234 – Pro255) は長く、触媒部位のトンネルの一部を形成し、酵素活性に影響する Phe246 を含んでいる [81]。LC-Est1C*のβ8 鎖とαF ヘリックスの間のループ (Pro475 – His479) は短いのに対して Tm-EstA の対応するループ (Glu362 – His374) は比較的長く、η7 ヘリックスを含んでいる。

Table 3-2. Data collection and refinement statistics of LC-Est1C*

	Native	SeMet-peak	SeMet-inflection	SeMet-remote
Data collection statistics				
Wavelength (Å)	0.900	0.978769	0.979101	0.994813
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell parameters				
a, b, c (Å)	45.86, 55.45, 156.27	46.13, 55.73, 156.24	46.13, 55.78, 156.32	46.13, 55.79, 156.35
α=β=γ (°)	90	90	90	90
Molecules/asymmetric unit	2	2	2	2
Resolution range (Å)	50.00-1.53 (1.56-1.53)[a]	50.00-2.30 (2.34-2.30)[a]	50.00-2.30 (2.34-2.30)[a]	50.00-2.30 (2.34-2.30)[a]
Reflections measured	392,694	123,978	122,975	121,933
Unique reflections	60,179	17,756	17,643	17,581
Redundancy	6.5 (7.2)[a]	7.0 (7.3)[a]	7.0 (7.3)[a]	6.9 (7.3)[a]
Completeness (%)	98.5 (100.0)[a]	94.6 (99.7)[a]	93.9 (99.5)[a]	93.6 (99.5)[a]
R_{merge} (%)[b]	18.1 (28.3)[a]	11.5 (33.4)[a]	9.5 (36.5)[a]	10.3 (42.2)[a]
Average I/σ (I)	14.5 (3.6)[a]	23.3 (5.0)[a]	23.5 (4.5)[a]	22.2 (3.8)[a]
Refinement statistics				
Resolution range (Å)	78.14-1.53			
No. of atoms				
Protein/Water	2992/451			
R_{work} (%)	19.1			
R_{free} (%)[c]	22.5			
Rms deviations from ideal values				
Bond lengths (Å)	0.014			
Bond angles (°)	1.439			
Average B factors (*Å*^{2})				
Protein/Water	18.3/31.6			
Ramachandran plot statistics				
Favored region (%)	97.1			
Allowed region (%)	2.9			

[a] Values in parentheses are for the highest-resolution shell.
[b] $R_{merge} = \sum|I_{hkl} - \langle I_{hkl}\rangle|/\sum I_{hkl}$, where I_{hkl} is an intensity measurement for reflection with indices *hkl* and $\langle I_{hkl}\rangle$ is the mean intensity for multiply recorded reflections.
[c] Free *R*-value was calculated using 5% of the total reflections chosen randomly and omitted from refinement.

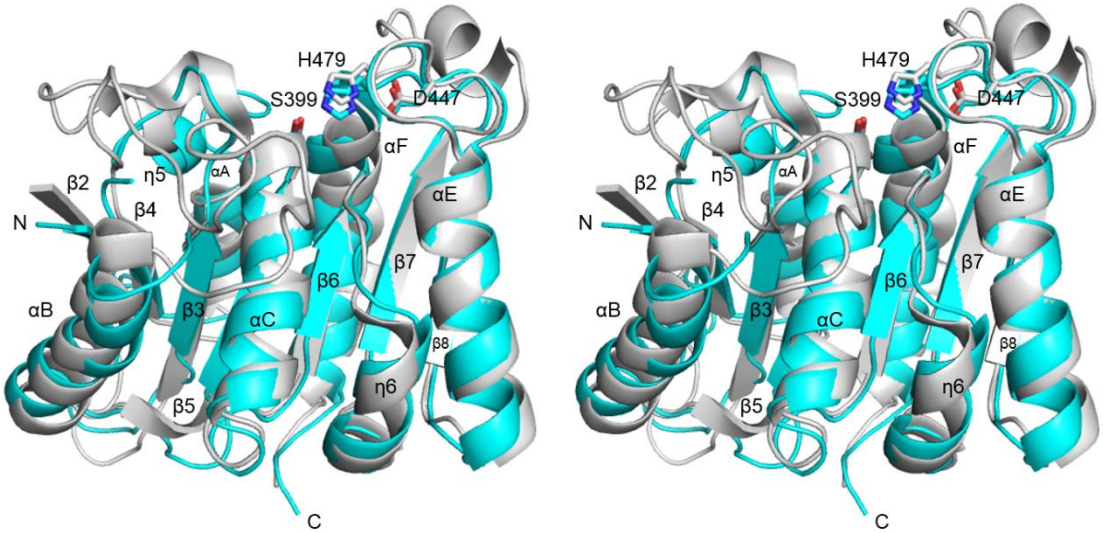


Figure 3-4. Crystal structure of LC-Est1C*. The stereo view of the structure of LC-Est1C* (cyan) is superimposed on that of the C-terminal esterase domain of Tm-EstA (PDB entry 3DOH). Three active-site residues of LC-Est1C* (Ser399, Asp447 and His479) and the corresponding residues of Tm-EstA (Ser286, Asp334 and His374) are shown by stick models. N and C represent N- and C-termini.

3-3-6. LC-Est1 と LC-Est1C の酵素活性

LC-Est1C*はβ1 鎖を含まないため、今回 LC-Est1C を諸特性解析に使用した。まず、*p*NP-acetate (C2) から *p*NP-myristate (C14) まで鎖長の異なる基質を用いて、pH 7.5、30°C の条件下で LC-Est1 と LC-Est1C の酵素活性を測定した。その結果を Figure 3-5 に示す。LC-Est1 は C4 基質を最も良く分解し、C6 基質でも同程度の活性を示した。しかし C2 あるいは C8 基質を用いたときは、最大活性の 15–25%程度の活性しか示さず、C10 以上の基質はほとんど分解しなかった。LC-Est1 同様、LC-Est1C も C6 基質を最も良く分解し、C4 基質でも同程度の活性を示した。また、いずれもオリーブオイルは分解しなかった。ゆえに LNTE の欠損は、LC-Est1 の基質特異性にはほとんど影響しない。しかし、C4 基質を用いて 30°C、pH 7.5 で測定した LC-Est1 と LC-Est1C の比活性はそれぞれ 255 units/μmol、101 units/μmol であったので、LC-Est1 の活性は LNTE の欠損により約 60%低下した (Table 3-3)。従って、LNTE は LC-Est1 の活性を高めるのに必要である。

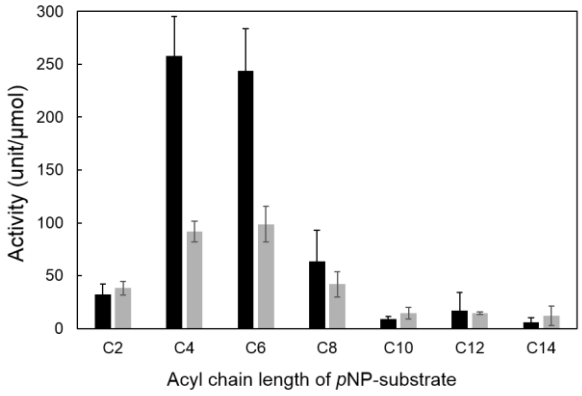


Figure 3-5. Substrate selectivity of LC-Est1 and LC-Est1C. Specific activities of LC-Est1 (solid bar) and LC-Est1C (grey bar) toward triglyceride substrates are shown. C2~C14 represent the acyl chain lengths of the substrates. The experiment was carried out in duplicate. Each value represents the average value and errors from the average values are shown.

Table 3-3. Specific activities, kinetic parameters, and $T_{1/2}$ values of LC-Est1 and LC-Est1C [a]

Protein	Specific activity (unit/μmol)	Relative activity (%)	K_m (mM)	k_{cat} (s^{-1})	$T_{1/2}$ (°C)	$\Delta T_{1/2}$[b] (°C)
LC-Est1	255 ± 31	100	0.30 ± 0.01	5.8 ± 0.3	67.2 ± 0.4	-
LC-Est1C	101 ± 19	40	0.48 ± 0.04	2.4 ± 0.1	63.9 ± 0.1	-3.3

[a] The enzymatic activity was determined at 30°C for 10 min, in 100 μl of 20 mM Tris-HCl (pH 7.5) containing 10% acetonitrile, by using *p*-nitrophenyl butyrate (C4) as a substrate. The substrate concentration was 1 mM. For determination of the kinetic parameters, the substrate concentration was varied from 0.25 to 6.0 mM. The relative activity was calculated by dividing the specific activity of LC-Est1C by that of LC-Est1. Experiments were carried out at least twice and the average values are shown together with the errors. The temperature of the midpoint of the thermal denaturation transition, $T_{1/2}$, was determined by monitoring the change in the CD value at 222 nm.

[b] $\Delta T_{1/2} = T_{1/2}$(LC-Est1C) – $T_{1/2}$(LC-Est1).

LC-Est1 と LC-Est1C の酵素活性の温度依存性を、C4 基質を用いて pH 7.5、20 - 80°C の範囲で測定した (Fig. 3-6)。いずれも 40°C で最大活性を示した。同様に pH 依存性を、C4 基質を用いて 30°C、 pH 4.0 – 9.0 の範囲で測定した (Fig. 3-7)。いずれも pH 7.5 で最大活性を示した。ただし Tris-HCl を用いて pH 7.0 – 9.0 の範囲で LC-Est1C の酵素活性を測定した場合は、pH 8.0 の方が pH 7.5 に比べて僅かに活性が高かった。これらの結果より、LNTE の欠損は LC-Est1 の至適温度や至適 pH にはほとんど影響しないことが示唆された。なお、LC-Est1C*の比活性や温度依存性は LC-Est1C とほぼ同程度であり、LC-Est1C の N 末端側 20 残基の欠損が、LC-Est1C の安定性に僅かに影響している可能性は排除できない

ものの、活性や安定性にはほとんど影響しないと考えられた。

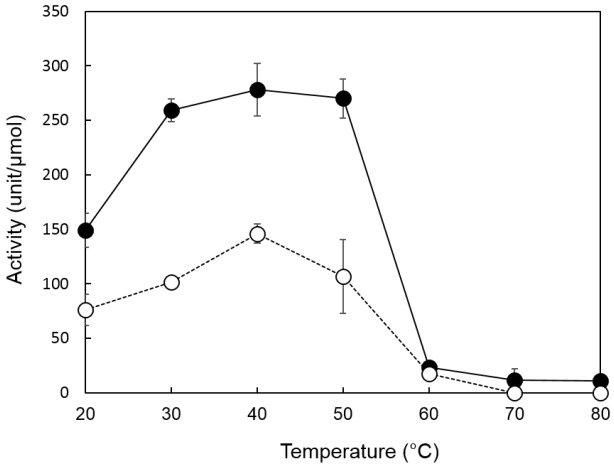


Figure 3-6. Optimum temperatures for activities of LC-Est1 and LC-Est1C. The temperature dependencies of the enzymatic activities of LC-Est1 (solid circle, solid line) and LC-Est1C (open circle, broken line) are shown. The activity was determined at pH 7.5 and the temperatures indicated using *p*NP-butyrate (C4) as a substrate, as described in Materials and Methods. The experiment was carried out at least twice, and errors from the average values are indicated by *vertical lines*.

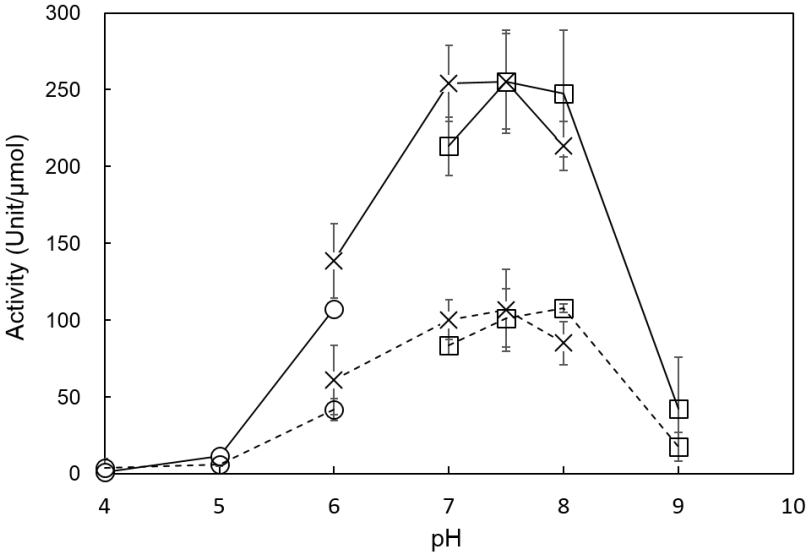


Figure 3-7. Optimum pH for activities of LC-Est1 and LC-Est1C. The pH dependencies of the enzymatic activities of LC-Est1 (solid line) and LC-Est1C (broken line) are shown. The activity was determined at 30°C and the pH values indicated using *p*NP-butyrate (C4) as a substrate, as described in Materials and Methods. The buffers used to analyze the pH dependence of activity were 20 mM sodium citrate (pH 4.0-6.0) (open circle), 20 mM sodium phosphate (pH 6.0-8.0) (cross) and 20 mM Tris-HCl (pH 7.0-9.0) (solid square). The experiment was carried out at least twice, and errors from the average values are indicated by *vertical lines*.

LC-Est1C の酵素活性が LC-Est1 よりも低い要因が、基質結合親和性の減少によるものか、ターンオーバー数の減少によるものかを調べるために、C4 基質を用いて、30°C、 pH 7.5 の条件下で、LC-Est1 と LC-Est1C のキネティックパラメーターを決定した。いずれも Michaelis-Menten kinetics に従い、キネティックパラメーターを Lineweaver-Burk plot により決定した (Table 3-3)。LC-Est1C の K_m 値は LC-Est1 よりも僅かに高く、k_{cat} 値は LC-Est1 よりも 60%程度低かった。これらの結果より、LNTE の欠損は基質結合親和性よりも、LC-Est1 のターンオーバー数を減少させることが示唆された。おそらく、LNTE の欠損により、活性部位のコンフォメーションが僅かに変化すると考えられる。

3-3-7. LC-Est1 と LC-Est1C の耐熱性

LNTE の欠損による LC-Est1 の耐熱性への影響を調べるために、pH 7.5 で LC-Est1 と LC-Est1C の熱変性曲線を測定した。この条件ではこれらの熱変性は不可逆的に起こる。10 mM Tris-HCl (pH 7.5) の条件で測定した LC-Est1 と LC-Est1C の熱変性曲線を Figure 3-8 に示し、熱変性曲線の中点温度 ($T_{1/2}$) を Table 3-3 にまとめる。LC-Est1C の $T_{1/2}$ 値は LC-Est1 よりも 3.3°C 低かったので、LNTE は LC-Est1 の安定化にやや寄与することが示唆された。

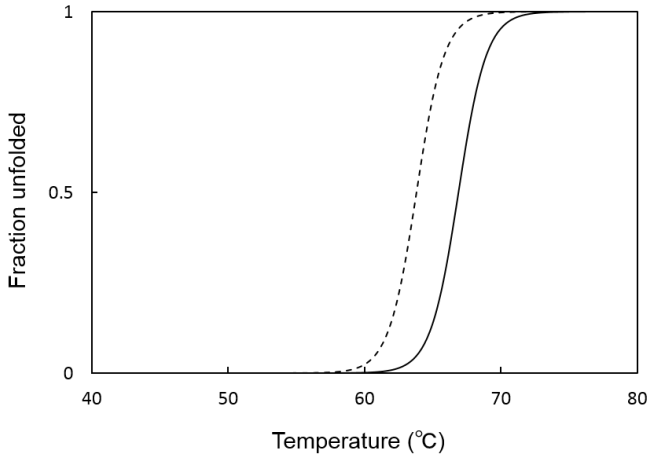


Figure 8. Thermal denaturation LC-Est1 and LC-Est1C. The thermal denaturation curves of LC-Est1 (solid line) and LC-Est1C (broken line) measured in 10 mM Tris-HCl (pH 7.0) are shown. The curves were recorded by monitoring the change in CD values at 222 nm as described in Materials and Methods.

3-3-8. LC-Est1C*と Tm-EstAC の構造的特徴の比較

Tm-EstA は 95°C 以上で最大活性を示すのに対して、Tm-EstAC は 60°C で最大活性を示し、Ig-like ドメインの欠損により耐熱性が大幅に低下している [81]。それでも、40°C で最大活性を示す LC-Est1C*に比べればまだ大幅に耐熱性は高い。一般的にイオン対の数 [82]、水素結合の数 [83]、ループ領域中のプロリン残基の数 [84]、ジスルフィド結合の数 [85] が多いほど、また内部の疎水性度が高い [86] ほど、また第 1 章に示した N 末端の固定化や C 末端の固定化 [87] により、タンパク質の安定性は高くなる傾向にある。そこで LC-Est1C*と Tm-EstAC を比較すると、イオン対の数、水素結合の数、ループ中のプロリン残基の数は Tm-EstAC の方が LC-Est1C*よりも多い (Table 3-4)。LC-Est1C*と Tm-EstAC のイオン対の数はそれぞれ 11、14、水素結合の数はそれぞれ 169、206、ループ中のプロリン残基の数はそれぞれ 10、12 であった。しかしいずれの酵素もジスルフィド結合は存在していない。反対に、内部の疎水性度は LC-Est1C*が 71.3%であるのに対し、Tm-EstAC は 65.4%であった。これらの結果より、イオン対や水素結合の数、ループ中のプロリン残基の数の違いが、LC-Est1C*と Tm-EstAC の安定性の違いをもたらすと考えられた。

Table 3-4. Structural features of LC-Est1C* and Tm-EstAC

	LC-Est1C*	Tm-EstAC
No. of amino acids	199	222
Opt. Temp. for activity (°C)[a]	40	60
Content of the residues (%)		
Buried polar	15.6	20.2
Buried apolar	38.7	38.3
Buried apolar/buried total	71.3	65.4
No. of ion pairs (≤4.0 Å)	11	14
No. of hydrogen bonds	169	206
No. of disulfide bond	0	0
No. of Pro in loop	10	12

[a] Data from Ref. 81 for Tm-EstAC.

3-4. 考察

枝葉コンポストからメタゲノム法により単離された、N 末端に機能未知のドメイン (LNTE) を持つ新規エステラーゼ LC-Est1 について解析を行った。LC-Est1 と LC-Est1C の Far-UV CD スペクトルや LC-Est1C*の結晶構造より、LNTE は触媒ドメインのコンフォメーションには影響を及ぼさないことが示唆された。LC-Est1 に比べて LC-Est1C の活性は約 60%低く、k_{cat} 値も約 60%低いことから、LNTE が加水分解効率の向上に寄与していることが示唆された。また $T_{1/2}$ 値は LC-Est1 よりも LC-Est1C の方が 3.3°C 低いため、LNTE が僅かに耐熱化にも寄与していると考えられた。

LC-Est1 の LNTE と相同性が高く、その機能が知られているタンパク質はこれまでに見つかっていない。今回の研究で、活性や安定性における役割は検討されたものの、LC-Est1 の全体構造は解かれていないため、LC-Est1 における LNTE の役割は未知な点が多い。Tm-EstA は Ig-like ドメインの欠損によりオリゴマー形成能を失い、その活性や安定性も大きく低下する [81]。Tm-EstA は溶液中において 6 量体で存在し、Tm-EstA の至適温度が 95°C 以上、100°C での半減期は 1.5 時間であることから、極めて熱安定な酵素である。一方で Tm-EstAC は単量体として存在し、Tm-EstA の 5%の活性しか示さず、その至適温度は 60°C であり、90°C での半減期が 15 分である。これらの結果より、Tm-EstA の Ig-like ドメインは多量体化を誘導して活性や安定性を高める役割を果たしていると考えられている。しかし、LC-Est1 の LNTE は Tm-EstA の Ig-like ドメインとは異なる役割を持つと考えられる。なぜなら LC-Est1 と LC-Est1C はいずれも溶液中で単量体として存在し、LC-Est1 における LNTE の欠損は、Tm-EstA における Ig-like ドメインの欠損ほど活性や安定性に影響を与えないためである。LC-Est1 の LNTE は恐らく独立したドメインとして存在し、エステラーゼドメインとは弱い相互作用しかしていないと考えられる。これらの相互作用を取り除くことにより、LC-Est1 を僅かに不安定化し、同時に活性部位のコンフォメーションを僅かに変えることにより、活性を低下させると予想される。本酵素が活性を示す基質の発見、またそれに基づく活性や基質認識、安定性における LNTE の役割を調べるためには、全体の結晶構造の解明が待たれる。

【総括】

耐熱性の高い新規セルラーゼと新規エステラーゼ/リパーゼの取得を目指して、枝葉コンポストからメタゲノム法により 10 種類の新規セルラーゼ遺伝子と 6 種類の新規エステラーゼ/リパーゼ遺伝子を取得した。その中でも耐熱性に優れ、産業利用への応用が期待できる 2 種類の新規セルラーゼと、N 末端に長い伸長領域を有する新規エステラーゼを研究の対象とした。セルラーゼやエステラーゼの基質認識や触媒機能に必須な触媒ドメインの役割に関してはこれまでの研究でかなり明らかにされている。一方、付随するドメインやリンカー領域の中にはその役割がまだ明らかにされていないものも多い。そこで本研究ではメタゲノム法で単離した酵素が有する付随領域の役割を明らかにすることを目的として構造機能解析を試みた。

第 1 章では、至適温度が 95°C 以上である *Rm*Cel12A と高いアミノ酸配列の相同性を示すセルラーゼ LC-CelA を研究の対象とした。*Rm*Cel12A の N 末端 FL が安定性に寄与すると考えられたことから、FL を含んだ LC-CelA の結晶構造を初めて解析した。その結果、FL の C 末端 2 残基 (Glu34、Pro35) が中心領域 (Leu72 – Thr74) と 3 つの水素結合を介して相互作用していることが明らかになった。この水素結合が安定性に寄与するかどうかを調べるために、FL 欠損変異体、E34A 変異体を構築し、その活性や安定性を解析した。すると、これら変異体の活性は野生型とほぼ同程度有していたが、耐熱性が大幅に低下しており、高い安定性に必要であることを明らかにした。また、野生型酵素の安定性が DTT 存在下では DTT 非存在下と比べて大きく低下することから、N 末端付近の分子内ジスルフィド結合も耐熱化に大きく寄与することを明らかにした。今回発見した N 末端残基の水素結合は LC-CelA と *Rm*Cel12A 以外では保存されていない。従って、N 末端領域に水素結合を導入する方法は、他の GHF12 セルラーゼの安定化法として有効であることが期待される。実際、GHF12 セルラーゼと構造が類似している GHF11 キシラナーゼにおいて、N 末端のユニークな伸長配列や N 末端付近ジスルフィド結合が安定化に寄与する例がいくつか報告されている。

第 2 章では、N 末端に Ig-like ドメインを有するセルラーゼ LC-CelG の結晶構造を決定し、諸特性を解析した。すでに結晶構造の決定されている 3 種類の相同タンパク質と比較することにより、全体構造および触媒残基や基質結合ポ

ケットを形成する残基の構造には大きな違いがないことを明らかにした。一方、金属イオンの結合部位には違いが見られた。これらの部位は、触媒機構には必須ではないが、構造の安定化には寄与することが示唆された。Ig-like ドメインの役割を調べるために、His-ΔIg-CelG と、ドメイン間の相互作用に関わる Gln40、Asp99 をそれぞれ Ala に置換した変異体、および両方を Ala に置換した変異体を構築し、諸特性を解析した。His-ΔIg-CelG の二次構造は His-LC-CelG と類似していることが CD スペクトルから示唆された。しかし、His-ΔIg-CelG はほとんどすべて不溶性タンパク質として菌体内に蓄積し、可溶性タンパク質として蓄積した His-ΔIg-CelG の活性は His-LC-CelG の 1/100 に減少した。また、その $T_{1/2}$ 値は His-LC-CelG より 6.3°C 低下した。一方、His-Q40A/D99A-CelG の $T_{1/2}$ 値は His-LC-CelG より 5.0°C 低下したものの、大半が可溶性タンパク質として得られ、活性も His-LC-CelG と同程度であった。LC-CelG の結晶構造によると、Ig-like ドメインは活性部位や基質結合ポケットを形成する残基を有する長いループ領域と相互作用している。従って、これらの結果より、Ig-like ドメインは分子内シャペロン機能を有しており上述の長いループ領域のフォールディングに必要であることが示唆された。ただし Gln40、Asp99 の変異のみでは、この長いループ領域のコンフォメーションを変化させるのに十分ではないと考えられる。GHF9 酵素の中には LC-CelG のように N 末端に Ig-like ドメインを有するものの他に、C 末端に Family 3 carbohydrate-binding module を有するもの、触媒ドメインのみで機能するものが存在する。ゆえにこれらのドメインの有無による活性や基質結合、安定化機構の違いを分子レベルで解き明かすことは大変興味深い。

第 3 章では、N 末端に長い伸長領域 (LNTE) を持つ新規エステラーゼ LC-Est1 の構造と機能の解析を行った。LC-Est1 と LC-Est1C の Far-UV CD スペクトルや LC-Est1C*の結晶構造より、LNTE は触媒ドメインのフォールディングには必要のないことが示唆された。LC-Est1 に比べて LC-Est1C の活性は約 60%低く、k_{cat} 値も約 60%低いことから、LNTE が加水分解効率の向上に寄与することが示唆された。また、LC-Est1C の $T_{1/2}$ 値は LC-Est1 より 3.3°C 低いので、LNTE は僅かに安定化にも寄与していると考えられた。LC-Est1C と 29%のアミノ酸配列相同性を示し、結晶構造が決定されている Tm-EstA は N 末端に Ig-like ドメインを有するが、この Ig-like ドメインはオリゴマー化に必要で、Ig-like ドメインを除去すると酵素活性や安定性は大きく低下することから、LNTE はこの Ig-like ドメインとは異なる役割を持つことが示唆された。LNTE の役割を明らかにするた

めに、LNTE を含む LC-Est1 の結晶構造の解明が待たれる。

本研究を通して、メタゲノム法により身近な枝葉コンポストから、新規のセルラーゼとエステラーゼを単離し、結晶構造解析、諸特性解析を行うことに成功した。このうち、LC-CelA と LC-CelG は高度な熱安定性を有することから産業上有用であると期待される。一方、LC-Est1 は機能未知の N 末端伸長領域を有することから、更に構造や機能解析を進めることで、新規エステラーゼ/リパーゼとしての産業利用も期待できる。本研究では、触媒部位や基質結合部位から離れた末端のリンカーやドメインの役割に焦点を当て、それらの活性や安定性における役割について調べた。その結果、酵素機能上重要な領域からは大きく離れた場所に位置するこれらのリンカーやドメインが活性や安定性に大きな影響を及ぼすことを明らかにした。一方、LC-CelG の Ig-like ドメインや LC-Est1 の LNTE などの役割を分子レベルで解明できれば、これらのドメインを利用して酵素の機能を向上させる新たな方法を見つけ出すことができると期待される。

参考文献

1. Steele, H.L., Jaeger, K.E., Daniel, R. and Streit, W.R. (2009) Advances in recovery of novel biocatalysts from metagenomes. *J. Mol. Microbiol. Biotechnol*. **16**, 25-37.
2. Uchiyama, T. and Miyazaki, K. (2009) Functional metagenomics for enzyme discovery: challenges to efficient screening. *Curr. Opin. Biotechnol*. **20**, 616-622.
3. Tuffin, M., Anderson, D., Heath, C. and Cowan, D.A. (2009) Metagenomic gene discovery: how far have we moved into novel sequence space? *Biotechnol. J.* **4**, 1671-1683.
4. Duan, C.J. and Feng, J.X. (2010) Mining metagenomes for novel cellulase genes. *Biotechnol. Lett.* **32**, 1765–1775.
5. Simon, C. and Daniel, R. (2011) Metagenomic analyses: past and future trends. *Appl. Environ. Microbiol.* **77**, 1153-1161.
6. Iqbal, H.A., Feng, Z. and Brady, S.F. (2012) Biocatalysts and small molecule products from metagenomic studies. *Curr. Opin. Chem. Biol*. **16**, 109-116.
7. Kanaya, E., Sakabe, T., Nguyen, N.T., Koikeda, S., Koga, Y., Takano, K. and Kanaya, S. (2010) Cloning of the RNase H genes from a metagenomic DNA library: identification of a new type 1 RNase H without a typical active-site motif. *J. Appl. Microbiol.* **109**, 974–983.
8. Sulaiman, S., Yamato, S., Kanaya, E., Kim, J.J., Koga, Y., Takano, K. and Kanaya, S. (2011) Isolation of a novel cutinase homolog with polyethylene terephthalate degrading activity from leaf-branch compost using a metagenomic approach. *Appl. Environ. Microbiol.* **78**, 1556-1562.
9. Nguyen, TN., You, D.-J., Kanaya, E., Koga, Y. and Kanaya, S. (2013) Crystal structure of metagenome-derived LC9-RNase H1 with atypical DEDN active site motif. *FEBS Lett.* **587**, 1418–1423.
10. Nguyen, T.N., Angkawidjaja, C., Kanaya, E., Koga, Y., Takano, K. and Kanaya, S. (2012) Activity, stability, and structure of metagenome-derived LC11-RNase H1, a homolog of *Sulfolobus tokodaii* RNase H1. *Protein Sci.* **21**, 553-561.
11. Nguyen, T.N., You, D.-J., Matsumoto, H., Kanaya, E., Koga, Y. and Kanaya, S. (2013) Crystal structure of metagenome-derived LC11-RNase H1 in complex with RNA/DNA hybrid. *J. Struct. Biol.* **182**, 144–154.

12. Sulaiman, S., You, D.-J., Kanaya, E., Koga, Y. and Kanaya, S. (2014) Crystal structure and thermodynamic and kinetic stability of metagenome-derived LC-cutinase. *Biochemistry* **53**, 1858–1869.
13. Wilson, D.B. (2009) Cellulases and biofuels. *Curr. Opin. Biotechnol.* **20**, 295-299.
14. Lynd, L.R., van Zyl, W.H., McBride, J.E. and Laser, M. (2005) Consolidated bioprocessing of cellulosic biomass: an update. *Curr. Opin. Biotechnol.* **16**, 577-583.
15. la Grange, D.C., den Haan, R. and van Zyl, W.H. (2010) Engineering cellulolytic ability into bioprocessing organisms. *Appl. Microbiol. Biotechnol.* **87**, 1195-1208.
16. Yanase, S., Yamada, R., Kaneko, S., Noda, H., Hasunuma, T., Tanaka, T., Ogino, C., Fukuda, H. and Kondo, A. (2010) Ethanol production from cellulosic materials using cellulase-expressing yeast. *Biotechnol. J.* **5**, 449–455.
17. Yamada, R., Nakatani, Y., Ogino, C. and Kondo, A. (2013) Efficient direct ethanol production from cellulose by cellulase- and cellodextrin transporter-coexpressing *Saccharomyces cerevisiae*. *AMB Express* **3**, 34.
18. Hu, J., Arantes, V., Pribowo, A., Gourlay, K. and Saddler, J.N. (2014) Substrate factors that influence the synergistic interaction of AA9 and cellulases during the enzymatic hydrolysis of biomass. *Energy Environ. Sci.* **7**, 2308–2315.
19. Lynd, L.R., Weimer, P.J., van Zyl, W.H. and Pretorius, I.S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. *Mol. Biol. Rev.* **66**, 506–577.
20. López-López, O., Cerdán, M.E. and Gonzalez-Siso, M.I. (2014) New extremophilic lipases and esterases from metagenomics. *Curr. Protein Pept. Sci.* **15**, 445-455.
21. Akoh, C.C., Chang, S.W., Lee, G.C. and Shaw, J.F. (2007) Enzymatic approach to biodiesel production. *J. Agric. Food Chem.* **55**, 8995-9005.
22. Holmquist, M. (2000) Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr. Protein Pept. Sci.* **1**, 209-235.
23. Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G. and Goldman, A. (1992) The α/β hydrolase fold. *Protein Eng.* **5**, 197-211.
24. Verger, R. (1997) 'Interfacial activation' of lipases, facts and artifacts. *Tibtech.* **15,** 32–38.
25. Bornscheuer, U.T. (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. *FEMS Microbiol. Rev.* **26**, 73-81.

26. Jochens, H., Hesseler, M., Stiba, K., Padhi, S.K., Kazlauskas, R.J. and Bornscheuer, U.T. (2011) Protein engineering of α/β-hydrolase fold enzymes. *Chembiochem.* **12**, 1508-1517.
27. Bruins, M.E., Janssen, A.E. and Boom, R.M. (2001) Thermozymes and their applications: a review of recent literature and patents. *Appl. Biochem. Biotechnol.* **90**, 155–186.
28. Haki, G.D. and Rakshit, S.K. (2003) Developments in industrially important thermostable enzymes: a review. *Bioresour. Technol.* **89**, 17–34.
29. Unsworth, L.D., van der Oost, J. and Koutsopoulos, S. (2007) Hyperthermophilic enzymes – stability, activity and implementation strategies for high temperature applications. *FEBS J.* **274**, 4044–4056.
30. Halldorsdottir, S., Thorolfsdottir, E.T., Spilliaert, R., Johansson, M., Thorbijarnardottir, S.H., Palsdottir, A., Hreggvidsson, G.O., Kristjansson, J.K., Holst, O. and Eggertsson, G. (1998) Cloning, sequencing and overexpression of a *Rhodothermus marinus* gene encoding a thermostable cellulase of glycosyl hydrolase Family 12. *Appl. Microbiol. Biotechnol.* **49**, 277–284.
31. Crennell, S.J., Hreggvidsson, G.O. and Nordberg Karlsson, E. (2002) The structure of *Rhodothermus marinus* Cel12A, a highly thermostable family 12 endoglucanase, at 1.8 Å resolution. *J. Mol. Biol.* **320**, 883–897.
32. Crennell, S.J., Cook, D., Minns, A., Svergun, D., Andersen, R.L. and Nordberg Karlsson, E. (2006) Dimerisation and an increase in active site aromatic groups as adaptations to high temperatures: X-ray solution scattering and substrate-bound crystal structures of *Rhodothermus marinus* endoglucanase Cel12A. *J. Mol. Biol.* **356**, 57–71.
33. Jo, W.S., Bae, S.H., Choi, S.Y., Park, S.D., Yoo, Y.B. and Park, S.C. (2010) Development of detection methods for cellulolytic activity of *Auricularia auricula-judae*. *Mycobiology* **38**, 74–77.
34. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
35. Goodwin, T.W. and Morton, R.A. (1946) The spectrophotometric determination of tyrosine and tryptophan in proteins. *Biochem. J.* **40**, 628-632.
36. Miller, G.L., Blum, R., Glennon, W.E. and Burton, A.L. (1960) Measurement of

carboxymethylcellulase activity. *Anal. Biochem.* **2**, 127–132.

37. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
38. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763.
39. Vagin, A. and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25.
40. Langer, G., Cohen, S.X., Lamzin, V.S. and Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171–1179.
41. Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255.
42. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132.
43. Gilkes, N.R., Henrlssat, B., Kilburn, D.G., Miller Jr., R.C. and Warren, R.A.J. (1991) Domains in microbial β-l,4-glycanases: sequence conservation, function, and enzyme families. *Microbiol. Rev.* **55**, 303–315.
44. Wicher, K.B., Abou-Hachem, M., Halldórsdóttir, S., Thorbjarnadóttir, S.H., Eggertsson, G., Hreggvidsson, G.O., Nordberg Karlsson, E. and Holst, O. (2001) Deletion of a cytotoxic, N-terminal putative signal peptide results in a significant increase in production yields in Escherichia coli and improved specific activity of Cel12A from *Rhodothermus marinus*. *Appl. Microbiol. Biotechnol.* **55**, 578–584.
45. Pace, C.N., Scholtz, J.M. and Grimsley, G.R. (2014) Forces stabilizing proteins. *FEBS Lett.* **588**, 2177–2184.
46. Pace, C.N., Grimsley, G.R., Thomson, J.A. and Barnett, B.J. (1988) Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.* **263**, 11820–11825.
47. Piatek, R., Bruz´dziak, P., Wojciechowski, M., Zalewska-Piatek, B.M. and Kur, J. (2010) The noncanonical disulfide bond as the important stabilizing element of the immunoglobulin fold of the Dr fimbrial DraE subunit. *Biochemistry* **49**, 1460–1468.
48. Schulenburg, C., Weininger, U., Neumann, P., Meiselbach, H., Stubbs, M.T., Sticht,

H., Balbach, J., Ulbrich-Hofmann, R. and Arnold, U. (2010) Impact of the C-terminal disulfide bond on the folding and stability of onconase. *ChemBioChem* **11**, 978–986.

49. Mason, J.M., Bendall, D.S., Howe, C.J. and Worrall, J.A.R. (2012) The role of a disulfide bridge in the stability and folding kinetics of *Arabidopsis thaliana* cytochrome c6A. *Biochim. Biophys. Acta* **1824**, 311–318.
50. Sandgren, M., Gualfetti, P.J., Shaw, A., Gross, L.S., Saldajeno, M., Day, A.G., Jones, T.A. and Mitchinson, C. (2003) Comparison of family 12 glycoside hydrolases and recruited substitutions important for thermal stability. *Protein Sci.* **12**, 848–860.
51. Cheng, Y.S., Chen, C.C., Huang, C.H., Ko, T.P., Luo, W., Huang, J.W., Liu, J.R. and Guo, R.T. (2014) Structural analysis of a glycoside hydrolase family 11 xylanase from *Neocallimastix patriciarum*: insights into the molecular basis of a thermophilic enzyme. *J. Biol. Chem.* **289**, 11020–11028.
52. Dumon, C., Varvak, A., Wall, M.A., Flint, J.E., Lewis, R.J., Lakey, J.H., Morland, C., Luginbühl, P., Healey, S., Todaro, T., DeSantis, G., Sun, M., Parra-Gessert, L., Tan, X., Weiner, D.P. and Gilbert, H.J. (2008) Engineering hyperthermostability into a GH11 xylanase is mediated by subtle changes to protein structure. *J. Biol. Chem.* **283**, 22557–22564.
53. Fenel, F., Leisola, M., Jänis, J. and Turunen, O. (2004) A de novo designed Nterminal disulphide bridge stabilizes the *Trichoderma reesei* endo-1,4-bxylanase II. *J. Biotechnol.* **108**, 137–143.
54. Jänis, J., Turunen, O., Leisola, M., Derrick, P.J., Rouvinen, J. and Vainiotalo, P. (2004) Characterization of mutant xylanases using fourier transform ion cyclotron resonance mass spectrometry: stabilizing contributions of disulfide bridges and N-terminal extensions. *Biochemistry* **43**, 9556–9566.
55. Wang, Y., Fu, Z., Huang, H., Zhang, H., Yao, B., Xiong, H. and Turunen, O. (2012) Improved thermal performance of *Thermomyces lanuginosus* GH11 xylanase by engineering of an N-terminal disulfide bridge. *Bioresour. Technol.* **112**, 275–279.
56. Li, H., Kankaanpää, A., Xiong, H., Hummel, M., Sixta, H., Ojamo, H. and Turunen, O. (2013) Thermostabilization of extremophilic *Dictyoglomus thermophilum* GH11 xylanase by an N-terminal disulfide bridge and the effect of ionic liquid [emim] OAc on the enzymatic performance. *Enzyme Microb. Technol.* **53**, 414–419.
57. Tishkov, V.I., Gusakov, A.V., Cherkashina, A.S. and Sinitsyn, A.P. (2013)

Engineering the pH-optimum of activity of the GH12 family endoglucanase by site-directed mutagenesis. *Biochimie* **95**, 1704–1710.

58. Sandgren, M., Ståhlberg, J. and Mitchinson, C. (2005) Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. *Prog. Biophys. Mol. Biol.* **89**, 246–291.
59. Parsiegla, G., Belaïch, A., Belaïch, J.P. and Haser, R. (2002) Crystal structure of the cellulase Cel9M enlightens structure/function relationships of the variable catalytic modules in glycoside hydrolases. *Biochemistry* **41**, 11134-11142.
60. Kesavulu, M.M., Tsai, J.Y., Lee, H.L., Liang, P.H. and Hsiao, C.D. (2012) Structure of the catalytic domain of the *Clostridium thermocellum* cellulase CelT. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 310-320.
61. Arimori, T., Ito, A., Nakazawa, M., Ueda, M. and Tamada, T. (2013) Crystal structure of endo-1,4-β-glucanase from *Eisenia fetida*. *J. Synchrotron Radiat.* **20**, 884-889.
62. Khademi, S., Guarino, L.A., Watanabe, H., Tokuda, G. and Meyer, E.F. (2002) Structure of an endoglucanase from termite, *Nasutitermes takasagoensis*. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 653-659.
63. Pereira, J.H., Sapra, R., Volponi, J.V., Kozina, C.L., Simmons, B. and Adams, P.D. (2009) Structure of endoglucanase Cel9A from the thermoacidophilic *Alicyclobacillus acidocaldarius*. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 744-750.
64. Chauvaux, S., Souchon, H., Alzari, P.M., Chariot, P. and Beguin, P. (1995) Structural and functional analysis of the metal-binding sites of *Clostridium thermocellum* endoglucanase CelD. *J. Biol. Chem.* **270**, 9757-62.
65. Schubot, F.D., Kataeva, I.A., Chang, J., Shah, A.K., Ljungdahl, L.G., Rose, J.P. and Wang, B.C. (2004) Structural basis for the exocellulase activity of the cellobiohydrolase CbhA from *Clostridium thermocellum*. *Biochemistry* **43**, 1163-1170.
66. Mandelman, D., Belaich, A., Belaich, J.P., Aghajari, N., Driguez, H. and Haser, R. (2003) X-Ray crystal structure of the multidomain endoglucanase Cel9G from *Clostridium cellulolyticum* complexed with natural and synthetic cello-oligosaccharides. *J. Bacteriol.* **185**, 4127-4135.
67. Sakon, J., Irwin, D., Wilson, D.B. and Karplus, P.A. (1997) Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. *Nat. Struct. Biol.* **4**, 810-818.

68. Zhou, W., Irwin, D.C., Escovar-Kousen, J. and Wilson, D.B. (2004) Kinetic studies of *Thermobifida fusca* Cel9A active site mutant enzymes. *Biochemistry* **43**, 9655-9663.
69. Kataeva, I.A., Uversky, V.N., Brewer, J.M., Schubot, F., Rose, J.P., Wang, B.C. and Ljungdahl, L.G. (2004) Interactions between immunoglobulin-like and catalytic modules in *Clostridium thermocellum* cellulosomal cellobiohydrolase CbhA. *Protein Eng. Des. Sel.* **17**, 759–769.
70. Eckert, K., Vigouroux, A., Lo Leggio, L. and Moréra, S. (2009) Crystal structures of *A. acidocaldarius* endoglucanase Cel9A in complex with cello-oligosaccharides: strong -1 and -2 subsites mimic cellobiohydrolase activity. *J. Mol. Biol.* **394**, 61–70.
71. Eckert, K., Zielinski, F., Lo Leggio, L. and Schneider, E. (2002) Gene cloning, sequencing, and characterization of a family 9 endoglucanase (CelA) with an unusual pattern of activity from the thermoacidophile *Alicyclobacillus acidocaldarius* ATCC27009. *Appl. Microbiol. Biotechnol.* **60**, 428–436.
72. Liu, H., Pereira, J.H., Adams, P.D., Sapra, R., Simmons, B.A. and Sale, K.L. (2010) Molecular simulations provide new insights into the role of the accessory immunoglobulin-like domain of Cel9A. *FEBS Lett.* **584**, 3431–3435.
73. Ueda, M., Ito, A., Nakazawa, M., Miyatake, K., Sakaguchi, M. and Inouye, K. (2014) Cloning and expression of the cold-adapted endo-1,4-β-glucanase gene from *Eisenia fetida*. *Carbohydr. Polym.* **101**, 511– 516.
74. Kurokawa, J., Hemjinda, E., Arai, T., Kimura, T., Sakka, K. and Ohmiya, K. (2002) *Clostridium thermocellum* cellulase CelT, a family 9 endoglucanase without an Ig-like domain or family 3c carbohydrate-binding module. *Appl. Microbiol. Biotechnol.* **59**, 455–461.
75. Mingardon, F., Bagert, J.D., Maisonnier, C., Trudeau, D.L. and Arnold, F.H. (2011) Comparison of family 9 cellulases from mesophilic and thermophilic bacteria. *Appl. Environ. Microbiol.* **77**, 1436-1442.
76. Hendrickson, W.A., Horton, J.R. and LeMaster, D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9**, 1665-1672.
77. Pape, T. and Schneider, T.R. (2004) HKL2MAP: a graphical user interface for

macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843–844.

78. Sheldrick, G.M. (2008) A short history of SHELX. *Acta Crystallogr. A Found Crystallogr.* **64**, 112–122.
79. Arpigny, J.L. and Jaeger, K.E. (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem. J.* **343**, 177-183.
80. Challacombe, J.F., Eichorst, S.A., Hauser, L., Land, M., Xie, G. and Kuske, C.R. (2011) Biological consequences of ancient gene acquisition and duplication in the large genome of *Candidatus Solibacter usitatus* Ellin6076. *PLoS ONE* **6**, e24882.
81. Levisson, M., Sun, L., Hendriks, S., Swinkels, P., Akveld, T., Bultema, J.B., Barendregt, A., van den Heuvel, R.H., Dijkstra, B.W., van der Oost, J. and Kengen, S.W. (2009) Crystal structure and biochemical properties of a novel thermostable esterase containing an immunoglobulin-like domain. *J. Mol. Biol.* **385**, 949-962.
82. Karshikoff, A. and Ladenstein, R. (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a “traffic rule” for hot roads. *Trends Biochem. Sci.* **26**, 550– 556.
83. Shirley, B.A., Stanssens, P., Hahn, U. and Pace, C.N. (1992) Contribution of hydrogen bonding to the conformational stability of ribonuclease T1. *Biochemistry* **31**, 725-732.
84. Watanabe, K., Chishiro, K., Kitamura, K. and Suzuki, Y. (1991) Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermoglucosidasius* KP1006. *J. Biol. Chem.* **266**, 24287–24294.
85. Boutz, D.R., Cascio, D., Whitelegge, J., Perry, L.J. and Yeates, T.O. (2007) Discovery of a thermophilic protein complex stabilized by topologically interlinked chains. *J. Mol. Biol.* **368**, 1332–1344.
86. Pace, C.N., Fu, H., Fryar, K.L., Landua, J., Trevino, S.R., Shirley, B.A., Hendricks, M.M., Iimura, S., Gajiwala, K., Scholtz, J.M. and Grimsley, G.R. (2011) Contribution of hydrophobic interactions to protein stability. *J. Mol. Biol.* **408**, 514–528.
87. You, D.J., Chon, H., Koga, Y., Takano, K. and Kanaya, S. (2007) Crystal structure of type 1 ribonuclease H from hyperthermophilic archaeon *Sulfolobus tokodaii*: role of arginine 118 and C-terminal anchoring. *Biochemistry* **46**, 11494–11503.

本研究に関する論文

1. **Okano, H.**, Ozaki, M., Kanaya, E., Kim, JJ., Angkawidjaja, C., Koga, Y. and Kanaya, S. (2014) Structure and stability of metagenome-derived glycoside hydrolase family 12 cellulase (LC-CelA) a homolog of Cel12A from *Rhodothermus marinus*. *FEBS Open Bio* **4**, 936–946.
2. **Okano, H.**, Hong, X., Kanaya, E., Angkawidjaja, C. and Kanaya, S. (2015) Structural and biochemical characterization of a metagenome-derived esterase with a long N-terminal extension. *Protein Sci*. **24**, 93-104.
3. **Okano, H.**, Ozaki, M., Kanaya, E., Angkawidjaja, C. and Kanaya, S. (2015) Structure, activity and stability of metagenome-derived glycoside hydrolase family 9 endoglucanase with an N-terminal Ig-like domain. *Protein Sci*., in press. doi: 10.1002/pro.2632.

謝辞

本研究は、大阪大学大学院工学研究科生命先端工学専攻　金谷茂則教授のもとにおいて行われたものであり、先生より多大なる御指導、御鞭撻を賜りましたことを心より感謝致します。

本研究に際し、暖かい激励と御指導を賜りました大阪大学大学院工学研究科生命先端工学専攻　古賀雄一准教授、Clement Angkawidjaja 特任助教、京都府立大学大学院生命環境科学研究科応用生命科学専攻　高野和文教授に謹んで御礼申し上げます。

本論文をまとめるにあたり、有益な御助言を賜りました大阪大学大学院工学研究科生命先端工学専攻　福住俊一教授、菊地和也教授に深く感謝致します。日々様々な便宜を図って戴きました松本玲子女史、森尾雅江女史に深く感謝致します。また、6 年間の研究生活において様々な面で温かい御配慮を戴いた金谷研究室の皆様に感謝致します。

最後に、在学中も様々な面で支援して下さった家族の皆様に深く感謝致します。

岡野 啓志