

博士論文

エキソ型ガラクトサン分解酵素の  
基質認識機構

松山佳織



# 目次

<b>第一章 序論</b> .....	<b>1</b>
1.1. 植物細胞壁 .....	1
1.1.1. 一次細胞壁と二次細胞壁 .....	1
1.1.2. ペクチンの局在と機能 .....	5
1.1.3. ガラクタンの局在と種類 .....	6
1.2. 糖質関連酵素 .....	10
1.2.1. 糖質加水分解酵素の分類 .....	10
1.2.2. GHs の反応メカニズム .....	12
1.3. ガラクタン分解酵素の局在とその酵素特性 .....	14
1.3.1. GH35 $\beta$ -ガラクトシダーゼ (EC3.2.1.23) .....	17
1.3.2. GH43 イキソ- $\beta$ -1,3-ガラクターナーゼ (EC 3.2.1.145) .....	20
1.4. 本研究の目的 .....	24
<b>第二章 トマト果実由来 GH35 <math>\beta</math>-ガラクトシダーゼの酵素基質複合体の構造解析</b> .....	<b>25</b>
2.1. 緒言 .....	25
2.2. 材料および方法 .....	27
2.2.1. 酵素生産 .....	27
2.2.2. 基質の調製 .....	27
2.2.3. X線結晶構造解析実験および構造精密化 .....	27
2.2.4. ドッキングシミュレーション .....	29
2.3. 結果 .....	30
2.3.1. ドメインアノテーション .....	30
2.3.2. X線結晶構造解析 .....	31
2.3.3. アンサンブルリファインメント .....	37
2.3.4. ドッキングシミュレーション .....	40

2.4. 考察	42
2.4.1. TBG4 の基質認識機構	42
2.4.2. 成熟中のトマト果実細胞壁における TBG 4 の役割	46
2.5. 本章の結論	47
<b>第三章 担子菌由来 GH43 エキソ-<math>\beta</math>-1,3-ガラクターゼの酵素基質複合体の構造解析</b>	<b>49</b>
3.1. 緒言	49
3.2. 材料および方法	50
3.2.1. 酵素生産	50
3.2.2. 酵素活性の解析	51
3.2.3. X 線結晶構造解析実験および構造精密化	52
3.3. 結果	55
3.3.1. X 線結晶構造解析	55
3.3.2. アンサンブルリファインメント	64
3.4. 考察	73
3.4.1. <i>Pc1,3Gal43A</i> の加水分解機構	73
3.4.2. 側鎖バイパス機構	79
3.4.3. CBM35 のリガンド認識機構	82
3.5. 本章の結論	84
<b>第四章 総括</b>	<b>85</b>
<b>参考文献</b>	<b>94</b>
<b>謝辞</b>	<b>104</b>
<b>出版論文</b>	<b>107</b>

## 略語一覽

AG-I, II	arabinogalactan I, II, respectively
AGP	arabinogalactan proteins
CBM	carbohydrate binding module
Gal <sub>-1</sub> , Gal <sub>+1</sub> , Gal <sub>+2</sub>	Gal residue occupied subsite -1, +1, +2
Gal <sub>site</sub> 1, Gal <sub>site</sub> 2, Gal <sub>site</sub> 3	the non-reducing terminal, middle, and reducing terminal Gal residue of Gal3 bound to <i>PcCBM35</i>
Gal2, Gal3	galactobiose, galactotriose
GH	glycoside hydrolase
GH43 <sub>sub</sub> 24	GH family 43 subfamily 24
PEG	polyethylene glycol
SeMet	selenomethionine
<b>Enzymes</b>	
TBG4	Tomato $\beta$ -galactosidase 4
E181A $\beta$ -1,3-Gal2	TBG4 E181A bound with $\beta$ -1,3-galactobiose
E181A $\beta$ -1,4-Gal2	TBG4 E181A bound with $\beta$ -1,4-galactobiose
E181A $\beta$ -1,6-Gal2	TBG4 E181A bound with $\beta$ -1,6-galactobiose
<i>Pc</i> 1,3Gal43A	exo- $\beta$ -1,3-galactanase from <i>Phanerochaete chrysosporium</i>
<i>PcCBM35</i>	CBM35 domain of <i>Pc</i> 1,3Gal43A
E208Q <sub>Gal3</sub>	<i>Pc</i> 1,3Gal43A E208Q bound with Gal3
E208A <sub>Gal3</sub>	<i>Pc</i> 1,3Gal43A E208A bound with Gal3
WT <sub>Gal</sub>	TBG4 or <i>Pc</i> 1,3Gal43A WT bound with Gal
BT3683	$\beta$ -1,3-galactosidase from <i>Bacteroides thetaiotaomicron</i> VPI-5482
<i>Ct</i> 1,3Gal43A	exo- $\beta$ -1,3-galactanase from <i>Clostridium thermocellum</i>
<i>Cte</i> <sub>2137</sub>	CBM35 of <i>C. thermocellum</i> cellulosomal protein
<i>HjCel3A</i>	GH3 $\beta$ -1,3-glucosidase from <i>Hypocrea jecorina</i>
<i>HvExo1</i>	GH3 $\beta$ -D-glucan glucohydrolase from <i>Hordeum vulgare</i> subsp. <i>vulgare</i>
<i>PcCel45A</i>	endoglucanase V from <i>P. chrysosporium</i>
<i>PcLam55A</i>	exo- $\beta$ -1,3-glucanase from <i>P. chrysosporium</i>
<i>SacteLam55A</i>	GH55 exo- $\beta$ -1,3-glucanase from <i>Streptomyces</i> sp.



# 第一章 序論

## 1.1. 植物細胞壁

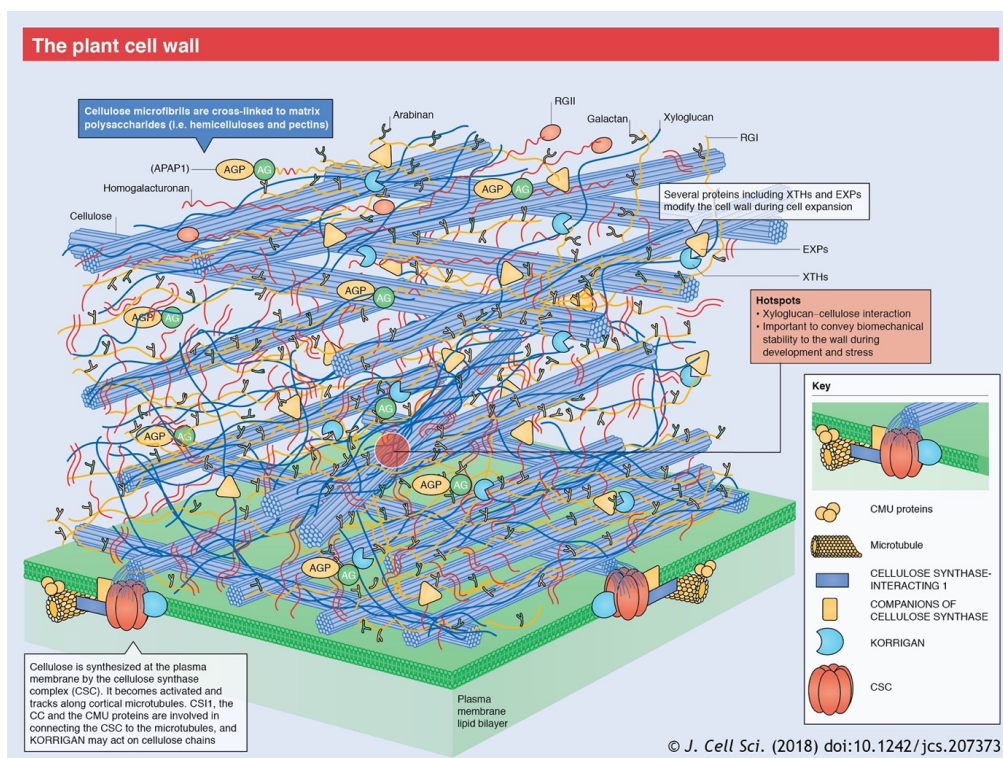
植物の生活環において、成長過程は配偶子形成から胚の形成にいたるまでの単相世代の過程と、胚から種子形成にいたる胚発生過程、発芽後に頂端分裂組織の増殖により行われる栄養成長過程、次世代を残すために花成に始まる生殖過程の 4 つの過程に分類される (Huijser and Schmid, 2011)。植物細胞の外側に特徴的にみられる細胞壁とよばれる構造体は、これらの過程において細胞分裂、細胞接着、多細胞体形成、細胞分化・細胞増殖の制御、組織強度の支持、シグナル伝達、生体防御、修復、水理機能、炭水化物と一部のイオンの貯蔵など、非常に重要な役割を担っている (Albersheim *et al.*, 2010; Taiz and Zeiger, 2010)。また、植物は成長に応じてこの細胞壁構造を構築および再構築する (Lampugnani *et al.*, 2018) ため、植物の生態を研究したり利用したりするためにはそれに関わる酵素の機能に関する理解が必要不可欠である。一方で、細菌や真菌などの微生物は多様な酵素を産生して植物細胞壁を分解し、炭素をはじめとする様々な栄養を得ている (Kubicek *et al.*, 2014; Rytioja *et al.*, 2014; Berlemont and Martiny, 2016)。微生物が産生する酵素は、すでに植物細胞壁の構造解析や化成品の原料や食品の調製などの用途で産業に利用されていて大変有用であるが、まだまだ機能やその作用機序が未解明な酵素も多い。また、植物バイオマスの大部分は細胞壁である。したがって植物の生態を理解して望ましい特性を有するように改変したり、バイオマスとして有効活用したりするためには、細胞壁構造やその構築・再構築に関与する植物自身の酵素だけでなく、微生物由来の酵素についても詳細に研究を行うことが非常に重要である。

### 1.1.1. 一次細胞壁と二次細胞壁

植物の細胞壁は、その構造と機能に基づき一次細胞壁 (primary wall) と二次細胞壁 (secondary wall) の 2 つに分類される (Somerville *et al.*, 2004; Burton *et al.*, 2010)。一次細胞壁は細胞膜の外側に最初に形成される、細胞分裂に不可欠な細胞壁であり、セルロース、ヘミセルロース、ペクチンと呼ばれる大きく分けて 3 つのドメインにより構成される (Zabackis *et al.*, 1995; Somerville *et al.*, 2004; Keegstra, 2010)。一次細胞壁の構造モデルとしては、不溶性のホモ多糖類であるセルロースマイクロフィブリルをヘテロ多糖類であるヘミセルロースが

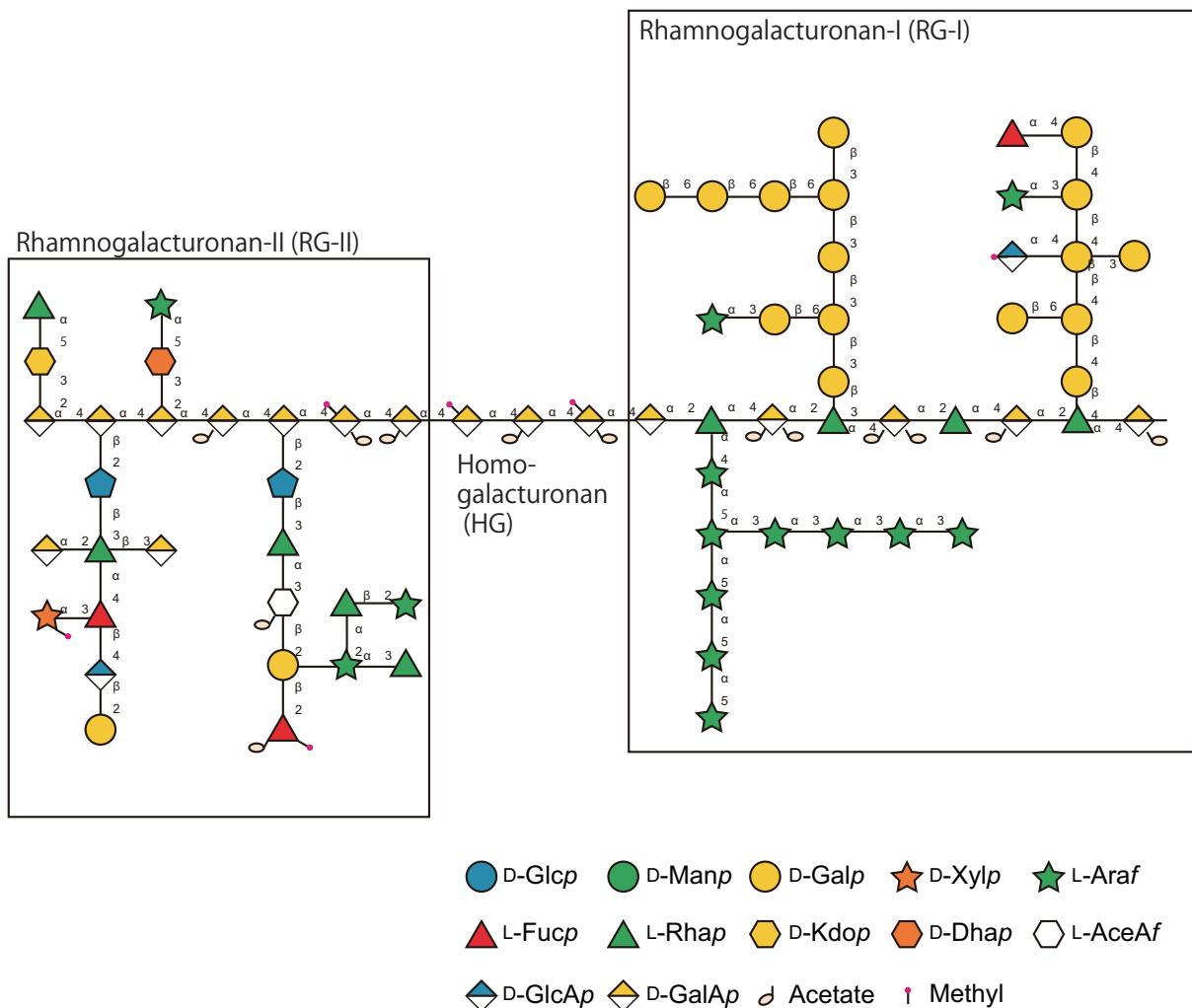
架橋し (Rose *et al.*, 2002; Scheller and Ulvskov, 2010; Pauly *et al.*, 2013)、ゲル状のペクチンがその間隙を充填するというモデルが提唱されている (Fig. 1-1) (Carpita and Gibeaut, 1993; Harholt *et al.*, 2010)。

一方、二次細胞壁は細胞が伸長した後に一次細胞壁の内側に肥厚する、植物に強度と剛性を与える構造体であり (Speck and Burgert, 2011)、主にセルロース、ヘミセルロース、リグニンによって構成される (Keegstra, 2010; Li and Chapple, 2010; Speck and Burgert, 2011)。セルロースマイクロフィブリルをヘミセルロースが架橋し、二次細胞壁に特徴的に見られるフェノール性ポリマーであるリグニンがセルロース-ヘミセルロースを被覆する構造モデルが提唱されている (Speck and Burgert, 2011)。リグニンは、細胞強度を高めたり水の通道性に寄与したりするため、植物が陸上で巨大化し、長期間にわたって繁茂するために重要な要素である (Li and Chapple, 2010; Neutelings, 2011; Speck and Burgert, 2011)。また、二次細胞壁にも微量ながらペクチンが存在する (Gorshkova *et al.*, 2018; Torode *et al.*, 2018)。したがって、ペクチンはセルロースやヘミセルロースと同様に一次細胞壁だけでなく二次細胞壁においても非常に重要な役割を担う多糖類であると考えられる。



**Fig. 1-1. Structure model of the primary cell wall in growing cells from *Arabidopsis thaliana* (Lampugnani *et al.*, 2018).**





**Fig. 1-2. Schematic diagram of the structure of pectin. The structures are illustrated based on O'Neill and York (2003).**

**Table 1-1. The ratio <sup>\*</sup> of polysaccharides and lignin in the cell wall.**

Polymer	Cell wall % (w/v)			
	Primary wall		Secondary wall <sup>**</sup>	
	Eudicots	Poaceae	Angiosperms	Gymnosperms
Cellulose	20-30	20-30	37-57	38-52
Hemicellulose <sup>***</sup>	25-30	30-70	20-37	16-27
Pectin <sup>****</sup>	30-35	1-5	<10	<10
Lignin	0	0	17-30	26-36

<sup>\*</sup>The ratio is average of species and organs. It does not include protein amount because of the rate depends on species.

<sup>\*\*</sup>Secondary wall also includes a few amount of primary wall.

<sup>\*\*\*</sup>Hemicellulose includes xyloglucan, xylan, and  $\beta$ -1,3/1,4-glucan.

<sup>\*\*\*\*</sup>Pectin includes HG, RG-I, and RG-II.

The data is based on Rose et al. 2003; Albersheim et al. 2011; 西谷 and 梅澤 2013.

### 1.1.2. ペクチンの局在と機能

ペクチンは植物の成長過程において、細胞強度や細胞接着や生体防御などの生理学的に重要な役割を担う多糖類の1つである。古くからジャムやゼリーなどのゲル化剤や酸性乳飲料のタンパク質分散安定剤、増粘剤など、食品添加物として食品分野に広く用いられている。また、化粧品や医薬品にも用いられているため、産業的にも非常に重要な多糖類であるといえる。

ペクチンは、一次細胞壁から熱水や酸、キレートにより抽出される多様な糖類から構成されるヘテロ多糖類の総称であり、D-ガラクトツロン酸 (GalA) を多く含有する。ペクチンはホモガラクトツロナン (homogalacturonan; HG)、ラムノガラクトツロナン-I (rhamnogalacturonan-I; RG-I)、ラムノガラクトツロナン-II (rhamnogalacturonan-II; RG-II) の3つのドメインにより構成される (Fig. 1-2) (Harholt *et al.*, 2010)。HG は、GalA が $\beta$ -1,4-結合した直鎖状のホモポリマーで、一般に一次細胞壁の中で最大の割合を占める (Mohnen, 2008)。多くの場合は6位のカルボニル基がメチルエステル化、2位または3位のヒドロキシ基がアセチル化されている (Ishii, 1997; Saulnier and Thibault, 1999; Rose *et al.*, 2003; Albersheim *et al.*, 2010)。RG-I は、 $\alpha$ -1,4-GalA- $\alpha$ -1,2-L-ラムノース (Rha) の二糖繰り返し構造の主鎖を持ち (Lau *et al.*, 1985; Carpita and Gibeaut, 1993; Ulvskov *et al.*, 2005)、GalA の2位または3位のヒドロキシ基がアセチル化されており (Komalavilas and Mort, 1989; Carpita and Gibeaut, 1993; Ishii, 1997)、さらに20~80%のRha残基に $\beta$ -1,4-ガラクトン (type I ガラクトン) やアラビノガラクトン (type II ガラクトン)、 $\alpha$ -1,3/1,5-アラビナンなどの中性糖側鎖が結合した分岐ポリマーである (Lau *et al.*, 1985; Carpita and Gibeaut, 1993; Ulvskov *et al.*, 2005)。RG-II はペクチンのマイナードメインであり (Matsunaga *et al.*, 2004)、主鎖はHGと同様に $\alpha$ -1,4-結合したGalAによって構成される (Whitcombe *et al.*, 1995)。側鎖には12種類もの単糖を含み、20種類以上のグリコシド結合から構成される非常に複雑な構造をもつ (Ridley *et al.*, 2001; O'Neill and York, 2003; Matsunaga *et al.*, 2004; Harholt *et al.*, 2010)。さらに、細胞壁中に存在するハウ素が二本のRG-II鎖を共有結合的に架橋してハウ酸ジオールエステルとなる (Ishii and Matsunaga, 1996; Kobayashi *et al.*, 1996)。シロイヌナズナ (*Arabidopsis thaliana*) の葉では、ペクチンは一次細胞壁の50% (w/v) 程度を占める (Zabackis *et al.*, 1995) が、細胞壁中においてペクチンが占める割合や各構成多糖類の割合および構造は植物種、器官、成長段階、組織、生育環境によって異なる (Albersheim *et al.*, 2010; Harholt *et al.*, 2010)。一方、

二次細胞壁においてはあて材や G 層などに RG-I や type I ガラクタンが特異的にみられることが報告されている (Roach *et al.*, 2011; Gorshkova *et al.*, 2015, 2018)。

### 1.1.3. ガラクタンの局在と種類

ガラクトタンは植物体を構成する様々な組織に局在する多糖類であり、その構造からアラビノガラクトタン-I (AG-I) とアラビノガラクトタン-II (AG-II) の 2 つに大きく分類される。AG-I は $\beta$ -1,4-ガラクトタン主鎖に L-フコピラノース (Fuc)、L-アラビノフラノース (Ara)、4-O-メチル-D-グルクロン酸 (4-O-Me-GlcpA) や Gal が側鎖として結合した構造をとる。ペクチンの主要領域の 1 つである RG-I の側鎖を構成する (Fig. 1-2) ことからペクチン性ガラクトタン (pectic galactan) とよばれ、ジャガイモ (*Solanum tuberosum*) の RG-I の主要な中性糖側鎖は AG-I であると知られている (Øbro *et al.*, 2004)。AG-I は葉や花卉、茎、果実、種子などの特に成長が盛んな組織にみられる多糖類で、セルロース-ペクチン相互作用に関与するため、細胞の成長のあらゆる段階において重要である (McCartney *et al.*, 2003; Ulvskov *et al.*, 2005; Zykwincka *et al.*, 2007, 2008; Corral-Martínez *et al.*, 2019)。シロイヌナズナにおいて $\beta$ -ガラクトシダーゼを過剰発現させて $\beta$ -1,4-ガラクトタンが細胞壁中に堆積しないようにし、 $\beta$ -1,4-ガラクトタンの影響を調査した研究では、 $\beta$ -1,4-ガラクトタンは細胞成長の停止や二次細胞壁形成の際にキシログルカンやグルクロナキシランなどのヘミセルロースに分類される多糖類を適切に配置し、正しく組織化するために重要な構造的役割を担うと明らかになった (Moneo-Sánchez *et al.*, 2019, 2020)。また、広葉樹の引張あて材や針葉樹の圧縮あて材では正常材よりも Gal 含量が多く、 $\beta$ -1,4-ガラクトタン主鎖に $\beta$ -1,6-ガラクトタンが側鎖として結合した状態で存在すると報告されており (Mast *et al.*, 2009; Gorshkova *et al.*, 2015; 吉澤, 2016)、特に圧縮あて材においては主要なヘミセルロースとして存在している (Mast *et al.*, 2009)。AG-I の $\beta$ -1,4-ガラクトタン主鎖は水溶液中において右巻き 6 回螺旋構造をとる (Ryttersgaard *et al.*, 2002; Le Nours *et al.*, 2009)。

AG-II はペクチンの RG-I 側鎖、アラビノガラクトタンプロテイン (AGP) の糖鎖領域、カラムツの心材などにみられ、 $\beta$ -1,3-ガラクトタンからなる主鎖と $\beta$ -1,6-ガラクトタンからなる側鎖により基本構造が構成される (Swenson *et al.*, 1969; Fincher and Stone, 1983; Willför *et al.*, 2002; 西谷 and 梅澤, 2013; Tsumuraya and Kotake, 2017)。側鎖の $\beta$ -1,6-ガラクトタンには

さらに Ara や 4-O-Me-D-GlcA が結合することもある (Swenson *et al.*, 1969; Willför *et al.*, 2002)。β-1,3-ガラクトサン主鎖は水溶液中において右巻き 6~8 回螺旋構造をとり、さらに 3 本の β-1,3-ガラクトサン鎖が絡まりあった三重螺旋構造をとる。β-1,6-側鎖が結合している場合には側鎖部分が三重螺旋の外側に出る構造をとるため、側鎖が結合していても糖鎖全体のコンフォメーションを保つ上での障害にならないと考えられている (Chandrasekaran and Janaswamy, 2002; Kitazawa *et al.*, 2013)。なお、各結合様式のガラクトサン (β-1,3-、β-1,4-、β-1,6-ガラクトサン) の構造を Fig. 1-3 に示した。

すなわち、AG-I や AG-II はセルロースやヘミセルロースと相互作用したり別の RG-I 分子の中性糖側鎖と相互作用したり二次細胞壁の形成に関与したりするため、細胞壁の強度の保持に非常に重要である。しかしながら主鎖の RG-I への結合の頻度や中性糖側鎖の構成糖、その結合様式は植物種や組織、成長段階により異なるため、その構造や役割についてはいまだ明らかになっていない点も多い。

AGP は植物にみられるプロテオグリカンの 1 つであり、細胞壁や細胞の表層に普遍的に存在し、細胞外情報分子としてシグナル伝達や細胞分化、細胞形態の制御などの様々な生理学的現象に関与している (Fincher and Stone, 1983; Seifert and Roberts, 2007)。AGP のコアタンパク質は主に Ser、Thr、Hyp から構成され、AG-II 糖鎖が Hyp 残基に O-グリコシド結合している (Tsumuraya *et al.*, 1984; Majewska-Sawka and Nothnagel, 2000; Ellis *et al.*, 2010) (Fig. 1-4)。AGP の多くは N 末端のシグナルペプチドにより細胞外に輸送され、C 末端のグリコシルホスホスファチジルイノシトール (glycosylphosphatidylinositol; GPI) アンカーにより細胞膜上に結合している (Oxley and Bacic, 1999; Majewska-Sawka and Nothnagel, 2000; Gaspar *et al.*, 2001; Seifert and Roberts, 2007; Ellis *et al.*, 2010)。GPI アンカーがリパーゼにより切断されると AGP は細胞壁中に分泌される (Oxley and Bacic, 1999; Majewska-Sawka and Nothnagel, 2000; Gaspar *et al.*, 2001)。同種の植物内に多数の AGP が存在し、コアタンパク質や糖鎖の構造には多くのバリエーションがあるため、いまだに機能や構造の全容は明らかになっていない (Gaspar *et al.*, 2001; Seifert and Roberts, 2007; Ellis *et al.*, 2010)。

このようにガラクトサン (AG-I、AG-II、AGP 糖鎖) は植物にとって非常に重要な糖分子のひとつであるため、植物の成長機構を理解するためにはガラクトサンの構造や機能を詳細に理解することが肝要である。また一方でバイオマスの大部分は細胞壁であるため、細胞壁の主要な構成成分

であるセルロースやヘミセルロースをより有効活用するためにもこれらの多糖類と相互作用するガラクトランの構造を理解することは重要である。また、多糖類の構造解析を行う際にはその多糖類に作用する糖質加水分解酵素を用いることが一般的であるため、本研究においてはガラクトランを加水分解する酵素に着目した。

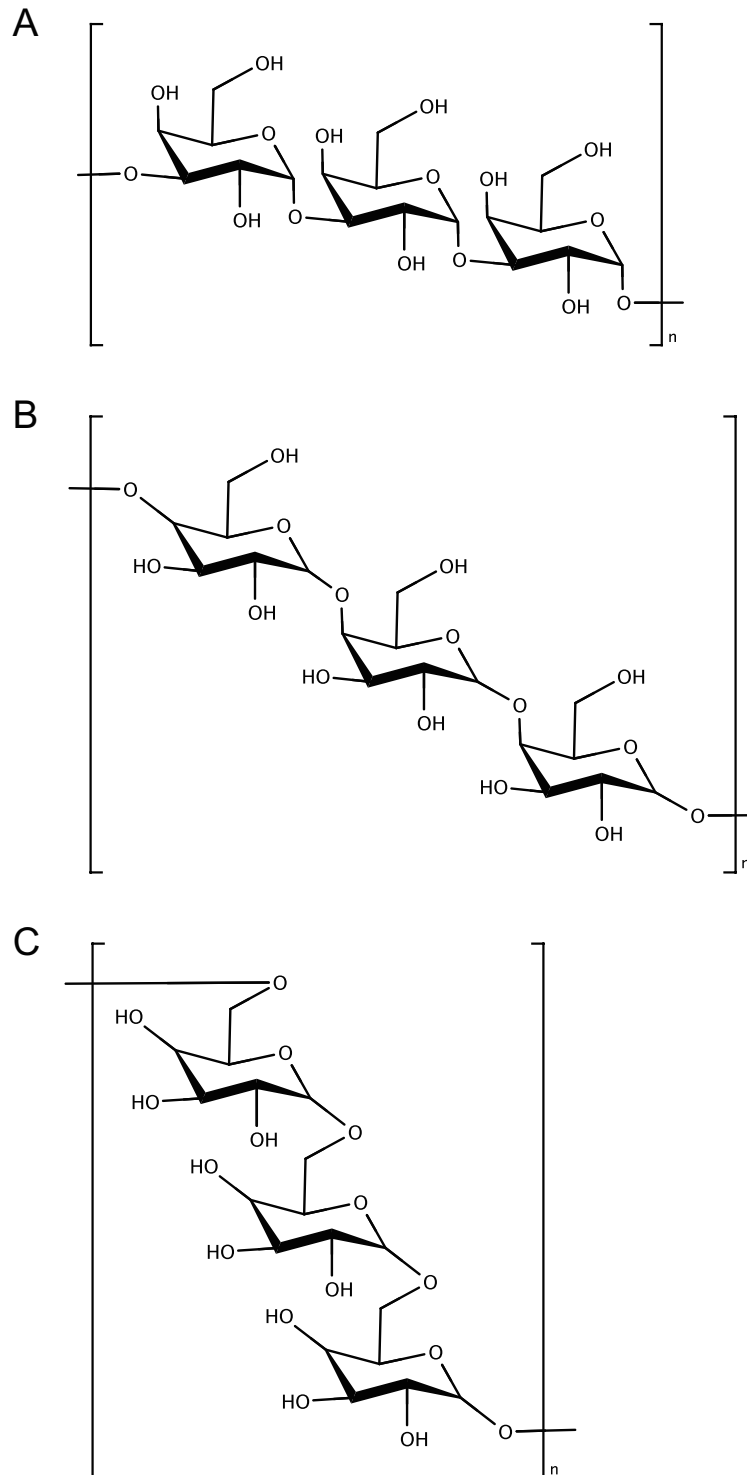
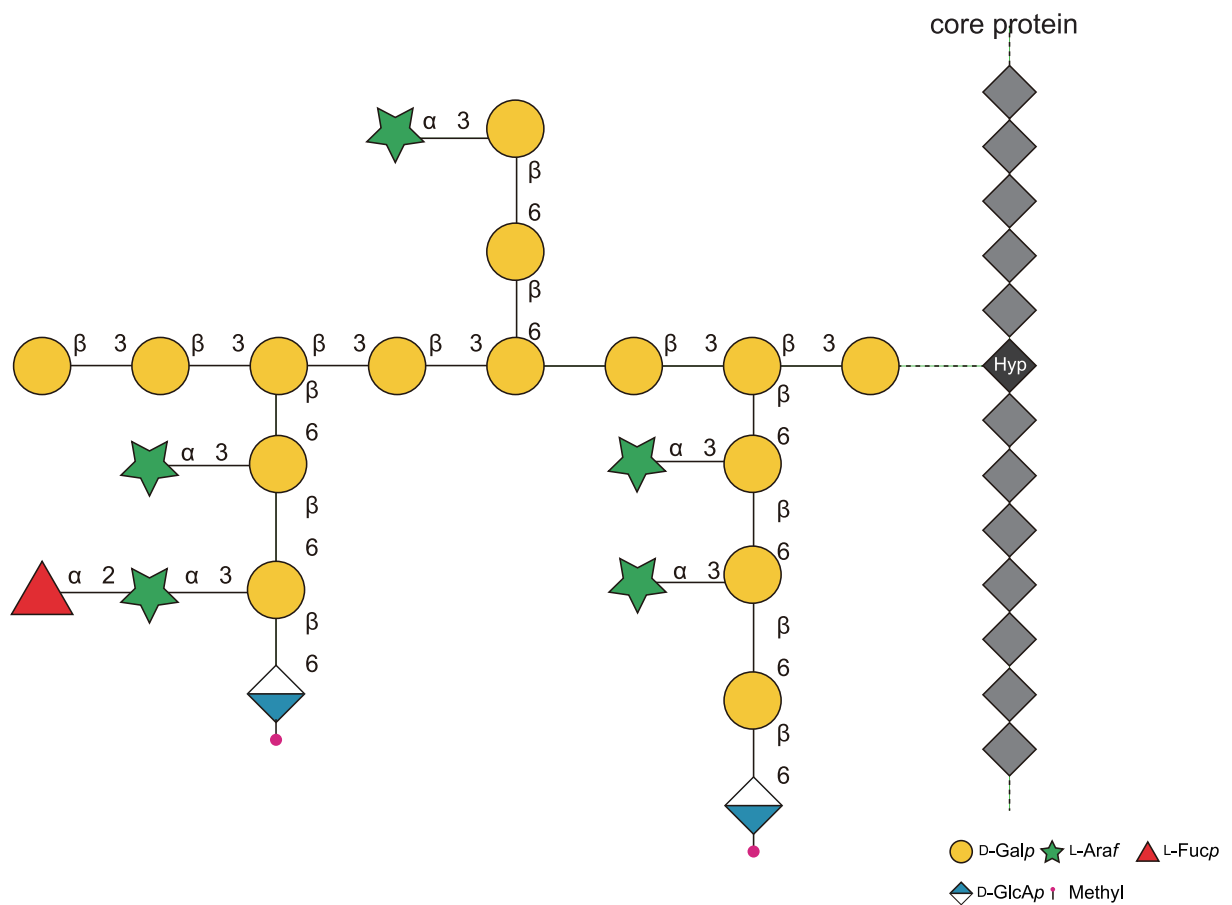


Fig. 1-3. The structure of  $\beta$ -1,3- (A),  $\beta$ -1,4- (B), and  $\beta$ -1,6-galactan (C).



**Fig. 1-4. Schematic diagram of the structure of arabinogalactan protein.**

The structures are illustrated based on the sugar ratio in Tsumuraya and Kotake (2017).

## 1.2. 糖質関連酵素

植物細胞壁構成多糖類の植物自身による構築および再構築、微生物による分解には非常に多くのタンパク質が関与している。これらのタンパク質には多糖類の合成（糖転移酵素; Glycosyl Transferases; GTs）、分解（糖質加水分解酵素、Glycoside Hydrolases, GHs; 多糖リアーゼ、Polysaccharides Lyases、PLs; 炭水化物エステラーゼ、Carbohydrate Esterases、CEs）に関連する酵素が含まれ、これらは総称して糖質関連酵素（Carbohydrate Active enZymes; CAZymes）と呼ばれている（<http://www.cazy.org/>）（Henrissat and Grenoble, 1991; Henrissat and Bairoch, 1993, 1996; Davies and Henrissat, 1995; Henrissat and Davies, 1997）。CAZymes の各酵素群は、アミノ酸配列の相同性に基づいて GHs、GTs、PLs、CEs がそれぞれ 170、114、41、18 のファミリーに分類されている（2021年2月22日現在、欠番含む）。またさらに、CAZymes には GHs、GTs、PLs、CEs と協調して作用する酸化還元酵素（Auxiliary Activities; AAs）の 16 ファミリーと炭水化物に結合するモジュール（Carbohydrate Binding Modules; CBMs）の 88 ファミリーも含まれる（Lombard *et al.*, 2014）（2021年2月22日現在）。すべての生物がこれらの酵素を産生し、生命活動に用いているが、なかでも植物は CAZymes を細胞壁の構築や再構築に活用している（Minic and Jouanin, 2006; Minic, 2008）。一方で、微生物はこれらの CAZymes を菌体外に分泌して植物細胞壁を分解し、分解した糖類を菌体内に取り込んで栄養源として用いることが知られている（Kubicek *et al.*, 2014; Rytioja *et al.*, 2014; Berlemont and Martiny, 2016）。

### 1.2.1. 糖質加水分解酵素の分類

GHs は糖のグリコシド結合の加水分解反応を触媒し、ヘミアセタールまたはヘミケタールと対応するアグリコンを遊離する酵素群であり、GHs は O-、N-、S-結合糖鎖を加水分解できる。これらはアミノ酸配列の相同性に基づいたファミリーだけではなく、立体構造のフォールディングの種類からファミリーの上位分類である clan というグループにも分類される。さらに、GHs は配列や構造だけではなく、分解タイプ（エキソ型またはエンド型）、反応メカニズム（立体保持型または立体反転型）、酵素番号（Enzyme Commission; EC numbers）によっても分類できる。また、GHs は触媒部位に基質のピラノースまたはフラノースと相互作用するサブサイトと呼



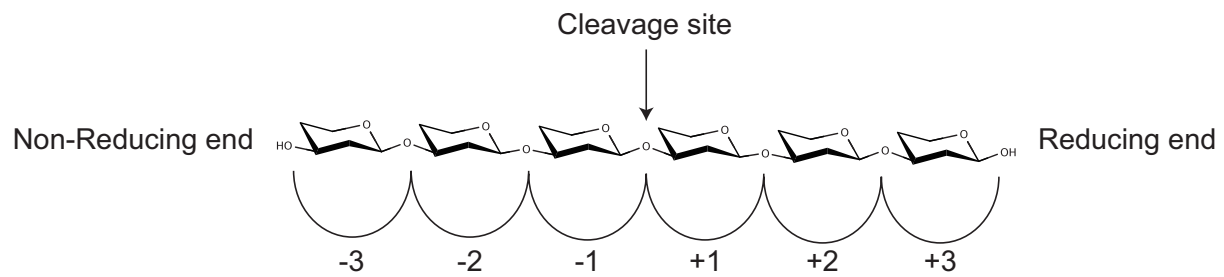
ばれるサイトを有しており、加水分解反応が起こるグリコシド結合を起点として非還元末端側をマイナス、還元末端側をプラスとし、各糖残基を認識して相互作用する残基群を基質のグリコシド結合から近い順にそれぞれサブサイト-1位、-2位、-3位…、+1位、+2位、+3位…と呼称する (Fig. 1-5) (Davies *et al.*, 1997)。

エキソ (exo) 型およびエンド (endo) 型は GHs の糖鎖の分解様式に基づく分類であり、エキソ型は糖鎖の末端、エンド型は糖鎖の内部をそれぞれ分解する (Davies and Henrissat, 1995; Henrissat and Davies, 1997)。オリゴ糖や多糖類は単糖が複数結合した構造をとるため、還元性を示すのは末端に位置する 1 つのアノマー炭素のみである。この還元性を示すアノマー炭素が位置する末端を還元末端 (reducing terminal)、他方の末端を非還元末端 (non-reducing terminal) と呼び (Henrissat and Davies, 1997)、エキソ型の酵素には非還元末端から作用するものと還元末端から作用するものがある。

反応メカニズムによる分類では反応後の還元末端のアノマー炭素の向きが反応前と反転するものを立体反転型 (Inverting mechanism)、一致するものを立体保持型 (Retaining mechanism) と称する。いずれの反応メカニズムに分類される酵素においても触媒反応に重要となるアミノ酸残基は 2 つ一組の酸性アミノ酸残基 (Asp または Glu) であることが多いが、いずれかの酸性残基を欠損した酵素や外因性の塩基を用いる酵素などの例外もみられ、またどのようなメカニズムで加水分解反応を触媒するのか明らかになっていないファミリーも多い (Koshland, 1953; Henrissat and Grenoble, 1991; Henrissat and Bairoch, 1993, 1996; McCarter and Withers, 1994; Davies and Henrissat, 1995; Henrissat and Davies, 1997)。

EC 番号 (Enzyme Commission numbers) は、国際生化学分子生物学連合酵素委員会 (NC-IUBMB) が酵素が触媒する反応の種類に基づいて分類して付与している番号であり、アミノ酸配列に相同性がない酵素でも同じ反応を触媒する場合は同じ番号が割り当てられる (McDonald *et al.*, 2009)。また一方で、配列の相同性があっても同じ反応を触媒するとは限らず、1 つの GH ファミリーに色々な EC 番号をもつ酵素が分類されている場合もある。さらに、例えば GH3 のシロイヌナズナの  $\alpha$ -L-アラビノフラノシダーゼ活性と  $\beta$ -1,4-キシロシダーゼ活性の両方を示す酵素 (*At*BX3) では EC 3.2.1.37 と EC 3.2.1.55 の 2 つの EC 番号が与えられているように、複数の反応を触媒する酵素には複数の EC 番号が割り当てられる場合もある。GHs の大部分を占める糖鎖の O-または S-結合を分解する酵素は EC 3.2.1.X、N-グリコシド化合物を分解する酵素には

EC 3.2.2.X、S-グリコシド化合物を分類する酵素には EC 3.2.3.X という番号がそれぞれ割り当てられている (<http://enzyme-database.org>) (McDonald *et al.*, 2009)。



**Fig. 1-5. The schematic diagram of subsite.**

The arrow indicates the cleavage site, left and right side are non-reducing end and reducing end, respectively.

### 1.2.2. GHs の反応メカニズム

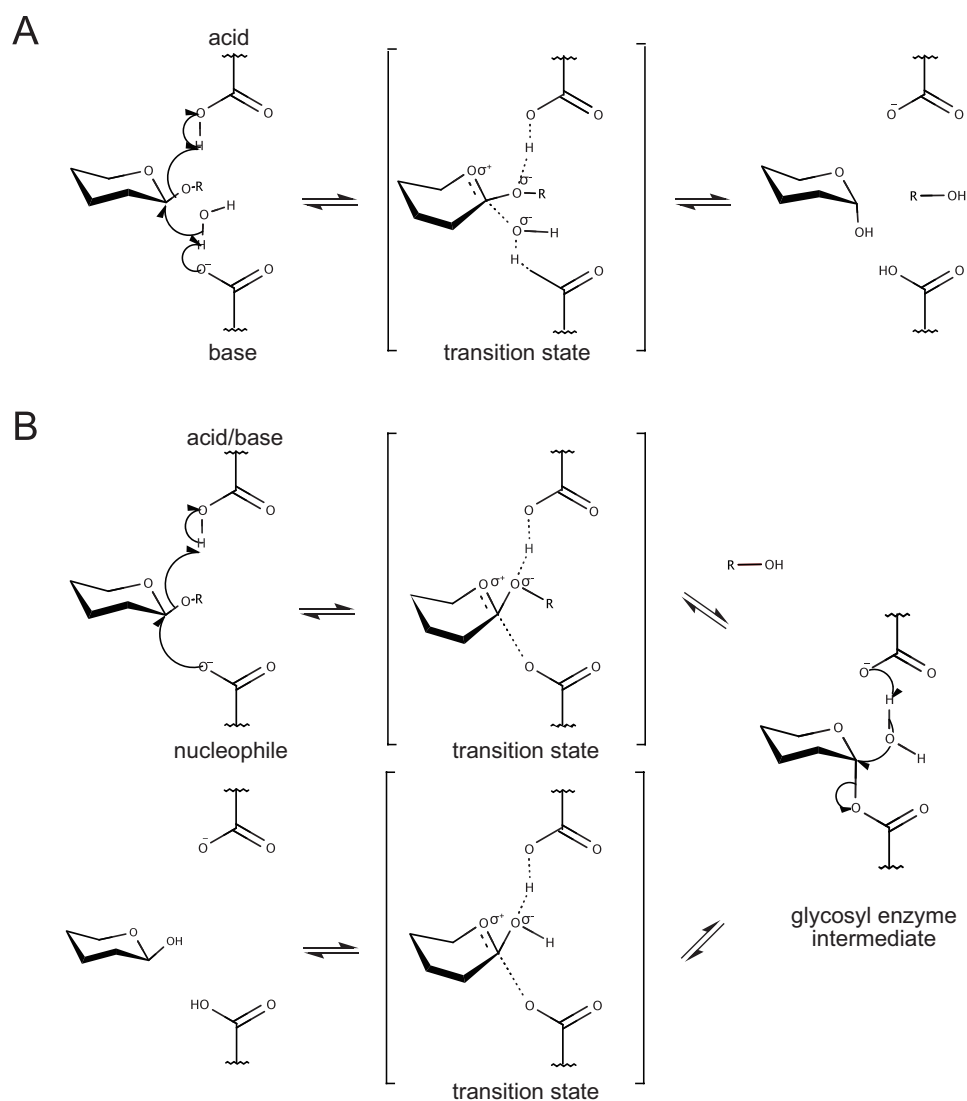
#### **Inverting 型酵素の反応メカニズム**

Inverting 型酵素では、通常 Asp または Glu の一般酸触媒残基 (general acid) および一般塩基触媒残基 (general base) と呼ばれる 2 つの酸性残基が 6~11 Å 離れたところに位置しており、これらが関与して置換反応が一度起こって基質のグリコシド結合を切断する。この反応メカニズムではグリコシド結合が general acid からプロトンを受け取り、general base によって活性化された求核水から求核攻撃を受け、オキソカルベニウムイオン中間体と呼ばれる中間体を経てアノマーが反転し、反応が完了する過程を経る (Fig. 1-6. A) (McCarter and Withers, 1994; Davies and Henrissat, 1995; Rye and Withers, 2000; Vasella *et al.*, 2002)。

#### **Retaining 型酵素の反応メカニズム**

Retaining 型酵素の反応メカニズムは “Classical Koshland retaining mechanism” と呼ばれるメカニズムで、約 5.5 Å 離れたところに位置し、一般酸/塩基触媒残基 (general acid/base) および求核性触媒残基 (nucleophile) とよばれる 2 つの酸性残基が関与して二重置換反応を経て加水分解反応を触媒するメカニズムである (Fig. 1-6. B)。まずグリコシル化と呼ばれる反応の第1ステップでは nucleophile がアグリコンのアノマー中心を求核攻撃してアグリコンを脱離し、

グリコシル-酵素中間体を形成する。同時に、general acid/base が酸として機能し、グリコシル酸素をプロトン化する。この際、アノマー炭素が一度反転する。脱グリコシル化と呼ばれる第 2 ステップでは general acid/base が塩基として作用して水分子を求核的に攻撃して脱プロトン化して活性化し、グリコシル-酵素中間体がこの活性化した水分子に求核攻撃を受けることで再度アノマー炭素が反転し、加水分解反応が完了する。このようにアノマー炭素の反転が二度おこるので、反応の前後でのアノマーの向きは保持される (Fig. 1-6. B)。この第 2 ステップでグリコシル-酵素中間体を求核攻撃するのが水分子ではなく糖分子であった場合には逆反応である糖転移反応が生じる (Koshland, 1953; McCarter and Withers, 1994; Davies and Henrissat, 1995; Rye and Withers, 2000; Vasella *et al.*, 2002)。



**Fig. 1-6. The reaction mechanisms of glycoside hydrolases.**

A and B show "Inverting type" and "Retaining type", respectively.

### 1.3. ガラクタン分解酵素の局在とその酵素特性

ガラクトタンを分解する酵素は $\beta$ -ガラクトシダーゼあるいはガラクタナーゼと呼ばれる。エキソ型のガラクトタン分解酵素のうち $\beta$ -ガラクトシダーゼはガラクトオリゴ糖の結合位特異性が低く、様々な結合様式のガラクトオリゴ糖や *p*-ニトロフェニル- $\beta$ -D-ガラクトピラノシド (*p*NP- $\beta$ -Gal) のような人工基質やラクトースなどのさまざまな結合様式の配糖体に対しても分解活性を示すのに対し、エキソ-ガラクタナーゼは *p*NP- $\beta$ -Gal やラクトースに対しては加水分解活性を全く示さないか示しても微弱であるが、ガラクトオリゴ糖の結合様式に対する特異性が非常に高いという違いがある (Sakamoto and Ishimaru, 2013)。 $\beta$ -ガラクトシダーゼ (EC 3.2.1.23) は糖鎖の非還元末端から Gal を遊離するエキソ型の反応を示し、GH ファミリーの 1、2、3、35、39、42、59、147、163、165 の 10 ファミリーに分類される (<http://www.cazy.org/Glycoside-Hydrolases.html>) が、これらの酵素の多くがラクトースを特異的に分解する酵素であり、ガラクタンの分解活性が報告されているのは GH2、35、42、147 に属する酵素のみである (Ichinose *et al.*, 2008; Sakamoto and Ishimaru, 2013; Luis *et al.*, 2018)。一方でガラクタナーゼと呼ばれる酵素はエキソ型とエンド型、分解対象とする結合様式の違いからエキソ- $\beta$ -1,3-ガラクタナーゼ (3.2.1.145)、エキソ- $\beta$ -1,4-ガラクタナーゼ (EC 3.2.1.-)、エンド- $\beta$ -1,3-ガラクタナーゼ (EC 3.2.1.181)、エンド- $\beta$ -1,4-ガラクタナーゼ (EC 3.2.1.89、EC 3.2.1.102)、エンド- $\beta$ -1,6-ガラクタナーゼ (EC 3.2.1.164、3.2.1.213) に分類され、GH ファミリーの 43、35、16、53 および 147、5 および 30 にそれぞれ属する (Fig 1-8、Table 1-2、<http://www.cazy.org/>)(Ichinose *et al.*, 2008; Sakamoto and Ishimaru, 2013)。

これらのガラクトタン分解酵素は clan A、clan B、clan F の 3 つに属しており、それぞれ TIM ( $\beta/\alpha$ )<sub>8</sub> バレル構造、 $\beta$ -ゼリーロール構造、5 枚羽根プロペラ構造と称される構造をとる (Fig 1-8、Table 1-2)。ガラクトタン分解酵素の中で最も頻繁にみられる構造は clan A の TIM ( $\beta/\alpha$ )<sub>8</sub> バレル構造であり、GH2、5、30、35、42、53、147 に属する分解酵素が TIM バレル構造をとる (<http://www.cazy.org/>)。ガラクトタン分解酵素のうち、clan B の $\beta$ -ゼリーロール構造と clan F の 5 枚羽根プロペラ構造をもつファミリーはそれぞれ GH16 と GH43 である (<http://www.cazy.org/>)。また、これまでに立体構造が明らかになっているガラクトタン分解酵素の基質結合部位近傍の表面構造は奥に深い“ポケット”状の構造と浅い溝のような“クレフト”状の構造に分類される (Fig. 1-7)。この基質結合部位の表面構造は各酵素の反応様式と相関があ

り、エキソ型の反応を示すものはポケット状の、エンド型の反応を示すものはクレフト状の基質結合部位をそれぞれ有する (Fig. 1-7)。この特徴はセルラーゼをはじめとする他のオリゴ糖や多糖類に反応性を示す糖質加水分解酵素で見られる特徴と同様である (Davies and Henrissat, 1995; Hrmova and Fincher, 2001)。

このように自然界には多様なガラクトサン分解酵素が存在するが、植物は GH35  $\beta$ -ガラクトシダーゼしか細胞外に産生しないのに対し、真菌や細菌はエキソ型のガラクトサン分解酵素を有する種が多い (van den Brink and de Vries, 2011; Rytioja *et al.*, 2014; Cartmell *et al.*, 2018; Fujita *et al.*, 2019b)。しかしながら、エンド-ガラクタナーゼを産生する種は限られていたり、産生されるエンド型酵素が側鎖を有するガラクトサンには分解活性を示さなかったりする (Kotake *et al.*, 2011)。したがって、自然界において主にガラクトサン分解の役割を担うのはエキソ型のガラクトサン分解酵素であると考えられる。また、ガラクタンの構造解析を行う上では、主鎖を構成する結合様式を含む様々な重合度のオリゴ糖が遊離するエンド型酵素よりも、主鎖の Gal 単位の単糖あるいはオリゴ糖が遊離するエキソ型酵素の方が有用である。そこで本研究ではエキソ型のガラクトサン分解酵素に着目することとした。

	Bacterial		Fungal	Plant
GH35	$\beta$ -1,2	$\beta$ -1,3	$\beta$ -1,3 $\beta$ -1,4	$\beta$ -1,3 $\beta$ -1,4 $\beta$ -1,6
	<i>C. japonicus</i> ; 4D1J	<i>S. pneumoniae</i> ; 4E8C	<i>T. reesei</i> ; 3OG2	<i>S. lycopersicum</i> ; 6IK6
GH43 sub24 ( $\beta$ -1,3)			-	-
	<i>C. thermocellum</i> ; 3VZF	<i>P. chrysosporium</i> ; 7BYV		Exo
GH53 ( $\beta$ -1,4)			-	Endo
	<i>B. licheniformis</i> ; 2J74	<i>A. aculeatus</i> ; 6Q3R		

**Fig. 1-7. Structural comparison of galactan degrading enzymes.**

Species and PDB IDs for each enzyme are shown below the structure. Their scientific names are *C. japonicus*, *Cellvibrio japonicus*; *S. pneumoniae*, *Streptococcus pneumoniae*; *T. reesei*, *Trichoderma reesei*; *S. lycopersicum*, *Solanum lycopersicum*; *C. thermocellum*, *Clostridium thermocellum*; *P. chrysosporium*, *Phanerochaete chrysosporium*; *B. licheniformis*, *Bacillus licheniformis*; *A. aculeatus*, *Aspergillus aculeatus*. 6IK6 and 7BYV are solved in this study.

**Table. 1-2. GH family, reaction characteristics, and origin species of galactan degrading enzymes.**

Family	Sub-family	Clan	EC number	Exo/Endo	Specificity	Organisms
2	-	A	3.2.1.23	Exo	$\beta$ -1,4	bacteria
5	16	A	3.2.1.-	Endo	$\beta$ -1,6	fungi
16	10	B	3.2.1.181	Endo	$\beta$ -1,3	bacteria, fungi
30	5	A	3.2.1.164	Endo	$\beta$ -1,6	bacteria, fungi
			3.2.1.233	Exo	$\beta$ -1,6	bacteria
35	-	A	3.2.1.23	Exo	$\beta$ -1,3/1,4/1,6	bacteria, fungi, plant, animal
42	-	A	3.2.1.23	Exo	$\beta$ -1,3/1,4	bacteria
43	24	F	3.2.1.145	Exo	$\beta$ -1,3	bacteria, fungi
53	-	A	3.2.1.89	Endo	$\beta$ -1,4	bacteria, fungi
147	-	A	3.2.1.23	Exo	$\beta$ -1,4	bacteria

### 1.3.1. GH35 $\beta$ -ガラクトシダーゼ (EC3.2.1.23)

GH35 に分類される $\beta$ -ガラクトシダーゼはガラクトタンまたはガラクトース誘導体の非還元末端側から Gal を遊離する酵素であり、細菌、真菌、植物、動物に広くみられ、現在までに数万の配列情報と 11 の立体構造が登録されている (<http://www.cazy.org/>)。各酵素の加水分解活性に関する研究は数多く行われており、ガラクトタンを分解する酵素だけでなくラクトースに作用する酵素も多く含まれると知られている (Wang *et al.*, 2009; Talens-Perales *et al.*, 2016; Liu *et al.*, 2017; Rico-Díaz *et al.*, 2017; Chanalia *et al.*, 2018)。また、糖転移活性を示す酵素の存在も報告されており、現在では糖転移活性を用いてガラクトオリゴ糖を生産する研究が数多く行われている (Rico-Díaz *et al.*, 2017; Chanalia *et al.*, 2018; Gao *et al.*, 2019)。さらに、相同性の高いホモログ同士は基質特異性が類似する (Kotake *et al.*, 2005; Sakamoto and

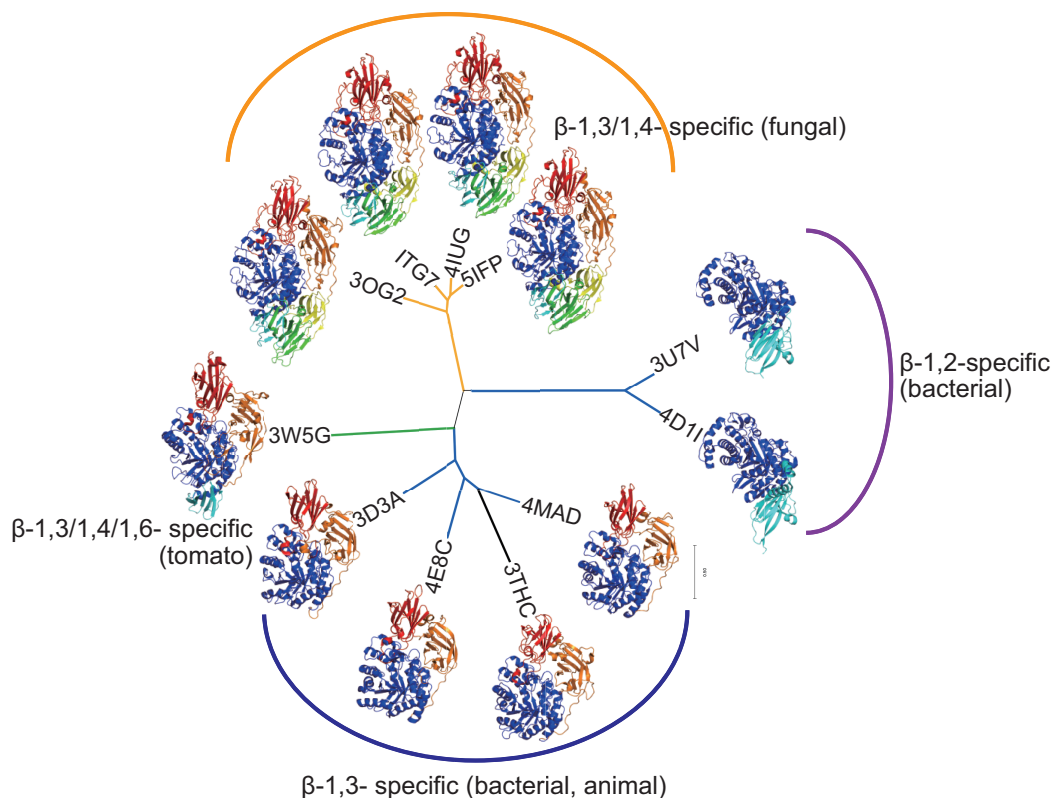
Ishimaru, 2013; Chandrasekar and van der Hoorn, 2016)。GH35 に属する $\beta$ -ガラクトシダーゼの触媒残基の general acid/base、nucleophile はいずれも Glu であり、このファミリーに保存されている。また、GH35  $\beta$ -ガラクトシダーゼは触媒部位を含む TIM バレルドメインと複数の $\beta$ -サンドウィッチドメインから構成され、各酵素を構成する $\beta$ -サンドウィッチドメインの数は生物種および基質特異性によって異なる (Fig. 1-9) (Cheng *et al.*, 2012; Eda *et al.*, 2016)。

すでに反応特性が報告されている酵素について系統解析を行うと、系統樹上で同じクラスターに属する酵素間の基質特異性は類似しているが、離れたクラスターに属する酵素間の基質特異性はそれぞれ異なる (Kotake *et al.*, 2005; Sakamoto and Ishimaru, 2013; Chandrasekar and van der Hoorn, 2016)。例えば植物ではシロイヌナズナやトマト (*Solanum lycopersicum*)、モモ (*Prunus persica*) では各17個、イネ (*Oryza sativa*) では15個、コケ (*Physcomitrella patens*) では6個と同一の植物が複数の $\beta$ -ガラクトシダーゼ遺伝子を発現すると確認されており、これらの遺伝子群はmulti gene familyと呼ばれる (Ahn *et al.*, 2007; Tanthanuch *et al.*, 2008; Chandrasekar and van der Hoorn, 2016; Guo *et al.*, 2018)。植物由来の $\beta$ -ガラクトシダーゼの酵素特性に関する研究は特にシロイヌナズナやアスパラガス、トウガラシ、リンゴ、パイナップル、イネ、ダイコン、ブドウ、モモ、トマトなどの食用の青果物をはじめとする様々な植物の産生する酵素に関して研究が進められており (Pressey, 1983; Carey *et al.*, 1995; O'Donoghue *et al.*, 1998; Barnavon *et al.*, 2000; Lazan *et al.*, 2004; Kotake *et al.*, 2005; Chantarangsee *et al.*, 2007; Ogasawara *et al.*, 2007; Ishimaru *et al.*, 2009; Eda *et al.*, 2014; Yang *et al.*, 2018)、各植物の各アイソザイムはそれぞれ異なる基質特異性を示す (Chandrasekar and van der Hoorn, 2016) ことから植物細胞壁中において異なる役割を担うと考えられる。

一方、細菌や真菌は各生物種にみられる GH35 に属する $\beta$ -ガラクトシダーゼ遺伝子の数は比較的少なく (Sakamoto and Ishimaru, 2013)、細菌では *Bacteroides thetaiotamicron* VPI-5482 株で 3 個 (<http://www.cazy.org/b134.html>)、*Bifidobacterium longum* subsp. *longum* JDM30 株で 1 個 (<http://www.cazy.org/b1244.html>)、*Xanthomonas* sp. で 5 個 (<http://www.cazy.org/b18731.html>) であり、真菌では *Aspergillus oryzae* RIB40 株、*Trichoderma reesei*、*Phanerochaete chrysosporium* でそれぞれ 7 個、1 個、3 個である (<https://mycocosm.jgi.doe.gov/mycocosm/home>) (Wymelenberg *et al.*, 2005, 2006; Martinez *et al.*, 2008;



Arnaud *et al.*, 2012; Grigoriev *et al.*, 2014; Nordberg *et al.*, 2014)。これらの真菌の GH35  $\beta$ -ガラクトシダーゼの酵素特性としては $\beta$ -1,3-結合特異的、 $\beta$ -1,4-結合特異的、 $\beta$ -1,3/1,4-結合特異的、 $\beta$ -1,2-結合特異的（キシログルカン側鎖の末端に結合した Gal）などが報告されており、酵素によって異なる（Sakamoto and Ishimaru, 2013）。また、ラクトースを特異的に分解する酵素や、主に糖タンパク質の N 結合型糖鎖やミルクオリゴ糖の非還元末端から Gal を遊離する酵素も多くみられる（Jeong *et al.*, 2009; Terra *et al.*, 2010; Cheng *et al.*, 2012; Hu *et al.*, 2014）。立体構造既知の 11 個の GH35  $\beta$ -ガラクトシダーゼのうち、細菌由来は 5 個、真菌由来は 4 個、植物由来は 1 個、動物由来は 1 個であり、限られた数の酵素の構造情報しか明らかになっておらず、特に酵素基質複合体の立体構造は *Aspergillus niger* の酵素でしか報告されていなかった（[http://www.cazy.org/GH35\\_structure.html](http://www.cazy.org/GH35_structure.html)）（Rico-Díaz *et al.*, 2017）。ゆえに特に植物や動物の $\beta$ -ガラクトシダーゼの立体構造や、機能と構造との相関などに関する知見はまだまだ不十分であった。



**Fig. 1-9. Phylogenetic tree of structure known GH35  $\beta$ -galactosidases.**

The letters in each branch indicate the PDB ID. The domains of each enzyme are colored, from the N-terminal side, with the catalytic domain (blue) and the  $\beta$ -sheet domain (cyan, green, yellow, red, and orange). The same color indicates the corresponding domain.

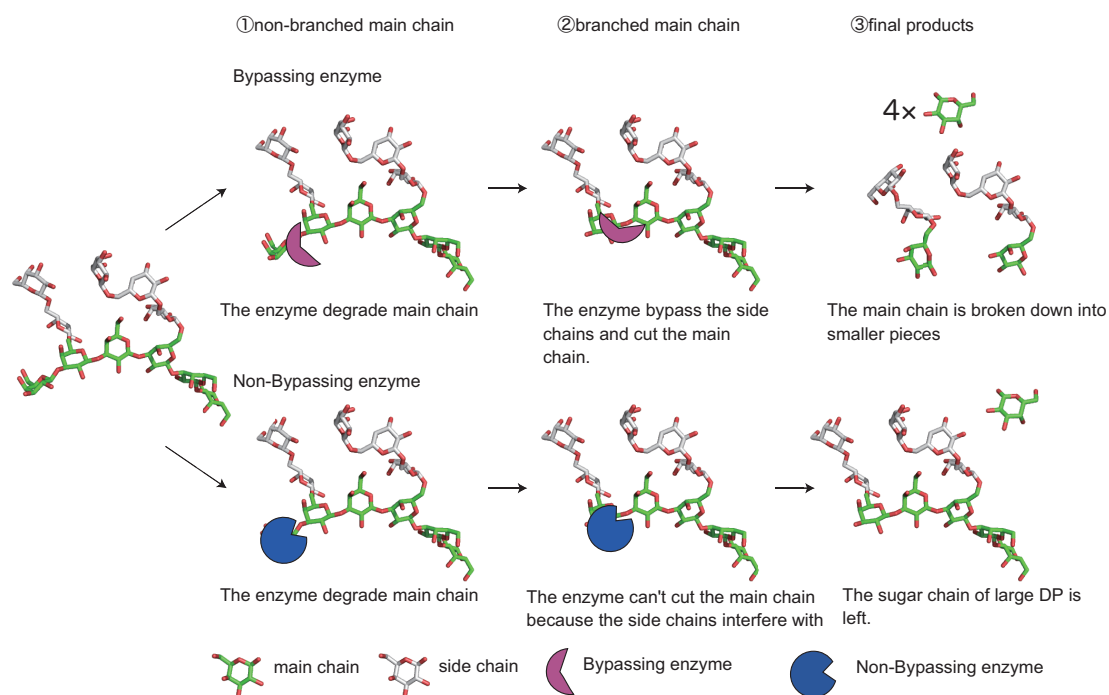
### 1.3.2. GH43 エキソ- $\beta$ -1,3-ガラクトナーゼ (EC 3.2.1.145)

GH43 は様々な基質特異性を示す Inverting 型の酵素が属する大きなファミリーであり、37 のサブファミリーに分類されている (Mewis *et al.*, 2016)。このうち、大半のサブファミリーは Glu を catalytic acid、Asp を catalytic base として用いる (Mewis *et al.*, 2016)。しかしながらエキソ- $\beta$ -1,3-ガラクトナーゼが属するサブファミリー24 (sub24) では他のサブファミリーに保存されている catalytic base の役割を担う Asp を欠損していることから、GH43\_sub24 の触媒残基については様々な議論がなされてきた (Jiang *et al.*, 2012; Mewis *et al.*, 2016; Cartmell *et al.*, 2018)。例えば *Clostridium thermocellum* の酵素 (Ct1,3Gal43A) では Glu112 が catalytic base であると推定された (Jiang *et al.*, 2012) が、*B. thetaiotamicron* の酵素 (BT3683) では Glu367 (Ct1,3Gal43A の Glu112 に相当する残基) は catalytic base ではなく非還元末端の Gal の 4 位ヒドロキシ基の認識に関与する残基であると結論づけられ、Gln577 が互変異性化してイミド酸型となって catalytic base の役割を担うと推定された (Cartmell *et al.*, 2018)。GH43\_sub24 と同様に GH45 の他のサブファミリーに保存された catalytic base の酸性残基を欠損した *P. chrysosporium* の GH45 エンドグルカナーゼ (PcCel45A; GH45\_sub C) では、中性子構造解析により互変異性化した Asn がイミド酸型となって catalytic base として機能すると明らかにされている (Nakamura *et al.*, 2015) ことから、BT3683 においても PcCel45A と同様に Gln が catalytic base となるか、または求核水を安定化することが可能であると推定されるためである。しかしながら、GH43\_sub24 では触媒部位に生産的に基質が結合した複合体構造が決定されていないため、この反応メカニズムは推定にすぎなかった。

植物も GH43 に分類されるエキソ- $\beta$ -1,3-ガラクトナーゼドメインを有するが、N 末端側に膜貫通ドメインを有するため、微生物の GH43 エキソ- $\beta$ -1,3-ガラクトナーゼとは機能が異なると考えられている (Ichinose *et al.*, 2008)。これまでに GH43 のエキソ- $\beta$ -1,3-ガラクトナーゼは *B. thetaiotamicron* VPI-5482 株、*B. longum* sub sp. *longum* JCM 1217 株、*C. thermocellum* ATCC 27405 株、*Streptomyces avermitillis* MA-4680 = NBRC 14893 株、*Streptomyces* sp. 19 株などの細菌由来の酵素や *Fusarium oxysporum* 12S 株、*Irpex lacteus* NBRC 5367 株、*P. chrysosporium* などの真菌由来の酵素において、いずれも $\beta$ -1,3-ガラクトタンに強い基質特異性を示し、非還元末端側から Gal を遊離すると報告されている (Tsumuraya *et al.*, 1990; Ichinose *et al.*, 2005, 2006a,b; Kotake *et al.*, 2009; Okawa *et al.*, 2013; Cartmell *et al.*,

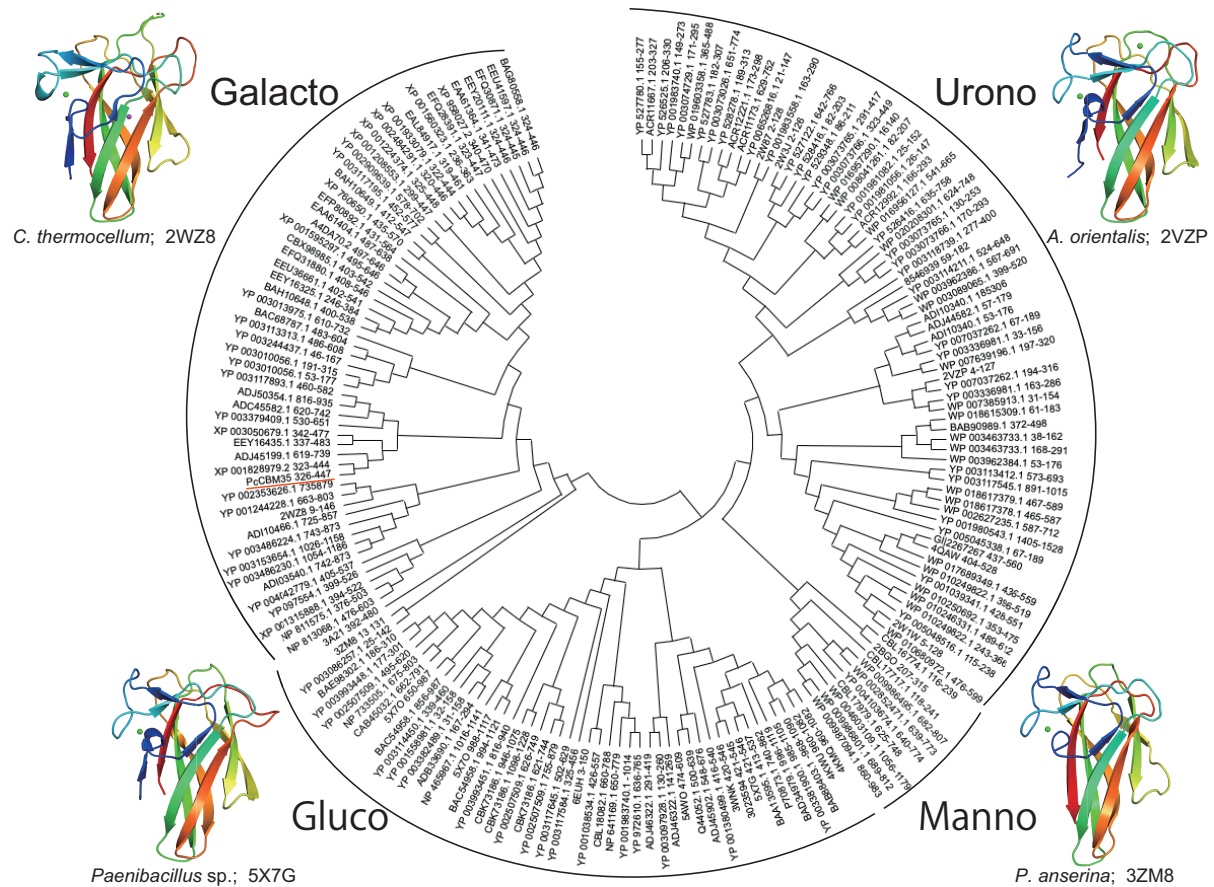
2018; Fujita *et al.*, 2019)。GH43\_sub24 の多くは $\beta$ -1,6-ガラクトサン側鎖が非還元末端の Gal に結合していても主鎖の  $\beta$ -1,3-ガラクトシル結合を分解できる“バイパス活性”と呼ばれる活性 (Fig. 1-10) を示すが、 $\beta$ -1,6-ガラクトサン側鎖が存在すると主鎖を分解できない酵素も存在する (Cartmell *et al.*, 2018)。また、AG-II の $\beta$ -1,6-ガラクトサン側鎖にはさらに Ara や Me-GlcA も結合しているが、これらの糖が結合している場合にはバイパス活性を示す酵素であっても直鎖の $\beta$ -1,3-ガラクトサンと比較すると主鎖の加水分解活性が顕著に低下する (Ichinose *et al.*, 2005, 2006a,b; Kotake *et al.*, 2009)。したがって、本ファミリーに属するガラクタナーゼのバイパス活性の有無と触媒部位の構造には相関があると考えられる。

GH43\_sub24 の多くは触媒ドメインに加え CBM13 または CBM35 の糖質結合モジュールを結合した 2 ドメインから構成される (Fujita *et al.*, 2014)。CBM13 を結合した酵素は *C. thermocellum* や *S. avermitilis* など主に細菌由来の酵素にみられるが (Ichinose *et al.*, 2006a,b)、CBM35 を結合した酵素は *P. chrysosporium* や *I. lacteus* などの真菌由来の酵素によくみられる (Ichinose *et al.*, 2005; Ishida *et al.*, 2009a; Kotake *et al.*, 2009)。CBM35 は 11 本の $\beta$ -シート構造によって構成される $\beta$ -サンドウィッチ構造をとる 140 アミノ酸残基程度の小さなドメインで、多くの場合は 1 つまたは 2 つの  $\text{Ca}^{2+}$  結合部位を有し、一部  $\text{Ca}^{2+}$  および  $\text{Mg}^{2+}$  の 2 つの金属イオンと結合するものもある (Montanier *et al.*, 2009; Correia *et al.*, 2010)。これまでに立体構造が報告された CBM35 はキシラン、グルカン、マンナンやガラクトグルコマンナンの $\alpha$ -Gal に結合性を示すものであり、系統樹上で同じクラスターに属する CBM の結合性は類似することが知られている (Correia *et al.*, 2010) (Fig. 1-11)。*P. chrysosporium* の GH43\_sub24 に属するエキソ- $\beta$ -1,3-ガラクタナーゼ (*Pc*1,3Gal43A) の CBM35 (*Pc*CBM35) はラクトースや $\beta$ -1,4-ガラクトサン、 $\beta$ -1,3-キシラン、ラミナリン ( $\beta$ -1,3-グルカン)、多様な側鎖が結合した AG-II には結合性を示さないが、Gal- $\beta$ -1,3-ラクトースや $\beta$ -1,3-Gal2-ラクトース、側鎖を除去した AG-II ( $\beta$ -1,3-ガラクトサン) には強い結合性を示したことから、*Pc*CBM35 は $\beta$ -1,3-ガラクトサンのうち少なくとも 2 つの連続した Gal 残基を認識すると推定されている (Ichinose *et al.*, 2005)。したがって *Pc*CBM35 は他の多糖類を認識する CBM35 とは異なるメカニズムでリガンドを認識すると考えられる。



**Fig. 1-10. Schematic diagram of AGP degradation.**

This figure schematically illustrates the process and end product of the enzymatic reaction between an enzyme that can bypass the side chain (upper) and an enzyme that cannot (lower). As shown in ①, both enzymes can degrade main chain (green). However, as shown in ②, when there is a branch, the bypassing enzyme (magenta) can successfully avoid the branch and degrade the main chain, while non-bypassing enzyme (blue) cannot degrade the main chain because it cannot avoid the branch well. Finally, as shown in ③, the bypassing enzyme degrade main chain into pieces, but non-bypassing enzyme degrade only at the non-reducing end and leave a large clump.



**Fig. 1-11. The phylogenetic tree of CBM35 and typical structures.**

Total 180 Sequences are extracted from GenBank, based on the conserved protein domain family cd04081, cd04082, and cd04083 and structure protein data bank (PDB). Sequences were aligned MUSCLE on MEGA X (Kumar *et al.*, 2018; Stecher *et al.*, 2020), and drawn the phylogenetic tree by using Neighbor-Joining method. The analyses were conducted in MEGA X. Each structure is shown in a gradient diagram with the N-terminal side in blue and the C-terminal side in red.

#### 1.4. 本研究の目的

バイオマスの大部分を占める細胞壁を効率的に利用するためにはセルロースやヘミセルロースだけでなくペクチンも分解する必要がある。また、ガラクトンは植物にとって生理学的に重要な多糖類であるためにガラクトンの役割を詳細に理解したり、バイオマスを有効活用するためにはガラクトンの構造や特性を理解することが重要である。しかしながら、ガラクトンの結合様式や側鎖の分岐の様式などの詳細な糖鎖構造やガラクトンに作用する酵素の機能に関する研究はまだ不十分であった。また一方で植物と微生物ではガラクトン分解酵素の局在が異なるため、それぞれ異なるメカニズムでガラクトンを分解していると考えられるが、まだまだ知見が少なかった。そこで本研究では自然界におけるガラクトン分解の主役であると考えられるエキソ型ガラクトン分解酵素に着目し、植物のガラクトン分解酵素が唯一属するファミリーである GH35 の $\beta$ -ガラクトシダーゼおよび糖鎖の構造解析に有用な担子菌の GH43 のエキソ- $\beta$ -1,3-ガラクタナーゼの 2 つの酵素の酵素特性と立体構造の相関を明らかにし、さらに生物種によるガラクトン分解戦略の差異を明らかにすることを目的とした。

第二章ではトマト果実にみられ、 $\beta$ -1,3-、 $\beta$ -1,4-、 $\beta$ -1,6-ガラクトシル結合に対して幅広く加水分解活性を示す GH35  $\beta$ -ガラクトシダーゼ (TBG4) と各結合様式のガラクトビオースとの酵素基質複合体の立体構造解析、ドッキングシミュレーションを行い、TBG4 の基質認識機構の解明を試みた。

第三章では AGP の $\beta$ -1,6-ガラクトン側鎖をバイパスして $\beta$ -1,3-ガラクトン主鎖を特異的に分解する担子菌 *P. chrysosporium* の GH43\_sub24 エキソ- $\beta$ -1,3-ガラクタナーゼ (*Pc1,3Gal43A*) と $\beta$ -1,3-ガラクトトリオースの酵素基質複合体の立体構造解析を行い、本酵素の側鎖バイパス機構と CBM におけるリガンド認識機構の解明を試みた。

第四章において第二章および第三章で得られた TBG4 と *Pc1,3Gal43A* に関する知見を比較検討し、これらのエキソ型ガラクトン分解酵素の酵素特性と立体構造との相関や植物と真菌のガラクトン分解戦略を考察した。以上の取り組みにより、ガラクトン分解に関与するエキソ型酵素の構造機能相関や生物種によるガラクトンの分解機構の差異を解明することを目指した。

## 第二章 トマト果実由来 GH35 $\beta$ -ガラクトシダーゼの 酵素基質複合体の構造解析

### 2.1. 緒言

#### 果実の軟化とガラクトサン分解酵素

果実や野菜の“硬さ”はそれらの価値に影響するだけでなく、収穫や輸送の方法、日持ちにも多大なる影響をおよぼすため、果実の硬度を維持することは非常に重要である。果実の細胞壁は主に一次細胞壁であり、構成要素であるペクチン、ヘミセルロース、セルロースの構造変化が果実の硬度に密接に関係している (Redgwell *et al.*, 1997)。この構造変化には細胞壁構成多糖類の可溶化や低分子化だけでなく果実硬度を低下させる再構成も含まれ、これらの変化は植物自身が産生する多様な CAZymes が関与して引き起こされる (Brummell, 2006; Gilbert, 2010)。特に成熟過程における細胞壁の中性糖含量の変化に着目してみると、トマト、キウイフルーツ (*Actinidia deliciosa*)、イチゴ (*Fragaria*  $\times$  *ananassa*)、ブラックベリー (*Rubus fruticosus*)、アボカド (*Persea americana*)、スイカ (*Citrullus lanatus*)、リンゴ (*Malus domestica*)、セイヨウナシ (*Pyrus communis*) の各成熟果では未熟果と比べ細胞壁中の Gal 含量の顕著な減少がみられたことから、果実の軟化には特に細胞壁中のガラクトシル含量の減少が関与すると考えられている (Gross and Wallner, 1979; Gross and Sams, 1984; Redgwell *et al.*, 1997)。植物が産生する細胞壁中のガラクトシル含量の減少に関与する酵素のうち、細胞壁中に分泌されるのは GH35 の  $\beta$ -ガラクトシダーゼのみである (Smith *et al.*, 1998; Sakamoto and Ishimaru, 2013) ため、植物細胞壁のガラクトース代謝を理解するには  $\beta$ -ガラクトシダーゼの機能を詳細に解析することが不可欠である。そこで本章ではナス科および果実のモデル植物であるトマト果実の軟化時に発現量が増加する  $\beta$ -ガラクトシダーゼに着目した。

#### トマト果実の軟化と $\beta$ -ガラクトシダーゼ

トマト果実の軟化に関する研究は数多く行われている。トマト果実細胞壁中のガラクトシル含量は開花後 10 日から Mature green と呼ばれる期間の間に半減し、Breaker から Red-Ripe の期間にさらに半減する (Kim *et al.*, 1991)。またトマト果実の成熟過程において細胞壁中のガラ

クトシル化合物含量の低下に伴いエキソ-ガラクトナーゼ活性が4~5倍に増加するが (Carey *et al.*, 1995)、トマト果実から精製した3つのβ-ガラクトシダーゼアイソザイム (β-ガラクトシダーゼ I、II、III) のうち、β-ガラクトシダーゼ II がβ-1,4-ガラクトタンに対する強い加水分解活性を示した (Pressey, 1983)。のちにトマト果実から合計17種類のβ-ガラクトシダーゼ (TBG) 遺伝子が単離され、そのうち TBG4 という酵素がβ-ガラクトシダーゼ II に相当すると示された (Smith *et al.*, 1998; Smith and Gross, 2000; Chandrasekar and van der Hoorn, 2016)。この TBG4 の mRNA 発現時期と果実の成熟時期が一致すること、成熟期のエチレン生成量の増加に伴い mRNA 発現量が上昇すること、アンチセンス鎖を導入し、TBG4 の発現を抑制した果実では硬度が約1.4倍維持されたことから、TBG4 はトマト果実の軟化に非常に重要であると考えられている (Smith and Gross, 2000; Smith *et al.*, 2002)。TBG4 を含むすべての TBG アイソザイムは GH35 に分類され、TBG1、TBG4、TBG5 は既に *Saccharomyces cerevisiae* や *Pichia pastoris* を用いた異宿主発現系により酵素特性が報告されている (Ishimaru *et al.*, 2009; Eda *et al.*, 2014)。これらの酵素の中でも TBG4 はβ-1,4-ガラクトシル結合だけでなくβ-1,3-およびβ-1,6-ガラクトシル結合も認識して加水分解する幅広い基質特異性を示すが、他の生物由来の GH35 β-ガラクトシダーゼの多くはβ-1,3-またはβ-1,4-ガラクトシル結合のみを比較的厳密に認識し加水分解するという特徴がある (Cheng *et al.*, 2012; Eda *et al.*, 2016) ため、TBG4 は他の酵素とは異なる基質認識機構を有すると考えられる。また、TBG4 の比活性は対β-1,4-ガラクトビオース (β-1,4-Gal2) を100%とすると対β-1,3-ガラクトビオース (β-1,3-Gal2) は約70%、対β-1,6-ガラクトビオース (β-1,6-Gal2) は約10%である (Eda *et al.*, 2016) ことから、各ガラクトシル結合の認識様式には違いがあると推察される。すでに TBG4 のアポ構造および反応産物である Gal との複合体の立体構造は決定されており、触媒残基の general acid/base と nucleophile はそれぞれ Glu181、Glu250 と同定されている (Eda *et al.*, 2016)。しかしながら、酵素基質複合体の立体構造は決定されておらず、基質認識機構は明らかになっていなかった。また、GH35 において天然の基質との複合体の構造は *A. niger* 由来の1種類の酵素のみでしか明らかにされておらず (Rico-Díaz *et al.*, 2017)、特に植物の酵素では複合体構造の報告がなく、その基質認識機構は不明であった。



## 本章における目的

本章では TBG4 とその変異体を用いて 3 つの結合様式のガラクトビオース ( $\beta$ -1,3-Gal2、 $\beta$ -1,4-Gal2、 $\beta$ -1,6-Gal2) との複合体の X 結晶構造解析、アンサンブルリファインメント、ドッキングシミュレーションを組み合わせて行うことにより、TBG4 の幅広い基質特異性の要因となる基質認識機構を解明することを目的とした。

## 2.2. 材料および方法

### 2.2.1. 酵素生産

TBG4 の野生型 (WT) および general acid/base の変異体 E181A の各酵素の生産は Eda *et al.* (2015) と同様の方法で行った。なお、第二章、第三章ともに試薬類は特に記載がない限り和光純薬 (大阪) から購入したものをを用いた。

### 2.2.2. 基質の調製

$\beta$ -1,4-Gal2 は Sigma Aldrich (ミズーリ州、USA) より購入したものをを用いた。 $\beta$ -1,3-Gal2、 $\beta$ -1,6-Gal2 は Kondo *et al.* (2020) と同様の方法で調製した。

### 2.2.3. X 線結晶構造解析実験および構造精密化

## 結晶化および回折データの収集

回折実験に用いる結晶の作出にあたり、WT の結晶は Eda *et al.* (2015) の条件を元に緩衝液の pH とポリエチレングリコール (PEG) 10000 の濃度を検討して結晶化条件を最適化した。16% (w/v) の PEG10000 (Hampton research、カリフォルニア州、USA) を含む 0.1 M HEPES (pH 7.3) 緩衝液 (Hampton research) を結晶化剤とし、0.9% (w/v) WT を結晶化剤と体積比 1:1 で混合し、シッティングドロップ蒸気拡散法で 4°C のインキュベーター中で静置して s 晶出した結晶を回折実験に供した。また、E181A の結晶化は WT と同じ結晶化剤を用い、1.55 % (w/v) E181A を結晶化剤と体積比 1:1 で混合し、シッティングドロップ蒸気拡散法により 4°C

のインキュベーター中で行った。

WTの結晶は9.4 mM  $\beta$ -1,4-Gal2 および 16% (w/v) PEG10000 を含む 0.1 M HEPES 緩衝液 (pH 7.3) に 4°C下で 2 時間ソーキングした後に 16% (w/v) PEG10000 および 25% (w/v) PEG400 (Hampton research、カリフォルニア州、USA) を含む 0.1 M HEPES (pH 7.5) 緩衝液から構成されるクライオプロテクトant溶液に数秒間浸漬した (WT\_Gal)。

E181A\_ $\beta$ -1,3-Gal2 の結晶は 0.146 mM  $\beta$ -1,3-Gal2 および 16% (w/v) PEG10000 を含む 0.1 M HEPES 緩衝液 (pH 7.3) 中で 4°Cのインキュベーター内に 1 ヶ月静置して晶出したものを用いた。E181A\_ $\beta$ -1,4-Gal2 は晶出した E181A 結晶を 10 mM  $\beta$ -1,4-Gal2 および 16% (w/v) PEG10000 を含む 0.1 M HEPES 緩衝液 (pH 7.3) に二晩ソーキングして複合体とし、Humid Air and Glue-coating (HAG) 法 (Baba *et al.*, 2013) により抗凍結処理を行った。HAG 法では 5% (v/v) エチレングリコールを含むポリビニルアルコール (PVA) 4500 を glue として用い、84% 相対湿度雰囲気下で X 線回折実験を行った。E181A\_ $\beta$ -1,6-Gal2 は E181A 結晶を 2.9 mM  $\beta$ -1,6-Gal2 および 16% (w/v) PEG10000 を含む 0.1 M HEPES 緩衝液 (pH 7.3) に一晩ソーキングして複合体とし、抗凍結処理は行わずにそのまま X 線を照射した。

X 線回折実験は大型放射光施設 SPring-8 の BL38B1 において 100 K の窒素気流下で行い、1.000000 Å の X 線を照射して回折データを収集した。収集した回折データは HKL-2000 スイートプログラムの DENZO と SCALPACK を用いてプロセッシングおよびスケーリング処理を行った (Otwinowski and Minor, 1997)。

### 構造の決定および精密化

各複合体の位相は CCP4 プログラムスイートの MOLREP プログラムを用い、WT と Gal の複合体 (PDB ID: 3W5G) をサーチモデルとして分子置換法 (Vagin and Teplyakov 2010) で決定した。手動モデル構築と分子リファインメントは Coot (ver. 0.8.9, University of Oxford, England)、CCP4 プログラムスイートの REFMAC5 (ver. 7.0,063, Science & Technology Facilities Council, England) (Vagin *et al.* 2004)、Phenix suite の phenix.refine (ver. 1.13-2998-000, Lawrence Berkeley National Laboratory, USA) (Adams *et al.*, 2010; Berkholtz *et al.*, 2011; Afonine *et al.*, 2012) を用いて R-factor 値が収束するまで繰り返し行った。精密化後のモデルは Phenix suite の ensemble\_refinement (Burnley *et al.*, 2012; Burnley and

Gros, 2013; Forneris *et al.*, 2014) によりアンサンブルリファインメントを行った。pTLS 値はいずれの構造においても 0.8% に設定した。なお、アンサンブルリファインメントとは、X 線構造の精密化と分子動力学を組み合わせることで回折データに適合したアンサンブルモデルを生成するプログラムで、異方性分布と非調和分布を考慮することができる方法である。これにより結晶中のタンパク質の“ゆらぎ”を可視化することが可能となる (Burnley *et al.*, 2012; Burnley and Gros, 2013; Forneris *et al.*, 2014)。モデルの立体構造や相互作用の模式図は LigPlot + (ver. 1.4.5) (Wallace *et al.*, 1995; Laskowski and Swindells, 2011) および PyMOL (ver. 2.2.3, Schrödinger, LLC) を用いて作成した。

## ドメインアノテーション

TBG4 の各ドメインは Pfam データベース (<https://pfam.xfam.org>) (El-Gebali *et al.*, 2019) に基づいてアノテーションを試みたが、アノテーションできないドメインもあった。そこで TBG4 の全長配列を用いて NCBI の conserved domain search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) でも相同性検索を行った。しかしながらすべてのドメインのアノテーションにはいたらなかったため、BioXGEM.3D-BLAST Protein Structure Search server (<https://3d-blast.life.nctu.edu.tw/dbsas.php>) (Yang and Tung, 2006; Tung *et al.*, 2007) を用い、TBG4 (PDB ID: 3W5F) の第 3 ドメイン (Asn416-Glu438 および Leu586-Arg724) と第 4 ドメイン (Glu439-Gly585) の構造をクエリとしてドメインアノテーションを行った。

### 2.2.4. ドッキングシミュレーション

基質との結合の様相をエネルギー的に比較するために Chimera プログラム (Pettersen *et al.*, 2004) にて AutoDock Vina を用いてドッキングシミュレーションを行った (Pettersen *et al.*, 2004; Trott and Olson, 2010)。本研究で決定した WT\_Gal 構造 (PDB ID: 6IK5) の chain A をレセプタータンパク質として設定し、SWEET2 (<http://www.glycosciences.de/modeling/sweet2/doc/index.php>) (Bohne *et al.*, 1998, 1999) を用いて構築した各結合様式の Gal2 のモデルをリガンドとした。レセプターサーチボリュームオプションは TBG4 の触媒部位に設定し、

結合モードの最大数を 10、検索の網羅性を 8 としてシミュレーションを行った。

## 2.3. 結果

### 2.3.1. ドメインアノテーション

TBG4 の全体構造はすでに決定されており、1 つの TIM バレルドメインと 3 つの  $\beta$ -シートドメインから構成されることが既に報告されていた (Eda *et al.*, 2016) が、各ドメインのアノテーションは行われていなかった。そこで本章ではまずドメインのアノテーションを行った。触媒ドメインである 8 つの  $\beta/\alpha$  repeats からなる TIM バレルドメインはすべての  $\beta$ -ガラクトシダーゼに保存されている (Kumar *et al.*, 2019)。しかしながら、TBG4 の TIM バレルドメイン (Ser24-Ala343) は 5 番目と 6 番目の  $\alpha$ -ヘリックスを欠損している (Fig. 2-1. AI) (Eda *et al.*, 2016)。第 1 の  $\beta$ -シートドメイン (Leu344-Val415) は GH35 の  $\beta$ -ガラクトシダーゼに保存されている GHD とよばれる  $\beta$ -サンドウィッチドメインであった (Fig. 2-1. AII)。残りの 2 つの  $\beta$ -シートドメインは  $\beta$ -Gal 結合レクチンであるガレクチンに相同性を有していた。第 2 の  $\beta$ -シートドメイン (Fig. 2-1. AIII) は 2 つの  $\beta$ -ストランドのループ領域を構成する Asn416-Glu438 と逆平行  $\beta$ -サンドウィッチ構造を構成する Leu586-Arg724 から成る。このドメインはフィブロネクチン III 型スーパーファミリーと相同性を示した (E-value 1e-07, SCOPe ID: 49265)。フィブロネクチン III 型ドメインは細菌、真菌、植物、動物に進化的に広く保存されているドメインであり、生物学的機能を必要な場所で発揮するためのスペーサーとして機能したり、細菌のセルラーゼでは基質との結合に作用する可能性も示唆されている (Campbell and Spitzfaden, 1994)。最後の第 3 の  $\beta$ -シートドメイン (Fig. 2-1. AIV) は Glu439-Gly585 によって構成されており、ガラクトース結合ドメイン様スーパーファミリー (galactose-binding domain-like superfamily; E-value: 2e-12, SCOPe ID:49785) と相同性を示した。このガラクトース結合ドメイン様スーパーファミリー構造は GH35 だけでなく他のファミリーの  $\beta$ -ガラクトシダーゼにも広くみられる。また、このドメインのうち Leu500-Val548 の領域は担子菌 *P. chrysosporium* 由来エキソ- $\beta$ -1,3-ガラクタナーゼ (Pc1,3Gal43A) の CBM35 ドメインと低い相同性を示した (E-value: 0.02)。

### 2.3.2. X線結晶構造解析

#### 全体構造

WT および E181A と各 Gal2 の複合体の X 線回折データは兵庫県の大型放射光施設 SPring-8 で収集し (Fig. 2-1)、4 つの酵素基質複合体の立体構造を決定した (Table 2-1、Fig. 2-2)。本章において決定した 4 つの複合体構造の主鎖構造には互いに大きな違いはみられず (Fig. 2-2. B)、最大の RMSD 値は  $0.364 \text{ \AA}$  (E181A $_{\beta-1,3}$ -Gal2 vs E181A $_{\beta-1,6}$ -Gal2, Table 2-2) であった。しかしながら、各結合様式の Gal2 の還元末端側の Gal はそれぞれ少しずつ異なる位置に結合していたため (Fig. 2-3. A-D)、触媒部位での酵素—基質の相互作用の違いが反応性の違いに影響することを明確に示した。

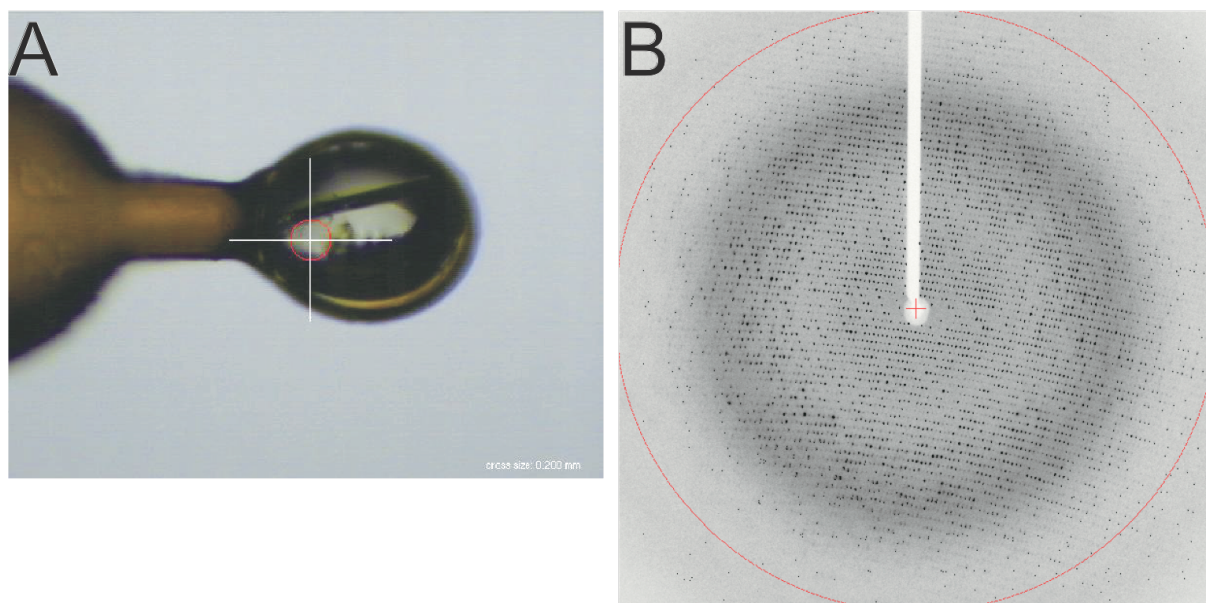
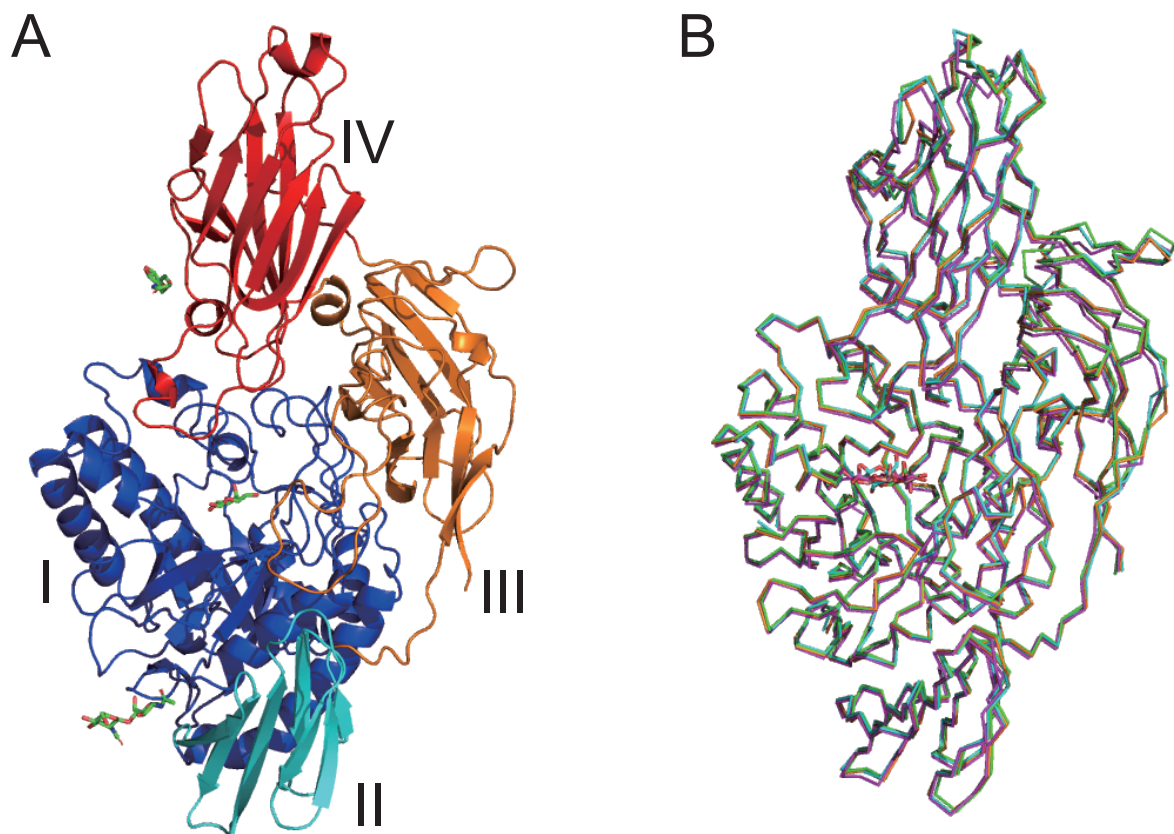


Fig. 2-1. The appearance of E181A $_{\beta-1,3}$ -Gal2 (A) and its diffraction (B).



**Fig. 2-2. Domain organization and comparison of Ca backbones.**

A: Overall structure of WT\_Gal. Bule (shown as I) is TIM barrel domain, cyan (II) is GHD domain, orange (III) is fibronectin III domain, and red (IV) shows a domain which is belonging to galactose binding domain like super family. B: Comparison of Ca backbones. WT, E181A\_β-1,3-Gal2, E181A\_β-1,4-Gal2, and E181A\_β-1,6-Gal2 are represented as green, orange, cyan, and magenta, respectively.

**Table 2-1. Summary of data collection and refinement statistics.**

	WT_Gal	E181A_β-1,3-Gal2	E181A_β-1,4-Gal2	E181A_β-1,6-Gal2
Wavelength	1.0000	1.0000	1.0000	1.0000
Resolution range	41.28 - 1.82 (1.885 - 1.82)	43.40 - 3.10 (3.21 - 3.10)	46.86 - 2.79 (2.89 - 2.79)	44.6 - 2.80 (2.90 - 2.80)
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell	92.136 95.335 158.327 90 90 90	93.608 97.954 161.252 90 90 90	94.032 110.73 162.141 90 90 90	92.94 95.95 158.959 90 90 90
Unique reflections	123244 (11826)	27468 (2691)	42501 (3935)	35651 (3480)
Multiplicity	7.1	7.3	7.0	6.6
Completeness (%)	98.34 (95.47)	99.88 (99.93)	99.17 (93.60)	99.76 (98.70)
Mean I/sigma(I)	23.4 (2.68)	9.8 (2.12)	11.0 (1.82)	11.5 (2.11)
Wilson B-factor (Å <sup>2</sup> )	23.09	45.66	59.09	50.12
R-merge	0.052 (0.615)	0.127 (0.615)	0.200 (0.800)	0.152 (0.755)
Reflections used in refinement	123232 (11825)	27468 (2691)	42460 (3935)	35638 (3479)
Reflections used for R-free	6167 (591)	1375 (124)	2148 (177)	1777 (170)
R-work (%)	17.31 (24.11)	19.22 (29.30)	19.62 (27.57)	17.78 (27.15)
R-free (%)	21.23 (29.15)	23.67 (32.04)	27.99 (36.98)	24.63 (33.01)
Number of non-hydrogen atoms	12055	11293	11285	11450
macromolecules	11062	11054	11054	11054
ligands	108	130	133	142
solvent	885	109	98	254
Protein residues	1410	1410	1410	1410
RMS (bonds)	0.011	0.012	0.01	0.008
RMS (angles)	1.71	1.69	1.34	1.57
Ramachandran favored (%)	96.23	92.18	92.67	93.88
Ramachandran allowed (%)	3.56	7.40	6.26	5.97
Ramachandran outliers (%)	0.21	0.43	1.07	0.14
Rotamer outliers (%)	0.51	0.34	0.25	1.10
Clash score	8.01	16.09	11.68	17.12
Average B-factor (Å <sup>2</sup> )	27.35	44.69	54.89	52.39
macromolecules (Å <sup>2</sup> )	26.61	44.66	54.66	52.13
ligands	46.38	64.87	76.46	90.17
solvent	34.39	23.77	51.00	42.45
PDB ID	6IK5	6IK7	6IK6	6IK8

**Table 2-2. Difference in the overall structures of each other (Ca, RMSD, Å).**

	WT_Gal	E181A_β-1,3- Gal2	E181A_β-1,4- Gal2	E181A_β-1,6- Gal2
WT_Gal		0.309	0.309	0.328
E181A_β-1,3- Gal2			0.329	0.364
E181A_β-1,4- Gal2				0.307
E181A_β-1,6- Gal2				

### 触媒部位での酵素と基質の相互作用

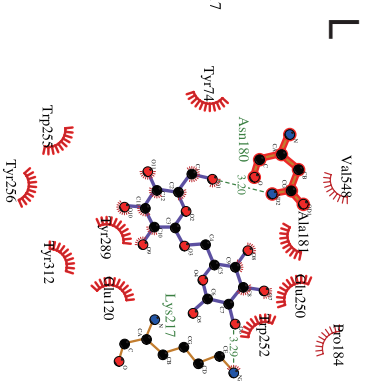
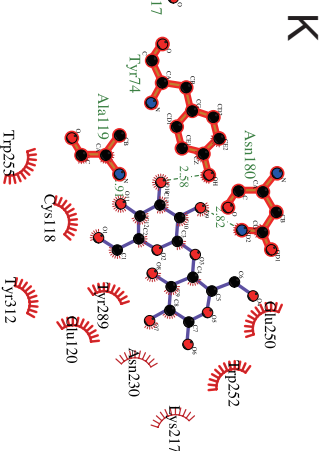
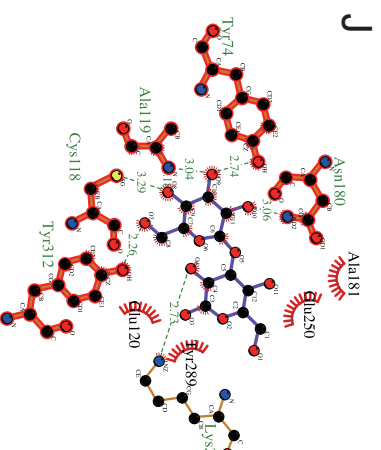
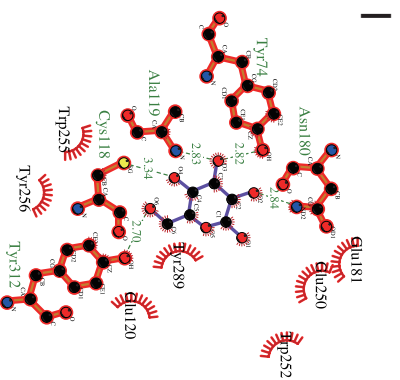
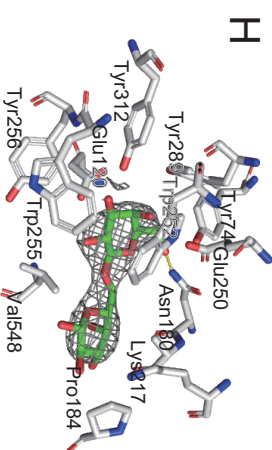
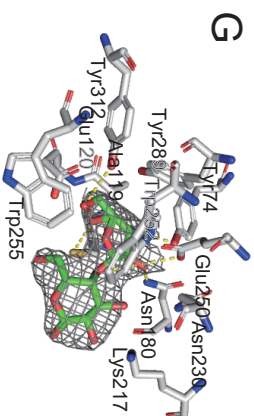
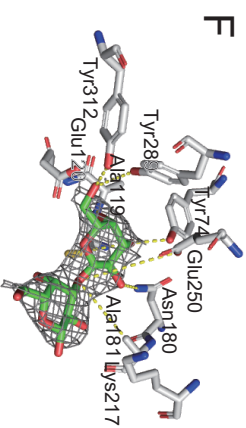
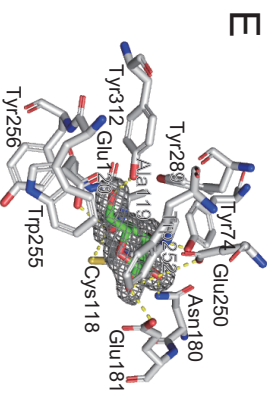
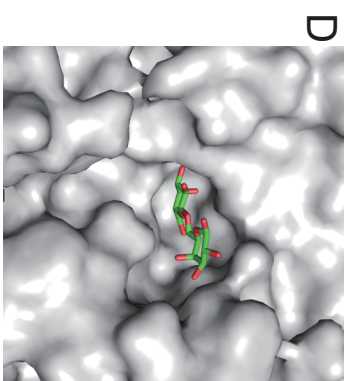
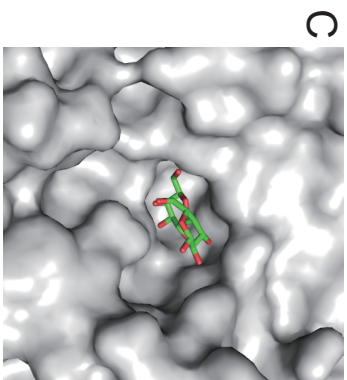
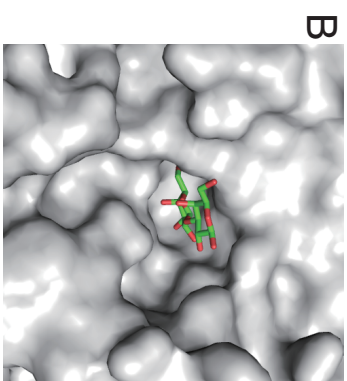
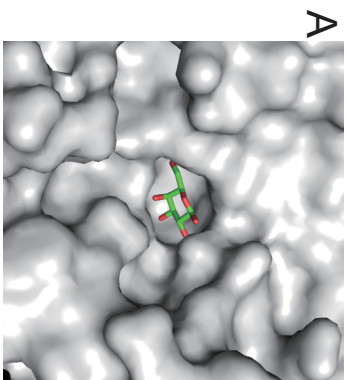
WT\_Gal では既報の構造 (PDB ID: 3W5G, 3.00 Å) と比べ分解能を改善し (PDB ID: 6IK5, 1.82 Å)、Gal が 3W5G と同じ位置に観測されたこと、全体構造に大きな違いがないことを確認した。3.00 Å 分解能では Gal や Gal と相互作用するアミノ酸残基側鎖の電子密度が電子雲としてぼんやりとしかみえていなかったのに対し、1.82 Å 分解能では Gal の電子密度が各ヒドロキシ基や六員環のリングまで見え、アミノ酸残基側鎖の配向も確認できたために、より正確に Gal 分子を構造モデルに配置することができた。また、Gal との相互作用に関与するアミノ酸残基や他の領域のアミノ酸残基に関してもより高精度で立体構造を決定することができた。本章において決定した構造より、Gal は 2 位ヒドロキシ基と Asn180-ND2、3 位ヒドロキシ基と Tyr74-OH および Ala119-N、4 位ヒドロキシ基と Cys118-SG、6 位メチロール基と Tyr312-OH の間でそれぞれ水素結合を介して相互作用しており (Fig. 2-3. E)、Glu120、Glu181、Glu250、Trp252、Trp255、Tyr256、Tyr289 が形成する疎水性表面に位置すると明らかになった (Fig. 2-3. E)。WT\_Gal の構造解析にあたっては WT 結晶をβ-1,4-Gal2 を含む溶液にソーキングしてβ-1,4-Gal2 との複合体の構造の取得を試みたが、還元末端側の Gal の電子密度は得られなかった。基質を含む溶液へのソーキング時間を 1 分間とし、窒素気流下で直ちに凍結した場合にも同様に単糖分の電子密度のみしか観測されなかったため、TBG4 は pH 7.3 の弱アルカリ性の溶液中においても即座にβ-1,4-Gal2 を分解したと考えられた。



E181A $_{\beta}$ -1,3-Gal2 のサブサイト-1 位では非還元末端の Gal (Gal $_{-1}$ ) 残基は WT\_Gal と同様に 2 位ヒドロキシ基と Asn180-ND2、3 位ヒドロキシ基と Tyr74-OH および Ala119-N、4 位ヒドロキシ基と Cys118-SG、6 位メチロール基と Tyr312-OH がそれぞれ水素結合を介して相互作用していた (Fig. 2-3. F)。サブサイト+1 位に位置する還元末端側の Gal (Gal $_{+1}$ ) は主に Glu181 (Ala)、Glu250、Tyr289 と疎水的に相互作用しており、水素結合を介した相互作用は 2 位ヒドロキシ基と Lys217 との間でのみみられた (Fig. 2-3. F)。

E181A $_{\beta}$ -1,4-Gal2 の Gal $_{-1}$  でみられた水素結合を介した相互作用の数は WT\_Gal と比較して少なかった。これは主に Tyr312 と 6 位ヒドロキシ基との水素結合がみられなかったことに起因する (Fig. 2-3. G)。一方で、Cys118、Glu120、Trp255、Tyr289、Tyr312 が Gal $_{-1}$  の結合サイトに疎水性表面を供していた。Gal $_{+1}$  と酵素との水素結合は観測されなかったが、Lys217、Asn230、Glu250、Trp252 は Gal $_{+1}$  と疎水的に相互作用していた (Fig. 2-3. G)。

これらの構造とは対照的に、E181A $_{\beta}$ -1,6-Gal2 の Gal $_{-1}$  は他の複合体の Gal $_{-1}$  とはピラノース環が上下反転し、2 位ヒドロキシ基と 6 位メチロール基が入れ替わった向きで触媒部位に結合していた (Fig. 2-3. H)。リガンドは触媒ポケットに固定されていたが、グリコシド結合が触媒ポケットの外側を向いていたため、今回観測された構造は触媒残基とグリコシド結合との間でプロトンの授受が正しくできずに加水分解反応を触媒できない非生産的な結合の様相であると推察される (Fig. 2-3. H)。このような非生産的な“反転”構造はドッキングシミュレーションでも確認された。



**Fig. 2-3. (previous page) Substrate-enzyme interactions of each complex.**

A-D: Surface structure and 2Fo-Fc density maps (1.0 sigma) of catalytic sites. E-H: The schematic diagrams of interaction mode between substrates and enzymes. The diagrams were drawn by using LigPlot + (ver. 1.4.5.)(Wallace *et al.*, 1995; Laskowski and Swindells, 2011). A and E, B and F, C and G, D and F show WT\_Gal, E181A\_β-1,3-Gal2, E181A\_β-1,4-Gal2, and E181A\_β-1,6-Gal2, respectively.

### 2.3.3. アンサンブルリファインメント

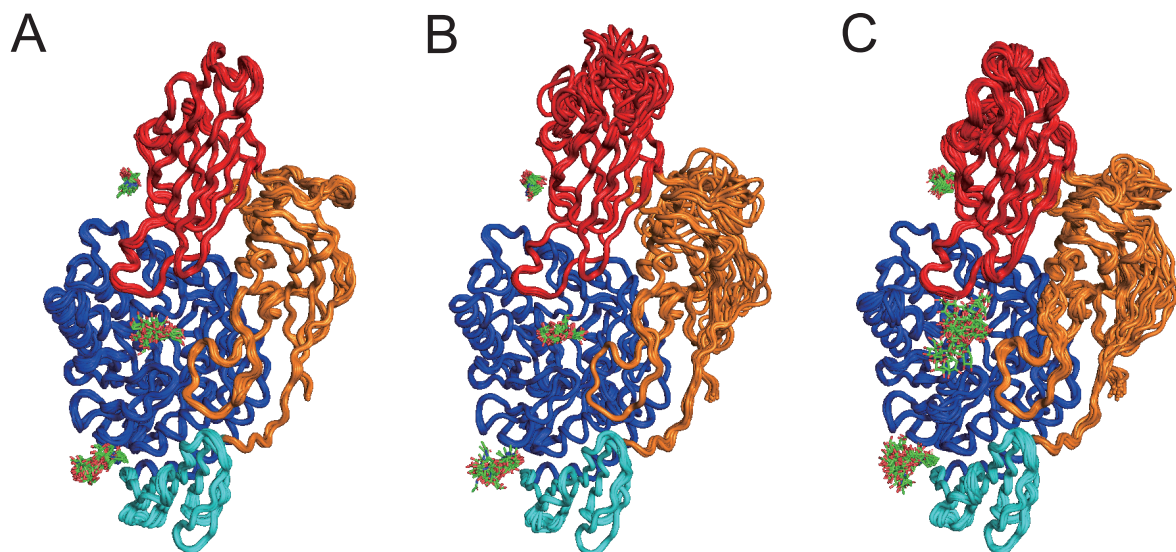
酵素は溶液中においては静止しているのではなく、常にゆらいている。したがって、酵素の反応機構や基質認識機構を明らかにするためには、立体構造を明らかにするだけでなく、酵素-基質複合体全体と触媒部位の構造ダイナミクスを解明する必要がある。そこで本研究においては、通常の X 線結晶構造解析における精密化と分子動力学 (MD) シミュレーションを組み合わせ、回折データにフィットしたアンサンブルモデルを作成し、局所的な分子の振動を可視化する手法であるアンサンブルリファインメントを行った。なお、ここでは可視化した分子の複数のコンフォメーションを“ゆらぎ”と表現する。アンサンブルリファインメントによって得られた構造モデルの統計値を Table. 2-3 に示した。主鎖構造のゆらぎの大きさの程度には 3 つの酵素基質複合体 (E181A\_β-1,3-Gal2、E181A\_β-1,4-Gal2、E181A\_β-1,6-Gal2) 間で差異が見られた (Fig. 2-4)。しかしながら、いずれも触媒部位から離れた領域のゆらぎの違いだったため、この差異は酵素活性に直接的に影響しないと推察された。触媒部位に着目してみると、分解能は同程度だったにも関わらず、β-1,6-Gal2 は他の Gal2 と比べて多くのコンフォメーション (アンサンブルモデル) を示した (Fig. 2-5)。これはβ-1,6-Gal2 はβ-1,3-Gal2 やβ-1,4-Gal2 と比較して酵素との相互作用が弱いためであるといえる。また、リガンドに着目すると X 線結晶構造解析の結果から予測された通り、Gal<sub>-1</sub> は Gal<sub>+1</sub> よりもゆらぎが小さかった (Fig. 2-5)。これは Gal<sub>-1</sub> が主に水素結合を介して相互作用しているのに対し、Gal<sub>+1</sub> は疎水性相互作用で比較的弱く相互作用していることを反映しているためであると考えられる。

**Table 2-3. Refinement statistics of ensemble refinement.**

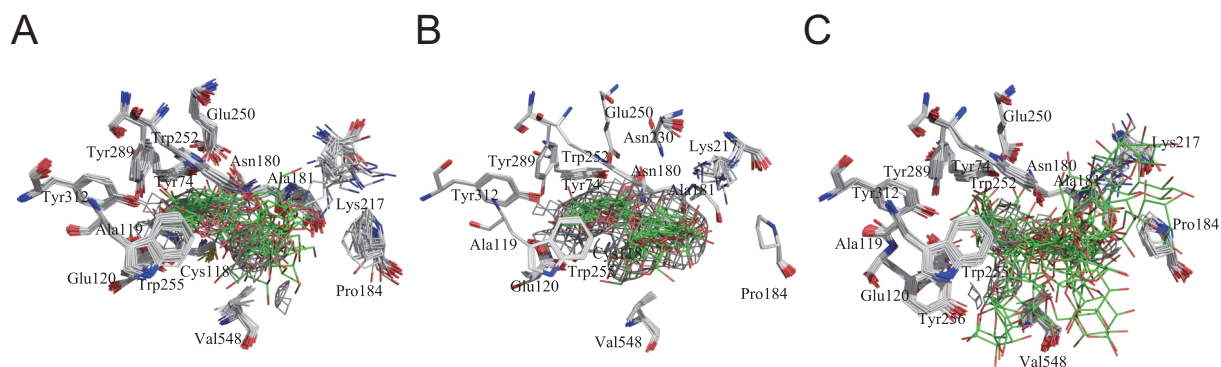
	E181A_β-1,3-Gal2	E181A_β-1,4-Gal2	E181A_β-1,6-Gal2
<b>Refinement parameters</b>			
Relaxation time (ps)	0.2	0.3	0.3
pTLS (%)	0.8	0.8	0.8
Conformers (#)	18	15	24
<b>Refinement and model statistics</b>			
Resolution range (Å)	43.40-3.10 (3.21-3.10)	48.58-2.79 (2.86-2.79)	44.60-2.80 (2.87-2.80)
R work (%)	18.71(29.77)	20.47(29.29)	17.09(26.04)
R free (%)	27.47(37.68)	27.26(33.59)	26.15(37.35)
ΔR work* (%)	0.51	-0.85	0.69
ΔR free** (%)	-3.8	0.73	-1.52
<b>Mean RMSD per structure</b>			
Bonds (Å)	0.007	0.008	0.008
Angles (°)	1.039	1.098	1.11
Dihedral (°)	16.1	15.72	16.81

\* ΔR work = [ensemble refinement's R work] – [X-Ray's R-work]

\*\* ΔR free = [ensemble refinement's R free] – [X-Ray's R-free]



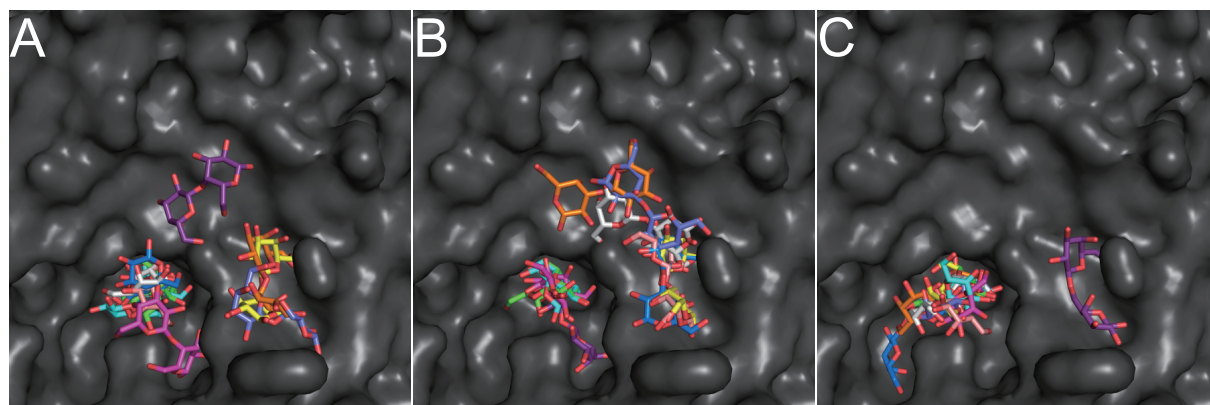
**Fig. 2-4. Overall structure of ensemble models obtained by ensemble refinement.** The colors show each domain as Fig. 2-1. The stick model show sugars, with each galactobiose seen in the center of the enzyme and *N*-glycans seen on the outside of the enzyme.



**Fig. 2-5. Ensemble models of catalytic center with 2Fo-Fc density maps (1.5 sigma).** Residues involved in the interaction with substrate in Fig. 2-2 are shown in this figure and galactobioses are represented as stick models. A, B, and C show E181A\_ $\beta$ -1,3-Gal2, E181A\_ $\beta$ -1,4-Gal2, and E181A\_ $\beta$ -1,6-Gal2, respectively.

#### 2.3.4. ドッキングシミュレーション

結合エネルギーを評価し、リガンドと TBG4 との相互作用様式を比較するために、AutoDock Vina を用いてドッキングシミュレーションを行った。各結合様式の Gal2 に関して得られたシミュレーションの結果を Table 2-4 に示した。各 Gal2 をリガンドとした際の最も低い結合エネルギーはそれぞれ-7.0、-7.1、-7.2 kcal/mol であり、同等の値であった。また、いずれの結合様式の Gal2 においても非還元末端側の Gal がサブサイト-1 位に位置し、グリコシド結合が触媒残基に面していた (Table 2-4、Fig. 2-6)。しかしながら、 $\beta$ -1,3-Gal2 の state 9 (-5.0 kcal/mol) や  $\beta$ -1,4-Gal2 の state 2 (-6.2 kcal/mol) はサブサイト+1 位および+2 位と相互作用していた。この結果は TBG4 において  $\beta$ -1,3-および  $\beta$ -1,4-ガラクトオリゴ糖と相互作用できる正のサブサイトが+1 位以外にも複数存在する可能性を示唆する。さらに触媒ポケットの周囲には  $\beta$ -1,3-および  $\beta$ -1,4-Gal2 を収容できる他の結合サイトの存在も示されたため、TBG4 が  $\beta$ -1,3-ガラクトンおよび  $\beta$ -1,4-ガラクトンから Gal を遊離するエキソ-ガラクタナーゼ活性を示すこと (Pressey, 1983; Carey *et al.*, 1995; Ishimaru *et al.*, 2009) を構造的な特徴からも説明できる。一方で  $\beta$ -1,6-Gal2 のドッキングシミュレーションの結果では、state 6 や 7 が X 線結晶構造解析で得られた結果と比較的類似した相互作用様式をとっており、ほとんどのモデル (例えば state 2 や 3) では生産的な結合の state 1 とは基質が触媒部位に反転して酵素に結合した反応できない非生産的な状態で相互作用している様子がみられた。



**Fig. 2-6. The interaction modes of simulated by AutoDock Vina.**

A, B, and C show E181A\_ $\beta$ -1,3-Gal2, E181A\_ $\beta$ -1,4-Gal2, and E181A\_ $\beta$ -1,6-Gal2, respectively. Each state is showed in different colors. The left binding site is catalytic site. The color shows green, state 1; cyan, state 2; magenta, state 3; yellow, state 4; wheat, state 5; white, state 6; slate, state 7; orange, state 8, purple, state 9; marine, state 10.

**Table 2-4. Simulated scores of AutoDock Vina.**

State	β-1,3-Gal2			β-1,4-Gal2			β-1,6-Gal2		
	binding energy (kcal/mol)	lower RMSD (kcal/mol)	upper RMSD (kcal/mol)	binding energy (kcal/mol)	lower RMSD (kcal/mol)	upper RMSD (kcal/mol)	binding energy (kcal/mol)	lower RMSD (kcal/mol)	upper RMSD (kcal/mol)
1	-7.0	0.000	0.000	-7.1	0.000	0.000	-7.2	0.000	0.000
2	-6.6	1.457	1.764	-6.2	1.935	4.049	-6.4	2.464	6.292
3	-5.9	1.724	3.385	-5.2	5.249	7.249	-6.2	2.605	6.559
4	-5.8	12.319	14.009	-5.2	11.914	13.264	-6.2	1.606	2.376
5	-5.4	11.757	13.308	-5.2	1.860	6.209	-6.0	2.358	6.926
6	-5.3	10.194	13.065	-5.2	2.192	6.048	-5.8	1.732	6.480
7	-5.1	12.050	14.361	-5.1	12.716	14.794	-5.6	1.637	6.121
8	-5.0	10.578	13.090	-5.0	12.599	13.712	-5.6	1.685	6.616
9	-5.0	3.898	6.714	-4.8	10.483	13.285	-5.5	13.310	15.231
10	-5.0	12.073	13.605	-4.8	2.728	4.803	-5.5	3.786	7.536

Small capitals indicate direction of ligand; a, the direction of ring O is opposite; b, reducing and non-reducing end is opposite; c, the direction of ring O and reducing and non-reducing end are opposite. RMSD is compared to the best model.

## 2.4. 考察

糖質加水分解酵素の中には基質特異性が広いものも数多くあり、その構造と機能の関係についてはこれまでも盛んに議論が行われてきた。しかしながら、植物が産生する GH35  $\beta$ -ガラクトシダーゼでは酵素の構造や特性に関する知見が少なく、果実の成熟期には明確に遺伝子発現量の増加がみられるものの、実際の機能については未解明な部分が非常に多い。本章では、TBG4 の立体構造がどのように基質特異性に影響するかを探るために、X 線結晶構造解析とアンサンブルリファインメント、ドッキングシミュレーションを組み合わせ、構造生物学的な観点から解析を行った。

### 2.4.1. TBG4 の基質認識機構

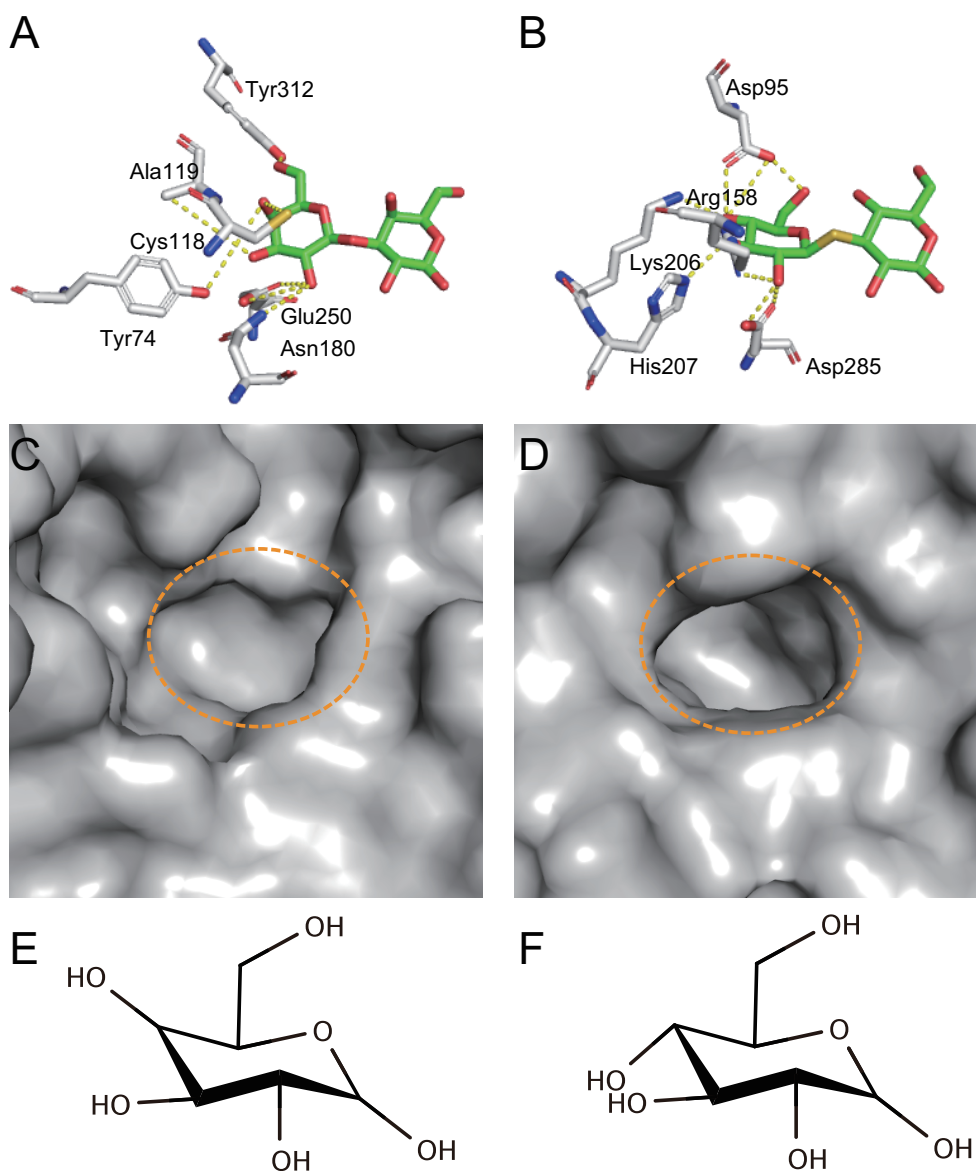
X 線結晶構造解析では、今回決定した 4 つの複合体の全体構造はリガンドとしたガラクトオリゴ糖の結合様式に関わらず既報のアポ構造 (PDB ID: 3W5F) や Gal との複合体構造 (PDB ID: 3W5G) と非常によく似ており、触媒部位への基質の結合の有無で大きな構造変化は起こらないことが示された (Fig. 2-2)。またさらに、基質認識に関与するアミノ酸残基は基質の結合様式によらずよく保存されていた (Fig. 2-3)。アンサンブルリファインメントを行い、局所的なゆらぎを可視化して検証したところ、全体構造のゆらぎの大きさには各モデルごとに違いがみられたが、大きくゆらいでいたのはループや N 末端、C 末端の領域であったため、酵素全体としては硬い構造であると推察された (Fig. 2-4)。一方で、触媒部位において基質との相互作用に関与する残基とリガンドのゆらぎに着目してみると、 $\beta$ -1,3-Gal2 および  $\beta$ -1,4-Gal2 は酵素と強く相互作用しており、小さなゆらぎしか示さなかったが、 $\beta$ -1,6-Gal2 は大きなゆらぎを示し、他のサイトに結合している様相もみられた (Fig. 2-5)。これは  $\beta$ -1,6-Gal2 を基質とする際には生産的な結合が生じにくいことを示唆する。したがって TBG4 の比活性は対  $\beta$ -1,3-Gal2 および  $\beta$ -1,4-Gal2 に比べ対  $\beta$ -1,6-Gal2 では著しく低い (約 1/10) (Eda *et al.*, 2016) という生化学実験の結果を説明できる。このアンサンブルリファインメントの結果とドッキングシミュレーションの結果より、TBG4 の  $\beta$ -1,6-Gal2 に対する反応性が他の結合様式の Gal2 と比べて低い理由は  $\beta$ -1,6-Gal2 が触媒部位に生産的に結合しにくいことに起因すると考えられる (Fig. 2-6、Table 2-4)。



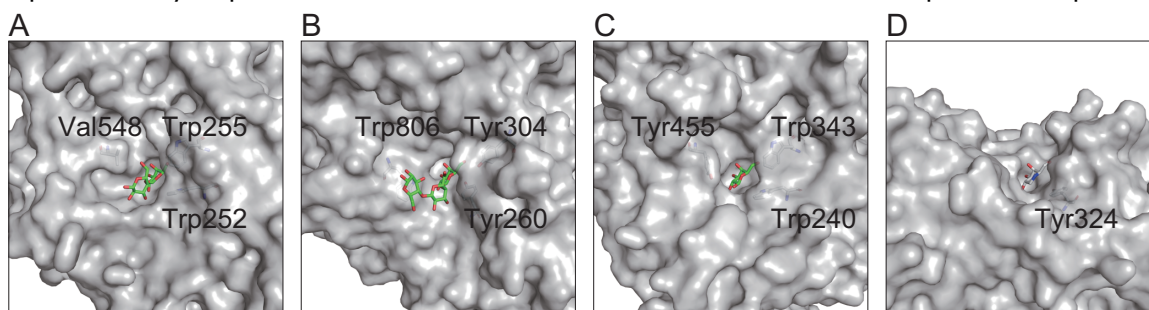
TBG4 のサブサイト-1 位では Gal は水素結合により厳密に認識されるのに対し、サブサイト +1 位では主に疎水性相互作用によって Gal の六員環が固定されるメカニズムをとっている。このような基質認識の特徴は他のエキソ型の糖質加水分解酵素と類似性がみられ、代表的な例として GH3 のオオムギ (*Hordeum vulgare* subsp. *vulgare*) 由来 $\beta$ -D-グルカングルコヒドロラーゼ (*HvExo1*) が挙げられる (Fig. 2-7)。*HvExo1* は $\beta$ -1,2-、 $\beta$ -1,3-、 $\beta$ -1,4-、 $\beta$ -1,6-グルコシド結合を分解して非還元末端側から Glc を遊離する (Hrmova *et al.*, 1996, 2002; Kotake *et al.*, 1997; Hrmova and Fincher, 1998)、基質特異性が広い酵素である。X 線結晶構造解析と触媒部位の分子モデリングの結果から、各基質の非還元末端側の Glc はサブサイト-1 位に位置しており、触媒ポケットの底に位置する 6 つのアミノ酸残基と水素結合を形成すると考えられている (Hrmova *et al.*, 2001, 2002)。TBG4 と *HvExo1* のサブサイト-1 位における相互作用様式を比較してみると、サブサイト-1 位を形成するポケットの表面構造は 2 つの酵素間で類似していたが、TBG4 の方が少し奥行きが広い形状となっていた (Fig. 2-7. A、B)。糖の認識に関わるアミノ酸残基の種類は異なるものの、いずれの酵素も各位のヒドロキシ基を認識しており、特に 2 位ヒドロキシ基と 3 位ヒドロキシ基の認識メカニズムは類似していたが、4 位ヒドロキシ基の認識は TBG4 と *HvExo1* で異なっており、それぞれヒドロキシ基の上側から押さえ込むような形で相互作用していた (Fig. 2-7. C、D)。Gal と Glc の構造の違いは 4 位ヒドロキシ基の向きである (Fig. 2-7. E、F) ため、この違いにより Gal と Glc を識別していると考えられる。一方、サブサイト+1 位でみられる相互作用の大部分は触媒ポケット入口に位置する疎水性残基との間の疎水性相互作用であり、水素結合による相互作用は比較的少なく (Hrmova *et al.*, 2001, 2002)、2 つの Trp 残基と疎水性相互作用により結合するため、サブサイト-1 位の Glc はサブサイト+1 位のように認識されず、むしろ疎水性残基間に挟まれて“固定”されるだけであると報告されている (Hrmova *et al.*, 2002)。TBG4 においてはサブサイト+1 位に位置する Gal 残基 ( $\text{Gal}_{+1}$ ) は 2 つの Trp 残基に挟まれてはいないが、主に疎水性相互作用によって酵素表面に位置しており、水素結合が少ないという点が *HvExo1* と同様であった。したがって、TBG4 においてもサブサイト+1 位の役割は基質の“認識”ではなく“固定”であると推察される。なお、多くの GH35  $\beta$ -ガラクトシダーゼではサブサイト+1 位に相当する位置のポケット入口を形成する壁の両側に合計 3 つの芳香族残基がみられ、これらの残基が基質認識に重要であると示されている (Fig. 2-8)(Maksimainen *et al.*, 2011; Cheng *et al.*, 2012; Eda *et al.*, 2016; Rico-Díaz *et al.*, 2017)。

しかしながら、TBG4をはじめとする植物の酵素はこれらの芳香族残基のうち2つしか有さず、もう1つの芳香族残基が疎水性残基に置換されているものが多い (Eda *et al.*, 2016)。このような酵素は TBG4 と同様にサブサイト+1 位では基質を“認識”というよりはむしろ“固定”していると考えられる。

TBG4 (PDB ID: 6IK7) と *HvExo1* (PDB ID: 1J8V) の触媒ポケットの構造を比較すると、TBG4 の触媒ポケットは入口が円形に広がった漏斗状の形状をとるのに対し *HvExo1* では入口が扁平な楕円形をしたノズル状の形状をとる (Fig. 2-9)。Glc と Gal の構造の違いは4位ヒドロキシ基の向きのみであり、 $\beta$ -1,4-Gal2 はアキシヤルな、 $\beta$ -1,4-グルコビオース (セロビオース) はエカトリアルな結合となるため、 $\beta$ -1,4-Gal2 の方がセロビオースよりも嵩高い構造となる。したがって、酵素の触媒ポケットの入口の大きさの構造的な違いはグルカンとガラクトランのコンフォメーションの違いに起因すると考えられる。両酵素ともにサブサイト+1 位における主な相互作用は疎水性相互作用であるが、活性部位の形状は主な基質の形状に最適化されているために異なるといえる。



**Fig. 2-7.** The difference of TBG4 (PDB ID: 6IK7, A, C) and *HvExo1* (PDB ID: 1J8V, B, D). A and B: The recognition modes of non-reducing terminal sugar. C and D: The shape of catalytic pocket bottom. E and F: The structural formula of  $\beta$ -Gal and  $\beta$ -Glc.



**Fig. 2-8.** Structural comparison of catalytic pocket entrance for TBG4 (A, PDB ID: 6IK6), *An* $\beta$ Gal (B, PDB ID: 5MGC, from *A. niger*), *BgaC* (C, PDB ID: 4E8C, from *Streptococcus pneumoniae*), and *CjBal35A* (D, PED ID: 4D1J, from *Cellovibrio japonicus*).

#### 2.4.2. 成熟中のトマト果実細胞壁における TBG4 の役割

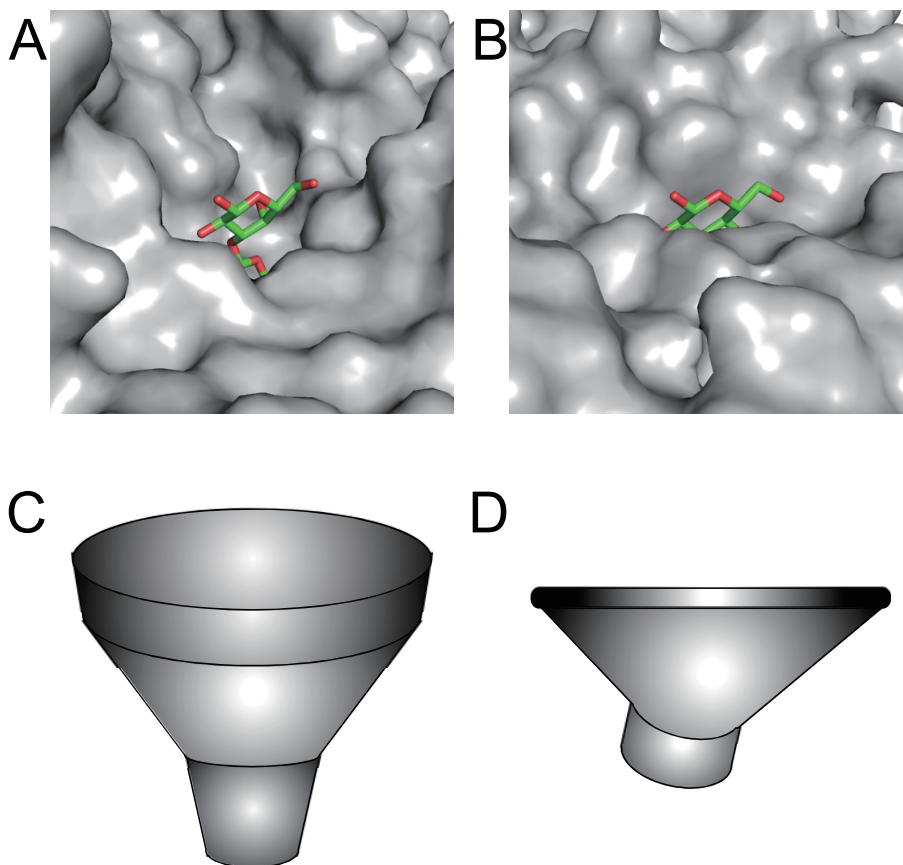
ドッキングシミュレーションの結果、 $\beta$ -1,3-Gal2 と  $\beta$ -1,4-Gal2 が結合する触媒部位の近傍にはより多くの正のサブサイトが存在する可能性が示された (Fig. 2-6、Table 2-4)。これは TBG4 が  $\beta$ -1,3-および  $\beta$ -1,4-ガラクトタンを分解する際に -1 位と +1 位の 2 つのサブサイトだけでなく、複数のサブサイトを用いて基質を認識する可能性を示唆した。担子菌 *P. chrysosporium* と放線菌 *Streptomyces* sp. SirexAA-E 由来の GH55 エキソ- $\beta$ -1,3-グルカナーゼ (ラミナリナーゼ、それぞれ PcLam55A、SacteLam55A) は  $\beta$ -1,3-グルカンの非還元末端から Glc を遊離する酵素であり、酵素表面には少なくとも 6 つのサブサイト (-1~+5) が存在する (Ishida *et al.*, 2009b; Bianchetti *et al.*, 2015)。PcLam55A では重合度 2 以上のオリゴ糖を分解可能であるが、より長鎖のオリゴ糖であるほど高い分解活性を示す (Ishida *et al.*, 2009b)。TBG4 は pNP- $\beta$ -Gal や ガラクトオリゴ糖だけでなくガラクトタンも分解する  $\beta$ -ガラクトシダーゼであると報告されている (Carey *et al.*, 1995; Carrington and Pressey, 1996; Ishimaru *et al.*, 2009) ため、TBG4 も複数の正のサブサイトをもつ GH55 エキソ- $\beta$ -1,3-グルカナーゼと同様に複数の正のサブサイトを有し、長鎖のガラクトタンに対してより高い加水分解活性を示す可能性がある。

GH35  $\beta$ -ガラクトシダーゼのうち、*A. niger* 由来の GH35  $\beta$ -ガラクトシダーゼである An $\beta$ Gal では、-1 位から +2 位の各サブサイトでの基質との相互作用様式が報告されている (Rico-Díaz *et al.*, 2017)。この酵素と比較すると TBG4 において  $\beta$ -1,4-ガラクトトリオースが相互作用する際には Fig. 2-10. A のオレンジ色の丸印の位置がサブサイト +2 位に相当すると推定される (Fig. 2-10)。An $\beta$ Gal ではサブサイト +2 位となりうる箇所がもう 1 つあると考えられていること (Rico-Díaz *et al.*, 2017)、ガラクトタンのコンフォメーションは結合様式 ( $\beta$ -1,3-、 $\beta$ -1,4-、 $\beta$ -1,6-結合) によって異なることから、TBG4 も複数のサブサイト +2 位を有する可能性がある。さらに TBG4 の触媒ポケットは入口が広く開いているため、分岐や側鎖をもつ多糖類が基質となった場合であっても加水分解活性を示す可能性がある。つまり TBG4 はトマト果実の成熟過程において AG-I だけでなく、RG-I の AG-II 側鎖や AGP の糖鎖領域も分解すると考えられる。

## 2.5. 本章の結論

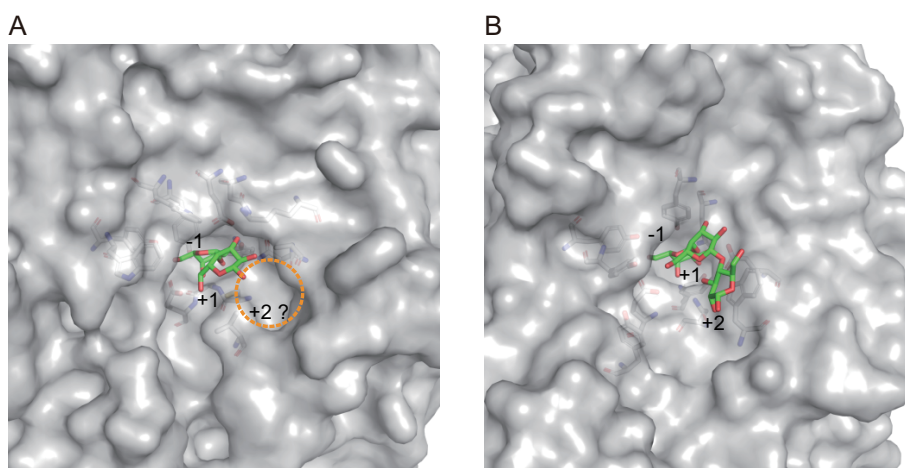
本章では $\beta$ -1,3、 $\beta$ -1,4、 $\beta$ -1,6-ガラクトンに対して加水分解活性を示す TBG4 の基質認識機構の解明を目指し、Gal および各結合様式の Gal2 との複合体の立体構造を決定した。その結果、TBG4 はサブサイト-1 位では強く、サブサイト+1 位では比較的弱く各 Gal 残基と相互作用し、触媒ポケット入口が広く開いた構造をとるために幅広く基質を認識して加水分解できると明らかにした。また、各結合様式の Gal2 への反応性の高さは触媒部位への基質の生産的な結合のしやすさと相関があることを明らかにした。

細胞壁の構造や組成は同じ植物の同じ器官であっても均質ではないため、トマト果実の細胞壁中のガラクトンの状態を完全に再現することは困難である。また、ガラクトオリゴ糖のなかでも特に重合度の大きいガラクトオリゴ糖は現在は市販されておらず、自前で重合度毎に分画し、高純度のオリゴ糖を調製することは簡単ではないため、TBG4 をはじめとする TBG アイソザイムのトマト果実細胞壁中での実際の反応を解析することは非常にハードルが高い。またさらに、真核生物のタンパク質を異宿主発現することが比較的得意な *P. pastoris* を用いても依然として植物のタンパク質の異宿主発現は難しい (Matsuyama *et al.*, 2020)。したがって、トマト果実の成熟メカニズムの完全な解明に向けてはいまだ多くの障壁が残るが、本章で得られた知見を足掛かりに 1 つずつ各アイソザイムの機能を解明することで、トマト果実の成熟メカニズムの中でも特に軟化時におけるガラクトンの役割やガラクトンの分解メカニズムを解明できると期待できる。また、トマトはモデル植物であるため、トマト果実の成熟機構を解明することは他のナス科の植物や果菜類、果実類の $\beta$ -ガラクトシダーゼの機能や成熟機構の解明につながると期待される。



**Fig. 2-9. Comparison of surface structure of catalytic pocket and the appearance of "funnel" and "wide nozzle".**

The upper structures are E181A\_β-1,3-Gal2 (A) and *HvExo1* (B, PDB ID: 1J8V). Stick models indicate β-1,3-Gal2 and β-1,3-laminaribiose (Glc-β-1,3-Glc), respectively. The lower figures are appearance models of "funnel" (C) and "wide nozzle" (D). Comparing the top and bottom diagrams, it is possible to see the similarities in their shapes.



**Fig. 2-10. Subsite comparison between TBG4 (A, PDB ID: 6IK6) and *AnβGal* (B, PDB ID: 5MGC). The dashed orange circle indicates estimated subsite +2 of TBG4.**

## 第三章 担子菌由来 GH43 エキソ- $\beta$ -1,3-ガラクトナーゼの 酵素基質複合体の構造解析

### 3.1. 緒言

#### アラビノガラクトタンプロテイン

アラビノガラクトタンプロテイン (AGP) は高等植物の細胞膜、細胞壁、細胞間層に特徴的に局在するプロテオグリカンの 1 つであり (Tsumuraya *et al.*, 1984)、植物の成長および発達の過程で生理学的に重要な役割を担う (Majewska-Sawka and Nothnagel, 2000)。AGP の糖鎖領域は主に $\beta$ -1,3-ガラクトタン主鎖と $\beta$ -1,6-ガラクトタン側鎖によって構成され、さらに Ara、Fuc、GlcA などが側鎖として結合した構造をとる AG-II 糖鎖である (Tsumuraya *et al.*, 1984; Majewska-Sawka and Nothnagel, 2000)。主鎖や側鎖の長さや構成糖、主鎖に対する側鎖の結合頻度は植物種や器官、発達段階によって異なるため、AGP の糖鎖領域の構造の全容は未解明である (Ellis *et al.*, 2010)。反応特性既知の酵素を用いて多糖類を分解することは、その構造や植物における機能を理解するために非常に強力な手法である。エキソ- $\beta$ -1,3-ガラクトナーゼ (EC 3.2.1.145) は $\beta$ -1,3-ガラクトタンの $\beta$ -1,3-ガラクトシル結合を特異的に認識して分解し、非還元末端側から Gal を遊離する酵素である。この酵素は AGP の糖鎖領域を分解する際に Gal とともに側鎖由来の $\beta$ -1,6-ガラクトオリゴ糖を遊離する特性を示す (Tsumuraya *et al.*, 1990; Pellerin and Brillouet, 1994) ため、AGP の糖鎖の側鎖の構造解析に用いる上では非常に協力的なツールとなる酵素である。

#### *P. chrysosporium* のエキソ- $\beta$ -1,3-ガラクトナーゼ

担子菌 *P. chrysosporium* が産生するエキソ- $\beta$ -1,3-ガラクトナーゼ (*Pc1,3Gal43A*; GenBank accession No. BAD98241) (Ichinose *et al.*, 2005) は CAZy データベースで GH43\_sub24 に分類される触媒ドメインと CBM ファミリー35 に分類される糖質結合ドメインの 2 ドメインから構成される (Ichinose *et al.*, 2005; Ishida *et al.*, 2009a; Mewis *et al.*, 2016)。*Pc1,3Gal43A* の酵素特性および CBM ドメインの結合特性に関する研究はすでにメタノール産性酵母 *P. pastoris* を用いて産生した組換え酵素を用いて行われており、GH43\_sub24

ではエキソ- $\beta$ -1,3-ガラクトナーゼ活性を示す酵素として初めて、CBM ドメインでは $\beta$ -1,3-ガラクトタンに結合性を示す CBM35 として初めてその特徴が報告されている (Ichinose *et al.*, 2005)。Pc1,3Gal43A は $\beta$ -1,3-ガラクトオリゴ糖や長鎖の $\beta$ -1,3-ガラクトタン、AG-II の主鎖の $\beta$ -1,3-結合のみを特異的に認識し、最終産物として Gal と $\beta$ -1,6-ガラクトオリゴ糖を遊離するが、 $\beta$ -1,4 および $\beta$ -1,6-ガラクトタンは分解しない (Ichinose *et al.*, 2005)。したがって、Pc1,3Gal43A は $\beta$ -1,3-結合の Gal を特異的に認識するだけでなく、 $\beta$ -1,6-ガラクトタン側鎖をバイパスして主鎖の $\beta$ -1,3-結合を分解できる酵素であると考えられている。しかしながら本酵素の立体構造は未解明であり、 $\beta$ -1,3-ガラクトタンの認識メカニズムや側鎖をバイパスすることが可能となる要因は明らかになっていない。

## 本章における目的

Pc1,3Gal43A のアポ構造およびホロ構造の構造解析を行い、Pc1,3Gal43A の加水分解機構および側鎖バイパス機構、CBM35 のリガンド認識機構を解明することを目的とした。

### 3.2. 材料および方法

#### 3.2.1. 酵素生産

Pc1,3Gal43A の 6 つの変異体の発現ベクターは pPICZαA-Pc1,3Gal43A を鋳型とし、Table 3-1 に示すプライマーを用いて PrimeSTAR MAX (Takara、東京) でインバース PCR を行い、点変異を導入して作製した。作製した発現ベクターの塩基配列は DNA シーケンスを行って確認した。結晶化には Ichinose *et al.* (2005) および Ishida *et al.* (2009a) と同様に Pc1,3Gal43A 遺伝子を組み込んだ *P. pastoris* KM71H 株を大量培養して培養上清を回収し、二段階のカラムクロマトグラフィーにより精製した酵素を用いた。酵素活性測定には 500 mL 三角フラスコを用いて目的遺伝子を導入した *P. pastoris* を培養し、培養上清を 1 M 硫酸アンモニウムを含む 20 mM 酢酸ナトリウム緩衝液 (pH 4.0) で平衡化した SkillPak TOYOPEARL Phenyl-650M (c.v.=5ml) に供し、0.7 M 硫酸アンモニウムを含む 20 mM 酢酸ナトリウム緩衝液 (pH 4.0) で溶出する疎水性クロマトグラフィーによる一段精製を行った酵素を用いた。なお、精製度は SDS-PAGE を行って確認した。



**Table 3-1. Mutagenesis primer sequences.**

Primer name		Primer Sequence
E102Q	Fw	GTTGTCC <b>CA</b> ACGGCCCAAGGTCATCTTC
	Rv	GGGCC <b>GTT</b> GGACAACATTCGATGTCTGA
E102A	Fw	GTTGTCC <b>GCT</b> CGGCCCAAGGTCATCTTC
	Rv	GGGCC <b>AGCG</b> ACAACATTCGATGTCTGA
E208Q	Fw	ACCCTG <b>CA</b> AGCACCCGGTATCGTGAAG
	Rv	GGGTGC <b>TTG</b> CAGGGTCGCGCCGTTTCAT
E208A	Fw	ACCCTG <b>GCT</b> GCACCCGGTATCGTGAAG
	Rv	GGGTGC <b>AGCC</b> CAGGGTCGCGCCGTTTCAT
Q263E	Fw	TACTCG <b>GAG</b> AACGCGTTTCGACCTGCCG
	Rv	CGCGTT <b>CTCC</b> GAGTACCACGTCCGCGT
Q263A	Fw	TACTCG <b>GCT</b> AACGCGTTTCGACCTGCCG
	Rv	CGCGTT <b>AGCC</b> GAGTACCACGTCCGCGT

The mutation sites are shown in italics.

### 3.2.2. 酵素活性の解析

*Pc1,3Gal43A* WT および各変異体の $\beta$ -1,3-Gal2 および $\beta$ -1,3-ガラクトトリオース ( $\beta$ -1,3-Gal3)に対する反応性を評価するために、100  $\mu$ Lの反応溶液中に終濃度 20 nMの各酵素と 0.263 mM Gal2 または 0.266 mM Gal3、20 mM 酢酸ナトリウム緩衝液 (pH 5.0) を添加して 30°C で 30 分間インキュベートし、95°Cで 5 分間加熱して反応を停止した。反応液の上清は 100% (v/v) アセトニトリルで 4 倍に希釈し、0.22  $\mu$ m PES フィルターで濾過して不溶物を除去した。濾過後の溶液をコロナ帯電エアロゾル検出器 (ESA Biosciences、現 Thermo Fisher Scientific Corporation、マサチューセッツ州、USA) を用いた高速液体クロマトグラフィー (HPLC ; LC-2000 シリーズ ; 日本分光、東京) に供し、溶出ピークのピーク面積に基づいて遊離した Gal を定量した。分離カラムは Shodex Asahipak NH2P-50 4E (昭和電工、東京) を用い、カラム温度 40 °C、流速 1 ml/ min において 75% (v/v) アセトニトリルで溶出した。酵素活性の 1 unit は実験条件下で 1 分間に 1  $\mu$ mol の Gal を遊離する酵素量と定義した。

### 3.2.3. X線結晶構造解析実験および構造精密化

酵素活性の測定および構造解析に用いた Gal2 および Gal3 は Ichinose *et al.* (2005) と同様の方法で調製したものを埼玉大学理工学研究科円谷陽一名誉教授および小竹敬久教授に譲渡していただいて用いた。精製後の酵素は Abs 280 値が 10~15 となるように調製し、結晶化セットアップを行った。野生型データ収集に使用した平板状の結晶は、2.1 M 硫酸アンモニウム、0.1 M クエン酸ナトリウム緩衝液 pH 5.5 の結晶化剤組成の条件から得た (WT)。他の野生型結晶は、16% (w/v) PEG10000、0.1 M 硫酸アンモニウム、0.1 M ビストリス緩衝液 pH 5.5、および 5.0% (v/v) グリセロールの条件で晶出した。セレノメチオニン (SeMet) 標識結晶は 16% (w/v) PEG10000、95 mM 硫酸アンモニウム、95 mM ビストリス緩衝液 pH 5.0、4.8% (v/v) グリセロールにおいて結晶化を行い、同一条件から薄板状結晶 (空間群  $P2_1$ ) と棒状結晶 (空間群  $P2_12_12_1$ ) の2種類の結晶を得た。E208Q 変異体は、16% (w/v) PEG10000、95 mM 硫酸アンモニウム、95 mM ビストリス緩衝液 pH 5.5、4.8% (v/v) グリセロールの条件で 10 mM Gal3 を用いて共結晶化させたところ、薄板結晶が得られた (E208Q\_Gal3)。E208A 変異体は、0.2 M 硝酸カリウム、15% (w/v) PEG6000、20 mM クエン酸ナトリウム緩衝液 pH 4.5、5% (v/v) グリセロール、10 mM Gal3 の条件で共結晶化し、二錐体結晶を得た (E208A\_Gal3)。

$Pc1,3Gal43A$  結晶の回折実験は、茨城県の高エネルギー加速器研究機構のフォトンファクトリー (PF) またはフォトンファクトリーアドバンスドリング (PF-AR) のビームラインで行った (Table 3-2)。回折データは CCD 検出器 (Area Detector Systems Corp., Poway, カリフォルニア州, USA) を用いて収集した。また、野生型酵素と Gal3 との複合体のデータ収集のために、WT の結晶を 1% (w/v) の Gal3 を含む結晶化剤に 10 分間浸漬した (WT\_Gal)。回折データは、HKL2000 プログラムスイートのプログラム DENZO と SCALEPACK を使用して統合し、スケーリングした (Otwinowski and Minor, 1997)。

位相は SeMet 標識結晶を用いて多波長異常分散 (MAD) 法により決定した (Table 3-2) (Ishida *et al.*, 2009a)。初期位相は、SOLVE/RESOLVE プログラムを用いて 5 つのセレン原子位置から計算した (Terwilliger and Berendzen, 1999)。その結果得られた座標から、CCP4 プログラムスイートに設定されたオートモデリング ARP/wARP プログラムで構造モデルを構築した (Perrakis *et al.*, 1999; Winn *et al.*, 2011)。野生型およびリガンド結合構造の位相は、SeMet 標識結晶で決定した構造またはリガンドフリー構造をモデル構造とし、CCP4 プログラムスイー

トの MOLREP プログラムを用いた分子置換法により決定した (Vagin and Teplyakov, 2010)。各酵素に結合した糖、水分子および結晶化剤は、観測された電子密度差マップにモデル化し、Ca<sup>2+</sup> は、電子密度マップと配位距離に基づいてモデル化した。本酵素には 3 つの N 結合型糖鎖が観測されたため、同定された糖をそれぞれモデル化した。手動モデル構築と分子リファインメントは第二章と同様、Coot (ver. 0.8.9, University of Oxford, Oxfordshire, England) (Winn *et al.*, 2011)、REFMAC5 (ver. 7.0.063, Science & Technology Facilities Council, England) (Vagin *et al.*, 2004)、Phenix suite の phenix.refine および ensemble\_refinement (ver. 1.13-2998-000, Lawrence Berkeley National Laboratory, USA) (Afonine *et al.*, 2012; Burnley Tom *et al.*, 2012; Burnley and Gros, 2013; Forneris *et al.*, 2014) を用いて行った。精密化後の統計値を Table 3-2 にまとめた。アンサンブルリファインメントにあたっては pTLS 値を複数の条件で設定し、全体的に最も良い統計値を示したモデルを最終モデルとした。なお、結晶の調製と X 線回折実験は当研究室卒業生の石田卓也博士、農研機構の藤本瑞博士に行っていた。

モデルの立体構造や相互作用の模式図は第二章と同様に LigPlot + (ver. 1.4.5)(Wallace *et al.*, 1995; Laskowski and Swindells, 2011) および PyMOL (ver. 2.2.3, Schrödinger, LLC) を用いて作成した。

**Table 3-2. Summary of data collection statistics.**

Data						WT_Gal	E208Q_Gal3	E208A_Gal3
	WT	SeMet (peak)	(edge)	(low remote)	(high remote)	Gal3 soaking	Gal3 co-crystal	Gal3 co-crystal
Space group	$P1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_12_12_1$	$P2_1$	$P3_221$
Unit-cell parameters ( $\text{\AA}$ , $^\circ$ )	$a=40.5$ $b=66.3$ $c=74.0$ $\alpha=72.0$ $\beta=84.7$ $\gamma=82.1$	$a=66.4$ $b=50.5$ $c=75.8$ $\alpha=90.0$ $\beta=111.9$ $\gamma=90.0$				$a=50.8$ $b=66.6$ $c=106.4$ $\alpha=90.0$ $\beta=90.0$ $\gamma=90.0$	$a=66.1$ $b=50.4$ $c=75.7$ $\alpha=90.0$ $\beta=111.3$ $\gamma=90.0$	$a=156.7$ $b=156.7$ $c=147.7$ $\alpha=90.0$ $\beta=120.0$ $\gamma=90.0$
Beam Line	PF BL-5	PF BL-6A	PF BL-6A	PF BL-6A	PF BL-6A	PF-AR NW12	PF-AR NE3	PF-AR NE3
Detector	ADSC Q315	ADSC Q4R				ADSC Q210	ADSC Q270	ADSC Q270
Wavelength ( $\text{\AA}$ )	0.90646	0.97882	0.97950	0.98300	0.96400	1.0000	1.0000	1.0000
Resolution ( $\text{\AA}$ )	50-1.40 (1.45-1.40)	50.0-1.80 (1.86-1.80)	50.0-2.00 (2.07-2.00)	50.0-2.00 (2.07-2.00)	50.0-2.00 (2.07-2.00)	100.0-1.50 (1.55-1.50)	50.0-2.50 (2.54-2.50)	100.0-2.30 (2.38-2.30)
$R_{\text{sym}}$	0.054 (0.370)	0.079 (0.672)	0.061 (0.307)	0.060 (0.303)	0.062 (0.307)	0.046 (0.109)	0.143 (0.399)	0.167 (0.627)
Completeness (%)	95.6 (89.0)	100.0 (99.9)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	97.5 (94.9)	96.2 (83.0)	99.1 (92.0)
Multiplicity	3.8 (3.1)	14.0 (12.6)	7.2 (6.9)	7.2 (6.9)	7.2 (7.0)	9.2 (8.9)	4.4 (3.0)	9.7 (5.1)
Average $I/\sigma(I)$	24.4 (2.8)	36.6 (4.7)	30.9 (8.3)	30.8 (8.2)	31.3 (8.2)	48.9 (21.0)	13.5 (2.7)	17.9 (2.7)
Unique reflections	136692 (12747)	43643 (4353)	31744 (3139)	31760 (3144)	31780 (3146)	57278 (5493)	16007 (702)	92497 (8510)
Observed reflections	520085	613162	227158	228381	228595	524957	69939	900469
Z	2	1				1	1	4

### 3.3. 結果

#### 3.3.1. X線結晶構造解析

本章では SeMet 変異体を用いた多波長異常分散法により真核生物由来の GH43\_sub24 の酵素として、またガラクトタンに結合性を示す CBM35 として初めて *Pc1,3Gal43A* のリガンドと結合していない野生型 (WT)、WT を $\beta$ -1,3-Gal3 にソーキングして得られた Gal との複合体 (WT\_Gal)、catalytic acid の Glu208 の2種の変異体 (E208Q および E208A) と $\beta$ -1,3-Gal3 の共結晶の E208Q\_Gal3、E208A\_Gal3 の4つの構造を決定した。データ収集および精密化で得られた統計値は Table. 3-3 にまとめた。

**Table 3-3. Summary of refinement statistics.**

Data	WT	WT_Gal	E208Q_Gal3	E208A_Gal3
Resolution range	7.997 - 1.398	41.56 - 1.500	29.79 - 2.499	30.66 - 2.300
	(1.448 - 1.398)	(1.554 - 1.500)	(2.588 - 2.499)	(2.382 - 2.300)
Completeness (%)	95.46 (87.82)	97.51 (94.80)	96.41 (85.67)	98.78 (92.17)
Wilson B-factor	12.76	10.11	29.91	30.40
Reflections used in refinement	136655 (12497)	57105 (5474)	15762 (1381)	92011 (8507)
Reflections used for R-free	6862 (630)	2884 (272)	799 (64)	4568 (441)
R-work (%)	15.47 (22.50)	13.43 (12.71)	16.62 (25.54)	16.10 (22.39)
R-free (%)	18.56 (26.28)	16.00 (17.93)	24.39 (42.53)	21.43 (28.28)
Number of non-hydrogen atoms				
macromolecules	7966	3923	3576	14570
ligands	6615	3290	3235	12886
solvent	109	121	114	678
Protein residues	1242	512	227	1006
RMS (bonds)	2106	427	428	1708
RMS (angles)	0.008	0.006	0.008	0.011
Ramachandran favored (%)	1.22	0.87	0.94	1.05
Ramachandran allowed (%)	97.29	97.41	94.13	95.76
Ramachandran outliers (%)	2.71	2.59	5.87	4.24
Rotamer outliers (%)	0	0	0	0
Clash score	0.81	0.55	0.29	0.36
Average B-factor (Å <sup>2</sup> )	2.06	1.95	6.94	3.50
macromolecules	17.21	12.45	30.48	32.98
ligands	14.97	10.57	29.77	31.60
solvent	29.38	23.33	52.26	56.11
PDB ID	28.09	22.02	29.74	35.03
	7BYS	7BYT	7BYV	7BYX

## Pc1,3Gal43A の全体構造

*P. pastoris* を用いて産生した組換え Pc1,3Gal43A には Gln21-Tyr448 の 428 アミノ酸残基に加え、制限酵素サイトに由来する Glu19 と Phe20 の 2 残基が付加されていたが、この 2 残基の電子密度は観測されなかった。本酵素は Asn79、Asn194、Asn389 の 3 箇所に N 結合型糖鎖修飾部位を有しており、観測された N 結合型糖鎖の糖残基数はサイトにより異なっていた。Pc1,3Gal43A は 2 つのドメインにより構成されており、ソーキング法または共結晶法により結合した基質は触媒部位または CBM35 のリガンド結合サイトで観測された (Fig. 3-1)。N 末端側の触媒ドメイン (Gln21-Gly325) は他の GH clan F に分類される酵素と同様に 5 枚羽根プロペラ構造をとっており、C 末端側の CBM35 に分類される糖質結合ドメイン (PcCBM35; Thr326-Tyr448) は、同ファミリーに典型的にみられるβ-ゼリーロール構造をとっていた。PcCBM35 はリガンド結合面とは異なる表面に位置する 1 つ目のβ-ストランドの C 末端側に Ca<sup>2+</sup>結合サイトを有しており、これは CBM35 で保存されている 2 つの Ca<sup>2+</sup>結合部位のうち一方に相当した。他分子と相互作用する領域の表面積は 686 Å<sup>2</sup> であり、本酵素は結晶化の影響により単位格子中に単量体としてみられた場合 (WT\_Gal、E208Q\_Gal3) と二量体 (WT) や四量体 (E208A\_Gal3) としてみられた場合があったが、PDBePISA サーバー (<https://www.ebi.ac.uk/pdbe/pisa/>) (Krissinel and Henrick, 2007) で解析すると、溶液中では単量体として存在すると示唆された。

## 触媒ドメイン複合体構造

5 枚羽根プロペラ構造の触媒ドメインは球状に近い形状で中央にくぼみが 2 つある構造をしており、そのくぼみは 5 枚羽根先端が集まる中央部に存在していた (Fig. 3-1)。他の GH43 の酵素と同様、空洞のうち的一方は触媒部位であった。各羽根はそれぞれ Gln21 または Asn22-Leu87 (Fig. 3-1I)、Ser88-Asp155 (Fig. 3-1II)、Ser156-Gly204 (Fig. 3-1III)、Ala205-Ser257 (Fig. 3-1IV)、Ala248-Asp297 (Fig. 3-1V) から構成されていた。

共結晶法にて得られた E208Q\_Gal3 複合体ではβ-1,3-Gal3 が触媒部位にみられ、サブサイト -1 位から +2 位を占めていた。本章ではサブサイトにみられた Gal を非還元末端から順に Gal<sub>-1</sub>、Gal<sub>+1</sub>、Gal<sub>+2</sub> と呼称する。Gal<sub>-1</sub> は触媒部位の底に位置しており、Gal<sub>+1</sub>、Gal<sub>+2</sub> は酵素の外側へ続いていた (Fig. 3-2)。Gal<sub>+1</sub> の半分ほどは触媒ポケットに埋まっており、Gal<sub>+2</sub> は完全に酵素表

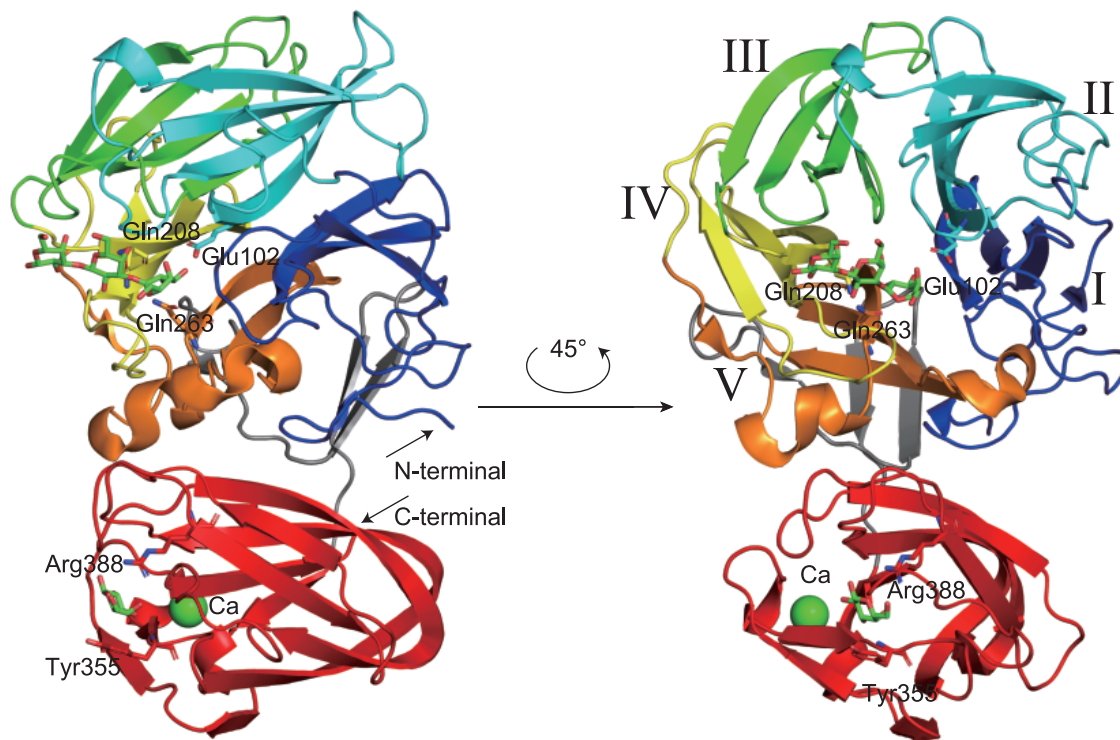
面に露出していた (Fig. 3-2)。

Gal<sub>-1</sub> は <sup>1</sup>S<sub>3</sub> skew boat form (ねじれ舟形) 配座をとっており、複数の残基と水素結合を介して相互作用していた。Fig. 3-2. B に示したように、Gal<sub>-1</sub> の 2 位ヒドロキシ基は Arg103 の NH<sub>2</sub> および水分子を介して Gln263 の OE1 とそれぞれ水素結合を形成し、さらにこの水分子は Gly228 の O と結合して安定化していた。3 位ヒドロキシ基は 2 位ヒドロキシ基とは異なる水分子を介して Glu57 の OE2 と水素結合していた。また、Glu102、Tyr126、Asp158、Gln208、Thr226、Trp229、Gln263 は疎水性相互作用に関与していた。特に Trp229 は Gal<sub>-1</sub> の C3-C4-C5-C6 平面の支持、Tyr126 は 4 位ヒドロキシ基および 6 位メチロール基、Glu102 は 3 位および 4 位ヒドロキシ基の認識にそれぞれ関与していた (Fig. 3-2)。Gal<sub>+1</sub> では、2 位ヒドロキシ基は Gln208 の NE2 および Gly228 の N と、環内酸素 (Fig. 3-2. "O5") は Asn180 の NE2 と、6 位メチロール基は水分子を介して Asn179 の OD1 とそれぞれ水素結合を形成していた。また、Tyr126、Arg157、Asn180、Gln208 が疎水的に相互作用していた。Gal<sub>+2</sub> では、2 位および 3 位ヒドロキシ基が Thr226 の OG1 および Asn180 の ND2 とそれぞれ水素結合しており、Thr226 はさらに疎水的な相互作用にも関与していた。またさらに、Gal<sub>+1</sub> と Gal<sub>+2</sub> の間のグリコシド結合の O は Asn180 の ND2 と水素結合を形成していた (Fig. 3-2)。

β-1,3-Gal3 をソーキングした WT の構造では、サブサイト-1 位に <sup>4</sup>C<sub>1</sub> いす型コンフォメーションをとった α-Gal がみられた。結合部位における周囲のアミノ酸残基との相互作用様式は E208Q\_Gal3 とおおよそ同様であったが、1 位ヒドロキシ基が Gly228 および Gln263 と水素結合していた。この Gal はソーキングに用いた β-1,3-Gal3 とはアノマーが反転して α-アノマーとなっていたことから、ソーキング中に加水分解反応がおこって生じた反応産物であると考えられた。E208A\_Gal3 では、触媒部位には Gal3 の電子密度は観測されなかった。

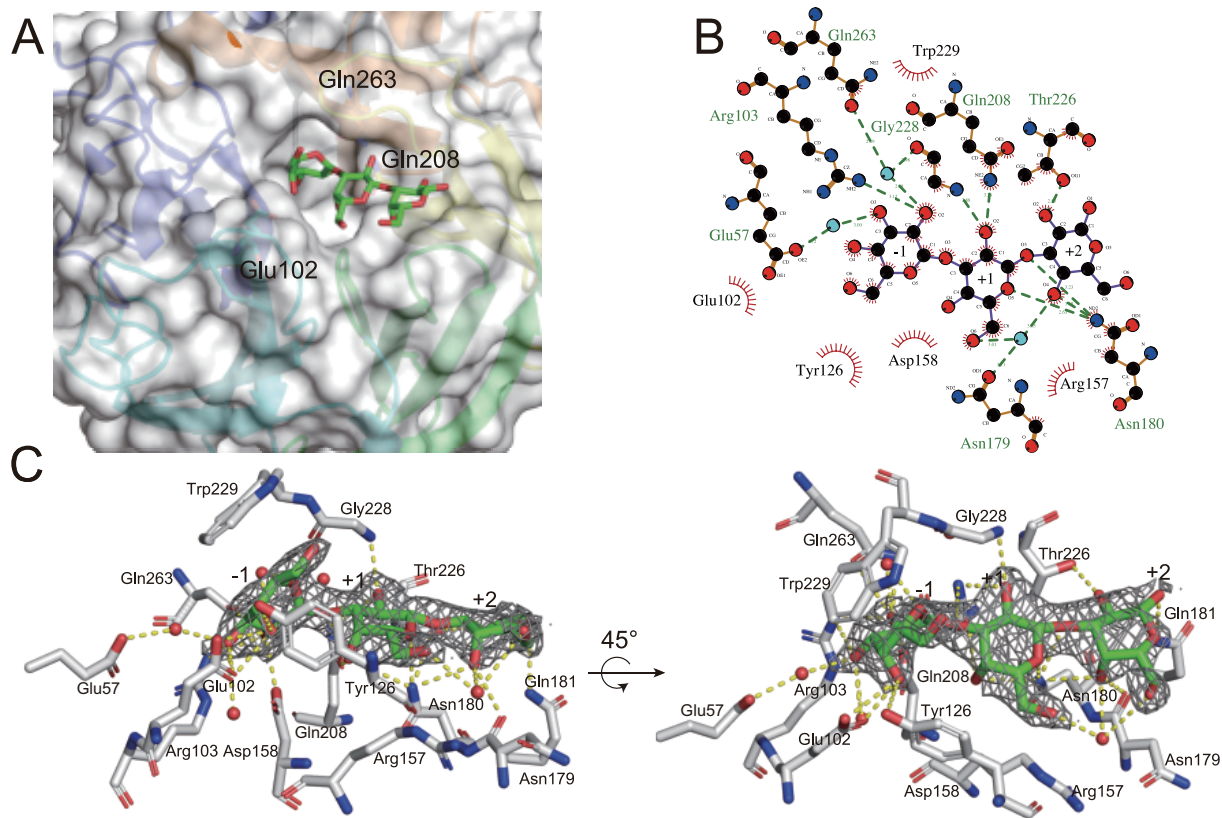
触媒残基を同定するために WT と触媒反応に重要であると推察される 3 つのアミノ酸残基の合計 6 つの変異体の Gal2 および Gal3 に対する分解活性を測定した。WT は Gal2 に対して 111.60±6.92 units、Gal3 に対して 223.10±7.89 units (1 unit は 1 分間に 1 μmol の Gal を遊離するために必要な酵素量) を示したのに対し、変異体はいずれも Gal を遊離しなかった。この結果はこれらの 3 つの残基は加水分解反応に必須な残基であることを示した。





**Fig. 3-1. Overall structure of *Pc1,3Gal43A*.**

In the three-dimensional structure of *Pc1,3Gal43A*, the five blades of the catalytic domain are shown in blue (Gln21-Leu87), cyan (Ser88-Asp155), green (Ser156-Gly204), yellow (Ala205-Ser247), and orange (Ala248-Asp297) with successive roman numerals. The CBM (The326-Val448) is shown in red. The linker connecting the two domains (Phe298-Gly325) is shown in gray.

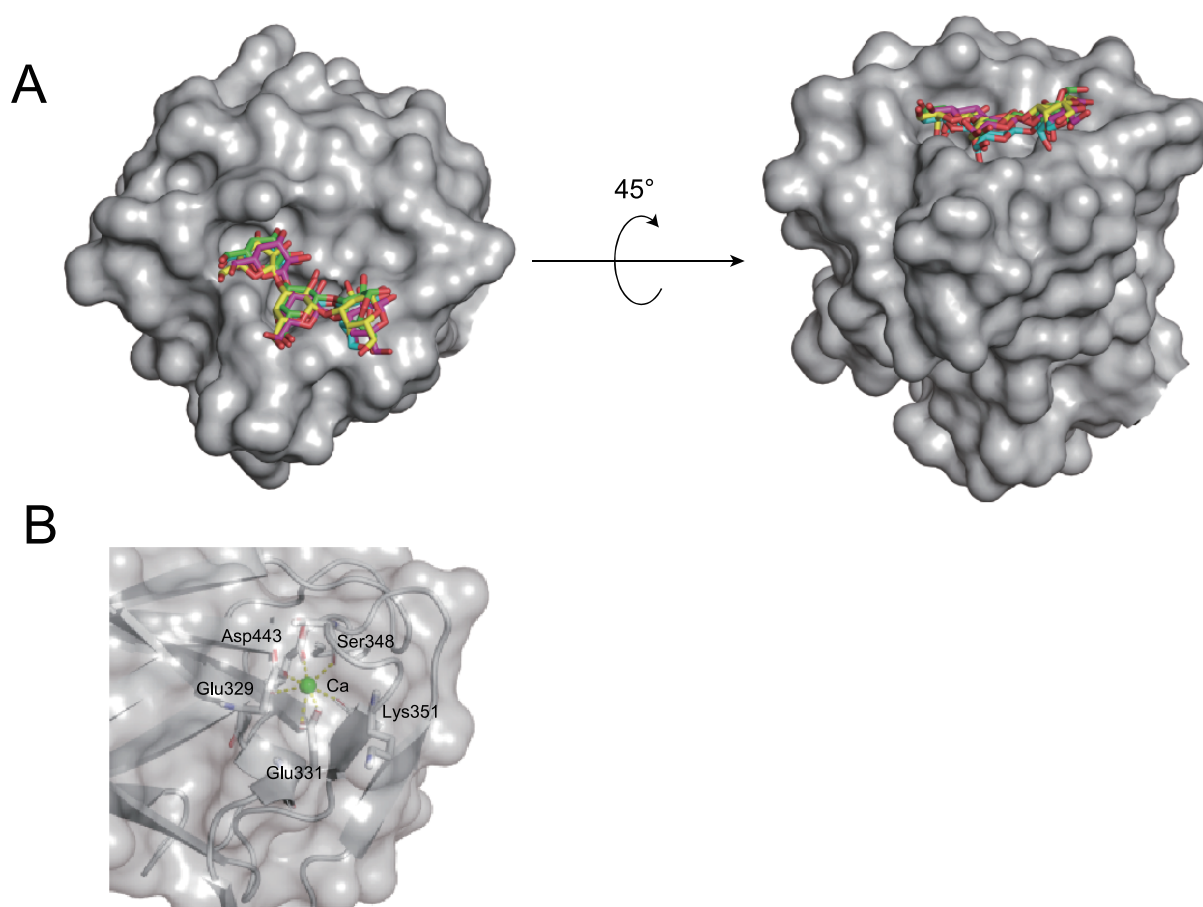


**Fig. 3-2.  $\beta$ -1,3-Gal3 binding mode at the catalytic site.**

A: The surface structure of the catalytic center.  $\beta$ -1,3-Gal3 is represented as green (carbon) and red (oxygen) sticks. B: Schematic diagram shows the interaction mode at the catalytic center. Black, red, and blue show carbon, oxygen, and nitrogen, respectively. Red lines indicate the hydrophobically interacting residues. This diagram was drawn with LigPlot + (ver. 1.4.5)(Wallace *et al.*, 1995; Laskowski and Swindells, 2011). C: The  $2F_o - F_c$  density map is drawn as a gray mesh (0.8 sigma). Residues are shown in white (carbon), red (oxygen), and blue (nitrogen). Gal3 is shown in green (carbon) and red (oxygen). Yellow dots indicate hydrogen bonds and/or hydrophobic interaction and red spheres show water molecules interacting with ligands or residues.

## CBM 複合体構造

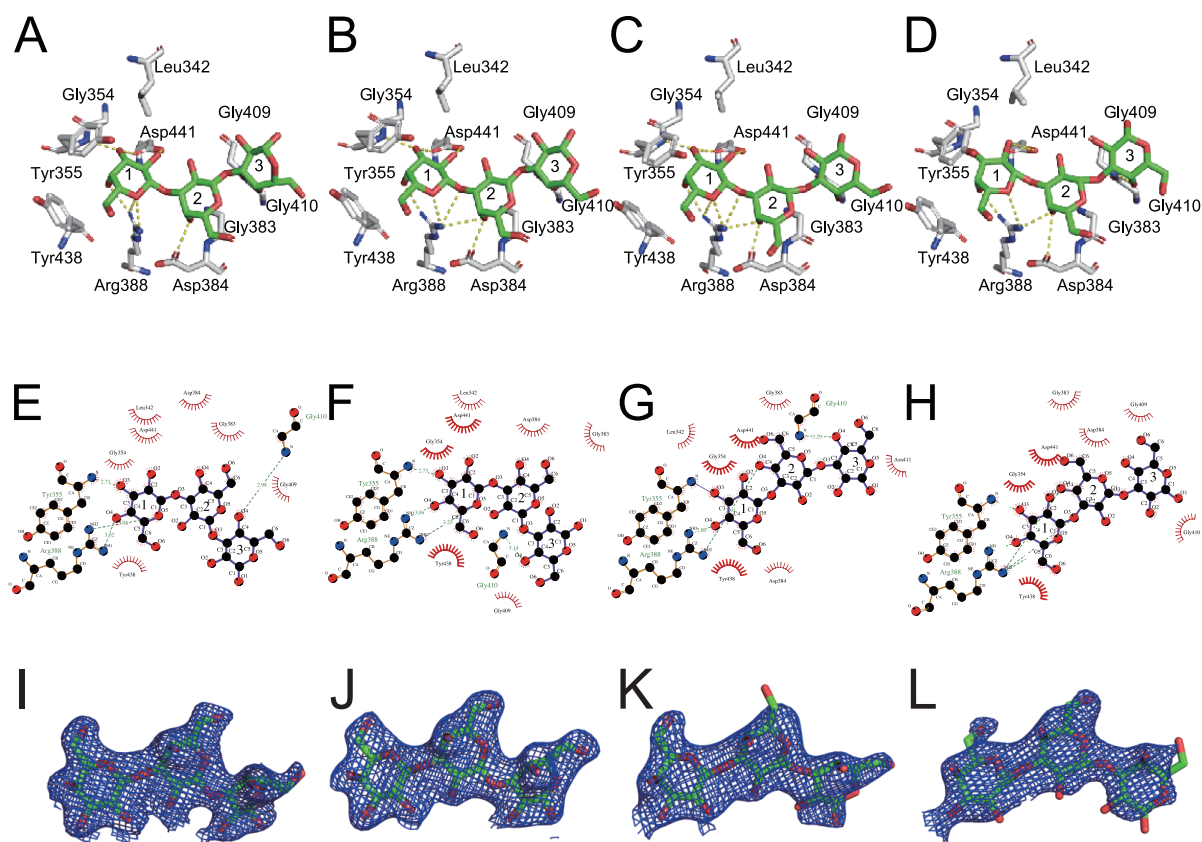
*Pc*1,3Gal43A は C 末端にβ-ゼリーロールフォールドによって構成される 1 つの CBM35 ドメインを有している。このドメインはリガンドが結合する面とは別表面に位置する 1 つ目のβ-ストランドの C 末端部分に  $\text{Ca}^{2+}$  結合部位を有しており、これは CBM35 に保存された  $\text{Ca}^{2+}$  結合部位に相当していた (Fig. 3-3)。CBM35 のうち、主にウロン酸を含む基質と結合する CBM35 ではリガンド結合部位にも  $\text{Ca}^{2+}$  が結合することが報告されているが、*Pc*CBM35 はリガンド結合部位では  $\text{Ca}^{2+}$  の結合はみられなかった (Fig. 3-3)。



**Fig. 3-3. Surface structures of *Pc*CBM35.**

A: Substrate binding mode at CBM35. Green, cyan, magenta, and yellow indicate ligands' carbons of chains A, B, C, and D of E208A\_Gal3, respectively, and red shows oxygen. The left side is the non-reducing end of Gal3, and the right side is the reducing end. B:  $\text{Ca}^{2+}$  binding mode at CBM35.  $\text{Ca}^{2+}$  is represented as green spheres and interacting residues are shown as stick models. Yellow dots indicate interaction.

E208A\_Gal3 では、PcCBM35 のリガンド結合部位に $\beta$ -1,3-Gal3 の電子密度がみられた。Fig. 3-4 のように、PcCBM35 はリガンドの非還元末端と強く相互作用する type C CBM (Boraston *et al.*, 2004; Gilbert *et al.*, 2013; Armenta *et al.*, 2017) の相互作用様式を示した。E208A\_Gal3 の非対称性単位中には 4 分子の Pc1,3Gal43A が含まれており、いずれの分子も同様の様式で Gal3 の非還元末端側の Gal (Gal\_site 1) と結合していた。各分子において相互作用に参与するアミノ酸残基は同じであったが、中央の Gal (Gal\_site 2) および還元末端側の Gal (Gal\_site 3) は大きく分けて 2 つの位置に存在しており、2 パターンの結合サイトを有する可能性が考えられた。Gal\_site 1 は Tyr355 および Arg388 と水素結合を形成し、Leu342、Gly354、Tyr438、Asp441 と疎水的に相互作用していた。Gal\_site 2 は Gly383 および Asp384 と疎水的に相互作用した。Gal\_site 3 で主にみられた相互作用は Gly409 との疎水性相互作用および Gly410 との水素結合による相互作用であったが、Chain C ではさらに Asn411 との疎水性相互作用もみられた (Fig. 3-4)。



**Fig. 3-4. Ligand interaction mode at *PccBM35*.**

A, (E, I), B, (F, J), C, (G, K), D, (H, L) are each chain of E208A, respectively.

A to D: Interaction modes between ligand and CBM35 residues. Atoms are indicated in the same colors as shown in Fig. 3-2. E to H: Schematic diagram showing the interaction mode at CBM35. Sugar binding sites are named Gal\_site 1, Gal\_site 2, and Gal\_site 3 from the non-reducing end of the sugar and in this figure, they are labeled 1, 2, and 3, respectively. I to L: 2Fo-Fc density maps (1.0 sigma) of Gal3 of each chain.

### 3.3.2. アンサンブルリファインメント

Pc1,3Gal43A の反応溶液中でのゆらぎを可視化するために、精密化したモデルを用いてアンサンブルリファインメントを行った。各モデルは 4 つの異なる pTLS 値 (%; 0.6、0.8、0.9、1.0) を設定して解析し、各モデルの最適な pTLS 値を設定した際に得られたモデルの統計値を Table. 3-4 に示した。また、触媒ドメインの触媒部位と CBM のリガンド結合部位の拡大図をそれぞれ Fig. 3-5 および Fig. 3-6 にそれぞれ示した。なお、非対称単位中に複数の分子を含む構造 (WT および E208A\_Gal3) は、すべての分子の結果をまとめて図示した。

触媒部位では、基質との結合に関与する残基のうち一部の残基ではゆらぎの大きさがアポ体とホロ体で有意に異なっていた。リガンドと結合した構造 (Fig. 3-5. B、C、F、G、J、K) の Tyr126、Arg157、Asp158、Asn179、Asn180、Gln181、Trp229、Gln263 のゆらぎはリガンドと結合していないアポ構造 (Fig. 3-5. A、D、E、H、I、L) と比べて非常に小さかった。この結果は、リガンドの触媒部位への結合により側鎖のゆらぎが収束したことを示唆した。また、Gal が結合した構造 (WT\_Gal) と Gal3 が結合した構造 (E208Q\_Gal3) を比較すると、E208Q\_Gal3 の Asn179、Gln181、Glu(Gln)208、Thr226 のゆらぎは WT\_Gal よりも小さかった (Fig. 3-5. B、C、F、G) ことから、これらの残基はプラス側サブサイトでリガンドの認識に関与すると推察された。また、WT および WT\_Gal において catalytic acid の Glu208 がとったコンフォメーションは大きく 2 つに分けられたが、これらのコンフォメーションは BT3683 の結晶構造 (PDB ID: 6EUG) で報告されているものと一致した (Cartmell *et al.*, 2018)。したがって、この Glu208 の二方向への動き (2 つのコンフォメーション) が触媒作用に重要であると示唆された。Gln263 は、PcCel45A において互変異性化によりイミド酸型となって catalytic base として機能する Asn92 のアンサンブルリファインメントの結果と同様のゆらぎを示した (Fig. 3-5、3-7)。Glu102 は Gal<sub>1</sub> のアキシナルな 4 位ヒドロキシ基と相互作用するために非還元末端の Gal の認識に重要である。また、この Glu102 のゆらぎの大きさの程度は WT と変異体で異なっており、Glu102 のゆらぎはリガンドの存在の有無ではなく Glu208 の存在と関係があるようであった (Fig. 3-5)。すなわち、Glu102 は Glu208 と相互作用していると考えられる。WT と E208A\_Gal3 の Asp158 は、WT\_Gal および E208Q\_Gal3 に比べゆらぎが大きかった。Asp158 は GH43\_sub24 で pKa モジュレーターであると考えられている Asp に相当する (Cartmell *et al.*, 2018) が、アンサンブルリファインメントの結果でみられた Asp158 のゆらぎとその機能には相関があると考えられ

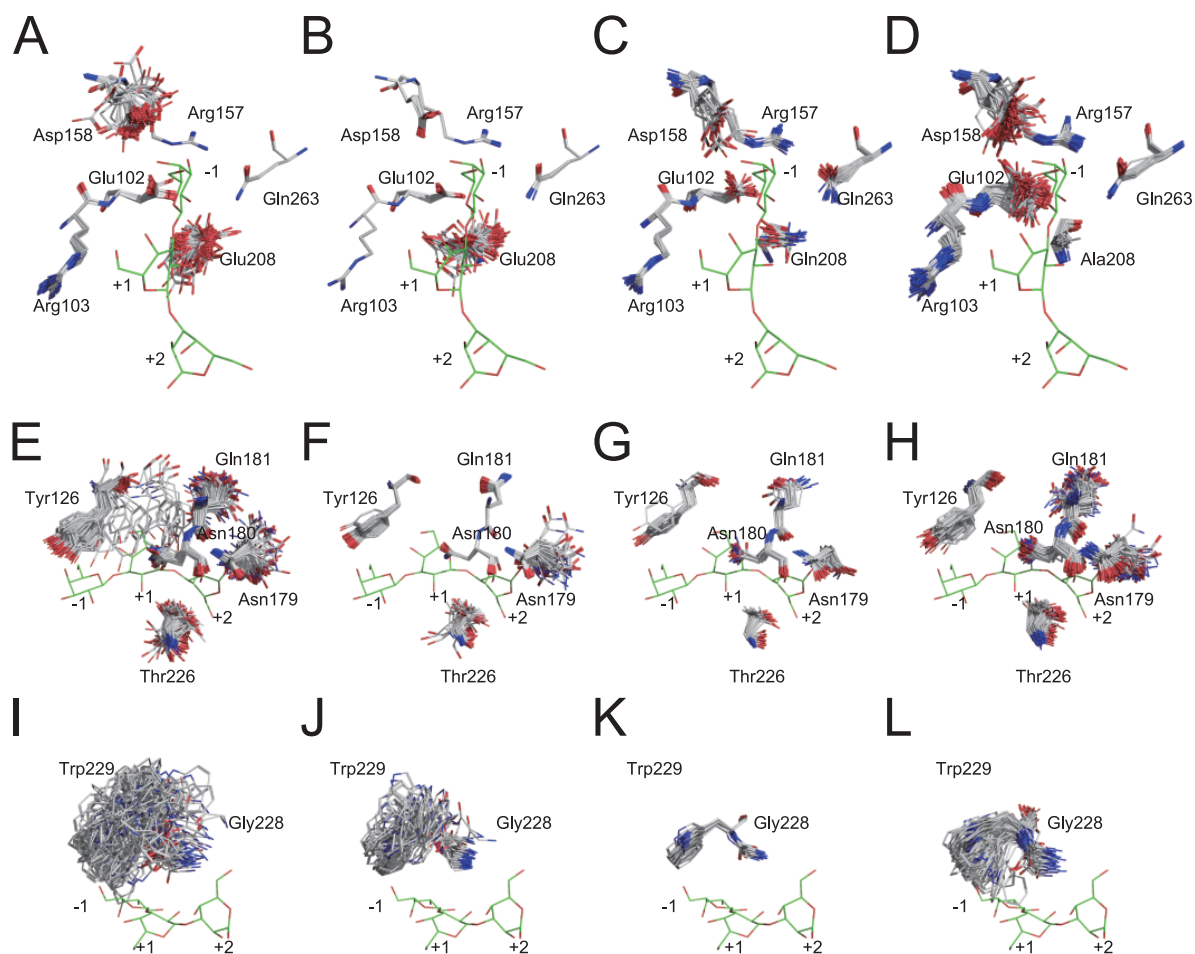
る。Fig. 3-5. I-M に着目すると、Trp229 のゆらぎの程度に大きな差がみられた。E208Q\_Gal3 (Fig. 3-5. K) では Trp229 のゆらぎは小さいが、他の構造 (Fig. 3-6. I, J, L) ではより大きなゆらぎがみられた。特に顕著な例は Fig. 3-5. I であり、Trp229 は非常に大きなゆらぎを示していた。これらの結果から、この Trp229 は常に動いており、リガンドが活性中心に到達すると  $\pi$ - $\pi$ 相互作用により適切な位置に固定するメカニズムで基質を認識すると推察された。これらの残基のゆらぎの程度を定量的にみるために各残基の二面角のヒストグラムを Fig. 3-8 に示した。

E208A\_Gal3 の非対称単位中に含まれる 4 分子の酵素のそれぞれのゆらぎを比較したところ、CBM においてリガンド認識に関与する残基のゆらぎの程度に大きな差はみられなかった (Fig. 3-6)。しかしアンサンブルリファインメントの結果、リガンドである Gal3 の Gal\_site 1 と Gal\_site 2 は比較的小さなゆらぎを示したのみだったのに対し、Gal\_site 3 は大きなゆらぎを示しており (Fig. 3-6)、このゆらぎのパターンは触媒部位におけるリガンドのゆらぎのパターンとは異なっていた (Fig. 3-9)。触媒部位においては、Gal<sub>-1</sub> は強く酵素に固定され、Gal<sub>+2</sub> は周囲の残基と相互作用しながら平面的にゆらぎながら結合していると考えられた (Fig. 3-9)。一方で CBM においては、Gal\_site 1 は酵素表面に強く固定されているものの、Gal\_site 3 は立体的にゆらぎながら吸着していると考えられた (Fig. 3-9)。すなわち、触媒部位においてはゆらいでいる基質の各 Gal 残基がサブサイトを構成するアミノ酸残基とそれぞれ相互作用して認識されるのに対し、CBM ではリガンドと酵素が面吸着するように結合するという方式をとるため、基質と酵素表面との相互作用の様式に差異がみられると考えられた。

**Table 3-4. Refinement statistics of ensemble refinement.**

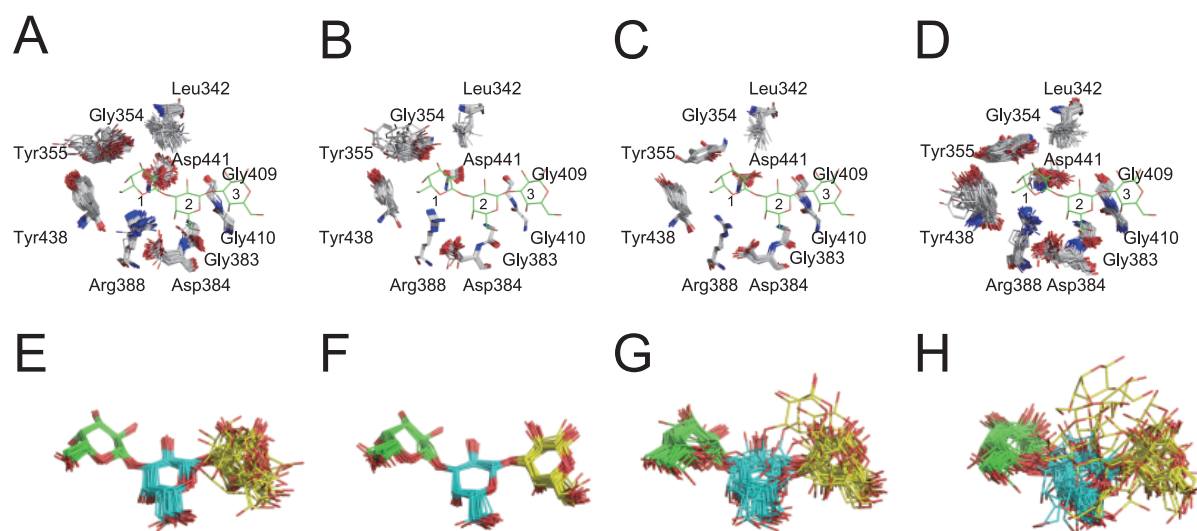
Data	WT	WT_Gal	E208Q_Gal3	E208A_Gal3
Resolution range	7.997-1.398	41.56-1.500	29.79-2.499	30.66-2.300
	(1.448-1.398)	(1.554-1.500)	(2.588-2.499)	(2.382- 2.300)
Completeness (%)	95.97(82)	97.52 (95)	96.47 (88)	98.93(87)
pTLS (%)	0.9	0.9	0.9	1.0
Tx (ps)	1.0	0.9	0.3	0.4
Wilson B-factor	12.8	10.1	29.9	30.4
Reflections used in refinement	136649	57112	15759	91994
Reflections used for R-free	6862	2885	799	4569
R-work (%)	13.81 (24.36)	12.08 (10.68)	17.82 (24.73)	15.92 (22.56)
$\Delta$ R-work (%)	-1.66 (-1.86)	-1.35 (-2.03)	1.2 (0.81)	-0.18 (0.17)
R-free (%)	17.08 (26.30)	15.29 (17.10)	23.33 (32.75)	20.71 (28.84)
$\Delta$ R-free (%)	-1.48 (-0.02)	-0.71 (-0.83)	-1.06 (-9.78)	-0.72 (0.56)
RMS(bonds)	0.008	0.010	0.007	0.008
RMS(angles)	1.171	1.312	1.078	1.090
Ramachandran favored (%)	94.06	95.39	88.98	92.62
Ramachandran allowed (%)	5.08	4.03	9.19	7.24
Ramachandran outliers (%)	0.86	0.58	1.83	0.74
Rotamer outliers (%)	7.45	7.00	11.05	7.85
Clash score	0	0	0	0
Average B-factor ( $\text{\AA}^2$ )	13.65	9.55	28.32	32.83
macromolecules	13.63	9.54	28.30	32.66
ligands	14.97	9.82	28.98	36.05
Molprobit score	1.56	1.45	1.87	1.64
Model number	100	103	20	34





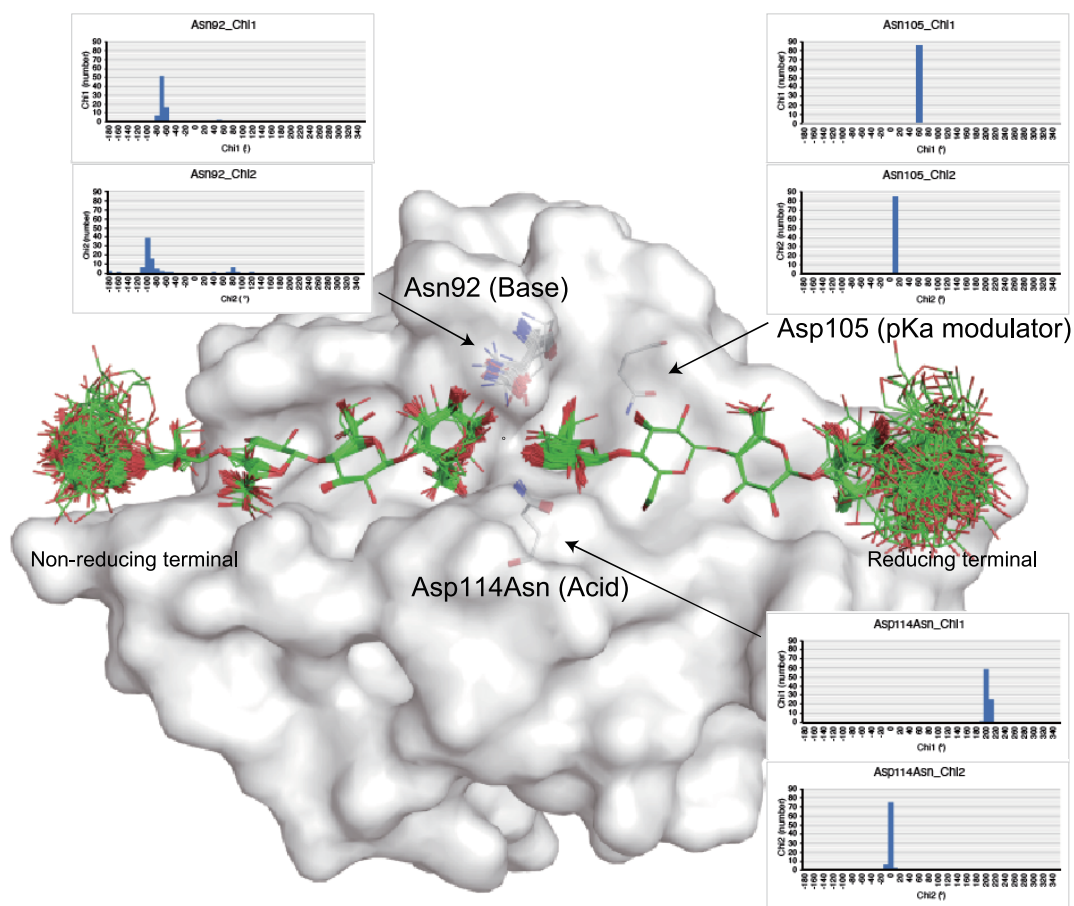
**Fig. 3-5. Results of ensemble refinement at the catalytic site.**

Each model is divided into three parts for clarity. A (E, I), B (F, J), C (G, K), and D (H, L) show WT, WT\_Gal, E208Q\_Gal3, and E208A\_Gal3, respectively. Although WT and E208A\_Gal3 contained multiple molecules in an asymmetric unit, the results obtained with multiple molecules were considered as an ensemble of one molecule in the present study. Atoms are indicated in the same colors as Fig. 3-2. Gal3 of the structure of E208Q\_Gal3 obtained by X-ray crystallography is arranged in each figure to maximize ease of comparison.



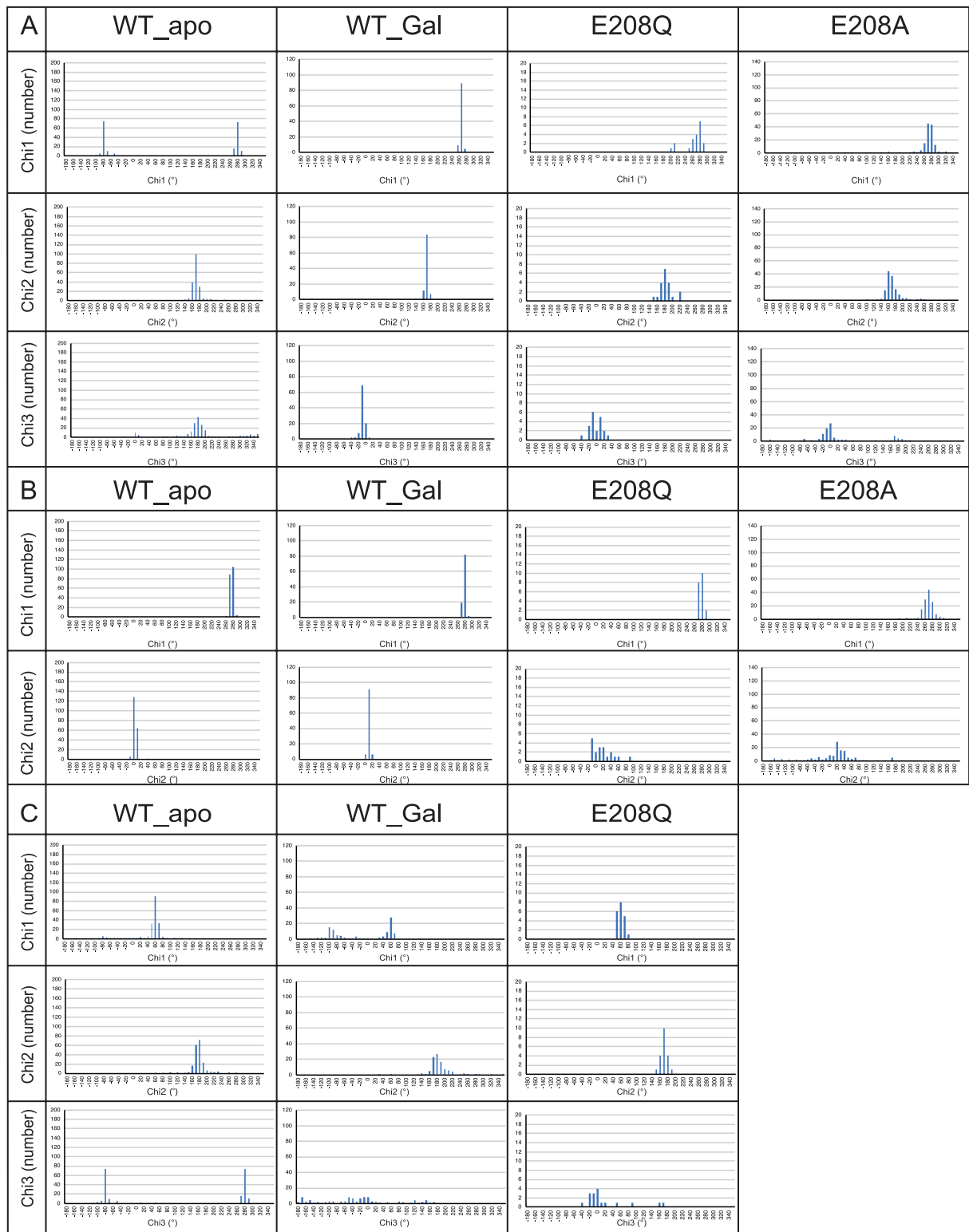
**Fig. 3-6. Results of ensemble refinement at the CBM ligand-binding site.**

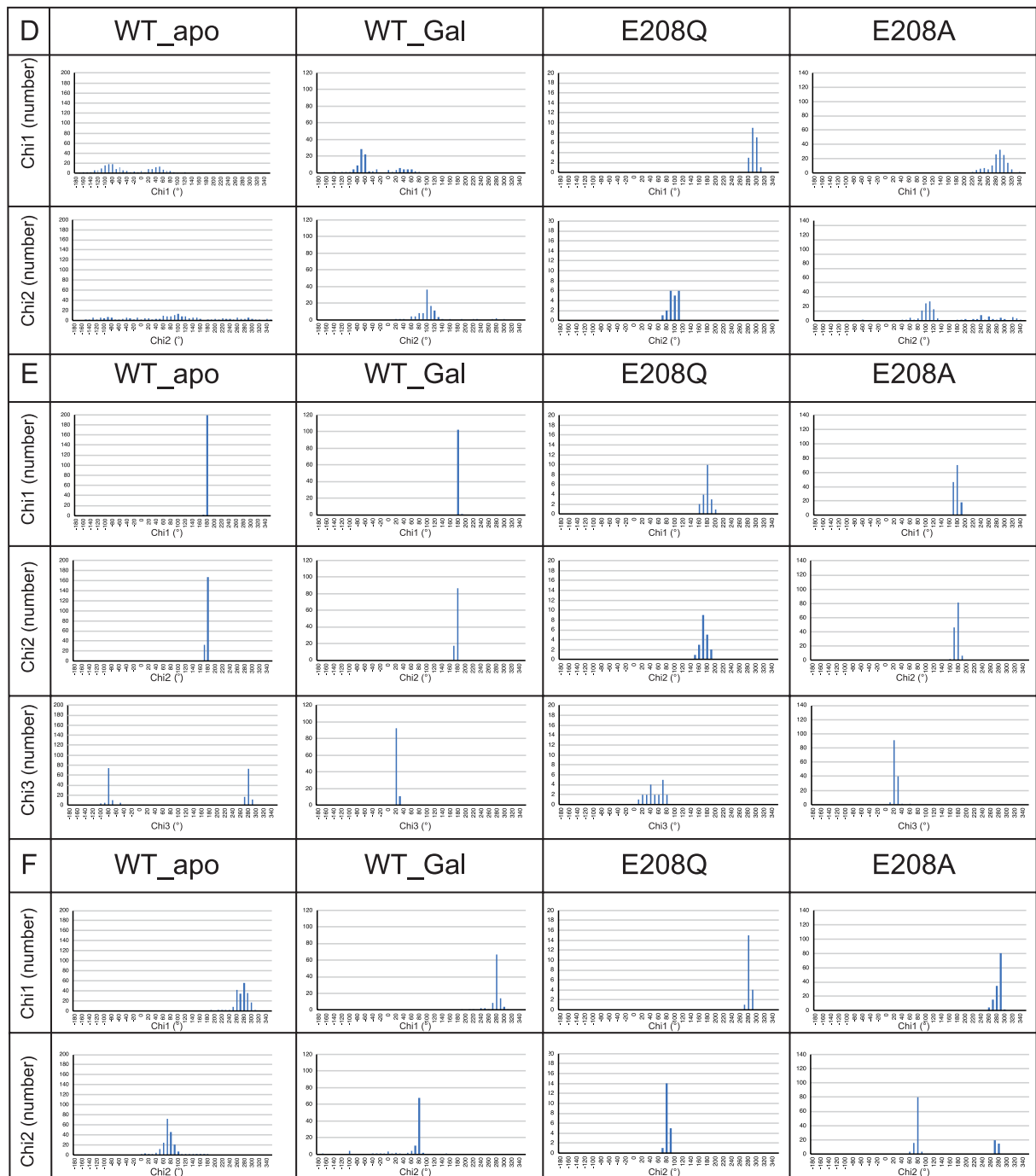
A, B, C, and D: Residues related to ligand interaction. In this figure, Gal3 of chain A of refined E208A\_Gal3 is drawn for comparison. E, F, G and H: The ligands of each chain. Green, cyan, and yellow show in order from the non-reducing terminal Gal. A and E, B and F, C and G, D and H represent chain A, B, C, and D of E208A\_Gal3, respectively. Atoms are indicated in the same colors as Fig. 3-2.

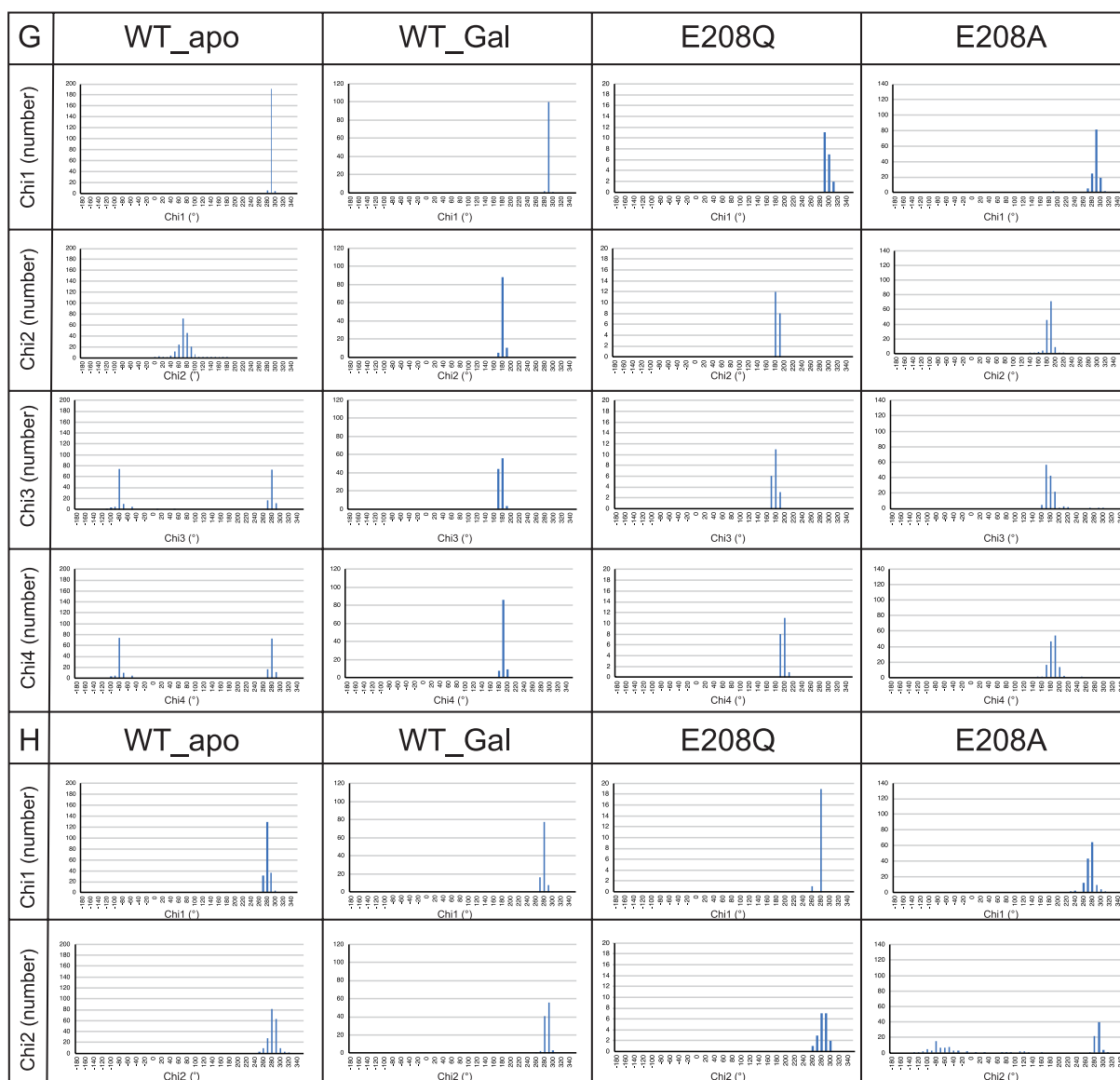


**Fig. 3-7. Fluctuation of catalytic residues in *PcCel45A*.**

Ensemble refinement was performed for *Pc*1,3Gal43A with *PcCel45A* D114N complexed with cellopentaose (PDB ID: 3X2K). In total, 86 states were calculated, representing all catalytic residues and ligands (two molecules of cellopentaose). Histograms of the dihedral angles of the three residues involved in the catalytic reaction are shown in the diagram.

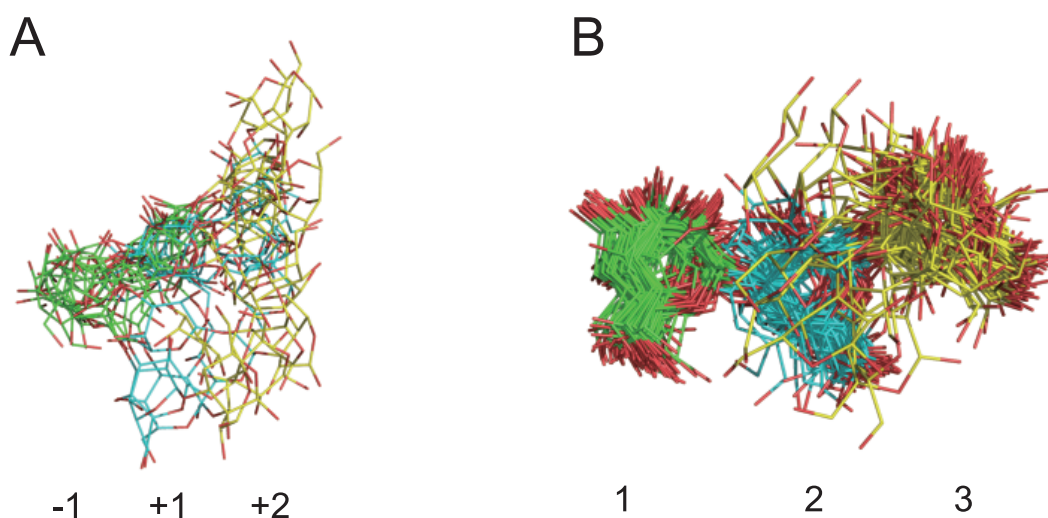






**Fig. 3-8. Histograms of the dihedral angles of each conformation obtained by ensemble refinement of amino acid residues involved in catalysis and/or substrate recognition.**

A-H indicate Glu102, Asp158, Glu(Gln)208, Trp229, Gln263, Tyr355, Arg388, and Tyr438, respectively.



**Fig. 3-9. Ligand conformation of ensemble refinement at a glance.**

A: Ligand conformation of E208Q\_Gal3 ensemble model. B: Ligand conformation of E208A\_Gal3 ensemble models with four chains aligned. Green, cyan, and yellow are used in order from the non-reducing terminal Gal.

### 3.4. 考察

GH43\_sub24 に分類されるエキソ- $\beta$ -1,3-ガラクトナーゼの多くが CBM35 または CBM13 に分類される糖質結合モジュールを有する (Fujita *et al.*, 2014)。本章では $\beta$ -1,3-ガラクトタンに結合性を示す CBM35 を有する GH43\_sub24 のアポ構造およびホロ構造を初めて決定することに成功した。またさらに、今回明らかにした構造は真核生物由来の GH43\_sub24 としてかつ $\beta$ -1,3-ガラクトタンに結合性を示す CBM35 として最初の構造であった。

#### 3.4.1. Pc1,3Gal43A の加水分解機構

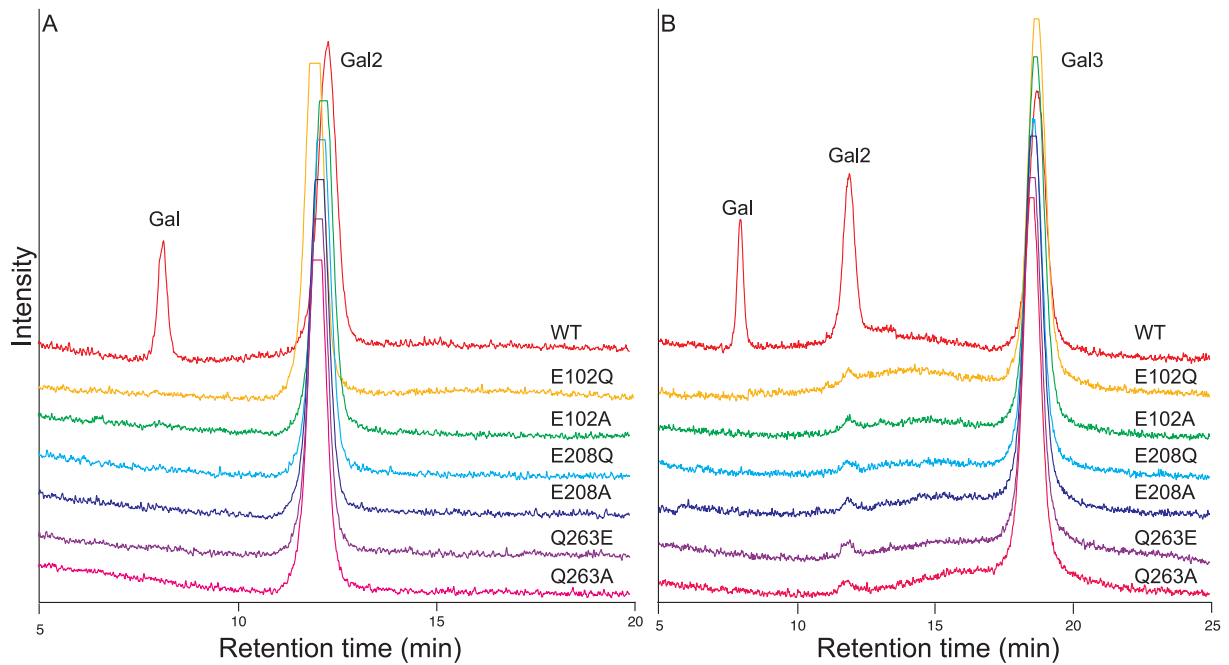
GH43 は Inverting 型酵素であり、Glu および Asp をそれぞれ catalytic acid、catalytic base として用いるが、GH43\_sub24 はこのうち catalytic base に相当する Asp を欠損している (Cartmell *et al.*, 2011; Mewis *et al.*, 2016)。Cartmell *et al.* (2018) によって GH43\_sub24 は互変異性化してイミド酸型となった Gln が catalytic base として機能するか、あるいはこの Gln が求核水を保持してプロトンを受受することで加水分解反応を触媒する可能性が示された

(Cartmell *et al.*, 2018)。Pc1,3Gal43A において触媒反応に関与すると推定される 3 つのアミノ酸残基の 6 つの変異体の酵素活性を調査した結果、反応後の溶液を解析した HPLC のクロマトグラム上で Gal のピークは検出されなかった (Fig. 3-10) ため、これらの残基は加水分解反応の触媒に必要不可欠であり、Glu102、Glu208、Gln263 の役割はそれぞれ 4 位ヒドロキシ基の認識、catalytic acid、catalytic base であると推定した。結晶構造中にみられた水分子と各アミノ酸残基の水素結合可能な距離より、本酵素は Fig. 3-11. A のような水素結合ネットワークを形成しており、特に Arg103、Asp158、Asp279、Trp281 はイミド酸型となった Gln263 の安定に関与していると考えられる。また、この推定水素結合ネットワークに基づく加水分解反応の触媒時には Fig. 3-11. B のような経路でプロトンの伝達を行っていると考えられた。しかしながら、X 線を用いた結晶構造解析では水素原子は観測できない。したがって実際にこのメカニズムでプロトンの授受を行っているかを調査し、触媒メカニズムを解明するためには水素原子を可視化することができる中性子線を用いた結晶構造解析を行う必要がある。なお、Pc1,3Gal43A で機能を同定した 3 つの残基 (Glu102、Glu208、Gln263) は GH43\_sub24 で保存されているため、GH43\_sub24 は広くこのメカニズムで加水分解反応を触媒すると推察される。

GH43\_sub24 に分類される 150 の酵素のアミノ酸配列および構造既知の BT3683 および Ct1,3Gal43A の構造を比較したところ、触媒反応に関わるアミノ酸残基だけではなく基質との相互作用に関与するアミノ酸残基の多くが GH43\_sub24 で広く保存されており、触媒ポケットの表面構造も類似しているとわかった (Fig. 3-12)。特にサブサイト-1 位で相互作用に関与するすべての残基 (Pc1,3Gal43A の Glu57、Glu102、Arg103、Tyr126、Asp158、Glu208、Trp229、Gln263) は保存されていた (Fig. 3-12)。さらに、真菌由来の GH43\_sub24 のアミノ酸配列のマルチプルアライメントの結果 (Fig. 3-13) から、GH43\_sub24 の相同性は高く、多くの残基が保存されているといえる。すなわち、サブサイト-1 位における相互作用様式は GH43\_sub24 で保存されていると推察される。また、アンサンブルリファインメントの結果に基づく、特に基質と結合していない構造において Trp229 は大きなゆらぎを示した (Fig. 3-5. I-L)。BT3683 の Trp541 (Pc1,3Gal43A の Trp229 に相当) は Gal と疎水的に相互作用すると報告されている (Cartmell *et al.*, 2018) ことから、Pc1,3Gal43A の Trp229 は溶液中でゆらいであり、リガンドが触媒部位に結合する際に疎水性相互作用により基質を押さえ込む役割を担っていると考えられる。一方で Asn179 および Thr226 は BT3683 では Asp210 と Cys258 に、Ct1,3Gal43A で

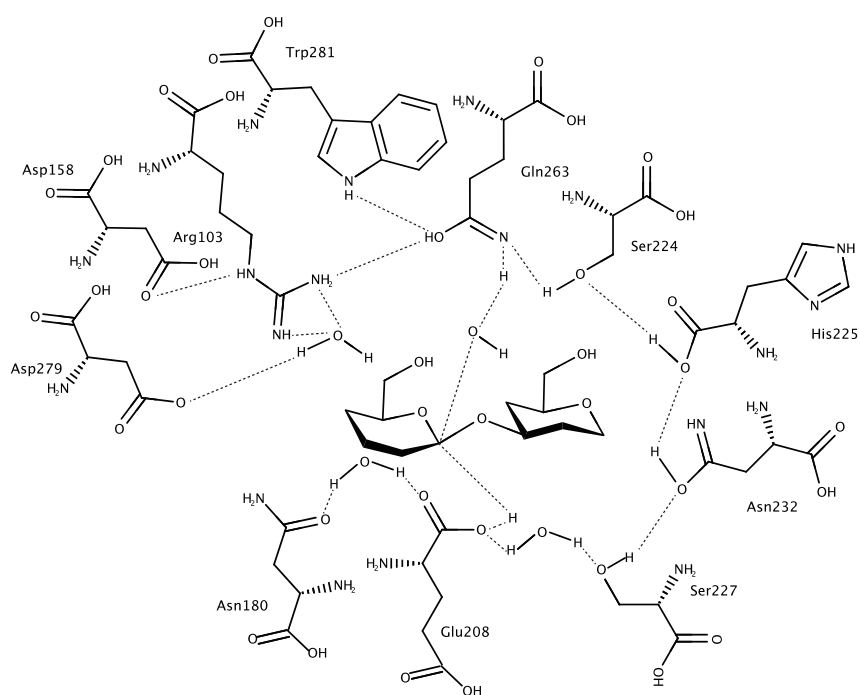
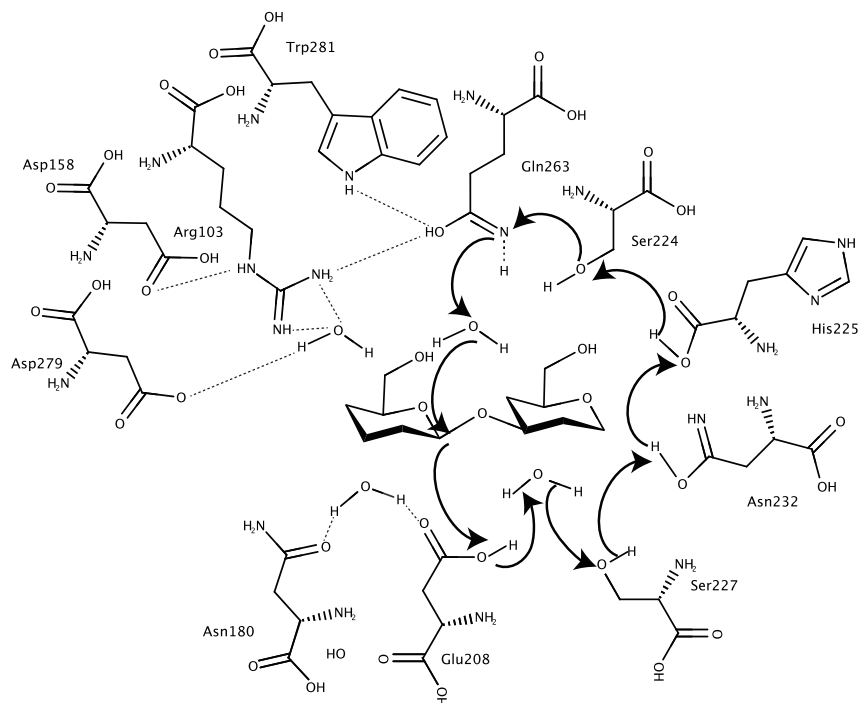


は Glu199 と Cys247 にそれぞれ置換されていた。

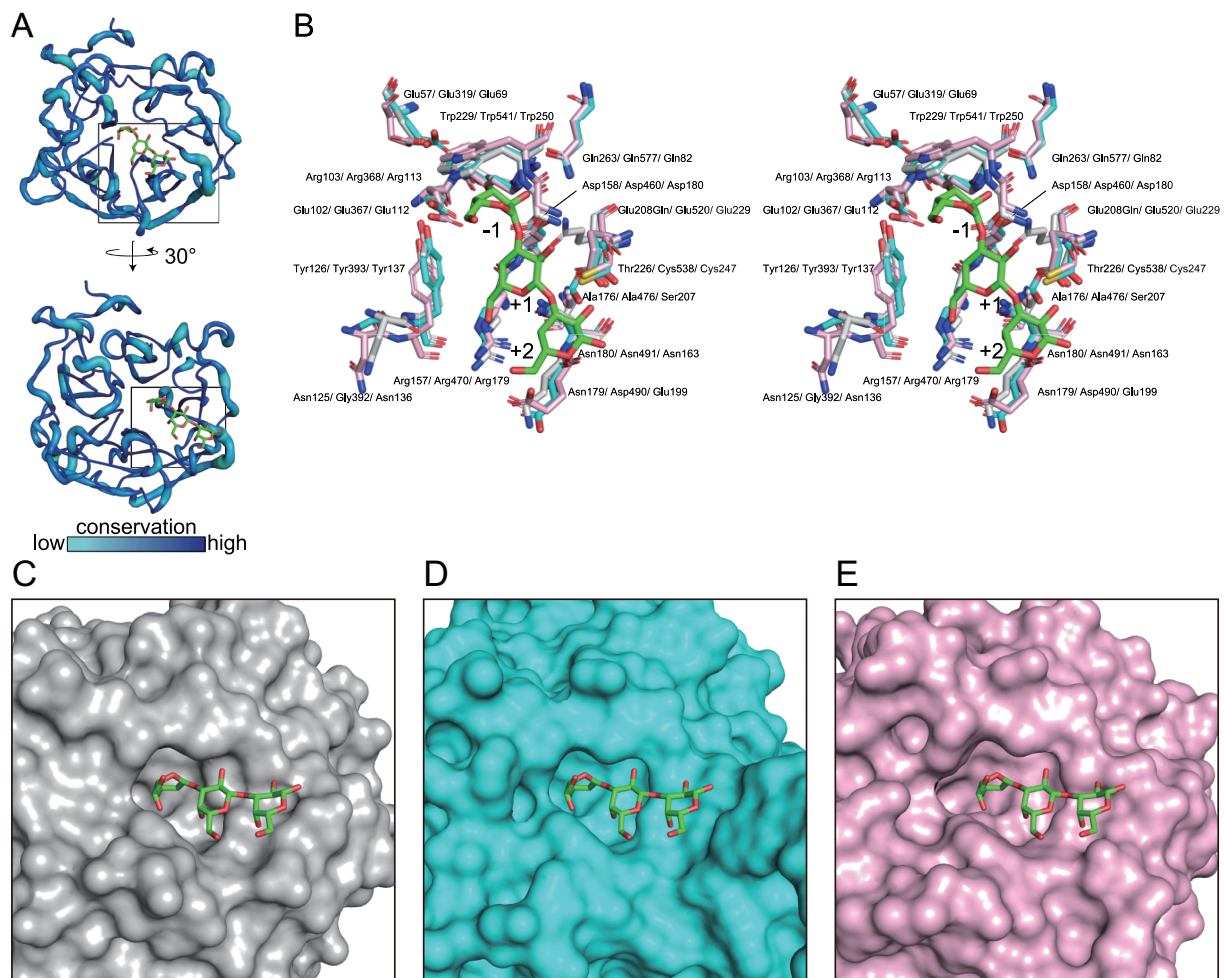


**Fig. 3-10. HPLC analysis of hydrolysis products formed from  $\beta$ -1,3-Gal2 (A) and  $\beta$ -1,3-Gal3 (B) by *Pc1,3Gal43A*\_WT and its mutants.**

Chromatograms of incubation mixtures of 0.263 mM  $\beta$ -1,3-Gal2 (A) and 0.266 mM  $\beta$ -1,3-Gal3 (B) with WT and each mutant. The enzyme (20 nM) was incubated with each substrate in 20 mM sodium acetate, pH 5.0, for 30 min at 30 °C. The reaction mixtures were subjected to HPLC. The amount of released Gal was calculated based on the peak area.

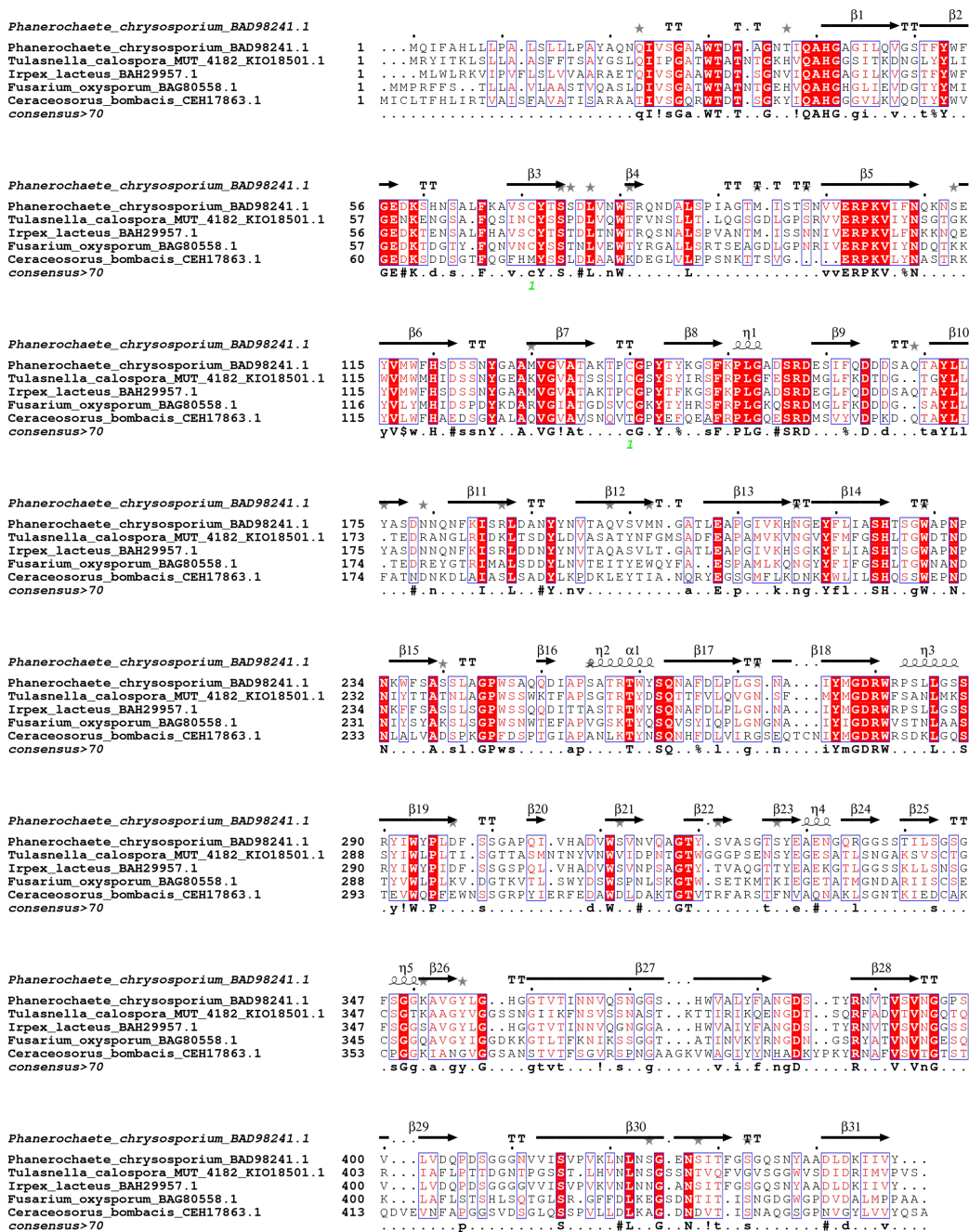
**A****B**

**Fig. 3-11. Estimated hydrogen bond networks (A) and proton transfer pathway (B) of *Pc1,3Gal43A*.**



**Fig. 3-12. Catalytic domain structure comparison.**

A: Visualization of the degree of preservation of GH43\_sub24. The degree of conservation for amino acid residues in the catalytic domain of GH43\_sub24 was visualized using the ConSurf server (<https://consurf.tau.ac.il>) (Glaser *et al.*, 2003; Landau *et al.*, 2005; Ashkenazy *et al.*, 2010, 2016; Celniker *et al.*, 2013). The BLAST query was set to *Pc1,3Gal43A* and the degree of conservation was analyzed based on 150 sequences in the ConSurf server. Conservation degrees are shown in a gradient with cyan for the lowest degree of preservation and blue for the highest. B: Catalytic domain comparison of *Pc1,3Gal43A* and two GH43\_sub24 galactanases. Catalytic center of E208Q\_Gal3 of *Pc1,3Gal43A* (white, PDB ID: 7BYV), BT3683 (cyan, PDB ID: 6EUI), and *Ct1,3Gal43A* (pink, PDB ID: 3VSF). Red, blue and yellow represent oxygen atoms, nitrogen atoms, and sulfur atoms, respectively. Residue names mean *Pc1,3Gal43A*/ BT3683/ *Ct1,3Gal43A*. C, D, E: Surface structure of *Pc1,3Gal43A* (C), BT3683 (D), and *Ct1,3Gal43A* (E). For easily comparison,  $\beta$ -1,3-Gal3 of *Pc1,3Gal43A* are shown in all three models in each structure.



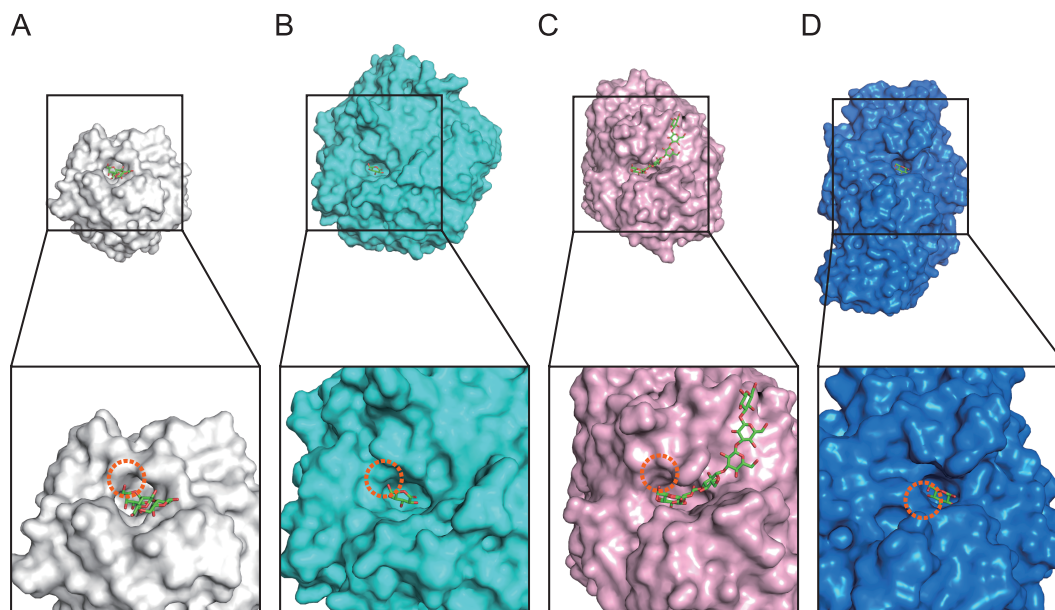
**Fig. 3-13. Sequence alignment of GH43\_sub24 galactanases of two basidiomycetes, one mold, and two bacteria.** The alignment was built by using MUSCLE on MEGAX: Molecular Evolutionary Genetics Analysis (Kumar *et al.*, 2018; Stecher *et al.*, 2020), and the figure was generated with ESPrnt 3.0 (<http://esprnt.ibcp.fr/ESPrnt/ESPrnt/>)(Robert and Gouet, 2014). The rate of conservation of the GH43\_sub24 region (up to approximately the sixth row in this alignment) is high, although C-terminal region is less well conserved because of the differences of the CBM.

### 3.4.2. 側鎖バイパス機構

$\beta$ -1,6-ガラクトサン側鎖が結合していても $\beta$ -1,3-ガラクトサン主鎖を切断できる *Pc1,3Gal43A* の側鎖バイパス機構はサブサイトにおける $\beta$ -1,3-Gal3 の 6 位メチロール基の配向から説明できる。 $\text{Gal}_{-1}$  の 6 位メチロール基は溶媒に露出しており、さらに  $\text{Gal}_{+1}$  と  $\text{Gal}_{+2}$  の 6 位メチロール基も溶媒にさらされているため、側鎖が結合していても酵素と立体障害をおこさない (Fig. 3-2) 。したがって、*Pc1,3Gal43A* は連続的に $\beta$ -1,6-ガラクトサン側鎖が結合していても $\beta$ -1,3-ガラクトサン主鎖を分解できると推察される。GH55 に分類される $\beta$ -1,3-エキソ-グルカナーゼは $\beta$ -1,6-グルカン側鎖をバイパスし、 $\beta$ -1,3-グルカン主鎖を分解して非還元末端側から Glc を遊離する酵素である。*Pc1,3Gal43A* の触媒部位の表面構造をこれらの酵素と比較すると、*Pc1,3Gal43A* は *PcLam55A* と同様に  $\text{Gal}_{-1}$  の 6 位に結合した側鎖を受容できる小さなポケットを有していた (Fig. 3-14. A、B) (Ishida *et al.*, 2009b)。また、*Pc1,3Gal43A* の  $\text{Gal}_{+1}$ 、 $\text{Gal}_{+2}$  でみられた 6 位メチロール基が溶媒に露出している特徴は *SacteLam55A* でも同様であった (Fig. 3-14. A、C) (Bianchetti *et al.*, 2015)。さらに、非還元末端の Gal に $\beta$ -1,6-結合した側鎖を許容できる構造的特徴は GH43\_sub24 に広くみられた (Fig. 3-12、3-13) 。一方、GH3 に分類される非バイパス型の *Trichoderma reesei* (*Hypocrea jecorina*) 由来 $\beta$ -グルコシダーゼ (*HjCel3A*) では、非還元末端の Glc の 6 位メチロール基が触媒ポケットの壁側を向いており、側鎖をもつ糖鎖が結合しようとする場合には立体障害が生じる (Fig. 3-14. D) ため、 $\beta$ -1,6-グルカン側鎖をバイパスできないと推察される。すなわち、側鎖をバイパスする酵素では非還元末端の糖の 6 位メチロール基の隣にある程度の大きさの空間があり、正のサブサイトに位置する主鎖の糖が溶媒側に露出することによって側鎖が結合した基質であっても触媒ポケットに収容できるようになっている。対照的に、側鎖をバイパスできない酵素ではサブサイト-1 位の近傍に側鎖を許容できるような空間がないために側鎖を受容できない (Fig. 3-14) 。一方で、本章において *Pc1,3Gal43A* と構造を比較した BT3683 や *Ct1,3Gal43A* はいずれもサブサイト-1 位近傍に空間を有するため、同様の機構で側鎖をバイパスしていると推察される (Fig. 3-12) 。しかし、*Ct1,3Gal43A* は *Pc1,3Gal43A* と比較するとバイパス活性が弱い (Ichinose *et al.*, 2005, 2006b)。触媒ポケットの構造は *Pc1,3Gal43A* と *Ct1,3Gal43A* で大きな違いはないものの、サブサイト+1 位、+2 位に存在する *Pc1,3Gal43A* の Asn179 および Thr226 が、*Ct1,3Gal43A* では Glu199 と Cys247 にそれぞれ置換されていることや *Pc1,3Gal43A* では CBM35 を有するのに対し、*Ct1,3Gal43A* では CBM13

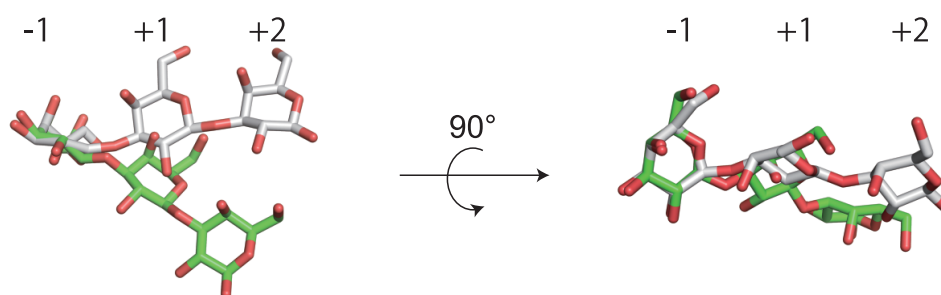
を有するという違いがある。したがって、バイパス活性の強さには正のサブサイトを構成するアミノ酸残基や触媒ドメインに結合した CBM の特性が影響をおよぼすと考えられる。

本章において複合体構造を決定した E208Q\_Gal3 では、 $\beta$ -1,3-Gal3 の電子密度が触媒部位に観測された。しかし、少なくとももう 1 つのサブサイトとなりうる構造的余裕があることから、本酵素は先行研究で述べられたように-1 から+3 の 4 つのサブサイトを有する (Ichinose *et al.*, 2005) ものと推察される。上述した通り本酵素は非還元末端 Gal の 6 位に結合した側鎖を受容できる構造をとるが、 $\beta$ -1,3/1,6-ガラクトンに対する分解活性は直鎖状の $\beta$ -1,3-ガラクトンの約 1/5 である (Ichinose *et al.*, 2005)。溶液中の $\beta$ -1,3-ガラクトンは $\beta$ -1,6-結合した Gal 側鎖との衝突を防ぐために 6-8 残基の Gal が 1 ターンを形成する右巻き三重螺旋構造をとると考えられている (Chandrasekaran and Janaswamy, 2002; Kitazawa *et al.*, 2013)。しかしながら、酵素と複合体となるとときにはガラクトンが酵素表面に固定されるために螺旋の状態が Fig. 3-15 のように通常の溶液中とは異なる状態になる。ゆえに側鎖同士、あるいは側鎖と主鎖で糖鎖の衝突が生じてしまう可能性が高い。したがって、 $\beta$ -1,3/1,6-ガラクトンを基質とした際には主鎖の螺旋構造の変化に伴う $\beta$ -1,6-ガラクトン(Gal)側鎖同士、あるいは側鎖と主鎖の間で干渉が生じ、それに伴い直鎖状の $\beta$ -1,3-ガラクトンを基質とした場合よりも分解活性が低下すると推察される。したがって直鎖状の基質と側鎖を有する分岐した基質に対する反応性の違いは酵素の触媒部位の構造ではなく、基質である糖鎖の構造的特徴に起因すると考えられる。



**Fig. 3-14. Structure comparison of the catalytic sites of *Pc1,3Gal43A* (A), GH55 exo- $\beta$ -1,3-glucanase from *P. chrysosporium* (B; *PcLam55A*; PDB ID: 3EQO), GH55 exo- $\beta$ -1,3-glucanase from *Streptomyces* sp. *SirexAA-E* (C; *SacteLam55A*; PDB ID 4PF0), GH3  $\beta$ -glucosidase from *H. jecorina* (D, PDB ID: 3ZYZ).**

A, B, and C hydrolyze the main chain of  $\beta$ -1,3-galactan (A) or  $\beta$ -1,3-glucan (B), bypassing  $\beta$ -1,6-branched side chains (Ichinose *et al.*, 2005; Ishida *et al.*, 2009b; Bianchetti *et al.*, 2015). D hydrolyzes four types of  $\beta$ -glucoside bonds, and it does not bypass side chains (Korotkova *et al.*, 2009; Karkehabadi *et al.*, 2014). The upper panels show the overall surface structure and the lower panels show an enlarged view of the catalytic region. Orange dashed circles indicate the space near the C-6 position of Gal or Glc at the non-reducing end.



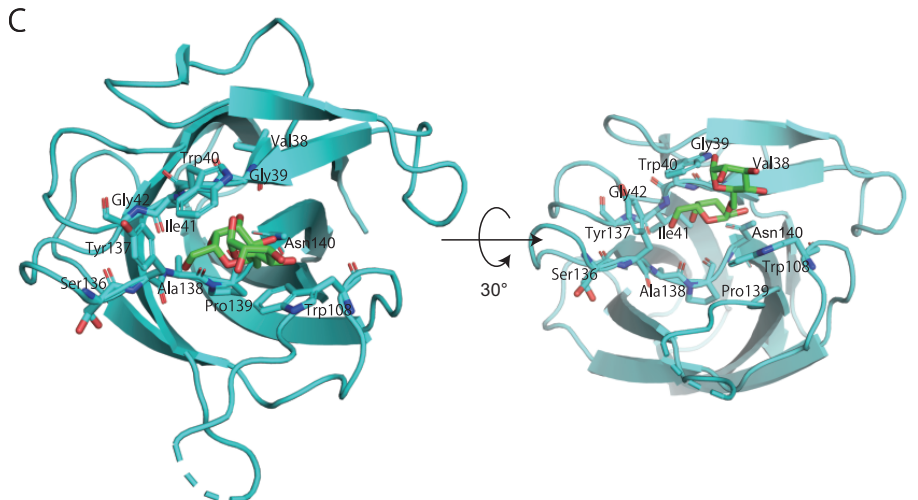
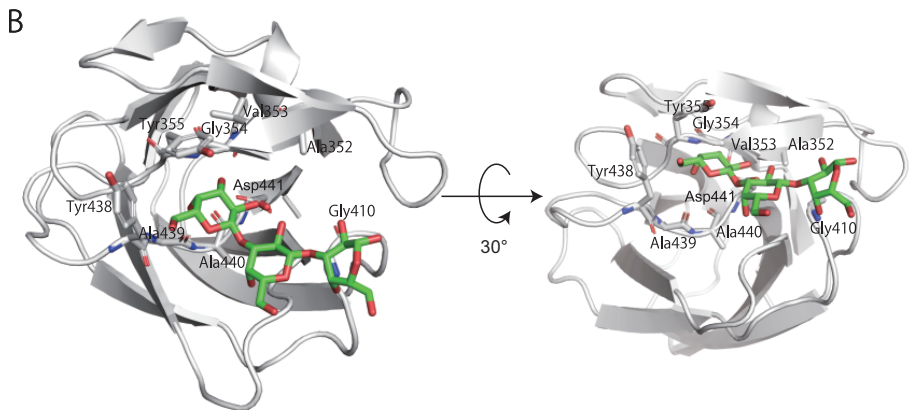
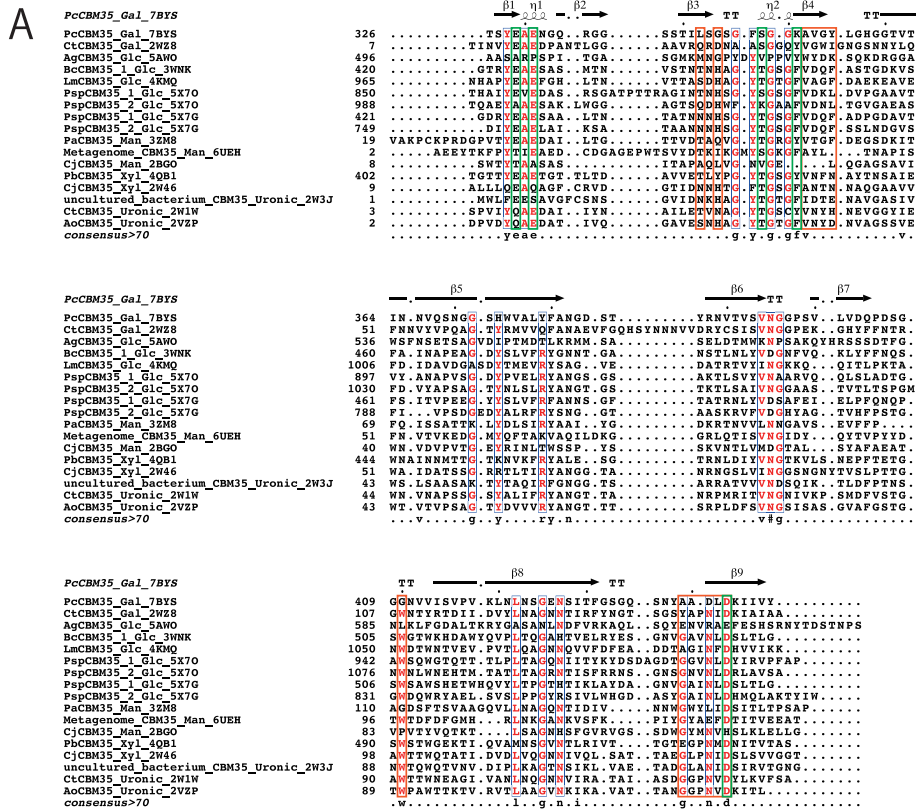
**Fig. 3-15. Comparison in the conformation of galactan.**

Comparison of Gal3 conformations between catalytic-site-bound Gal3 of *Pc1,3Gal43A*\_E208Q (white) and the "ideal" conformation of Gal3 (green). The ideal model was built by using web tool "SWEET2" (<http://www.glycosciences.de/modeling/sweet2/doc/index.php>) (Bohne *et al.*, 1998, 1999).

### 3.4.3. CBM35 のリガンド認識機構

CBM35 全体のアミノ酸配列の相同性は高くはないが、リガンド認識に関与するアミノ酸残基はよく保存されている。CBM35 に属する CBM は系統解析により 4 つのクレードに分けられるが、各クレードがそれぞれ異なる特異性を示すため、異なるクレード間でのリガンドおよび  $\text{Ca}^{2+}$  を配位する残基の多様性からリガンド特異性を説明できる (Correia *et al.*, 2010)。しかしながら、PcCBM35 のリガンド認識に関与する残基は同じクレードに属するガラクトマンナンの $\alpha$ -Gal 側鎖に結合する CBM35 のリガンド認識に関与する残基とは異なっている。この $\alpha$ -Gal に結合性を示す構造既知の CBM35 は *C. thermocellum* の cellulosomal protein の $\beta$ -キシロシダーゼであると予測されるタンパク質 (Cte\_2137; Fig. 3-15) の一部である (Correia *et al.*, 2010)。Cte\_2137 の $\alpha$ -Gal と相互作用する残基と PcCBM35 の $\beta$ -Gal と相互作用する残基にはいくつかの違いがある。例えば、PcCBM35 の Ala352-Tyr355 および Tyr438-Asp441 はリガンドの特異性に関係するが、この領域は Cte\_2137 では Val39-Gly42 および Ser136-Asn140 に相当する (Fig. 3-15. A)。特に Cte\_2137 の Asn140 はリガンド結合部位の底に位置する残基で、PcCBM35 では Asp に置換されている。またさらに、ピラノース環をスタックする役割を担う Cte\_2137 の Trp108 は PcCBM35 では Gly に置換されている (Fig. 3-15. B, C)。つまり、PcCBM35 は Cte\_2137 と同じクラスターに属するが、リガンドの認識に関与する残基は異なっており、この差異が $\beta$ -Gal と $\alpha$ -Gal、ガラクトタンとガラクトマンナンの識別に関与していると考えられる。さらに、PcCBM35 では Gal 結合サイトが 3 つあるが、Cte\_2137 に Gal- $\alpha$ -1,6-Man をモデリングしてみると、非還元末端の Gal を認識するサイトは有するが、Man の認識サイトは存在しないか認識しても微弱であると考えられた (Fig. 3-15. B, C)。したがって、PcCBM35 は複数の認識サイトを用いて $\beta$ -Gal を特異的に認識しているのに対して Cte\_2137 は主に非還元末端の $\alpha$ -Gal を認識しているといえる。





**Fig. 3-16 (previous page). Sequence alignment of structure known CBM35s (A) and structure comparison between CBM35s of *Pc1,3Gal43A* (B) and *Cte\_2137* (C).**

A: Taxon names are shown as scientific names, ligand specificity and PDB ID. When the same enzyme contains two CBM35 domains, the number indicates with 1 on the N-terminal and 2 on the C-terminal CBMs. Gal, Glc, Man, Xyl and Uronic means ligand-binding specificities. Among these, 3ZM8, 6UEH, and 2BGO, which bind to Man, are Type B CBMs which show endo-type binding, while the other 14 are all Type C CBMs, which show exo-type binding. The alignment was built by using MUSCLE on MEGAX: Molecular Evolutionary Genetics Analysis (Kumar *et al.*, 2018; Stecher *et al.*, 2020), and the figure was generated with ESPrnt 3.0 (<http://esprnt.ibcp.fr>)(Robert and Gouet, 2014). Orange and green boxes represent ligand binding and Ca<sup>2+</sup> binding residues, respectively. B and C: Ligand binding residues of *Pc1,3Gal43A* (chain A of E208A\_Gal3, PDB ID: 7BYX) and *Cte\_2137* (PDB ID: 2WZ8) with modeled ligand (Gal- $\alpha$ -1,6-Man). Red and blue mean oxygen and nitrogen, respectively. The green stick model represents  $\beta$ -1,3-Gal3.

### 3.5. 本章の結論

本章では側鎖をバイパス可能なガラクタナーゼのバイパスメカニズム CBM のリガンド認識様式の詳細な解明を目指し、*Pc1,3Gal43A* のアポ体、基質複合体の結晶構造を決定した。*Pc1,3Gal43A* は Glu208 を catalytic acid、Gln263 を catalytic base として用いる Inverting 型酵素であり、基質の側鎖が触媒反応を阻害しない位置に収容される構造を有している。したがって、本酵素は $\beta$ -1,6-ガラクトン側鎖が結合していても $\beta$ -1,3-ガラクトン主鎖を分解可能である。このように、ガラクトンをはじめとする天然に存在する多糖類には様々な分子修飾があるが、酵素はそれらの修飾に対応した構造を有し、認識できるようにしているようである。本章で得られた知見をもとに基質認識に関与する残基に変異を導入して基質認識特性を変化させた酵素を作製することにより、AGP 糖鎖の構造解析や有用なオリゴ糖の調製に役立つと期待される。

## 第四章 総括

ガラクトタンは植物の一次細胞壁および二次細胞壁、細胞間層に広く存在する多糖類であることから、植物の成長メカニズムを解明して成長をコントロールしたり、環境に適した植物に改変したり、バイオマスとして利用したりするためにはその構造や合成・分解に関与する酵素の機能を理解することが非常に重要である。そこで本研究では植物自身がペクチンの RG-I 側鎖のガラクトタンを再構成する際に用いる $\beta$ -ガラクトシダーゼと担子菌が産生する $\beta$ -1,3-ガラクトタン分解酵素の2種類の酵素に着目し、各酵素の酵素特性と立体構造との相関を調査した。

第二章では広い基質特異性を示すトマト果実由来の GH35  $\beta$ -ガラクトシダーゼ (TBG4) と3種類の結合様式のガラクトビオースの酵素基質複合体の結晶構造解析、アンサンブルリファインメント、ドッキングシミュレーションを行い、TBG4 が幅広く基質を認識するメカニズムを調査した。本研究において決定した複合体の立体構造は植物由来の $\beta$ -ガラクトシダーゼとして最初の構造であった ([http://www.cazy.org/GH35\\_structure.html](http://www.cazy.org/GH35_structure.html))。X線結晶構造解析の結果から、TBG4の基質との相互作用様式は基質の結合様式に関わらず、Gal<sub>-1</sub>では複数の水素結合を介して強く基質と相互作用していたのに対し、Gal<sub>+1</sub>とは主に疎水性相互作用を介して弱く相互作用しているのみであった (Fig. 2-3) ため、サブサイト+1 位における相互作用は弱いと明らかにした。また、TBG4は触媒ポケットの形状を入口が広い漏斗状とすることで3種類の結合様式の嵩高いガラクトタンを認識できるようにしていた。さらに、ドッキングシミュレーションの結果よりTBG4はサブサイト-1 位と+1 位だけではなく、複数の正のサブサイトを有して基質を認識する可能性が強く示された (Fig. 2-6)。したがって、TBG4はトマト果実の細胞壁中においてガラクトオリゴ糖のみではなく、長鎖の AG-I や AG-II、AGP の糖鎖領域の分解にも関与し、果実硬度の急激な低下を引き起こすと示唆された。

第三章では担子菌 *P. chrysosporium* が産生する GH43\_sub24 に分類されるエキソ- $\beta$ -1,3-ガラクタナーゼ (*Pc*1,3Gal43A) のアポ体および catalytic acid の Glu の変異体と $\beta$ -1,3-Gal3 との複合体の X線結晶構造解析およびアンサンブルリファインメントを行った。本研究により決定した構造は真核生物の GH43\_sub24 および触媒部位に Gal3 が結合した構造として新規の構造であった ([http://www.cazy.org/GH43\\_24\\_structure.html](http://www.cazy.org/GH43_24_structure.html))。他の GH43\_sub24 とのアミノ

酸配列の比較や触媒部位に結合した Gal3 と各残基との距離、アンサンブルリファインメントで得られた知見より触媒反応に重要なアミノ酸残基の変異体を作成し、それらの酵素活性の変化を調査した結果から、Pc1,3Gal43A の catalytic acid は Glu208、catalytic base は互変異性化によりイミド酸型となって機能する Gln263、非還元末端 Gal の 4 位ヒドロキシ基と水素結合を形成して Gal の認識に関与する残基は Glu102 であると特定した。また、触媒ポケットの表面構造を側鎖バイパス活性を示す GH55 エキソ-β-1,3-グルコシダーゼとバイパス活性を示さない GH3 β-グルコシダーゼと比較した結果、非還元末端の糖の 6 位ヒドロキシ基に隣接する位置に空間が存在するとβ-1,6-分岐した側鎖をバイパス可能であるが、その空間が存在しないと側鎖バイパスが不可能であると明らかにした (Fig. 3-14)。さらにβ-1,3-ガラクトランに結合する CBM35 として初めて、本酵素の CBM35 の構造を明らかにした。PcCBM35 の全体構造は他の CBM35 と類似していて、Ca<sup>2+</sup>結合サイトを 1 つ有しており (Fig. 3-1)、特に Gal3 の非還元末端の Gal を認識する type C に分類される CBM であると示した (Fig. 3-3、3-4)。

### ガラクトシダーゼとガラクタナーゼ

エキソ型ガラクトラン分解酵素には幅広くガラクトシル結合を認識して加水分解反応を触媒するβ-ガラクトシダーゼと特定の結合様式のガラクトランを特異的に認識して加水分解反応を触媒するエキソ-ガラクタナーゼの 2 種類がある。ここではこの反応性の違いを酵素の構造の違いから考察する。

さまざまな結合様式のガラクトランを加水分解可能な TBG4 とβ-1,3-結合に特異的に作用する Pc1,3Gal43A の基質認識様式を比較すると、サブサイト-1 位に位置する Gal (Gal<sub>-1</sub>) は主に水素結合によって強く認識される点は共通である (Fig. 2-3、3-2)。一方で、サブサイト+1 位の Gal (Gal<sub>+1</sub>) は TBG4 では主に疎水性相互作用による弱い相互作用でのみ認識されているに対し、Pc1,3Gal43A では疎水結合に加えて水素結合も介する比較的強い相互作用で認識されている。このようなエキソ型ガラクトラン分解酵素と同様の特徴はグルカン鎖を非還元末端から分解するβ-グルコシダーゼとエキソ-グルカナーゼでも報告されている。例えば、β-1,2、β-1,3、β-1,4、β-1,6-結合したグルカンを分解できる GH3 の HvExo1 や HjCel3A、GH1 のインドジャボク (*Rhauwolfia serpentina*) 由来ストリクトシジンβ-グルコシダーゼやトウモロコシ (*Zea mays*) 由来β-グルコシダーゼのサブサイト+1 位における酵素-基質間の相互作用は主に疎水性相互作用であると

報告されており、特に GH1 ではサブサイト+1 位を構成するアミノ酸残基の保存度は低いことから、GH1 の $\beta$ -グルコシダーゼはサブサイト+1 位を多様化させることにより、さまざまな配糖体基質に適した相互作用や触媒中心を形成していると考えられている (Czjzek *et al.*, 2001; Hrmova *et al.*, 2001, 2002; Barleben *et al.*, 2007; Karkehabadi *et al.*, 2014)。一方で、 $\beta$ -1,3-グルカンに特異的な GH55 の $\beta$ -グルコシダーゼである SacteLam55A は、サブサイト+1 位および+2 位において水素結合と芳香族残基による疎水的なスタッキング作用で基質と相互作用する (Bianchetti *et al.*, 2015)。また、イネの GH1 の $\beta$ -グルコシダーゼではサブサイト-1 位から+4 位の 5 つのサブサイトにおいてそれぞれ複数の水素結合を形成し、基質を強く認識すると報告されている。この酵素は $\beta$ -1,4-グルカン (セルロース) だけでなく $\beta$ -1,3-グルカン (ラミナリン) も分解するが、重合度の大きいラミナリンに対する分解活性は低下することが報告されている (Chuenchor *et al.*, 2008, 2011)。すなわち、ガラクトシダーゼやグルコシダーゼでは加水分解活性を示す際に非還元末端の Gal または Glc を水素結合によって強く固定するが、還元末端側の糖またはアグリコン部分での相互作用はあくまで基質を触媒ポケットの適切な位置に“固定”する役割にすぎないといえる。一方でガラクタナーゼやグルカナーゼではサブサイト-1 位だけでなく+1 位での基質との相互作用も基質認識に重要であり、さらに+2 位以降のサブサイトでも基質と相互作用する機構を有するために特異性が高くなっていると考えられる。以上より、エキソ型糖質加水分解酵素の基質特異性や結合特異性の厳密さはサブサイト+1 位以降の正のサブサイトにおける相互作用の強さと相関があり、ガラクトシダーゼ (グルコシダーゼ) であるかガラクタナーゼ (グルカナーゼ) であるかはサブサイト+1 位における基質との相互作用の強さが決定していると結論づけられる。なお、エンド型のガラクタン分解酵素では GH53 のエンド- $\beta$ -1,4-ガラクタナーゼの立体構造のみが明らかになっているが (Fig. 1-7)、GH53 のガラクタナーゼにおいても負のサブサイトだけでなくサブサイト+1、+2 位における相互作用が強い (Le Nours *et al.*, 2009) ことから、ガラクタンの結合位の認識には正のサブサイトの寄与が大きいといえる。

ガラクトシダーゼ (TBG4)、エキソ-ガラクタナーゼ (*Pc1,3Gal43A*)、*Aspergillus aculeatus* 由来 GH53 エンド-ガラクタナーゼ (*AaGal*) の表面構造を比較すると TBG4 の触媒ポケットは入口が広く開いた奥が深い漏斗状の構造をとるのに対し、*Pc1,3Gal43A* は触媒ポケット入口からポケットの底まで同程度の直径の筒状の構造をとり、さらに非還元末端の Gal ( $\text{Gal}_{-1}$ ) の 6 位の隣に小さな空間をもつ (Fig. 3-14. A, 4-1)。したがって TBG4 では触媒ポケット入口が広い

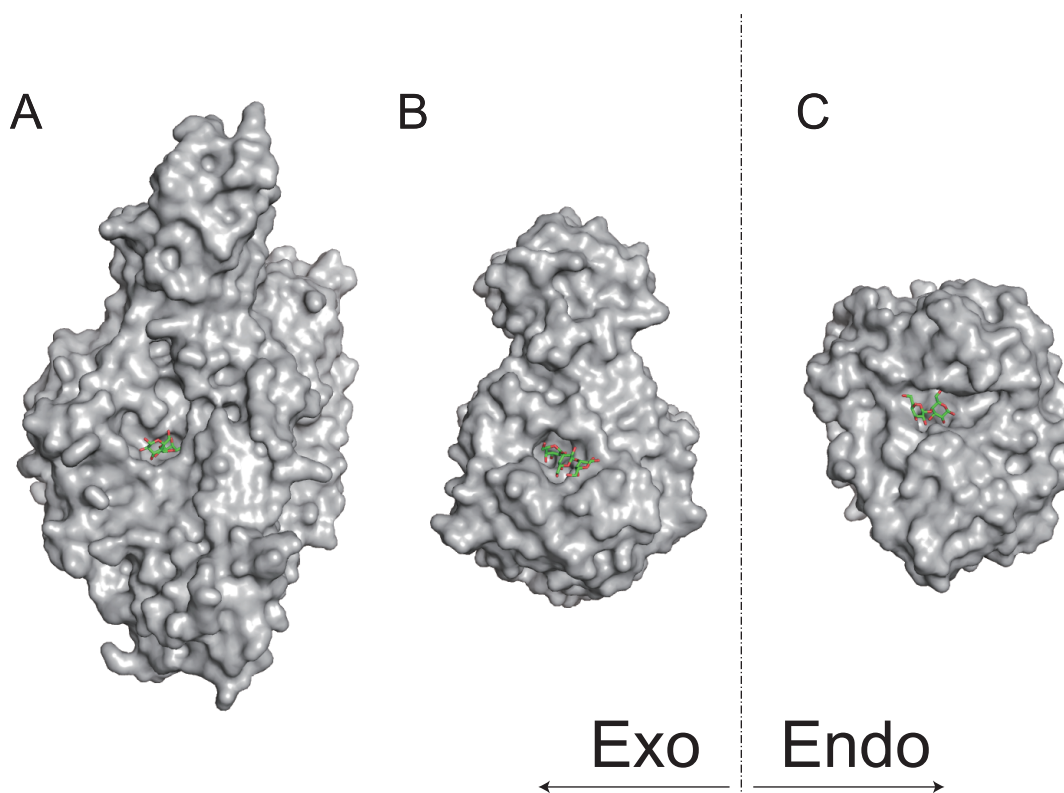
めに様々な結合様式の基質でも立体障害にならず、幅広く分解活性を示せる構造になっている。それに対して、*Pc1,3GalA* では $\beta$ -1,3-ガラクトサン主鎖の分解に特化した構造をとっている。このような構造的特徴によりTBG4が $\beta$ -1,3-ガラクトサンを分解する際には $\beta$ -1,6-ガラクトサン側鎖をバイパスせず、*Pc1,3Gal43A* は $\beta$ -1,3-結合以外のガラクトシル結合を分解しない理由が説明できる。また一方でエンド型の*AaGal* では触媒部位は浅く開いたクレフト構造をとっており、負のサブサイト側には壁となりうる構造が存在しない (Fig. 4-1. C) ためにGH53はポリマーの内部を分解することが可能であるといえる。以上より、ガラクトサン分解酵素の反応性は-2以降の負のサブサイトの存在の有無でエンド型かエキソ型が決定づけられ、そしてサブサイト+1位の認識の強さで結合位の基質特異性が高いか低いか、すなわちガラクトシダーゼかガラクターゼかが決定づけられる。

### 植物と微生物のガラクトサン分解機構

植物は $\beta$ -ガラクトシダーゼ遺伝子を複数有し、器官や成長段階により酵素を使い分けている (Smith and Gross, 2000; Tanthanuch *et al.*, 2008; Chandrasekar and van der Hoorn, 2016)。例えば、トマト果実で発現する17のTBGアイソザイムのうち酵素特性が特徴づけられているのはTBG1、TBG4、TBG5の3つである (Ishimaru *et al.*, 2009; Eda *et al.*, 2014)。このうちTBG1は果実の成熟期に恒常的に遺伝子発現量が増加し、TBG4は成熟期の中でも特に細胞壁中のガラクトシル化合物の含量が低下する時期に遺伝子発現量が顕著に増加しその後低下するのに対し、TBG5遺伝子は恒常的に発現しているが、特に結実期直後と成熟期に遺伝子の発現量が増加する。このようにこれら3つの酵素の遺伝子はいずれも成熟期に発現するために果実の軟化に関与すると考えられている (Smith and Gross, 2000)。TBG4は $\beta$ -1,3、 $\beta$ -1,4、 $\beta$ -1,6-結合のオリゴ糖に対して加水分解活性を示す (Ishimaru *et al.*, 2009) が、アイソザイムであるTBG1は $\beta$ -1,3、 $\beta$ -1,6-ガラクトオリゴ糖に対して加水分解活性を示し、さらに糖転移活性も示す (Eda *et al.*, 2014)。TBG5は $\beta$ -1,3、 $\beta$ -1,6-ガラクトオリゴ糖に対して加水分解活性を示す (Ishimaru *et al.*, 2009) ため、トマト果実では同時期に発現する特性の異なる酵素を使い分けているといえる。このような酵素の使い分けはトマトに限らず他の植物でも報告されており、例えばシロイヌナズナでもそれぞれ異なる機能を有する17の $\beta$ -ガラクトシダーゼを使い分けると報告されている (Ahn *et al.*, 2007; Dean *et al.*, 2007; Gantulga *et al.*, 2009; Sampedro

*et al.*, 2012; Chandrasekar and van der Hoorn, 2016)。

植物がエンド型のガラクトサン分解酵素を産生せず、GH35  $\beta$ -ガラクトシダーゼのみを複数産生する理由としては、細胞壁の力学的強度の維持や生理学的に重要なガラクトサンをエンド型の酵素で急激に分解するよりも、エキソ型酵素で徐々に分解する方が植物にとって望ましいからであると考えられる。特に果実の軟化時においてはペクチンメチルエステラーゼやポリガラクトツロナーゼ、 $\beta$ -ガラクトシダーゼ、アラビノフラノシダーゼなどの酵素の発現量が増加し、ペクチンの脱重合や可溶化、側鎖の中性糖の脱離を顕著に引き起こすことにより急速に細胞壁の力学的強度が低下する (Brummell, 2006; Prasanna *et al.*, 2007; Wang *et al.*, 2018)。細胞壁中では RG-I の中性糖側鎖であるガラクトサンやアラビナンなどの糖鎖がセルロースマイクロフィブリルやキシログルカンをはじめとするヘミセルロース、別の RG-I 分子と相互作用し細胞骨格の維持に関与する (Zykwinska *et al.*, 2007, 2008) ため、細胞壁の力学的強度と酵素の関係について明らかにする上では、ペクチン主鎖に作用する酵素だけでなく側鎖に作用する種々の酵素が重要である。トマト果実ではポリガラクトツロナーゼ遺伝子の発現を抑制し、ペクチンの脱重合を抑制しても果実の軟化の程度は野生型と比べ軟化の程度に大きな差はみられなかった (Kramer *et al.*, 1992) が、ガラクトシダーゼ遺伝子の発現を抑制すると果実の軟化が抑制された (Smith *et al.*, 2002) ため、特に軟化の過程においてはガラクトシダーゼ活性が非常に重要である。すなわち、植物はガラクトシダーゼ遺伝子の発現によりガラクタンの分解をコントロールすることで果実の硬度を制御しているといえる。したがって、このことから植物はガラクトサンをランダムに分解するエンド-ガラクタナーゼよりも末端から徐々に分解する $\beta$ -ガラクトシダーゼを好むと推察される。トマト果実以外にもモモ果実、カキ (*Diospyros kaki*) 果実、イチゴ果実などでも同様に果実の発達、軟化中に $\beta$ -ガラクトシダーゼ遺伝子発現量の増減がみられる (Trainotti *et al.*, 2001; Ban *et al.*, 2016; Guo *et al.*, 2018)、シロイヌナズナでは植物ホルモンの一種であるジャスモン酸に応答して遺伝子が多量に発現するアイソザイムと全く発現しないアイソザイムがある (Gantulga *et al.*, 2009)、イネでは器官により遺伝子発現量が増加するアイソザイムの種類と時期が異なる (Chantarangsee *et al.*, 2007) などといった報告がある。つまり植物は発達段階を調整したり環境に機敏に応答するために、あらゆる組織においてゆるやかにガラクトサンを分解するエキソ型の酵素を用いていると考えられる。



**Fig. 4-1. Surface structure comparison of exo- and endo-acting galactan degrading enzymes.**

A: GH35  $\beta$ -galactosidase, TBG4 (PDB ID: 6IK6), B: GH43\_sub24 exo- $\beta$ -1,3-galactatnase, Pc1,3Gal43A (PDB ID: 7BYV), C: GH53 endo- $\beta$ -1,4-galactanase, AaGal from *A. aculeatus* KSM 510, (PDB ID: 6Q3R).

真菌は様々なファミリーに属するエンド型、エキソ型のガラクトタン分解酵素を産生する (Table. 1-2) が、これはガラクトタンが存在する環境で菌が生育しようとする際には、分子量の大きい多糖類をなるべく早く菌体内に取り込める分子量にまで分解する必要があるため (Pakula *et al.*, 2005; Husain, 2010)、各結合位に特異的なエンド型およびエキソ型の酵素を用いて迅速に重合度を低下することにより、効率の良いガラクタンの分解と取り込みを可能にしていると推測される。担子菌が有するガラクトタン分解酵素遺伝子群は子のう菌が有するガラクトタン分解酵素遺伝子群に比べて種類も数も少ない (Table. 4-1)。また、真菌はガラクトタン分解酵素以外にもペクチンの主要構成成分である HG、RG-I、RG-II 主鎖を分解するエキソ-ポリガラクトツロナーゼやエンド-ポリガラクトツロナーゼ、ペクチンリアーゼや側鎖のアラビナンを分解するアラビナナーゼ、アラビノフラノシダーゼなどの様々なペクチン分解酵素を発現し (van den Brink and de Vries,



2011; Rytioja *et al.*, 2014)、これらは協調的に作用すると考えられている (Kondo *et al.*, 2020) が、やはり担子菌が産生するペクチン分解酵素遺伝子の種類や数は子のう菌に比べ非常に少ない (van den Brink and de Vries, 2011; Benoit *et al.*, 2012; Rytioja *et al.*, 2014)。また、細菌の中では例えば腸内細菌の *B. thetotaomicron* VPI-5482 株ではガラクトサン分解酵素として GH2、35、42、53、*Bifidobacterium longum* subsp. *longum* JCM1217 株では GH2、30、42、43\_sub24、53 に分類される酵素をそれぞれ有している。一方、土壌細菌の *Clostridium thermocellum* は GH30、43、53、*Streptomyces* sp. NA04227 は GH16\_sub10、35、43\_sub24、肺炎レンサ球菌の *S. pneumoniae* ATCC49619 株は GH35 に分類される酵素をそれぞれ有するなど、細菌も真菌と同様に有するガラクトサン分解酵素の種類や数は菌種によって異なる。また、ガラクトサン分解酵素群の中でも特に、GH16\_10 に分類されるエンド- $\beta$ -1,3-ガラクタナーゼは gum arabic やカラマツのアラビノガラクトサン (LWAG) のような $\beta$ -1,6-ガラクトサン側鎖の分岐頻度が高い AG-II には作用できないが、AGP の糖鎖領域を構成する AG-II のような分岐頻度の低い AG-II に対しては作用するという特徴があり (Kotake *et al.*, 2011; Kalomoiri *et al.*, 2019)、*P. chrysosporium* や *I. lactes* のような木材腐朽菌は GH16\_10 のガラクタナーゼを有さず (Table. 4-1)、*P. chrysogenum* のようなカビや、*Streptomyces* sp. NA04227 株のような土壌に生息して植物の病原菌となるような菌は GH16\_10 を有する (Table. 4-1)。また、GH35  $\beta$ -ガラクトシダーゼの中には側鎖が結合していると反応できない酵素もあり (Kondo *et al.*, 2020)、GH43\_sub24 のエキソ- $\beta$ -1,3-ガラクタナーゼのバイパス活性は細菌由来の酵素よりも真菌由来の酵素の方が高く (Ichinose *et al.*, 2005, 2006 a,b)、GH53 のエンド- $\beta$ -1,4-ガラクタナーゼの最終反応産物は細菌の酵素では様々な重合度のオリゴ糖を遊離するのにに対し真菌の酵素では単糖と二糖を遊離する (Sakamoto and Ishimaru, 2013) など、各酵素の特性は由来の菌の種類によって異なる。このように、さまざまな菌が有するガラクトサン分解酵素の種類や特性、その数や種類は各菌が生育している環境や主に炭素源としている基質の違いに起因すると考えられる。さらに各菌が有するペクチン分解酵素群 (ポリガラクトツロナーゼやペクチンリアーゼなど) や、ガラクトサンの側鎖の分解に関わる酵素 (アラビノフラノシダーゼやフコシダーゼなど) の種類や数に関しても同様の傾向がみられる (van den Brink and de Vries, 2011; Benoit *et al.*, 2012) ため、各菌のガラクトサン分解機構を明らかにするためにはガラクトシダーゼやガラクタナーゼ以外の酵素にも着目して研究を行う必要がある。したがって、食品分野で用いられるガラク

トオリゴ糖の生産やバイオマスの糖化などの産業的な用途に有用な酵素や、ガラクトランの構造解析に用いる酵素をスクリーニングする際には由来の生物種にも着目することが重要である。

以上のように本研究では GH35 に属する  $\beta$ -ガラクトシダーゼと 3 つの結合様式のガラクトビオースとの酵素基質複合体の構造 (第二章)、GH43\_sub24 に属するエキソ- $\beta$ -1,3-ガラクタナーゼのアポおよび酵素基質複合体の構造、CBM35 に属する糖質結合モジュールの立体構造 (第三章) をそれぞれ植物由来の $\beta$ -ガラクトシダーゼ、真菌由来のエキソ- $\beta$ -1,3-ガラクタナーゼ、 $\beta$ -1,3-ガラクトラン結合性を示す CBM35 としてそれぞれ初めて決定し、各酵素の基質認識機構および触媒機構と CBM のリガンド認識機構を明らかにした。ガラクトランは植物細胞壁の力学的強度を保持したり細胞壁を構築するうえで重要な多糖類であり (Moneo-Sánchez *et al.*, 2019, 2020)、特に果実や野菜などの青果物の残渣や草本バイオマスを分解・有効活用するためにはガラクトランをはじめとするペクチンの糖化も重要視されており (Edwards and Doran-Peterson, 2012)、さらにガラクトオリゴ糖は食品分野においても注目されている (Picard *et al.*, 2005; Husain, 2010) にも関わらず、これらの用途に非常に有用なガラクトラン分解酵素に関する知見はまだ十分とは言い難かった。そのような状況において本研究は植物や菌類の酵素の特性と立体構造との相関に関して基礎的な知見をもたらしたといえる。またさらに、各酵素の基質認識に関わるアミノ酸残基を特定し、認識機構を提唱したため、点変異を導入して目的の特性を有する酵素を作出するための足がかりとなる情報をもたらすことができたと考えられる。また、セルロース分解で提唱されている (Master *et al.*, 2004) のと同様にガラクトラン分解においても植物は少しずつ、微生物は急激に糖鎖の重合度を低下させる酵素を産生すると明らかにしたため、ガラクトラン分解に酵素を用いる際には目的に合わせた由来の酵素を用いることが有用であると示した。本研究で得られたこれらの知見が植物に含まれるガラクトランの機能や構造、利用に関する研究の発展に寄与し、将来的に持続可能な社会の実現の一端を担うことを願い総括とする。

**Table 4-1. Gene numbers of galactan degrading enzymes in fungi.**

	Organisms	GH5 _16	GH16 _10	GH30 _5	GH35	GH43 _24	GH53	Total
Ascomycete	<i>Aspergillus niger</i> CBS 513.88	1	0	0	5	1	1	8
	<i>Fusarium oxysporum</i> Fo47	2	1	1	6	0	2	12
	<i>Neurospora crassa</i> OR74A	1	1	1	3	4	1	11
	<i>Penicillium chrysogenum</i>	0	0	0	4	1	1	6
	<i>Saccharomyces cerevisiae</i> BJ4	0	0	0	0	0	0	0
	<i>Trichoderma reesei</i>	0	0	1	1	0	0	2
	Basidiomycete	<i>Armillaria mellea</i> DSM 3731	1	0	0	8	1	2
<i>Flammulina velutipes</i>		0	1	0	1	0	0	2
<i>Gloeophyllum trabeum</i>		1	0	0	2	1	2	6
<i>Irpex lactes</i> CCBAS Fr. 238 617/93		0	0	0	4	1	1	6
<i>Phanerochaete chrysosporium</i>		0	0	0	3	1	1	5
<i>Pleurotus eryngii</i> ATCC 90797		0	0	0	4	1	1	6
<i>Trametes versicolor</i>		0	0	0	2	1	1	4

The number of genes are based on CAZy database (<http://www.cazy.org/>) and The fungal genomes resource <https://mycocosm.jgi.doe.gov/mycocosm/home> (Grigoriev et al., 2012; Nordberg et al., 2014)

## 参考文献

- Adams PD, Afonine P V., Bunkóczi G, *et al.* 2010. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213–221.
- Afonine P V., Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwarta PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix. refine research papers. *Acta Crystallographica - Section D Biological Crystallography* **68**, 352–367.
- Ahn YO, Zheng M, Bevan DR, *et al.* 2007. Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 35. *Phytochemistry* **68**, 1510–1520.
- Albersheim P, Darvill A, Roberts K, Sederoff R, Staehelin A. 2010. *Plant Cell Walls*. Garland Science.
- Armenta S, Moreno-Mendieta S, Sánchez-Cuapio Z, Sánchez S, Rodríguez-Sanoja R. 2017. Advances in molecular engineering of carbohydrate-binding modules. *Proteins: Structure, Function and Bioinformatics* **85**, 1602–1617.
- Arnaud MB, Cerqueira GC, Inglis DO, *et al.* 2012. The Aspergillus Genome Database (AspGD): Recent developments in comprehensive multispecies curation, comparative genomics and community resources. *Nucleic Acids Research* **40**, 653–659.
- Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research* **44**, W344–W350.
- Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. 2010. ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research* **38**, 529–533.
- Baba S, Hoshino T, Ito L, Kumasaka T. 2013. Humidity control and hydrophilic glue coating applied to mounted protein crystals improves X-ray diffraction experiments. *Acta Crystallographica Section D: Biological Crystallography* **69**, 1839–1849.
- Ban Q, Han Y, Meng K, Hou Y, He Y, Rao J. 2016. Characterization of  $\beta$ -Galactosidase genes involved in persimmon growth and fruit ripening and in response to propylene and 1-methylcyclopropene. *Journal of Plant Growth Regulation* **35**, 1025–1035.
- Barleben L, Panjikar S, Ruppert M, Koepke J, Stöckigt J. 2007. Molecular architecture of strictosidine glucosidase: The gateway to the biosynthesis of the monoterpene indole alkaloid family. *Plant Cell* **19**, 2886–2897.
- Barnavon L, Doco T, Terrier N, Ageorges A, Romieu C, Pellerin P. 2000. Analysis of cell wall neutral sugar composition,  $\beta$ -galactosidase activity and a related cDNA clone throughout the development of *Vitis vinifera* grape berries. *Plant Physiology and Biochemistry* **38**, 289–300.
- Benoit I, Coutinho PM, Schols HA, Gerlach JP, Henrissat B, de Vries RP. 2012. Degradation of different pectins by fungi: correlations and contrasts between the pectinolytic enzyme sets identified in genomes and the growth on pectins of different origin. *BMC Genomics* **13**, 1.
- Berkholz DS, Shapovalov M V, Dunbrack RL, Karplus PA. 2011. Conformation dependence of backbone geometry in proteins. *Structure/Folding and Design* **17**, 1316–1325.
- Berlemont R, Martiny AC. 2016. Glycoside hydrolases across environmental microbial communities. *PLoS Computational Biology* **12**, 1–16.
- Bianchetti CM, Takasuka TE, Deutsch S, Udell HS, Yik EJ, Bergeman LF, Fox BG. 2015. Active site and laminarin binding in glycoside hydrolase family 55. *Journal of Biological Chemistry* **290**, 11819–11832.
- Bohne A, Lang E, Von Der Lieth CW Der. 1998. W3-SWEET: Carbohydrate modeling by internet. *Journal of Molecular Modeling* **4**, 33–43.
- Bohne A, Lang E, Von Der Lieth CW. 1999. SWEET - WWW-based rapid 3D construction of oligo- and polysaccharides. *Bioinformatics* **15**, 767–768.
- Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. 2004. Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochemical Journal* **382**, 769–781.

van den Brink J, de Vries RP. 2011. Fungal enzyme sets for plant polysaccharide degradation. *Applied Microbiology and Biotechnology* **91**, 1477–1492.

Brummell D a. 2006. Primary cell wall metabolism during fruit ripening. *New Zealand Journal of Forestry Science* **36**, 99–111.

Burnley BT, Gros P. 2013. phenix.ensemble\_refinement: a test study of apo and holo BACE1. *Computational crystallography newsletter* **4**, 51–58.

Burnley Tom B, Afonine P V., Adams PD, Gros P. 2012. Modelling dynamics in protein crystal structures by ensemble refinement. *eLife* **2012**, 1–29.

Burton RA, Gidley MJ, Fincher GB. 2010. Heterogeneity in the chemistry, structure and function of plant cell walls. *Nature Chemical Biology* **6**, 724–732.

Campbell ID, Spitzfaden C. 1994. Building proteins with fibronectin type III modules. *Structure* **2**, 333–337.

Carey AT, Holt K, Picard S, Wilde R, Tucker GA, Bird CR, Schuch W, Seymour CB. 1995. Tomato exo-(1→4)-β-D-galactanase. *Plant Physiology* **108**, 1099–1107.

Carpita NC, Gibeaut DM. 1993. Structural models of primary cell walls in flowering plants: Consistency of molecular structure with the physical properties of the walls during growth. *Plant Journal* **3**, 1–30.

Carrington CMS, Pressey R. 1996. β-galactosidase II activity in relation to changes in cell wall galactosyl composition during tomato ripening. *J. Amer. Soc. Hort. Sci.* **121**, 132–136.

Cartmell A, McKee LS, Pena MJ, et al. 2011. The structure and function of an arabinan-specific α-1,2- arabinofuranosidase identified from screening the activities of bacterial GH43 glycoside hydrolases. *Journal of Biological Chemistry* **286**, 15483–15495.

Cartmell A, Muñoz-Muñoz J, Briggs JA, et al. 2018. A surface endogalactanase in *Bacteroides thetaiotaomicron* confers keystone status for arabinogalactan degradation. *Nature Microbiology* **3**, 1314–1326.

Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, Mayrose I, Pupko T, Ben-Tal N. 2013. ConSurf: Using evolutionary data to raise testable hypotheses about protein function. *Israel Journal of Chemistry* **53**, 199–206.

Chanalía P, Gandhi D, Attri P, Dhanda S. 2018. Purification and characterization of β-galactosidase from probiotic *Pediococcus acidilactici* and its use in milk lactose hydrolysis and galactooligosaccharide synthesis. *Bioorganic Chemistry* **77**, 176–189.

Chandrasekar B, van der Hoorn RAL. 2016. Beta galactosidases in *Arabidopsis* and tomato—a mini review. *Biochemical Society Transactions* **44**, 150–158.

Chandrasekaran R, Janaswamy S. 2002. Morphology of Western larch arabinogalactan. *Carbohydrate Research* **337**, 2211–2222.

Chantarangsee M, Tanthanuch W, Fujimura T, Fry SC, Ketudat Cairns J. 2007. Molecular characterization of β-galactosidases from germinating rice (*Oryza sativa*). *Plant Science* **173**, 118–134.

Cheng W, Wang L, Jiang YL, Bai XH, Chu J, Li Q, Yu G, Liang QL, Zhou CZ, Chen Y. 2012. Structural insights into the substrate specificity of *Streptococcus pneumoniae* β(1,3)-galactosidase BgaC. *Journal of Biological Chemistry* **287**, 22910–22918.

Chuenchor W, Pengthaisong S, Robinson RC, et al. 2008. Structural insights into rice BGlu1 β-glucosidase oligosaccharide hydrolysis and transglycosylation. *Journal of Molecular Biology* **377**, 1200–1215.

Chuenchor W, Pengthaisong S, Robinson RC, Yuvaniyama J, Svasti J, Cairns JRK. 2011. The structural basis of oligosaccharide binding by rice BGlu1 beta-glucosidase. *Journal of Structural Biology* **173**, 169–179.

Corral-Martínez P, Driouich A, Seguí-Simarro JM. 2019. Dynamic changes in arabinogalactan-protein, pectin, xyloglucan and xylan composition of the cell wall during microspore embryogenesis in *Brassica napus*. *Frontiers in Plant Science* **10**, 1–17.

Correia MAS, Abbott DW, Gloster TM, et al. 2010. Signature active site architectures illuminate the molecular basis for ligand specificity in family 35 carbohydrate binding module. *Biochemistry* **49**, 6193–6205.

Czjzek M, Cicek M, Zamboni V, Burmeister WP, Bevan DR, Henrissat B, Esen A. 2001. Crymodel of its complex with *p*-nitrophenyl β-D-thioglucoside. *Society* **46**, 37–46.

Davies G, Henrissat B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859.

**Davies GJ, Wilson KS, Henrissat B.** 1997. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *The Biochemical journal* **321**, 557–559.

**Dean GH, Zheng H, Tewari J, et al.** 2007. The *Arabidopsis* MUM2 gene encodes a  $\beta$ -galactosidase required for the production of seed coat mucilage with correct hydration properties. *Plant Cell* **19**, 4007–4021.

**Eda M, Ishimaru M, Tada T.** 2015. Expression, purification, crystallization and preliminary X-ray crystallographic analysis of tomato  $\beta$ -galactosidase 4. *Acta Crystallographica Section F Structural Biology Communications* **71**, 153–156.

**Eda M, Ishimaru M, Tada T, Sakamoto T, Kotake T, Tsumuraya Y, Mort AJ, Gross KC.** 2014. Enzymatic activity and substrate specificity of the recombinant tomato  $\beta$ -galactosidase 1. *Journal of Plant Physiology* **171**, 1454–1460.

**Eda M, Matsumoto T, Sakamoto T, Ishimaru M, Tada T.** 2016. Structural and functional analysis of tomato  $\beta$ -galactosidase 4: insight into the substrate specificity of the fruit softening-related enzyme. *The Plant journal : for cell and molecular biology* **86**, 300–307.

**Edwards MC, Doran-Peterson J.** 2012. Pectin-rich biomass as feedstock for fuel ethanol production. *Applied Microbiology and Biotechnology* **95**, 565–575.

**El-Gebali S, Mistry J, Bateman A, et al.** 2019. The Pfam protein families database in 2019. *47*, 427–432.

**Ellis M, Egelund J, Schultz CJ, Bacic A.** 2010. Arabinogalactan-proteins: Key regulators at the cell surface? *Plant Physiology* **153**, 403–419.

**Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **66**, 486–501.

**Fincher GB, Stone BA.** 1983. Arabinogalactan-proteins: structure, biosynthesis, and function. *Annual review of plant physiology* **34**, 47–70.

**Forneris F, Burnley BT, Gros P.** 2014. Ensemble refinement shows conformational flexibility in crystal structures of human complement factor D. *Acta Crystallographica Section D Biological Crystallography* **70**, 733–743.

**Fujita K, Sakaguchi T, Sakamoto A, Shimokawa M, Kitahara K.** 2014. *Bifidobacterium longum* subsp. *longum* exo- $\beta$ -1,3-galactanase, an enzyme for the degradation of type II arabinogalactan. *Applied and Environmental Microbiology* **80**, 4577–4584.

**Fujita K, Sakamoto A, Kaneko S, Kotake T, Tsumuraya Y, Kitahara K.** 2019a. Degradative enzymes for type II arabinogalactan side chains in *Bifidobacterium longum* subsp. *longum*. *Applied Microbiology and Biotechnology* **103**, 1299–1310.

**Fujita K, Sasaki Y, Kitahara K.** 2019b. Degradation of plant arabinogalactan proteins by intestinal bacteria: characteristics and functions of the enzymes involved. *Applied Microbiology and Biotechnology* **103**, 7451–7457.

**Gantulga D, Ahn YO, Zhou C, Battogtokh D, Bevan DR, Winkel BSJ, Esen A.** 2009. Comparative characterization of the *Arabidopsis* subfamily a1  $\beta$ -galactosidases. *Phytochemistry* **70**, 1999–2009.

**Gao X, Wu J, Wu D.** 2019. Rational design of the beta-galactosidase from *Aspergillus oryzae* to improve galactooligosaccharide production. *Food Chemistry* **286**, 362–367.

**Gaspar Y, Johnson KL, Mckenna JA, Bacic A, Schultz CJ.** 2001. The complex structures of arabinogalactan-proteins and the journey towards understanding function. *Plant Molecular Biology* **47**, 161–176.

**Gilbert HJ.** 2010. The Biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiology* **153**, 444–455.

**Gilbert HJ, Knox JP, Boraston AB.** 2013. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Current Opinion in Structural Biology* **23**, 669–677.

**Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben-Tal N.** 2003. ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* **19**, 163–164.

**Gorshkova T, Chernova T, Mokshina N, Ageeva M, Mikshina P.** 2018. Plant ‘muscles’: fibers with a tertiary cell wall. *New Phytologist* **218**, 66–72.

**Gorshkova T, Mokshina N, Chernova T, et al.** 2015. Aspen tension wood fibers contain  $\beta$ -(1 $\rightarrow$ 4)-galactans and acidic arabinogalactans retained by cellulose microfibrils in gelatinous walls. *Plant Physiology* **169**, 2048–2063.

**Grigoriev I V., Nikitin R, Haridas S, et al.** 2014. MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Research* **42**, 699–704.

**Grigoriev I V., Nordberg H, Shabalov I, et al.** 2012. The genome portal of the department of energy joint genome institute. *Nucleic Acids Research* **40**, 26–32.

**Gross KC, Sams CE.** 1984. Changes in cell wall neutral sugar composition during fruit ripening: a species survey. *Phytochemistry* **23**, 2457–2461.

**Gross KC, Wallner SJ.** 1979. Degradation of cell wall polysaccharides during tomato fruit ripening. *Plant Physiology* **63**, 117–120.

**Guo S, Song J, Zhang B, Jiang H, Ma R, Yu M.** 2018. Genome-wide identification and expression analysis of beta-galactosidase family members during fruit softening of peach [*Prunus persica* (L.) Batsch]. *Postharvest Biology and Technology* **136**, 111–123.

**Harholt J, Suttangkakul A, Scheller HV.** 2010. Biosynthesis of pectin. *Plant Physiology* **153**, 384–395.

**Henrissat B, Bairoch A.** 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* **293**, 781–788.

**Henrissat B, Bairoch A.** 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal* **316**, 695–696.

**Henrissat B, Davies G.** 1997. Structural and sequence-based classification of glycoside hydrolases. *Carbohydrates and glycoconjugates 3D* **7**, 637–644.

**Henrissat B, Grenoble F.** 1991. A classification of glycosyl hydrolases based on sequence similarities amino acid. *Journal of Biochemistry* **280**, 309–316.

**Hrmova M, Fincher GB.** 1998. Barley  $\beta$ -D-glucan exohydrolases. Substrate specificity and kinetic properties. *Carbohydrate Research* **305**, 209–221.

**Hrmova M, Fincher GB.** 2001. Structure-function relationships of  $\beta$ -D-glucan endo- and exohydrolases from higher plants. *Plant Molecular Biology* **47**, 73–91.

**Hrmova M, Gori R De, Smith BJ, Fairweather JK, Driguez H, Varghese JN, Fincher GB.** 2002. Structural basis for broad substrate specificity in higher plant  $\beta$ -D-glucan glucohydrolases. *The Plant Cell* **14**, 1033–1052.

**Hrmova M, Harvey AJ, Wang J, Shirley NJ, Jones GP, Stone BA, Høj PB, Fincher GB.** 1996. Barley  $\beta$ -D-glucan exohydrolases with  $\beta$ -D-glucosidase activity. *The Journal of biological chemistry* **271**, 5277–5286.

**Hrmova M, Varghese JN, De Gori R, Smith BJ, Driguez H, Fincher GB.** 2001. Catalytic mechanisms and reaction intermediates along the hydrolytic pathway of a plant  $\beta$ -D-glucan glucohydrolase. *Structure* **9**, 1005–1016.

**Hu D, Zhang F, Zhang H, et al.** 2014. The  $\beta$ -galactosidase (BgaC) of the zoonotic pathogen *Streptococcus suis* is a surface protein without the involvement of bacterial virulence. *Scientific Reports* **4**, 1–9.

**Huijser P, Schmid M.** 2011. The control of developmental phase transitions in plants. *Development* **138**, 4117–4129.

**Husain Q.** 2010.  $\beta$  galactosidases and their potential applications: A review. *Critical Reviews in Biotechnology* **30**, 41–62.

**Ichinose H, Kotake T, Tsumuraya Y, Kaneko S.** 2006a. Characterization of an exo- $\beta$ -1,3-D-galactanase from *Streptomyces avermitilis* NBRC14893 acting on arabinogalactan-proteins. *Bioscience, Biotechnology and Biochemistry* **70**, 2745–2750.

**Ichinose H, Kotake T, Tsumuraya Y, Kaneko S.** 2008. Arabinogalactan-proteins degrading enzymes. *Journal of Applied Glycoscience* **55**, 149–155.

**Ichinose H, Kuno A, Kotake T, Yoshida M, Sakka K, Hirabayashi J, Tsumuraya Y, Kaneko S.** 2006b. Characterization of an exo- $\beta$ -1,3-galactanase from *Clostridium thermocellum*. *Applied and Environmental Microbiology* **72**, 3515–3523.

**Ichinose H, Yoshida M, Kotake T, Kuno A, Igarashi K, Tsumuraya Y, Samejima M, Hirabayashi J, Kobayashi H, Kaneko S.** 2005. An exo- $\beta$ -1,3-galactanase having a novel  $\beta$ -1,3-galactan-binding module from *Phanerochaete chrysosporium*. *Journal of Biological Chemistry* **280**, 25820–25829.

**Ishida T, Fujimoto Z, Ichinose H, Igarashi K, Kaneko S, Samejima M.** 2009a. Crystallization of selenomethionyl exo- $\beta$ -1,3-galactanase from the basidiomycete *Phanerochaete chrysosporium*. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **65**, 1274–1276.

**Ishida T, Fushinobu S, Kawai R, Kitaoka M, Igarashi K, Samejima M.** 2009b. Crystal structure of glycoside hydrolase family 55  $\beta$ -1,3-glucanase from the basidiomycete *Phenerochaete chrysosporium*. **The Journal of biological chemistry** **284**, 10100–10109.

**Ishii T.** 1997. O-acetylated oligosaccharides from pectins of potato tuber cell walls. **Plant Physiology** **113**, 1265–1272.

**Ishii T, Matsunaga T.** 1996. Isolation and characterization of a boron-rhamnogalacturonan-II complex from cell walls of sugar beet pulp. **Carbohydrate Research** **284**, 1–9.

**Ishimaru M, Smith DL, Mort AJ, Gross KC.** 2009. Enzymatic activity and substrate specificity of recombinant tomato  $\beta$ -galactosidases 4 and 5. **Planta** **229**, 447–456.

**Jeong JK, Kwon O, Lee YM, Oh DB, Lee JM, Kim S, Kim EH, Le TN, Rhee DK, Kang HA.** 2009. Characterization of the *Streptococcus pneumoniae* BgaC protein as a novel surface  $\beta$ -galactosidase with specific hydrolysis activity for the Gal $\beta$ 1-3GlcNAc moiety of oligosaccharides. **Journal of Bacteriology** **191**, 3011–3023.

**Jiang D, Fan J, Wang X, Zhao Y, Huang B, Liu J, Zhang XC.** 2012. Crystal structure of 1,3Gal43A, an exo- $\beta$ -1,3-galactanase from *Clostridium thermocellum*. **Journal of Structural Biology** **180**, 447–457.

**Kalomoiri P, Holck J, Coulomb L, Boos I, Enemark-Rasmussen K, Spodsberg N, Monrad RN, Clausen MH.** 2019. Substrate specificity of novel GH16 endo- $\beta$ -(1 $\rightarrow$ 3)-galactanases acting on linear and branched  $\beta$ -(1 $\rightarrow$ 3)-galactooligosaccharides. **Journal of Biotechnology** **290**, 44–52.

**Karkehabadi S, Helmich KE, Kaper T, et al.** 2014. Biochemical characterization and crystal structures of a fungal family 3  $\beta$ -glucosidase, Cel3A from *Hypocrea jecorina*. **Journal of Biological Chemistry** **289**, 31624–31637.

**Keegstra K.** 2010. Plant Cell Walls. **Plant Physiology** **154**, 483–486.

**Kim J, Gross KC, Solomos T.** 1991. Galactose metabolism and ethylene production during development and ripening of tomato fruit. **Postharvest Biology and Technology** **1**, 67–80.

**Kitazawa K, Tryfona T, Yoshimi Y, et al.** 2013.  $\beta$ -galactosyl Yariv reagent binds to the  $\beta$ -1,3-galactan of arabinogalactan proteins. **Plant Physiology** **161**, 1117–1126.

**Kobayashi M, Matoh T, Azuma JI.** 1996. Two chains of rhamnogalacturonan II are cross-linked by borate-diol ester bonds in higher plant cell walls. **Plant Physiology** **110**, 1017–1020.

**Komalavilas P, Mort AJ.** 1989. The acetylation of O-3 of galacturonic acid in the rhamnose-rich portion of pectins. **Carbohydrate Research** **189**, 261–272.

**Kondo T, Nishimura Y, Matsuyama K, Ishimaru M, Nakazawa M, Ueda M, Sakamoto T.** 2020. Characterization of three GH35  $\beta$ -galactosidases, enzymes able to shave galactosyl residues linked to rhamnogalacturonan in pectin, from *Penicillium chrysogenum* 31B. **Applied Microbiology and Biotechnology** **104**, 1135–1148.

**Korotkova OG, Semenova M V., Morozova V V., Zorov IN, Sokolova LM, Bubnova TM, Okunev ON, Sinitsyn AP.** 2009. Isolation and properties of fungal  $\beta$ -glucosidases. **Biochemistry (Moscow)** **74**, 569–577.

**Koshland DE.** 1953. Stereochemistry and the mechanism of enzymatic reactions. **Biological Reviews** **28**, 416–436.

**Kotake T, Dina S, Konishi T, Kaneko S, Igarashi K, Samejima M, Watanabe Y, Kimura K, Tsumuraya Y.** 2005. Molecular cloning of a  $\beta$ -galactosidase from radish that specifically hydrolyzes  $\beta$ -(1 $\rightarrow$ 3)- and  $\beta$ -(1 $\rightarrow$ 6)-galactosyl residues of arabinogalactan protein. **Plant physiology** **138**, 1563–1576.

**Kotake T, Hirata N, Degi Y, et al.** 2011. Endo- $\beta$ -1,3-galactanase from winter mushroom *Flammulina velutipes*. **Journal of Biological Chemistry** **286**, 27848–27854.

**Kotake T, Kitazawa K, Takata R, Okabe K, Ichinose H, Kaneko S, Tsumuraya Y.** 2009. Molecular cloning and expression in *Pichia pastoris* of a *Irpex lacteus* exo- $\beta$ -(1 $\rightarrow$ 3)-galactanase gene. **Bioscience, Biotechnology and Biochemistry** **73**, 2303–2309.

**Kotake T, Nakagawa N, Takeda K, Sakurai N.** 1997. Purification and characterization of wall-bound exo-1,3- $\beta$ -D-glucanase from barley (*Hordeum vulgare* L.) seedlings. **Plant and Cell Physiology** **38**, 194–200.

**Kramer M, Sanders R, Bolkan H, Waters C, Sheeny RE, Hiatt WR.** 1992. Postharvest evaluation of transgenic tomatoes with reduced levels of polygalacturonase: processing, firmness and disease resistance. **Postharvest Biology and Technology** **1**, 241–255.



- Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **372**, 774–797.
- Kubicek CP, Starr TL, Glass NL. 2014. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. *Annual Review of Phytopathology* **52**, 427–451.
- Kumar R, Henrissat B, Coutinho PM. 2019. Intrinsic dynamic behavior of enzyme:substrate complexes govern the catalytic action of  $\beta$ -galactosidases across clan GH-A. *Scientific Reports* **9**, 1–14.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**, 1547–1549.
- Lampugnani ER, Khan GA, Somssich M, Persson S. 2018. Building a plant cell wall at a glance. *Journal of Cell Science* **131**.
- Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. 2005. ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Research* **33**, 299–302.
- Laskowski RA, Swindells MB. 2011. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling* **51**, 2778–2786.
- Lau JM, McNeil M, Darvill AG, Albersheim P. 1985. Structure of the backbone of rhamnogalacturonan I, a pectic polysaccharide in the primary cell walls of plants. *Carbohydrate Research* **137**, 111–125.
- Lazan H, Ng SY, Goh LY, Ali ZM. 2004. Papaya  $\beta$ -galactosidase/galactanase isoforms in differential cell wall hydrolysis and fruit softening during ripening. *Plant Physiology and Biochemistry* **42**, 847–853.
- Li X, Chapple C. 2010. Understanding lignification: Challenges beyond monolignol biosynthesis. *Plant Physiology* **154**, 449–452.
- Liu Y, Chen Z, Jiang Z, Yan Q, Yang S. 2017. Biochemical characterization of a novel  $\beta$ -galactosidase from *Paenibacillus barengoltzii* suitable for lactose hydrolysis and galactooligosaccharides synthesis. *International Journal of Biological Macromolecules* **104**, 1055–1063.
- Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research* **42**, 490–495.
- Luis AS, Briggs J, Zhang X, et al. 2018. Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. *Nature Microbiology* **3**, 210–219.
- Majewska-Sawka A, Nothnagel EA. 2000. The multiple roles of arabinogalactan proteins in plant development. Update on extracellular matrix. *Plant Physiology* **122**, 3–9.
- Maksimainen M, Hakulinen N, Kallio JM, Timoharju T, Turunen O, Rouvinen J. 2011. Crystal structures of *Trichoderma reesei*  $\beta$ -galactosidase reveal conformational changes in the active site. *Journal of Structural Biology* **174**, 156–163.
- Martinez D, Berka RM, Henrissat B, et al. 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nature Biotechnology* **26**, 553–560.
- Mast SW, Donaldson L, Torr K, Phillips L, Flint H, West M, Strabala TJ, Wagner A. 2009. Exploring the ultrastructural localization and biosynthesis of  $\beta(1,4)$ -galactan in *Pinus radiata* compression wood. *Plant Physiology* **150**, 573–583.
- Master ER, Rudsander UJ, Zhou W, Henriksson H, Divne C, Denman S, Wilson DB, Teeri TT. 2004. Recombinant expression and enzymatic characterization of PttCe19A, a KOR homologue from *Populus tremula* x *tremuloides*. *Biochemistry* **43**, 10080–10089.
- Matsunaga T, Ishii T, Matsumoto S, Higuchi M, Darvill A, Albersheim P, O'Neill MA. 2004. Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in Pteridophytes, Lycophytes, and Bryophytes. Implications for the Evolution of Vascular Plants. *Plant Physiology* **134**, 339–351.
- Matsuyama K, Sunagawa N, Igarashi K. 2020. Mutation of cysteine residues increases heterologous expression of peach expansin in the methylotrophic yeast *Pichia pastoris*. *Plant Biotechnology* **37**, 1–7.
- McCarter JD, Withers SG. 1994. Mechanisms of enzymatic glycoside hydrolysis. *Current Opinion in Structural Biology* **4**, 885–892.

- McCartney L, Steele-King CG, Jordan E, Knox JP. 2003. Cell wall pectic (1→4)-β-D-galactan marks the acceleration of cell elongation in the *Arabidopsis* seedling root meristem. *Plant Journal* **33**, 447–454.
- McDonald AG, Boyce S, Tipton KF. 2009. ExplorEnz: The primary source of the IUBMB enzyme list. *Nucleic Acids Research* **37**, 593–597.
- Mewis K, Lenfant N, Lombard V, Henrissat B. 2016. Dividing the large glycoside hydrolase family 43 into subfamilies: A motivation for detailed enzyme characterization. *Applied and Environmental Microbiology* **82**, 1686–1692.
- Minic Z. 2008. Physiological roles of plant glycoside hydrolases. *Planta* **227**, 723–740.
- Minic Z, Jouanin L. 2006. Plant glycoside hydrolases involved in cell wall polysaccharide degradation. *Plant Physiology and Biochemistry* **44**, 435–449.
- Mohnen D. 2008. Pectin structure and biosynthesis. *Current Opinion in Plant Biology* **11**, 266–277.
- Moneo-Sánchez M, Alonso-Chico A, Knox JP, Dopico B, Labrador E, Martín I. 2019. β-(1,4)-galactan remodelling in *Arabidopsis* cell walls affects the xyloglucan structure during elongation. *Planta* **249**, 351–362.
- Moneo-Sánchez M, Vaquero-Rodríguez A, Hernández-Nistal J, Albornos L, Knox P, Dopico B, Labrador E, Martín I. 2020. Pectic galactan affects cell wall architecture during secondary cell wall deposition. *Planta* **251**, 1–15.
- Montanier C, Van Bueren AL, Dumon C, et al. 2009. Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 3065–3070.
- Nakamura A, Ishida T, Kusaka K, et al. 2015. “Newton’s cradle” proton relay with amide – imidic acid tautomerization in inverting cellulase visualized by neutron crystallography. *Science Advances* **1**:e1500263, 1–7.
- Neutelings G. 2011. Lignin variability in plant cell walls: Contribution of new models. *Plant Science* **181**, 379–386.
- Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, Smirnova T, Grigoriev I V., Dubchak I. 2014. The genome portal of the department of energy joint genome institute: 2014 updates. *Nucleic Acids Research* **42**, 26–31.
- Le Nours J, De Maria L, Welner D, Jørgensen CT, Christensen LLH, Borchert T V., Larsen S, Lo Leggio L. 2009. Investigating the binding of β-1,4-galactan to *Bacillus licheniformis* β-1,4-galactanase by crystallography and computational modeling. *Proteins: Structure, Function and Bioinformatics* **75**, 977–989.
- O’Donoghue EM, Somerfield SD, Sinclair BK, King GA. 1998. Characterization of the harvest-induced expression of β-galactosidase in *Asparagus officinalis*. *Plant Physiology and Biochemistry* **36**, 721–729.
- O’Neill MA, York WS. 2003. *The Plant Cell Wall* (JKC ROSE, Ed.). Blackwell Publishing.
- Øbro J, Harholt J, Scheller HV, Orfila C. 2004. Rhamnogalacturonan I in *Solanum tuberosum* tubers contains complex arabinogalactan structures. *Phytochemistry* **65**, 1429–1438.
- Ogasawara S, Abe K, Nakajima T. 2007. Pepper β-galactosidase 1 (PBG1) plays a significant role in fruit ripening in bell pepper (*Capsicum annuum*). *Bioscience, Biotechnology and Biochemistry* **71**, 309–322.
- Okawa M, Fukamachi K, Tanaka H, Sakamoto T. 2013. Identification of an exo-β-1,3-D-galactanase from *Fusarium oxysporum* and the synergistic effect with related enzymes on degradation of type II arabinogalactan. *Applied Microbiology and Biotechnology* **97**, 9685–9694.
- Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307–326.
- Oxley D, Bacic A. 1999. Structure of the glycosylphosphatidylinositol anchor of an arabinogalactan protein from *Pyrus communis* suspension-cultured cells. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 14246–14251.
- Pakula TM, Salonen K, Uusitalo J, Penttilä M. 2005. The effect of specific growth rate on protein synthesis and secretion in the filamentous fungus *Trichoderma reesei*. *Microbiology* **151**, 135–143.

- Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink A, Xiong G.** 2013. Hemicellulose biosynthesis. *Planta* **238**, 627–642.
- Pellerin P, Brillouet JM.** 1994. Purification and properties of an exo-(1→3)-β-D-galactanase from *Aspergillus niger*. *Carbohydrate Research* **264**, 281–291.
- Perrakis A, Morris R, Lamzin VS.** 1999. Automated protein model building combined with iterative structure refinement. *Nature Structural Biology* **6**, 458–463.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE.** 2004. UCSF Chimera — A visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**, 1605–1612.
- Picard C, Fioramonti J, Francois A, Robinson T, Neant F, Matuchansky C.** 2005. Review article: Bifidobacteria as probiotic agents - Physiological effects and clinical benefits. *Alimentary Pharmacology and Therapeutics* **22**, 495–512.
- Prasanna V, Prabha TN, Tharanathan RN.** 2007. Fruit ripening phenomena-an overview. *Critical Reviews in Food Science and Nutrition* **47**, 1–19.
- Pressey R.** 1983. β-galactosidases in ripening tomatoes. *Plant Physiology*, 132–135.
- Redgwell RJ, Fischer M, Kendal E, MacRae EA.** 1997. Galactose loss and fruit ripening: High-molecular-weight arabinogalactans in the pectic polysaccharides of fruit cell walls. *Planta* **203**, 174–181.
- Rico-Díaz A, Ramírez-Escudero M, Vizoso-Vázquez Á, Cerdán ME, Becerra M, Sanz-Aparicio J.** 2017. Structural features of *Aspergillus niger* β-galactosidase define its activity against glycoside linkages. *FEBS Journal* **284**, 1815–1829.
- Ridley BL, O'Neill MA, Mohnen D.** 2001. Pectins: Structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* **57**, 929–967.
- Roach MJ, Mokshina NY, Badhan A, Snegireva A V., Hobson N, Deyholos MK, Gorshkova TA.** 2011. Development of cellulosic secondary walls in flax fibers requires β-galactosidase. *Plant Physiology* **156**, 1351–1363.
- Robert X, Gouet P.** 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42**, 320–324.
- Rose JKC, Anderson M, Roberts J, et al.** 2003. *The Plant Cell Wall* (JKC ROSE, Ed.). Blackwell Publishing.
- Rose JKC, Braam J, Fry SC, Nishitani K.** 2002. The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: Current perspectives and a new unifying nomenclature. *Plant and Cell Physiology* **43**, 1421–1435.
- Rye CS, Withers SG.** 2000. Glycosidase mechanisms. *Current Opinion in Chemical Biology* **4**, 573–580.
- Rytioja J, Hildén K, Yuzon J, Hatakka A, de Vries RP, Mäkelä MR.** 2014. Plant-polysaccharide-degrading enzymes from basidiomycetes. *Microbiology and Molecular Biology Reviews* **78**, 614–649.
- Ryttersgaard C, Lo Leggio L, Coutinho PM, Henrissat B, Larsen S.** 2002. *Aspergillus aculeatus* β-1,4-galactanase: Substrate recognition and relations to other glycoside hydrolases in clan GH-A. *Biochemistry* **41**, 15135–15143.
- Sakamoto T, Ishimaru M.** 2013. Peculiarities and applications of galactanolytic enzymes that act on type I and II arabinogalactans. *Applied Microbiology and Biotechnology* **97**, 5201–5213.
- Sampedro J, Gianzo C, Iglesias N, Guitián E, Revilla G, Zarra I.** 2012. AtBGAL10 is the main xyloglucan β-galactosidase in *Arabidopsis*, and its absence results in unusual xyloglucan subunits and growth defects. *Plant Physiology* **158**, 1146–1157.
- Saulnier L, Thibault JF.** 1999. Ferulic acid and diferulic acids as components of sugar-beet pectins and maize bran heteroxylans. *Journal of the Science of Food and Agriculture* **79**, 396–402.
- Scheller HV, Ulvskov P.** 2010. Hemicelluloses. *Annual Review of Plant Biology* **61**, 263–289.
- Seifert GJ, Roberts K.** 2007. The biology of arabinogalactan proteins. *Annual Review of Plant Biology* **58**, 137–161.
- Smith DL, Abbott JA, Gross KC.** 2002. Down-regulation of tomato β-galactosidase 4 results in decreased fruit softening. *Plant Physiology* **129**, 1755–1762.

**Smith DL, Gross KC.** 2000. A family of at least seven  $\beta$ -galactosidase genes is expressed during tomato fruit development. *Plant Physiology* **123**, 1173–1183.

**Smith DL, Starrett DA, Gross KC.** 1998. A gene coding for tomato fruit  $\beta$ -galactosidase II is expressed during fruit ripening. *Plant Physiology* **117**, 417–423.

**Somerville C, Bauer S, Brininstool G, et al.** 2004. Toward a systems approach to understanding plant cell walls. **306**, 2206–2211.

**Speck T, Burgert I.** 2011. Plant stems: Functional design and mechanics. *Annual Review of Materials Research* **41**, 169–193.

**Stecher G, Tamura K, Kumar S.** 2020. Molecular evolutionary genetics analysis (MEGA) for macOS. *Molecular Biology and Evolution* **37**, 1237–1239.

**Swenson HA, Kaustinen HM, Bachhuber JJ, Carlson JA.** 1969. Fractionation and characterization of larchwood arabinogalactan polymers A and B. *Macromolecules* **2**, 142–145.

**Taiz L, Zeiger E.** 2010. *Plant physiology*. Sinauer Associates, Inc.

**Talens-Perales D, Polaina J, Marín-Navarro J.** 2016. Structural dissection of the active site of *Thermotoga maritima*  $\beta$ -galactosidase identifies key residues for transglycosylating activity. *Journal of Agricultural and Food Chemistry* **64**, 2917–2924.

**Tanthanuch W, Chantarangsee M, Maneesan J, Ketudat-Cairns J.** 2008. Genomic and expression analysis of glycosyl hydrolase family 35 genes from rice (*Oryza sativa* L.). *BMC Plant Biology* **8**.

**Terra VS, Homer KA, Rao SG, Andrew PW, Yesilkaya H.** 2010. Characterization of novel  $\beta$ -galactosidase activity that contributes to glycoprotein degradation and virulence in *Streptococcus pneumoniae*. *Infection and Immunity* **78**, 348–357.

**Terwilliger TC, Berendzen J.** 1999. Automated MAD and MIR structure solution. *Acta Crystallographica Section D: Biological Crystallography* **55**, 849–861.

**Torode TA, O'Neill R, Marcus SE, et al.** 2018. Branched pectic galactan in phloem-sieve-element cell walls: Implications for cell mechanics. *Plant Physiology* **176**, 1547–1558.

**Trainotti L, Spinello R, Piovan A, Spolaore S, Casadoro G.** 2001.  $\beta$ -galactosidases with a lectin-like domain are expressed in strawberry. *Journal of Experimental Botany* **52**, 1635–1645.

**Trott O, Olson A.** 2010. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry* **31**, 455–461.

**Tsumuraya Y, Hashimoto Y, Yamamoto S, Shibuya N.** 1984. Structure of L-arabino-D-galactan-containing glycoproteins from radish leaves. *Carbohydrate Research* **134**, 215–228.

**Tsumuraya Y, Kotake T.** 2017. Structure and functions of arabinogalactan-proteins, a family of plant proteoglycans. *Seikagaku* **89**, 498–507.

**Tsumuraya Y, Mochizuki N, Hashimoto Y, Kovac P.** 1990. Purification of an Exo- $\beta$ -(1-3)-D-galactanase of *Irpex lacteus* (*Polyporus tulipiferae*) and its action on arabinogalactan-proteins. *The Journal of biological chemistry* **265**, 7207–7215.

**Tung C, Huang J, Yang J.** 2007. Kappa-alpha plot derived structural alphabet and BLOSUM-like substitution matrix for rapid search of protein structure database. *Genome Biology* **8**, R31.

**Ulvskov P, Wium H, Bruce D, Jørgensen B, Qvist KB, Skjøt M, Hepworth D, Borkhardt B, Sørensen SO.** 2005. Biophysical consequences of remodeling the neutral side chains of rhamnogalacturonan I in tubers of transgenic potatoes. *Planta* **220**, 609–620.

**Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN.** 2004. REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallographica Section D: Biological Crystallography* **60**, 2184–2195.

**Vagin A, Teplyakov A.** 2010. Molecular replacement with MOLREP. *Acta Crystallographica Section D: Biological Crystallography* **66**, 22–25.

**Vasella A, Gideon J Davies, Böhm M.** 2002. Glycosidase mechanisms. *Current Opinion in Structural Biology* **6**, 619–629.

**Wallace AC, Laskowski RA, Thornton JM.** 1995. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering* **8**, 127–134.

- Wang H, Luo H, Bai Y, Wang Y, Yang P, Shi P, Zhang W, Fan Y, Yao B. 2009. An acidophilic  $\beta$ -galactosidase from bisporasp. MEY-1 with high lactose hydrolytic activity under simulated gastric conditions. *Journal of Agricultural and Food Chemistry* **57**, 5535–5541.
- Wang D, Yeats TH, Uluisik S, Rose JKC, Seymour GB. 2018. Fruit Softening: Revisiting the Role of Pectin. *Trends in Plant Science* **23**, 302–310.
- Whitcombe AJ, O'Neill MA, Steffan W, Albersheim P, Darvill AG. 1995. Structural characterization of the pectic polysaccharide, rhamnogalacturonan-II. *Carbohydrate Research* **271**, 15–29.
- Willför S, Sjöholm R, Laine C, Holmbom B. 2002. Structural features of water-soluble arabinogalactans from Norway spruce and Scots pine heartwood. *Wood Science and Technology* **36**, 101–110.
- Winn MD, Ballard CC, Cowtan KD, et al. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67**, 235–242.
- Wymelenberg A Vanden, Minges P, Sabat G, et al. 2006. Computational analysis of the *Phanerochaete chrysosporium* v2.0 genome database and mass spectrometry identification of peptides in ligninolytic cultures reveal complex mixtures of secreted proteins. *Fungal Genetics and Biology* **43**, 343–356.
- Wymelenberg A Vanden, Sabat G, Martinez D, Rajangam AS, Teeri TT, Gaskell J, Kersten PJ, Cullen D. 2005. The *Phanerochaete chrysosporium* secretome: Database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium. *Journal of Biotechnology* **118**, 17–34.
- Yang H, Liu J, Dang M, Zhang B, Li H, Meng R, Qu D, Yang Y, Zhao Z. 2018. Analysis of  $\beta$ -galactosidase during fruit development and ripening in two different texture types of apple cultivars. *Frontiers in Plant Science* **9**, 1–13.
- Yang J, Tung C. 2006. Protein structure database search and evolutionary classification. *Nucleic Acids Research* **34**, 3646–3659.
- Zabackis E, Huang J, Müller B, Darvill AG, Albersheim P. 1995. Characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves. *Plant Physiology* **107**, 1129–1138.
- Zykwinska A, Thibault JF, Ralet MC. 2007. Organization of pectic arabinan and galactan side chains in association with cellulose microfibrils in primary cell walls and related models envisaged. *Journal of Experimental Botany* **58**, 1795–1802.
- Zykwinska A, Thibault JF, Ralet MC. 2008. Competitive binding of pectin and xyloglucan with primary cell wall cellulose. *Carbohydrate Polymers* **74**, 957–961.
- 吉澤伸夫. 2016. あて材の科学 樹木の重力応答と生存戦略 (日本木材学会組織と材質研究会編, Ed.). 滋賀: 海青社.
- 西谷和彦, 梅澤俊明. 2013. 植物細胞壁. 東京: 講談社.

## 謝辞

本論文は近畿大学生物理工学部および近畿大学大学院生物理工学研究科博士前期課程での4年間と東京大学大学院農学生命科学研究科博士課程での3年間の合計7年間の研究内容をまとめたものです。

本研究を遂行するにあたり、実験や論文作成の際に熱心で温かいご指導を賜り、また博士課程からの入学であったにも関わらず修士課程までの研究を継続した上で森林化学研究室で勉強する機会をくださいました五十嵐圭日子准教授に厚く御礼申し上げます。ディスカッションや何気ないお話から研究に関して色々と気づきを与えてくださっただけでなく、様々な物事に関する考え方も教えていただき、とても大きな糧となりました。今後も先生にいただいたご指導を胸に精進し続けていきます。

実験操作や機器の扱い方を教えてください、また研究に関して熱心にディスカッションしてくださっただけでなく、様々な相談にも親身にのっていただきました砂川直輝特任助教に深く感謝いたします。砂川特任助教の研究知識の深さと広さ、そしてなにより人柄の温かさにたくさん助けられました。

研究室に温かく迎えてくださり、研究に関して沢山のご助言をくださった東京大学大学院農学生命科学研究科鮫島正浩名誉教授に厚く御礼申し上げます。いつも何かと気にかけていただき、研究との向き合い方をはじめ、たくさんのごことを教えていただきました。

学会等でいつも鋭いアドバイスやご助言をくださり、研究の進捗を気にかけてくださった徳安健特任准教授に深く感謝いたします。いつも研究室を支え、また温かく見守ってくださった寺田珠美助教に厚く御礼申し上げます。細かな実験作業や事務作業など多くの仕事をお引き受けいただきました古久保美樹研究員および中山暢子氏に深く感謝いたします。お仕事の合間に色々とお話しさせていただけたのもとても楽しく、嬉しかったです。

研究の世界の面白さを教えてくださり、また研究内容を継続したままでの東京大学への進学を認めていただきました近畿大学大学院生物理工学研究科の石丸恵教授に厚く感謝いたします。学部3年の頃から色々な大学に実験をしに伺う機会を作っていただいたおかげで博士課程への進学を決意できました。学位取得後もこれまでの気持ちを忘れずに努力します。遺伝子の扱いやタンパク質の生産や精製、酵素活性の測定方法など様々な実験技術を教えてくださいました近畿大学の先輩で現在京都大学大学院農学研究科生存圏研究所の近藤辰哉博士研究員に深く感謝申し上げます。一緒に実験していた頃も所属が離れてからも、色々と相談にのっていただき、沢山の確なご助言をいただきました。

主にガラクトオリゴ糖の調製や本研究とは別のテーマの酵素反応の反応産物の解析などでご指導を賜りました大阪府立大学生命環境科学研究科の阪本龍司教授に深く感謝いたします。阪本教授の研究室には学部生の頃からお世話になり、糖質関連酵素に関する研究の面白さを教えていただきました。また、博士課程進学後も大変気にかけてくださいました。

トマト果実の $\beta$ -ガラクトシダーゼの構造解析にあたりご指導を賜りました大阪府立大学大学院理学系研究科の多田俊治名誉教授、長浜バイオ科学技術大学バイオサイエンス学部の中江摂助手に深く感謝いたします。X線結晶構造解析の実験操作やデータの解析方法だけでなく理論についても基礎の基礎から熱心にご指導いただき、また Spring-8 では徹夜での実験に何度もご協力いただきました。

I would like to express my appreciation to Drs. Jerry Ståhlberg, Mats Sandran, Miao Wu, Mikael Gudmundsson, and lab members in Swedish University of Agricultural Sciences at Uppsala. They gave me a lot of advice on my research.

担子菌のガラクタナーゼの構造解析にあたりお世話になりました琉球大学農学部の金子哲教授、埼玉大学大学院理工学研究科の円谷陽一名誉教授、小竹敬久教授、農研機構高度解析センターの藤本瑞博士、岸根尚美氏、航空宇宙研究開発機構の石田卓也博士に深く感謝いたします。非常に質の高い回折データをご提供いただき、様々な解析を行うことができました。また、金子教授には論文の投稿にあたり、藤本博士には構造精密化にあたり非常にご尽力いただきましたことに御礼申し上げます。

埼玉大学大学院理工学研究科の小竹敬久教授、東京大学大学院農学生命科学研究科応用生命工学専攻の伏信進矢教授、生物材料科学専攻の岩田忠久教授、横山朝哉准教授には副査として本論文の審査をしていただきました。的確なご指摘や貴重なご助言を賜り、考察を深めることができましたことに感謝いたします。

7年間の研究室生活を通じ、公私にわたり大変お世話になりました東京大学大学院農学生命科学研究科生物材料科学専攻森林化学研究室の先輩方や学生のみなさま、近畿大学生物理工学部食品安全工学科食品保全学研究室の卒業生のみなさま、大阪府立大学大学院生命環境化学研究科応用生命科学専攻生物資源循環工学研究グループのみなさま、大阪府立大学大学院理学系研究科生物科学専攻藤井郁夫研究室のみなさまに厚く御礼申し上げます。また、学会や研究会などで活発な議論をしていただきました先生方、学生のみなさまに感謝いたします。

これまでの学生生活を通して京都大学大学院農学研究科生存圏研究所の今井友也准教授、大阪府立大学生命環境科学研究科の炭谷順一准教授、東京農工大学農学部の伴琢也准教授、和歌山県果樹試験場のみなさまをはじめとする多くの方々に大変ご尽力いただきましたことに御礼申し上げます。

いつも応援し、励ましてくださった友人や関わってくださったすべての方々に厚く感謝いたします。そしてどんな時もそばで支え、常に励ましてくださった奈良大輔さんに深く感謝いたします。最後に、当初の予定よりもはるかに長い9年間もの大学・大学院生活の期間中、私のやりたいことを応援し支えてくださった両親、弟、妹、祖母、叔母に感謝を表し、謝辞といたします。

2021年1月

松山 佳織



## 出版論文

### 【博士論文に関連する投稿論文】

1. Kaori Matsuyama, Tatsuya Kondo, Kiyohiko Igarashi, Tatsuji Sakamoto, Megumi Ishimaru. Substrate-recognition mechanism of tomato  $\beta$ -galactosidase 4 using X-ray crystallography and docking simulation. ***Planta***, 2020, Vol.252, No. 72.  
(第二章)
2. Kaori Matsuyama, Naomi Kishine, Zui Fujimoto, Naoki Sunagawa, Toshihisa Kotake, Yoichi Tsumuraya, Masahiro Samejima, Kiyohiko Igarashi, Satoshi Kaneko. Unique active site and subsite features in the arabinogalactan-degrading GH43 exo- $\beta$ -1,3-galactanase from *Phanerochaete chrysosporium*. ***Journal of the Biological Chemistry***, 2020, Vol. 295, No. 52, pp 18539-18552. (第三章)

### 【その他の投稿論文】

1. 松山佳織, 伴琢也, 石丸恵. 温風処理がブルーベリー果実の貯蔵中の品質に及ぼす影響. ***日本食品保蔵科学会誌***, 2019, Vol. 45, No. 4, pp 169-174.
2. Kaori Matsuyama, Naoki Sunagawa, Kiyohiko Igarashi. Mutation of cysteine residues increases heterologous expression of peach expansin in the methylotrophic yeast *Pichia pastoris*. ***Plant Biotechnology***, 2020, Vol. 37, No. 4, pp 397-403.
3. Tatsuya Kondo, Yuichi Nishimura, Kaori Matsuyama, Megumi Ishimaru, Masami Nakazawa, Mitsuhiro Ueda, Tatsuji Sakamoto. Characterization of GH35  $\beta$ -galactosidases, enzymes able to shave galactosyl residues linked to rhamnogalacturonan in pectin, from *Penicillium chrysogenum* 31B. ***Applied Microbiology and Biotechnology***, 2020, Vol. 104, pp 1135-1148.





# Substrate-recognition mechanism of tomato $\beta$ -galactosidase 4 using X-ray crystallography and docking simulation

Kaori Matsuyama<sup>1,3</sup> · Tatsuya Kondo<sup>2</sup> · Kiyohiko Igarashi<sup>1</sup> · Tatsuji Sakamoto<sup>2</sup> · Megumi Ishimaru<sup>3</sup>

Received: 8 June 2020 / Accepted: 22 September 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

**Main conclusion** TBG4 recognize multiple linkage types substrates due to having a spatially wide subsite + 1. This feature allows the degradation of AGI, AGII, and AGP leading to the fruit ripening.

**Abstract**  $\beta$ -galactosidase (EC 3. 2. 1. 23) catalyzes the hydrolysis of  $\beta$ -galactan and release of D-galactose. Tomato has at least 17  $\beta$ -galactosidases (TBGs), of which, TBG 4 is responsible for fruit ripening. TBG4 hydrolyzes not only  $\beta$ -1,4-bound galactans, but also  $\beta$ -1,3- and  $\beta$ -1,6-galactans. In this study, we compared each enzyme–substrate complex using X-ray crystallography, ensemble refinement, and docking simulation to understand the broad substrate-specificity of TBG4. In subsite – 1, most interactions were conserved across each linkage type of galactobioses; however, some differences were seen in subsite + 1, owing to the huge volume of catalytic pocket. In addition to this, docking simulation indicated TBG4 to possibly have more positive subsites to recognize and hydrolyze longer galactans. Taken together, our results indicated that during tomato fruit ripening, TBG4 plays an important role by degrading arabinogalactan I (AGI), arabinogalactan II (AGII), and the carbohydrate moiety of arabinogalactan protein (AGP).

**Keywords**  $\beta$ -galactosidase · Substrate-recognition · X-ray crystallography · Simulation

## Abbreviations

TBG4	Tomato $\beta$ -galactosidase 4
AGI	Arabinogalactan I
AGII	Arabinogalactan II
AGP	Arabinogalactan protein
GH	Glycoside hydrolase
WT_Gal	Wild type_D-galactose complex structure
E181A_ $\beta$ -1,3-Gal <sub>2</sub>	E181A_ $\beta$ -1,3-galactobiose complex
E181A_ $\beta$ -1,4-Gal <sub>2</sub>	E181A_ $\beta$ -1,4-galactobiose complex

E181A_ $\beta$ -1,6-Gal <sub>2</sub>	E181A_ $\beta$ -1,6-galactobiose complex
CBM	Carbohydrate-binding module
DP	Degree of polymerization

## Introduction

Plant primary cell wall has a very complicated structure and is mostly composed of cellulose, hemicellulose, pectin, and proteoglycans, with structural proteins filling the gaps. Especially during fruit development, the cell wall structure changes; some relationships between the content of galactosyl residue in the cell wall and fruit ripening have been reported previously (Redgwell et al. 1997). Redgwell et al. (1997) had reported that loss of galactose during fruit ripening mostly occurs from highly branched pectic polysaccharides. Galactose is present in pectin and arabinogalactan protein (AGP), which is a family of proteoglycan proteins. Pectin has two galactose regions, namely arabinogalactan I (AGI) and arabinogalactan II (AGII). AGI is composed of  $\beta$ -1,4-galactosyl main chain, while AGII is composed of  $\beta$ -1,3-galactosyl main chain and  $\beta$ -1,6-galactosyl side chain. AGP consists of hydroxyproline-rich core proteins

Communicated by Anastasios Melis.

✉ Megumi Ishimaru  
ishimaru@waka.kindai.ac.jp

- <sup>1</sup> Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Bunkyo-ku, Tokyo 113-8657, Japan
- <sup>2</sup> Division of Applied Life Sciences, Graduate School of Life and Environmental Sciences, Osaka Prefecture University, 1-1 Gakuencho, Naka-ku, Sakai, Osaka 599-8531, Japan
- <sup>3</sup> Faculty of Biology-Oriented Science and Technology, Kindai University, 930 Nishimitani, Kinokawa, Wakayama 649-6493, Japan

with attached carbohydrate moieties, whose main chain is  $\beta$ -1,3-galactan and side chain is  $\beta$ -1,6-galactan (Kotake et al. 2005), hence, indicating that galactose is released from these regions during fruit ripening.

$\beta$ -galactosidases (EC 3.2.1.23) are widespread glycoside hydrolases (GHs) characterized by their ability to hydrolyze non-reducing terminal galactosyl residues of  $\beta$ -D-galactans, galactosyl compounds including AGP, human milk oligosaccharide, lactose, and so on. The enzymes are currently classified into eight GH families, i.e., GH1, GH2, GH35, GH39, GH42, GH59, GH147, and GH165 in Carbohydrate-Active enZymes (CAZy) database (<https://www.cazy.org>; Henrissat and Grenoble 1991; Henrissat and Bairoch 1993, 1996; Davies and Henrissat 1995; Davies et al. 1997; Lombard et al. 2014). GH43 also contains enzymes that exhibit  $\beta$ -1,3-bond specific  $\beta$ -galactosidase activity (EC 3.2.1.145, CAZy database), but here we focus on the  $\beta$ -galactosidases in EC 3.2.1.23. While  $\beta$ -galactosidases of GH1, GH2, GH39, GH42, GH59, GH147, and GH165 have been found in microorganisms, those of GH35 are widely found in microorganisms, animals, and plants. Vast number of genes coding for  $\beta$ -galactosidase have been identified in plant genomes, and they are mainly assigned to GH35 (Ishimaru et al. 2009).

In plants,  $\beta$ -galactosidases play important roles in physiological events, including cell expansion and cell wall degradation, and in the turnover of signaling molecules during fruit ripening (Gilbert 2010). Especially in tomato, galactosyl group content changes significantly during fruit development, and  $\beta$ -galactosyl content in cell wall is reduced to half at 10 days post-pollination (dpp) till the mature green stage (approximately  $40 \pm 5$  days), and further reduced to half from the breaker stage to the red-ripe stage (approximately  $8 \pm 2$  days); exo- $\beta$ -galactanase activity is known to increase in the fruit developing stage (Kim et al. 1991; Redgwell et al. 1997). Physiological studies have suggested exo- $\beta$ -galactanase activity to increase by 4 to 5 times with increase in free galactose content, which may be related to the reduction in galactosyl content of the tomato fruit cell wall (Pressey 1983; Kim et al. 1991; Carey et al. 1995; Smith et al. 2002); however, actual mechanisms of these changes still remain unknown. Smith and Gross (2000) had reported seven  $\beta$ -galactosidases (EC 3.2.1.23) in tomato, and later, and Chandrasekar and van der Hoorn (2016) found ten additional  $\beta$ -galactosidases. Thus, at least 17  $\beta$ -galactosidase (TBG) genes have been identified in tomato. Smith and Gross (2000) revealed that transcriptional level of one of the genes, TBG4, is high during the breaker to red-ripe stage. Down regulation by antisense *tbg4* resulted in 1.4-times firmer tomato fruits than control, clearly indicating TBG4 to be responsible for tomato fruit softening (Smith et al. 2002).

Most of  $\beta$ -galactosidases belonging to GH35 recognize and hydrolyze  $\beta$ -1,3- and/or  $\beta$ -1,4-bound galactans (Cheng et al. 2012). On the contrary, TBG4 hydrolyzes not only

$\beta$ -1,4-galactan, but also  $\beta$ -1,3- and  $\beta$ -1,6-galactans (Ishimaru et al. 2009), suggesting TBG4 to adopt a mechanism to recognize and hydrolyze more various types of galactosyl linkages than other GH35  $\beta$ -galactosidases. Relative activity of TBG4, revealed against  $\beta$ -1,3-galactobiose and  $\beta$ -1,6-galactobiose, is approximately 0.67 and 0.10 times that against  $\beta$ -1,4-galactobiose, respectively (Eda et al. 2016).

Eda et al. (2015, 2016) heterologously expressed TBG4 using the yeast *Pichia pastoris*, characterized rTBG4 and solved the X-ray crystal structure of the enzyme with and without D-galactose. TBG4 was seen to comprise of a catalytic  $(\beta/\alpha)_8$ -TIM barrel domain, followed by three  $\beta$ -sandwich domains (Eda et al. 2016). While the core structure of the catalytic domain has a conserved GH35  $\beta$ -galactosidase, its substrate specificity varies from other enzymes (Eda et al. 2016). Furthermore, there is no structural evidence of how TBG4 recognize such various linkage types of ligands. Therefore, in this study, we compared three linkage types of ligand-substrate complexes using X-ray crystallography, ensemble refinement, and docking simulation to reveal the mechanisms of broad substrate-specificity of TBG4.

## Materials and methods

### Protein production, purification, and crystallization

The production and purification of recombinant TBG4 (wild type; WT) and E181A, which is the mutant of catalytic acid/base residue, were performed as described previously (Eda et al. 2015), with the crystallizing conditions being 12–16% (w/v) Polyethylene glycol 10,000 (PEG10000, HAMPTON RESEARCH, USA) with 0.1 M HEPES buffer (HEPES 1.0 M solution, HAMPTON RESEARCH, USA), pH 7.3–7.5. Protein concentrations were optimized for crystallization; 0.9% (w/v) TBG4 and 1.55% (w/v) E181A were used, respectively. Under these conditions, crystals grew to maximal dimensions of approximately  $0.2 \times 0.1 \times 0.05$  mm in one month. Protein concentration was assayed using Bradford method using Quick Start protein assay kit (BIO-RAD, Hercules, CA). For structural analysis of the liganded structures,  $\beta$ -1,4-galactobiose was purchased by Sigma-Aldrich (St. Louis, Missouri, USA) and the  $\beta$ -1,3-galactobiose and  $\beta$ -1,6-galactobiose were prepared in the same methods as described previously (Kondo et al. 2020).

### Data collection

The crystal of WT was first soaked for 2 h in a solution composed of 9.6 mM  $\beta$ -1,4-galactobiose, 0.1 M HEPES buffer, pH 7.3, and 16% (w/v) PEG10000 at 277 K. Thereafter, this crystal was soaked in a cryoprotectant solution consisting

of 0.1 M HEPES buffer, pH 7.5, 16% (w/v) PEG10000, and 25% (w/v) PEG400.

The crystal of E181A was grown in a solution composed of 0.146 mM  $\beta$ -1,3-galactobiose, 0.1 M HEPES buffer, pH 7.3, and 16% (w/v) PEG10000 at 277 K for one month. It was then soaked in cryoprotectant as described above. Crystal of E181A, without ligand, was soaked for two days in a solution composed of 10 mM  $\beta$ -1,4-galactobiose, 0.1 M HEPES buffer, pH 7.3, and 16% (w/v) PEG10000 at 277 K. It was then treated by Humid Air and Glue-coating (HAG) method (Baba et al. 2013). We used polyvinyl alcohol (PVA) 4500 (Japan VAM & POVAL Co. Ltd, Japan) containing 5% (v/v) ethylene glycol (Wako Pure Chemical Industries Ltd, Japan) as the glue and mounted the crystal at 84% relative humidity. The crystal of E181A was soaked overnight in a solution composed of 2.9 mM  $\beta$ -1,6-galactobiose, 0.1 M HEPES buffer, pH 7.3, and 16% (w/v) PEG10000 at 277 K. It was then directly mounted on the testing section.

All X-ray diffraction images were collected at a wavelength of 1.000000 Å on beamline BL38B1 of SPring-8 under cryogenic conditions at 100 K, and the diffraction data were processed and scaled using HKL-2000 (Otwinowski and Minor 1997).

### Structure determination and model refinement

Each complex structure was determined by molecular replacement (MR) using MOLREP (Science and Technology Facilities Council, England; Vagin and Teplyakov 2010) in CCP4 (ver 7.0.063, Science and Technology Facilities Council, England, Winn et al. 2011; Potterton et al. 2003). The structure of TBG4 and D-galactose (PDB ID: 3W5G) complex was used as the search model. Modeled structures were refined in stepwise cycles of manual model building using phenix.refine (Berkholz et al. 2011; Afonine et al. 2012) in Phenix (ver 1.13-2998-000, Lawrence Berkeley National Laboratory, USA, Adams et al. 2010), Coot (ver 0.8.9, University of Oxford, England, Emsley et al. 2010), and restrained refinement using Refmac5 (Science and Technology Facilities Council, England, Vagin et al. 2004) in CCP4 until the *R*-factor converged. These model structures were applied to ensemble\_refinement (Burnley Tom et al. 2012; Burnley and Gros 2013; Forneris et al. 2014) in Phenix. Ensemble refinement is a method that restricts the number of structures modeled and prevents over-fitting of the data. All figures of protein structures were prepared using PyMOL (ver 2.2.3, Schrödinger, LLC).

### Domain annotation

Each domain was annotated based on Pfam database (Protein ID: O081100\_SOLLC and [https://pfam.xfam.org/family/Glyco\\_hydro\\_35](https://pfam.xfam.org/family/Glyco_hydro_35)); however, all domains were not annotated.

Therefore, we used NCBI conserved domain search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>), based on full-length amino acid sequence of TBG4 (accession number: AAC25984.1). Search options were set as default. Since it was still found to be insufficient to annotate all domains, we used BioXGEM.3D-BLAST Protein Structure Search server (<https://3d-blast.life.nctu.edu.tw/dbsas.php>; Yang and Tung 2006; Tung et al. 2007). PDB files of third domain (Asn416 to Glu438 and Leu586 to Arg724) and fourth domain (Glu439 to Val415) were uploaded in the server, followed by 3D-BLAST; setting options were as default.

### Docking simulation

In order to compare the difference of binding modes, we performed docking simulation using AutoDockVina (Oleg and Olson 2010) on Chimera (ver 1.13.1, the University of California, USA, Pettersen et al. 2004). Chain A of WT\_Gal (PDB: 6IK5) was used as the receptor, and each linkage type of galactobiose made by SWEET2 (<https://www.glycosciences.de/modeling/sweet2/doc/index.php>, Bohne et al. 1998; Bohne et al. 1999) was used as ligand. Receptor search volume options were set on TBG4's catalytic region (center coordinates were  $x = -50.77$ ,  $y = -31.93$ , and  $z = 51.36$ , sizes were  $x = 29.38$ ,  $y = 39.12$ , and  $z = 36.97$ ). The maximum number of binding modes was set to 10, and exhaustiveness of search was set to 8.

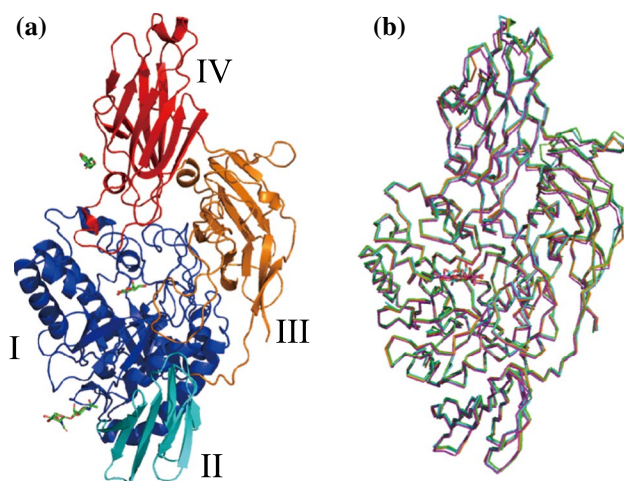
### Statistical analysis

The results obtained from protein crystallography include statistically processed results. The values are summarized in the tables in the “Results” section.

## Results

### Domain annotation

Although the overall structure of TBG4 had been revealed, each domain had not been annotated; therefore, we attempted to do so first. Catalytic TIM barrel domain is known to be conserved in all  $\beta$ -galactosidases. While a typical TIM barrel domain, shown in I of Fig. 1a, has eight  $\beta/\alpha$  repeats, that of TBG4 (Ser24 to Ala343) lacks the fifth and sixth  $\alpha$ -helices (Eda et al. 2016). The first  $\beta$ -sheet domain, shown in II of Fig. 1a, Leu344 to Val415, is a  $\beta$ -sandwich domain that is conserved across GH35 family  $\beta$ -galactosidases (GHD, Pfam ID: PF17834). The latter is similar to galectin, which is a  $\beta$ -galactose-binding lectin. Second  $\beta$ -sheet domain, shown in III of Fig. 1a, is formed by residues Asn416 to Glu438, which consists of loop regions and two  $\beta$ -strands, and residues Leu586 to Arg724, with anti-parallel  $\beta$ -sandwich



**Fig. 1** Domain organization and comparison of C $\alpha$  backbones **a** is overall structure of WT\_Gal. Colors show each domain. Blue (I) is TIM barrel domain, cyan (II) is GHD domain, orange (III) is fibronectin type III domain, and red (IV) shows a domain belonging to galactose binding domain like superfamily. **b** Shows compared C $\alpha$  backbones of each complex structures. WT\_Gal is shown in green, E181A\_ $\beta$ -1,3-Gal<sub>2</sub> is shown in orange, E181A\_ $\beta$ -1,4-Gal<sub>2</sub> is, magenta; E181A\_ $\beta$ -1,6-Gal<sub>2</sub>, respectively

structure. This domain is similar to that in fibronectin type III superfamily (E-value 1e-07, SCOPe ID: 49265), which is conserved across bacteria, fungi, plants, and animals. The last  $\beta$ -sheet domain, shown in IV of Fig. 1a, comprising of Glu439 to Gly585, is an anti-parallel  $\beta$ -sandwich domain. This domain belongs to galactose-binding domain-like superfamily (E-value: 2e-12, SCOPe ID: 49785, Yang and Tung 2006; Tung et al. 2007), which is found in  $\beta$ -galactosidases, and not only in GH35 family. Residues Leu500 to Val548 of TBG4 showed low similarity (E-value 0.02) with Carbohydrate-binding module (CBM) family 35, which includes galactan  $\beta$ -1,3-galactosidase (exo- $\beta$ -1,3-galactanase) from *Phanerochaete chrysosporium* ([https://www.cazy.org/CBM35\\_structure.html](https://www.cazy.org/CBM35_structure.html); Ichinose et al. 2005; Ishida et al. 2009a).

### Interaction between ligand and enzyme in the catalytic center of each complex

Sets of diffraction data were collected from WT and E181A crystals, containing ligands, using synchrotron radiation (Table 1), and four structures of substrate–ligand complexes were determined. In WT\_Gal, D-galactose was located at the same position as previously reported (PDB ID 3W5G), although the resolution was obviously higher (1.82 Å; previously 3.00 Å). As shown in Fig. 1b, superposition of all four ligand-bound TBG4 did not show any significant difference from each other; R.M.S.D. value was a maximum of 0.364 Å (E181A\_ $\beta$ -1,3-Gal<sub>2</sub> vs E181A\_ $\beta$ -1,6-Gal<sub>2</sub>) as listed

in Table 2. However, three galactobioses, i.e.,  $\beta$ -1,3-Gal<sub>2</sub>,  $\beta$ -1,4-Gal<sub>2</sub>, and  $\beta$ -1,6-Gal<sub>2</sub>, were accommodated quite differently at the subsites – 1 and + 1, as shown in Fig. 2b–d, clearly indicating that the difference in reactivity of the three substrates was due to the difference in recognition at the catalytic region.

In WT\_Gal, some residues interacted with D-galactose by making hydrogen bonds with O2 (Asn180-ND2), O3 (Tyr74-OH and Ala119-N), O4 (Cys118-SG), and O6 (Tyr312-OH), respectively, whereas Glu120, Glu181, Glu250, Trp252, Trp255, Tyr256, and Tyr289 provided a hydrophobic surface, as shown in Fig. 2e. We also tried to solve the WT structure with  $\beta$ -1,4-Gal<sub>2</sub> as a ligand, but failed to observe the galactose unit at the reducing terminal. This indicated that TBG4 hydrolyzed  $\beta$ -1,4-Gal<sub>2</sub> immediately when the crystal was soaked in the substrate-containing solution, even in a very short soaking time (1 min, data not shown).

In subsite -1 of E181A\_ $\beta$ -1,3-Gal<sub>2</sub>, non-reducing galactose unit was similarly recognized as in WT\_Gal, and it interacted via hydrogen bonds with O2 (Asn180-ND2), O3 (Tyr74-OH and Ala119-N), O4 (Cys118-SG), and O6 (Tyr312-OH). The clear density of reducing galactose unit was observed in subsite + 1, while only Lys217 made hydrogen bonds with O2, and majority of the interaction with Ala181, Glu250, and Tyr289 was hydrophobic, as shown in Fig. 2b, f.

In case of E181A\_ $\beta$ -1,4-Gal<sub>2</sub> complex, the number of hydrogen bonds was less than that at subsite -1, relative to that in WT\_Gal and E181A\_ $\beta$ -1,3-Gal<sub>2</sub>. This is mainly because of the lack of interaction of Tyr312 with the hydroxyl group at C6 position, as shown in Fig. 2g. On the other hand, Cys118, Glu120, Trp255, Tyr289, and Tyr312 provided a hydrophobic surface that interacted with the galactose residue. As in E181A\_ $\beta$ -1,3-Gal<sub>2</sub>, no hydrogen bond was observed in subsite + 1 of E181A\_ $\beta$ -1,4-Gal<sub>2</sub>, although Lys217, Asn230, Glu250, and Trp252 interacted hydrophobically with the reducing terminal galactose (Fig. 2f).

In contrast, non-reducing galactose residue was found to be completely upside down in E181A\_ $\beta$ -1,6-Gal<sub>2</sub> (Fig. 2d, h), compared to the residue in other structures. Although the ligand is fixed at the subsites of TBG4, it could be a disadvantage for hydrolysis, since the glycosidic linkage faces the other side of active site, thereby preventing any reaction of catalytic amino acids with the glycosidic bond. This “inverted” conformation was confirmed by docking simulation, as described below.

### Ensemble refinement of each linkage type of galactobiose complexes

To elucidate the dynamics of overall and catalytic region structure, each enzyme–substrate complex was refined

**Table 1** Summary of data-collection and refinement statistics

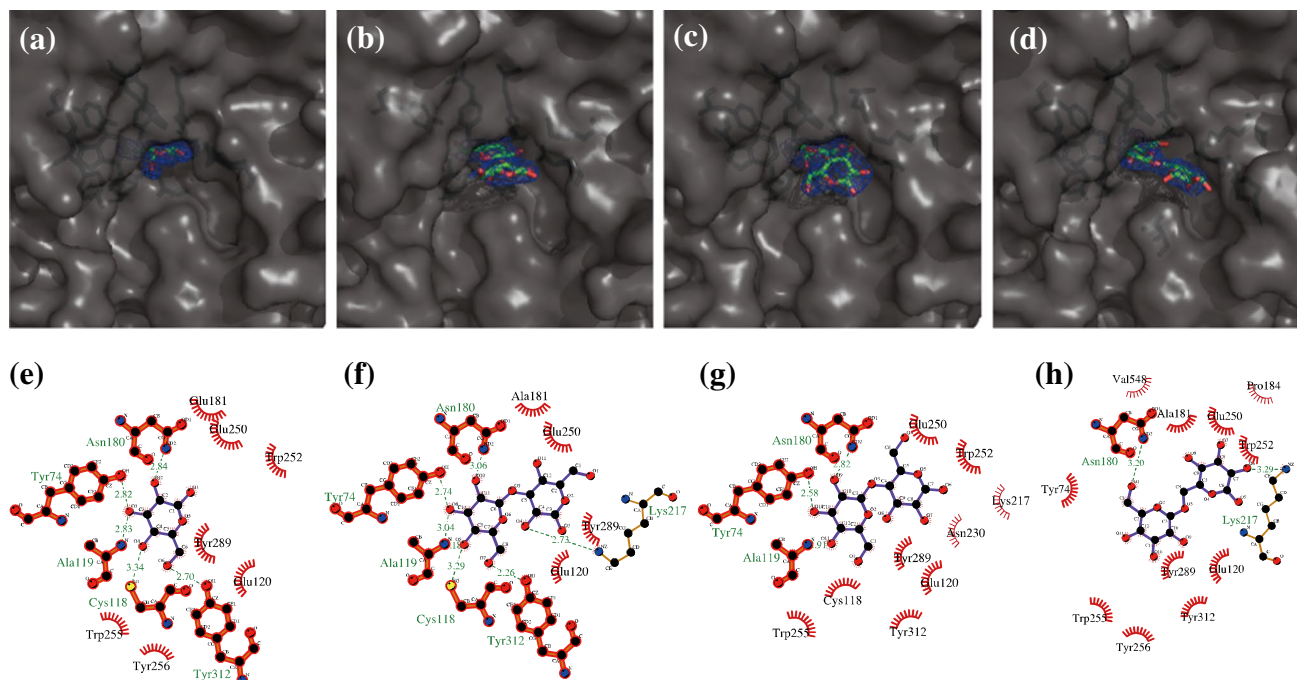
	WT_Gal	E181A_β-1,3-Gal2	E181A_β-1,4-Gal2	E181A_β-1,6-Gal2
Wavelength	1.0000	1.0000	1.0000	1.0000
Resolution range	41.28–1.82 (1.885–1.82)	43.40–3.10 (3.21–3.10)	46.86–2.79 (2.89–2.79)	44.6–2.80 (2.90–2.80)
Space group	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21
Unit cell	92.136 95.335 158.327 90 90 90	93.608 97.954 161.252 90 90 90	94.032 110.73 162.141 90 90 90	92.94 95.95 158.959 90 90 90
Unique reflections	123,244 (11,826)	27,468 (2691)	42,501 (3935)	35,651 (3480)
Multiplicity	7.1	7.3	7.0	6.6
Completeness (%)	98.34 (95.47)	99.88 (99.93)	99.17 (93.60)	99.76 (98.70)
Mean I/sigma (I)	23.4 (2.68)	9.8 (2.12)	11.0 (1.82)	11.5 (2.11)
Wilson B-factor (Å <sup>2</sup> )	23.09	45.66	59.09	50.12
R-merge	0.052 (0.615)	0.127 (0.615)	0.200 (0.800)	0.152 (0.755)
Reflections used in refinement	123,232 (11,825)	27,468 (2691)	42,460 (3935)	35,638 (3479)
Reflections used for R-free	6167 (591)	1375 (124)	2148 (177)	1777 (170)
R-work (%)	17.31 (0.2411)	19.22 (0.2930)	19.62 (27.57)	17.78 (27.15)
R-free (%)	21.23 (0.2915)	23.67 (32.04)	27.99 (36.98)	24.63 (33.01)
Number of non-hydrogen atoms	12,055	11,293	11,285	11,450
Macromolecules	11,062	11,054	11,054	11,054
Ligands	108	130	133	142
Solvent	885	109	98	254
Protein residues	1410	1410	1410	1410
RMS (bonds)	0.011	0.012	0.01	0.008
RMS (angles)	1.71	1.69	1.34	1.57
Ramachandran favored (%)	96.23	92.18	92.67	93.88
Ramachandran allowed (%)	3.56	7.40	6.26	5.97
Ramachandran outliers (%)	0.21	0.43	1.07	0.14
Rotamer outliers (%)	0.51	0.34	0.25	1.10
Clashscore	8.01	16.09	11.68	17.12
Average B-factor (Å <sup>2</sup> )	27.35	44.69	54.89	52.39
Macromolecules (Å <sup>2</sup> )	26.61	44.66	54.66	52.13
Ligands	46.38	64.87	76.46	90.17
Solvent	34.39	23.77	51.00	42.45
PDB ID	6IK5	6IK7	6IK6	6IK8

**Table 2** Difference of overall structures (RMSD, Å)

	WT_Gal	E181A_β-1,3-Gal2	E181A_β-1,4-Gal2	E181A_β-1,6-Gal2
WT_Gal		0.309	0.309	0.328
E181A_β-1,3-Gal2			0.329	0.364
E181A_β-1,4-Gal2				0.307
E181A_β-1,6-Gal2				

using ensemble\_refinement, that is, by combining X-ray structure refinement with molecular dynamic (MD) simulation in order to produce ensemble models fitted to diffraction data, thereby making local molecular vibrations (fluctuations) visible (Burnley Tom et al. 2012; Burnley

and Gros 2013; Forneris et al. 2014). The refinement statistics of ensemble refinement and improvements in statistics compared to the normal refinement are summarized in Table 3. There were several differences in the overall fluctuations across the three structures with β-1,3- (Fig. 3a), β-1,4- (Fig. 3b), and β-1,6-Gal<sub>2</sub> (Fig. 3c); they were basically at the surface of proteins, and hence not able to affect substrate specificity. On the other hand, when the view was zoomed into the catalytic region, β-1,6-Gal<sub>2</sub> showed various conformations (Fig. 4c) compared to other ligands, although the resolution of these structures was almost at a similar level. This indicates that β-1,6-Gal<sub>2</sub> fluctuated more in the enzyme while β-1,3- and β-1,4-Gal<sub>2</sub> were similarly fixed at the active site. As predicted from the static information of ligand binding mode in Fig. 2c,



**Fig. 2** Substrate-enzyme interactions of each complex. **a–d** are surface structure and 2Fo-Fc omit maps (1.0 sigma) of catalytic sites. **e–h** are the schematic diagrams interaction modes drawn by using Lig-

Plot (+ ver 1.4.5 Wallace et al. 1995; Laskowski and Swindells 2011). **a** and **e** show WT\_Gal, **b** and **f** show E181A\_β-1,3-Gal, **c** and **g** show E181A\_β-1,3-Gal<sub>2</sub>, and **d** and **h** show E181A\_β-1,6-Gal<sub>2</sub>, respectively

**Table 3** Refinement statistics of ensemble refinement

	E181A_β_1,3-Gal2	E181A_β_1,4-Gal2	E181A_β_1,6-Gal2
Refinement parameters			
Relaxation time (ps)	0.2	0.3	0.3
pTLS2 (%)	0.8	0.8	0.8
Conformers (#)	18	15	24
Refinement and model statistics			
Resolution range (Å)	43.40–3.00 (3.10–3.00)	48.58–2.79 (2.86–2.79)	44.60–2.80 (2.87–2.80)
R work (%)	18.71 (29.77)	20.47 (29.29)	17.09 (26.04)
R free (%)	27.47 (37.68)	27.26 (33.59)	26.15 (37.35)
ΔR work (%)	0.51	– 0.85	0.69
ΔR free (%)	– 3.8	0.73	– 1.52
Mean RMSD per structure			
Bonds (Å)	0.007	0.008	0.008
Angles (°)	1.039	1.098	1.11
Dihedral (°)	16.1	15.72	16.81

\* ΔR work = ensemble refinement's R work – X-Ray's R-work

\* ΔRfree = ensemble refinement's R free – X-Ray's R-free

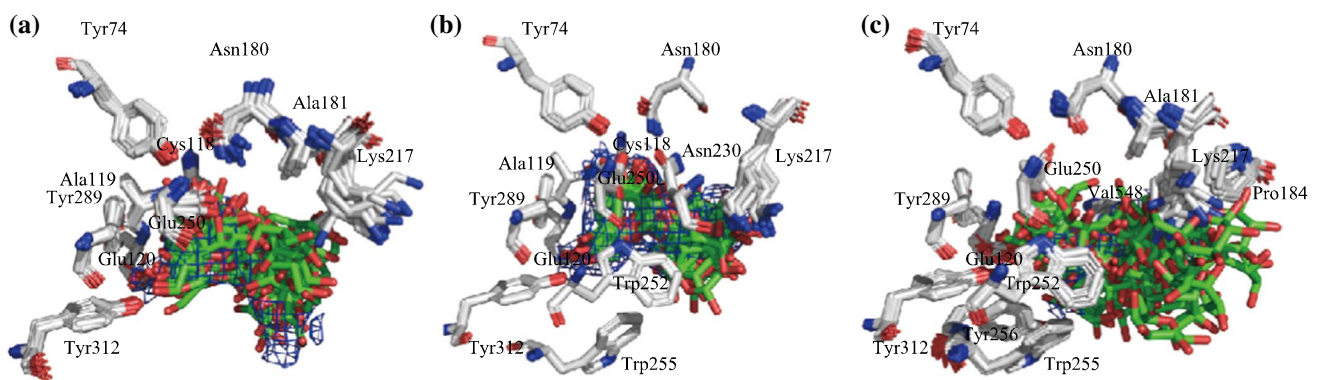
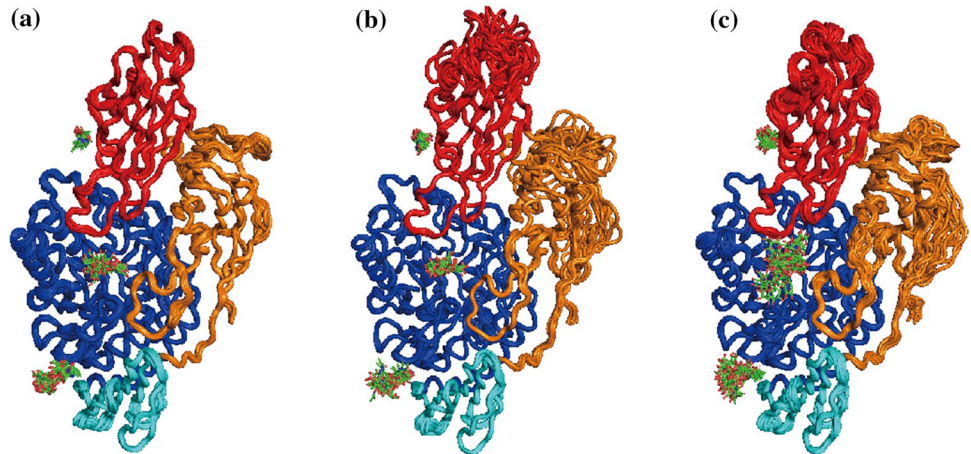
non-reducing galactose residue is tightly recognized by hydrogen bonds, although the reducing unit is rather fluctuating. This is mainly because of the fixation of reducing galactose unit mainly by hydrophobic interaction rather than by hydrogen bonds accepting various connections of galactobioses.

## Docking simulations

To evaluate the binding energy and to compare the interaction between a ligand and TBG4, docking simulations were performed using AutoDock vina (Oleg and Olson 2010). Ten types of ligand-enzyme interaction modes, for each linkage



**Fig. 3** Overall structures of each ensemble refinement model. **a** is E181A\_β-1,3-Gal<sub>2</sub>, **b** is E181A\_β-1,4-Gal<sub>2</sub>, **c** is E181A\_β-1,6-Gal<sub>2</sub>, respectively. The colors mean same domains. Sugars are represented as green (carbon) and red (oxygen)



**Fig. 4** 2Fo-Fc omit maps (1.5 sigma) of based on results from each ensemble refinement. **a** is E181A\_β-1,3-Gal<sub>2</sub>, **b** is E181A\_β-1,4-Gal<sub>2</sub>, **c** is E181A\_β-1,6-Gal<sub>2</sub>, respectively. Amino acid residues are

represented as white (carbon), red (oxygen), and blue (nitrogen). Sugars are represented as green (carbon) and red (oxygen)

type of galactobiose, are listed according to binding energy, in Table 4. In all the ligands, the lowest binding energy was similar,  $-7.0$ ,  $-7.1$ , and  $-7.2$  kcal/mol for  $\beta$ -1,3-,  $\beta$ -1,4-, and  $\beta$ -1,6-Gal<sub>2</sub>, respectively, and all of them retained the productive conformation, i.e., non-reducing galactose unit locates at  $-1$  subsite and glycosidic bond faces catalytic residue (Fig. 5; Table 4). However, in some cases, the ligand state 2 of  $\beta$ -1,4- ( $-6.2$  kcal/mol) and state 9 of  $\beta$ -1,3- ( $-5.0$  kcal/mol) interacted with subsite +1 and +2, suggesting the existence of more positive subsites interacting with  $\beta$ -1,3- or  $\beta$ -1,4-galactooligosaccharides in TBG4. There also exist other regions around the active site pocket that can accommodate  $\beta$ -1,3- and  $\beta$ -1,4-Gal<sub>2</sub>, suggesting that the enzyme hydrolyzes the terminal galactose residue even from polymer, like  $\beta$ -1,3-galactan or  $\beta$ -1,4-galactan (Fig. 5a, b). However, according to the docking simulation of  $\beta$ -1,6-Gal<sub>2</sub>, most simulated models possessed some limitation for proceeding with the reaction; some were inverted from the ideal position of galactose at subsite  $-1$ , causing difficulty of  $\beta$ -1,6-galactobiose in binding at the right position for

being hydrolyzed by the enzyme, while the non-reducing galactose residue remains conserved (Fig. 5c).

### Discussion

Various glycosidases are well known to have rather wide substrate specificities, and their structure–function relationship has been discussed previously. However, similar information is lacking in case of plant GH35 galactosidases, due to which, the actual function of galactosidase remains unclear, although apparent induction is observed during fruit ripening. Here, we tried the combination of X-ray crystallography, ensemble refinement, and docking simulation of TBG4 to unravel the molecular mechanisms of its substrate specificity.

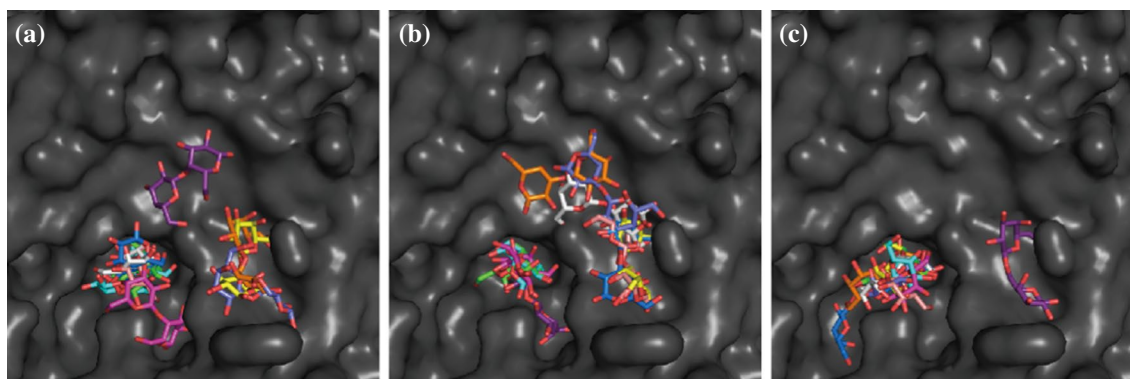
In X-ray crystallography, overall structures of the available four models were very similar to the apo (PDB ID: 3W5F) and holo structures, regardless of the ligand shape, indicating that no drastic conformational change occurs

**Table 4** Simulated scores of AutoDock Vina

State	$\beta$ -1,3-Gal2				$\beta$ -1,4-Gal2				$\beta$ -1,6-Gal2			
	Binding energy (kcal/mol)	Lower RMSD (kcal/mol)	Upper RMSD (kcal/mol)	Position	Binding energy (kcal/mol)	Lower RMSD (kcal/mol)	Upper RMSD (kcal/mol)	Position	Binding energy (kcal/mol)	Lower RMSD (kcal/mol)	Upper RMSD (kcal/mol)	Position
1	-7.0	0.000	0.000	Correct	-7.1	0.000	0.000	Correct	-7.2	0.000	0.000	Correct
2	-6.6	1.457	1.764	Correct	-6.2	1.935	4.049	Subsite+1+2	-6.4	2.464	6.292	Inverted <sup>b</sup>
3	-5.9	1.724	3.385	Inverted <sup>a</sup>	-5.2	5.249	7.249	2nd	-6.2	2.605	6.559	Inverted <sup>b</sup>
4	-5.8	12.319	14.009	2nd	-5.2	11.914	13.264	3rd	-6.2	1.606	2.376	Subsite+1 was not correct
5	-5.4	11.757	13.308	2nd	-5.2	1.860	6.209	Inverted <sup>b</sup>	-6.0	2.358	6.926	Inverted <sup>b</sup>
6	-5.3	10.194	13.065	2nd	-5.2	2.192	6.048	Inverted <sup>b</sup>	-5.8	1.732	6.480	Inverted <sup>c</sup>
7	-5.1	12.050	14.361	2nd	-5.1	12.716	14.794	3rd	-5.6	1.637	6.121	Inverted <sup>b</sup>
8	-5.0	10.578	13.090	2nd	-5.0	12.599	13.712	3rd	-5.6	1.685	6.616	Inverted <sup>b</sup>
9	-5.0	3.898	6.714	subsite+1+2	-4.8	10.483	13.285	4th	-5.5	13.310	15.231	2nd
10	-5.0	12.073	13.605	2nd	-4.8	2.728	4.803	Inverted <sup>c</sup>	-5.5	3.786	7.536	Subsite+1+2, inverted <sup>c</sup>

Small capitals indicate direction of ligand, a means the direction of ring O is opposite, b means reducing and non-reducing end is opposite, and c means the direction of ring O and reducing and non-reducing end are opposite

RMSD relative to the best model



**Fig. 5** The results of AutoDockVina. **a** is E181A\_β-1,3-Gal<sub>2</sub>, **b** is E181A\_β-1,4-Gal<sub>2</sub>, **c** is E181A\_β-1,6-Gal<sub>2</sub>, respectively. Colors show each simulated model. Each color means state 1: green, state 2: cyan,

state 3: magenta, state 4: yellow, state 5: wheat, state 6: white, state 7: slate, state 8: orange, state 9: purple, and state 10: marine, respectively

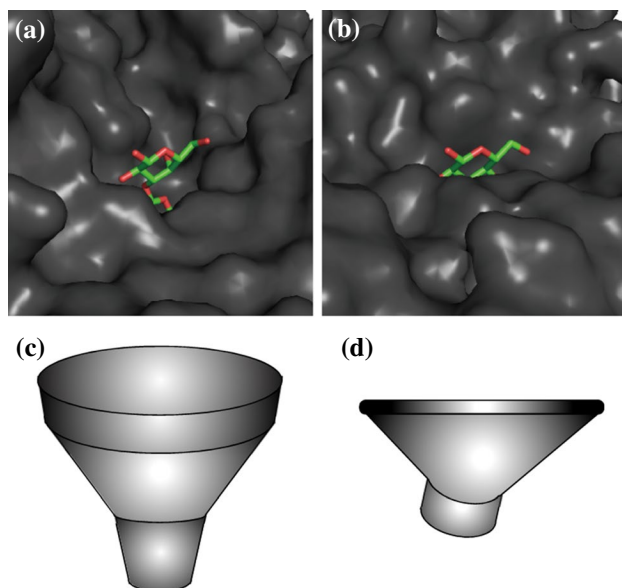
with and without substrate or with β-1,3-, β-1,4-, or β-1,6-galactobioses (Fig. 1). Moreover, side chain conformation of residues interacting with the ligands at the catalytic site was well conserved (Fig. 2). To visualize local fluctuations, we performed ensemble refinement and tested the variation of structures. Although there were some differences across the structures after accommodating each ligand, they were not conformational differences, suggesting a stiff structure of TBG4 (Fig. 3). On the other hand, regarding fluctuation of substrates, β-1,3- and β-1,4-Gal<sub>2</sub> were fixed well, while β-1,6-Gal<sub>2</sub> showed various patterns of binding at a subsite (Fig. 4). This supports the previous kinetics results reported by Eda et al. (2016) that β-1,6-Gal<sub>2</sub> is less active compared to the other substrates, β-1,3- and β-1,4-Gal<sub>2</sub>. According to the docking simulations in the present study, β-1,6-Gal<sub>2</sub> is often inserted in upside-down conformation, which might be the reason of low activity against the substrate (Fig. 5; Table 4).

Based on our results, we realized the similarity of TBG4 with other glycosidases that hydrolyze glycosidic bonds and produce monosaccharides from the non-reducing end; the interaction of hydrogen bonds at − 1 subsite is quite strict, while subsite + 1 is fixed by hydrophobic residues without recognizing the acceptance of various types of linkages by the hydroxyl residue of sugar unit. A typical example is β-D-glucan glucohydrolase (BGL) of GH family 3 from barley. This enzyme can hydrolyze β-1,2-, β-1,3-, β-1,4-, and β-1,6-linked β-glucosides (Hrmova et al. 1996, 2002; Kotake et al. 1997; Hrmova and Fincher 1998). Based on X-ray crystallography and molecular modeling of the active site, all linkages of non-reducing terminal glycosyl residue, occupying subsite − 1, form hydrogen bonds with six amino acid residues at the bottom of the active site pocket (Hrmova et al. 2001, 2002). At subsite + 1, most of the interactions occur with hydrophobic residues located at the entrance of catalytic pocket; however, there are relatively

fewer hydrogen bonds and it binds the reducing unit only when glucose dimer is presented (Hrmova et al. 2001, 2002). The reducing terminal glucose moiety shows hydrophobic π-interactions with Trp residues (Trp286 and Trp434); thus, subsite + 1 is not “recognized” like subsite − 1, rather just “fixed” by insertion between hydrophobic residues (Hrmova et al. 2002). This feature has an advantage of permitting various sugars and various types of glycosidic linkages while making it difficult for researchers to decide the actual substrate of the enzyme.

When the active site pockets of TBG4 (PDB ID: 6IK7) and barley β-D-glucan glucohydrolase (PDB ID: 1J8V) are compared in Fig. 6, apparent difference was seen only in the shape of the entrance. The entrance of the catalytic pocket of TBG4 is like a “funnel,” and subsite − 1 is located at the end of the funnel, whereas the catalytic-pocket of barley BGL looks like a “wide nozzle” with rather thinner entrance. This difference can be related to the conformation of glucans and galactans. β-1,4-galactobiose shows bulkier structure than β-1,4-glucose (cellobiose) since axial equatorial conformations of hydroxyl group at C4, respectively. Although the binding feature at + 1 site is commonly hydrophobic in both enzymes, shape of the active site is well designed to fit their substrates.

Results of docking simulation suggested more subsites around the active site, where β-1,3- and β-1,4-Gal<sub>2</sub> interact and proposed a possible role of TBG4 in hydrolyzing β-1,3-galactan and β-1,4-galactan polymers. In exo-β-1,3-glucanases (laminarinases) of GH family 55 from the basidiomycete *P. chrysosporium* (PcLam55A) and bacterium *Streptomyces* sp. SirexAA-E, the glucose moiety is released from non-reducing terminal of β-1,3-glucan, and there are at least six subsites (− 1 to + 5) that exist on the surface of the enzyme (Ishida et al. 2009b; Bianchetti et al. 2015). PcLam55A can hydrolyze oligosaccharides whose DPs are smaller than 6 (laminaribiose, laminaritriose,



**Fig. 6** Comparison of surface structures of catalytic pocket and the appearance of “funnel” and “wide nozzle”. **a** is E181A $\beta$ -1,3-Gal2 (PDB ID: 6IK7), **b** is barley  $\beta$ -D-glucan glucohydrolase (PDB ID: 1J8V). Green and red show carbon and oxygen of reducing terminal galactose (**a**) or glucose (**b**). **c** and **d** are pattern diagrams. **c** is “funnel,” and **d** is “wide nozzle,” respectively

laminaritetraose, and laminarihexaose), but the enzyme shows higher activity against the longer laminarioligosaccharides (Ishida et al. 2009b). Considering that GH family 55 laminarinases also carry the funnel-shaped active site, TBG4 may show higher activity against longer saccharides. TBG4 has been considered to have both  $\beta$ -galactanase and exo-galactanase activity (Carey et al. 1995) and hydrolyze galactan (Ishimaru et al. 2009; Eda et al. 2016). Therefore, our hypothesis might explain the reason from structural biology approach. Moreover, TBG4 may hydrolyze even if substrates have side chains, owing to the huge entrance of catalytic pocket, thus indicating that TBG4 may degrade not only AGI, but also AGII and AGP during fruit softening. Further experiments may disclose the actual mode of action of TBG4 in fruit ripening.

**Author contribution statement** KM and MI conceived and designed research. KM conducted all experiments supported by other authors. TK and TS helped preparation of substrate. KI supported structure refinement and in silico analysis. KM, KI, and MI wrote the manuscript. All authors read and approved the manuscript.

**Acknowledgements** The synchrotron radiation experiments were performed at SPring-8 under proposal number 2016B2500 and 2016B2728. We thank Dr. Toshiji Tada of the Graduate School of Science, Osaka Prefecture University, and Ms. Setsu Nakae, Faculty of Biosciences, Nagahama Institute of Bio-Science and Technology. We

thank Dr. Jerry Ståhlberg and Mikael Gudmundsson of the Department of Molecular Sciences, Swedish University of Agricultural Sciences, for refinement of the structures and for discussing the X-ray crystal structures. We thank the beamline staff at BL38B1 of SPring-8 (<https://www.spring8.or.jp>) for providing data collection facilities and support. This work was partially supported by Japan Society for the Promotion of Science (JSPS) a Grant-in-Aid for Scientific Research on Innovative Areas to KI (No. 18H05494).

## References

- Adams PD, Afonine PV, Bunkóczi G et al (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* 66:213–221. <https://doi.org/10.1107/S0907444909052925>
- Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwarta PH, Adams PD (2012) Towards automated crystallographic structure refinement with phenix. *refine research papers. Acta Crystallogr Sect D Biol Crystallogr* 68:352–367. <https://doi.org/10.1107/S0907444912001308>
- Baba S, Hoshino T, Ito L, Kumasaka T (2013) Humidity control and hydrophilic glue coating applied to mounted protein crystals improves X-ray diffraction experiments. *Acta Crystallogr Sect D Biol Crystallogr* 69:1839–1849. <https://doi.org/10.1107/S0907444913018027>
- Berkholtz DS, Shapovalov MV, Dunbrack RL, Karplus PA (2011) Conformation dependence of backbone geometry in proteins. *Struct Des* 17:1316–1325. <https://doi.org/10.1016/j.str.2009.08.012>
- Bianchetti CM, Takasuka TE, Deutsch S, Udell HS, Yik EJ, Bergeman LF, Fox BG (2015) Active site and laminarin binding in glycoside hydrolase family 55. *J Biol Chem* 290:11819–11832. <https://doi.org/10.1074/jbc.M114.623579>
- Bohne A, Lang E, Von Der Lieth CW (1998) W3-SWEET: carbohydrate modeling by internet. *J Mol Model* 4:33–43. <https://doi.org/10.1007/s008940050068>
- Bohne A, Lang E, Von Der Lieth CW (1999) SWEET—WWW-based rapid 3D construction of oligo- and polysaccharides. *Bioinformatics* 15:767–768. <https://doi.org/10.1093/bioinformatics/15.9.767>
- Burnley BT, Gros P (2013) phenix.ensemble\_refinement: a test study of apo and holo BACE1. *Comput Crystallogr Newsl* 4:51–58
- Burnley Tom B, Afonine PV, Adams PD, Gros P (2012) Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* 2012:1–29. <https://doi.org/10.7554/eLife.00311>
- Carey AT, Holt K, Picard S, Wilde R, Tucker CA, Bird CR, Schuch W, Seymour CB (1995) Tomato exo-(1 $\rightarrow$ 4)- $\beta$ -D-galactanase. *Plant Physiol* 108:1099–1107
- Chandrasekar B, van der Hoorn RAL (2016) Beta galactosidases in Arabidopsis and tomato—a mini review. *Biochem Soc Trans* 44:150–158. <https://doi.org/10.1042/BST20150217>
- Cheng W, Wang L, Jiang YL, Bai XH, Chu J, Li Q, Yu G, Liang QL, Zhou CZ, Chen Y (2012) Structural insights into the substrate specificity of *Streptococcus pneumoniae*  $\beta$ (1,3)-galactosidase BgaC. *J Biol Chem* 287:22910–22918. <https://doi.org/10.1074/jbc.M112.367128>
- Davies G, Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* 3:853–859. [https://doi.org/10.1016/S0969-2126\(01\)00220-9](https://doi.org/10.1016/S0969-2126(01)00220-9)
- Davies GJ, Wilson KS, Henrissat B (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J* 321(Pt 2):557–559. <https://doi.org/10.1007/s007920050009>
- Eda M, Ishimaru M, Tada T (2015) Expression, purification, crystallization and preliminary X-ray crystallographic analysis of tomato

- $\beta$ -galactosidase 4. *Acta Crystallogr Sect F Struct Biol Commun* 71:153–156. <https://doi.org/10.1107/S2053230X14027800>
- Eda M, Matsumoto T, Sakamoto T, Ishimaru M, Tada T (2016) Structural and functional analysis of tomato  $\beta$ -galactosidase 4: insight into the substrate specificity of the fruit softening-related enzyme. *Plant J* 86:300–307. <https://doi.org/10.1111/tbj.13160>
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* 66:486–501. <https://doi.org/10.1107/S0907444910007493>
- Forneris F, Burnley BT, Gros P (2014) Ensemble refinement shows conformational flexibility in crystal structures of human complement factor D. *Acta Crystallogr Sect D Biol Crystallogr* 70:733–743. <https://doi.org/10.1107/s1399004713032549>
- Gilbert HJ (2010) The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiol* 153:444–455. <https://doi.org/10.1104/pp.110.156646>
- Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 293:781–788. <https://doi.org/10.1042/bj2930781>
- Henrissat B, Bairoch A (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem J* 316:695–696. <https://doi.org/10.1042/bj3160695>
- Henrissat B, Grenoble F (1991) A classification of glycosyl hydrolases based on sequence similarities amino acid. *J Biochem* 280:309–316
- Hrmova M, Fincher GB (1998) Barley  $\beta$ -D-glucan exohydrolases. Substrate specificity and kinetic properties. *Carbohydr Res* 305:209–221
- Hrmova M, Harvey AJ, Wang J, Shirley NJ, Jones GP, Stone BA, Høj PB, Fincher GB (1996) Barley  $\beta$ -D-glucan exohydrolases with  $\beta$ -D-glucosidase activity. *J Biol Chem* 271:5277–5286. <https://doi.org/10.1074/jbc.271.9.5277>
- Hrmova M, Varghese JN, De Gori R, Smith BJ, Driguez H, Fincher GB (2001) Catalytic mechanisms and reaction intermediates along the hydrolytic pathway of a plant  $\beta$ -D-glucan glucohydrolase. *Structure* 9:1005–1016. [https://doi.org/10.1016/S0969-2126\(01\)00673-6](https://doi.org/10.1016/S0969-2126(01)00673-6)
- Hrmova M, De GR, Smith BJ, Fairweather JK, Driguez H, Varghese JN, Fincher GB (2002) Structural basis for broad substrate specificity in higher plant  $\beta$ -D-glucan glucohydrolases. *Plant Cell* 14:1033–1052. <https://doi.org/10.1105/tpc.010442>
- Ichinose H, Yoshida M, Kotake T, Kuno A, Igarashi K, Tsumuraya Y, Samejima M, Hirabayashi J, Kobayashi H, Kaneko S (2005) An  $\beta$ -1,3-galactanase having a novel  $\beta$ -1,3-galactan-binding module from *Phanerochaete chrysosporium*. *J Biol Chem* 280:25820–25829. <https://doi.org/10.1074/jbc.M501024200>
- Ishida T, Fujimoto Z, Ichinose H, Igarashi K, Kaneko S, Samejima M (2009a) Crystallization of selenomethionyl  $\beta$ -1,3-galactanase from the basidiomycete *Phanerochaete chrysosporium*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 65:1274–1276. <https://doi.org/10.1107/S1744309109043395>
- Ishida T, Fushinobu S, Kawai R, Kitaoka M, Igarashi K, Samejima M (2009b) Crystal structure of glycoside hydrolase family 55  $\beta$ -1,3-glucanase from the basidiomycete. *J Biol Chem* 284:10100–10109. <https://doi.org/10.1074/jbc.M808122200>
- Ishimaru M, Smith DL, Mort AJ, Gross KC (2009) Enzymatic activity and substrate specificity of recombinant tomato  $\beta$ -galactosidases 4 and 5. *Planta* 229:447–456. <https://doi.org/10.1007/s00425-008-0842-x>
- Kim J, Gross KC, Solomos T (1991) Galactose metabolism and ethylene production during development and ripening of tomato fruit. *Postharvest Biol Technol* 1:67–80. [https://doi.org/10.1016/0925-5214\(91\)90020-C](https://doi.org/10.1016/0925-5214(91)90020-C)
- Kondo T, Nishimura Y, Matsuyama K, Ishimaru M, Nakazawa M, Ueda M, Sakamoto T (2020) Characterization of three GH35  $\beta$ -galactosidases, enzymes able to shave galactosyl residues linked to rhamnogalacturonan in pectin, from *Penicillium chrysogenum* 31B. *Appl Microbiol Biotechnol* 104:1135–1148. <https://doi.org/10.1007/s00253-019-10299-y>
- Kotake T, Nakagawa N, Takeda K, Sakurai N (1997) Purification and characterization of wall-bound  $\beta$ -1,3- $\beta$ -D-glucanase from barley (*Hordeum vulgare* L.) seedlings. *Plant Cell Physiol* 38:194–200. <https://doi.org/10.1093/oxfordjournals.pcp.a029152>
- Kotake T, Dina S, Konishi T, Kaneko S, Igarashi K, Samejima M, Watanabe Y, Kimura K, Tsumuraya Y (2005) Molecular cloning of a  $\beta$ -galactosidase from radish that specifically hydrolyzes  $\beta$ -(1 $\rightarrow$ 3)- and  $\beta$ -(1 $\rightarrow$ 6)-galactosyl residues of arabinogalactan protein. *Plant Physiol* 138:1563–1576. <https://doi.org/10.1104/pp.105.062562>
- Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 51:2778–2786. <https://doi.org/10.1021/ci200227u>
- Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:490–495. <https://doi.org/10.1093/nar/gkt1178>
- Oleg T, Olson AJ (2010) Software news and update AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. <https://doi.org/10.1002/jcc.21334>
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326. [https://doi.org/10.1016/S0076-6879\(97\)76066-X](https://doi.org/10.1016/S0076-6879(97)76066-X)
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612. <https://doi.org/10.1002/jcc.20084>
- Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr Sect D Biol Crystallogr* 59:1131–1137. <https://doi.org/10.1107/S0907444903008126>
- Pressey R (1983)  $\beta$ -galactosidases in ripening tomatoes. *Plant Physiol* 71:132–135
- Redgwell RJ, Fischer M, Kendal E, MacRae EA (1997) Galactose loss and fruit ripening: High-molecular-weight arabinogalactans in the pectic polysaccharides of fruit cell walls. *Planta* 203:174–181. <https://doi.org/10.1007/s004250050179>
- Smith DL, Gross KC (2000) A family of at least seven  $\beta$ -galactosidase genes is expressed during tomato fruit development. *Plant Physiol* 123:1173–1183. <https://doi.org/10.1104/pp.123.3.1173>
- Smith DL, Abbott JA, Gross KC (2002) Down-regulation of tomato  $\beta$ -galactosidase 4 results in decreased fruit softening. *Plant Physiol* 129:1755–1762. <https://doi.org/10.1104/pp.011025>
- Tung C, Huang J, Yang J (2007) Kappa-alpha plot derived structural alphabet and BLOSUM-like substitution matrix for rapid search of protein structure database. *Genome Biol* 8:R31. <https://doi.org/10.1186/gb-2007-8-3-r31>
- Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr Sect D Biol Crystallogr* 66:22–25. <https://doi.org/10.1107/S0907444909042589>
- Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN (2004) REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr Sect D Biol Crystallogr* 60:2184–2195. <https://doi.org/10.1107/S0907444904023510>
- Wallace AC, Laskowski RA, Thornton JM (1995) Ligplot: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng Des Sel* 8:127–134. <https://doi.org/10.1093/protein/8.2.127>
- Winn MD, Ballard CC, Cowtan KD et al (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D*

Biol Crystallogr 67:235–242. <https://doi.org/10.1107/S0907444910045749>

Yang J, Tung C (2006) Protein structure database search and evolutionary classification. Nucleic Acids Res 34:3646–3659. <https://doi.org/10.1093/nar/gkl395>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# Unique active-site and subsite features in the arabinogalactan-degrading GH43 exo- $\beta$ -1,3-galactanase from *Phanerochaete chrysosporium*

Received for publication, September 22, 2020, and in revised form, October 20, 2020. Published, Papers in Press, October 22, 2020, DOI 10.1074/jbc.RA120.016149

Kaori Matsuyama<sup>1</sup>, Naomi Kishine<sup>2</sup>, Zui Fujimoto<sup>2</sup> , Naoki Sunagawa<sup>1</sup>, Toshihisa Kotake<sup>3</sup> , Yoichi Tsumuraya<sup>3</sup>, Masahiro Samejima<sup>1,4</sup> , Kiyohiko Igarashi<sup>1,5,\*</sup> , and Satoshi Kaneko<sup>6</sup>

From the <sup>1</sup>Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo, Japan, the <sup>2</sup>Advanced Analysis Center, National Agriculture and Food Research Organization (NARO), Ibaraki, Japan, the <sup>3</sup>Department of Biochemistry and Molecular Biology, Faculty of Science, Saitama University, Saitama, Japan, the <sup>4</sup>Faculty of Engineering, Shinshu University, Nagano, Japan, <sup>5</sup>VTT Technical Research Centre of Finland, Espoo, Finland, and the <sup>6</sup>Department of Subtropical Bioscience and Biotechnology, Faculty of Agriculture, University of the Ryukyus, Nishihara, Okinawa, Japan

Edited by Joseph M. Jez

Arabinogalactan proteins (AGPs) are plant proteoglycans with functions in growth and development. However, these functions are largely unexplored, mainly because of the complexity of the sugar moieties. These carbohydrate sequences are generally analyzed with the aid of glycoside hydrolases. The exo- $\beta$ -1,3-galactanase is a glycoside hydrolase from the basidiomycete *Phanerochaete chrysosporium* (Pc1,3Gal43A), which specifically cleaves AGPs. However, its structure is not known in relation to its mechanism bypassing side chains. In this study, we solved the apo and liganded structures of Pc1,3Gal43A, which reveal a glycoside hydrolase family 43 subfamily 24 (GH43\_sub24) catalytic domain together with a carbohydrate-binding module family 35 (CBM35) binding domain. GH43\_sub24 is known to lack the catalytic base Asp conserved among other GH43 subfamilies. Our structure in combination with kinetic analyses reveals that the tautomerized imidic acid group of Gln<sup>263</sup> serves as the catalytic base residue instead. Pc1,3Gal43A has three subsites that continue from the bottom of the catalytic pocket to the solvent. Subsite -1 contains a space that can accommodate the C-6 methylol of Gal, enabling the enzyme to bypass the  $\beta$ -1,6-linked galactan side chains of AGPs. Furthermore, the galactan-binding domain in CBM35 has a different ligand interaction mechanism from other sugar-binding CBM35s, including those that bind galactomannan. Specifically, we noted a Gly  $\rightarrow$  Trp substitution, which affects pyranose stacking, and an Asp  $\rightarrow$  Asn substitution in the binding pocket, which recognizes  $\beta$ -linked rather than  $\alpha$ -linked Gal residues. These findings should facilitate further structural analysis of AGPs and may also be helpful in engineering designer enzymes for efficient biomass utilization.

Arabinogalactan proteins (AGPs) are proteoglycans characteristically localized in the plasma membrane, cell wall, and intercellular layer of higher land plants (1), in which they play functional roles in growth and development (2). The carbohydrate moiety of AGPs is composed of a  $\beta$ -1,3-D-galactan main

chain and  $\beta$ -1,6-D-galactan side chain, decorated with arabinose, fucose, and glucuronic acid residues (1, 2). The chain lengths and frequencies of side chains are different among plant species, organs, and stages of development (3), and the overall structures of the carbohydrate moieties of AGPs are not yet fully understood. Degradation of polysaccharides using specific enzymes is one approach to investigate their structures and roles. In this context, exo- $\beta$ -1,3-galactanase (EC 3.2.1.145) specifically cleaves the nonreducing end  $\beta$ -1,3-linked galactosyl linkage of  $\beta$ -1,3-galactans to release D-galactose (Gal). In particular, it releases  $\beta$ -1,6-galactooligosaccharides together with Gal from AGPs (4, 5) and is therefore useful for structural analysis of AGPs.

The basidiomycete *Phanerochaete chrysosporium* produces an exo- $\beta$ -1,3-galactanase (Pc1,3Gal43A; GenBank<sup>TM</sup> accession no. BAD98241) that degrades the carbohydrates of AGPs when grown with  $\beta$ -1,3-galactan as a carbon source (6). Pc1,3Gal43A consists of a glycoside hydrolase (GH) family 43 subfamily 24 (GH43\_sub24) catalytic domain and a carbohydrate-binding module (CBM) belonging to family 35 (designated as PcCBM6 in (6)) based on the amino acid sequences in the Carbohydrate-Active enZymes (CAZy) database (RRID:SCR012935) (6–8). The properties of the enzyme have been analyzed using recombinant Pc1,3Gal43A expressed in the methylotrophic yeast *Pichia pastoris* (6). The CBM35 of Pc1,3Gal43A was characterized as the first  $\beta$ -1,3-galactan-binding module, and Pc1,3Gal43A showed typical GH43\_sub24 activity. The enzyme cleaves only  $\beta$ -1,3 linkages of oligosaccharides and polysaccharides but produces  $\beta$ -1,6-galactooligosaccharides together with Gal. Thus, Pc1,3Gal43A specifically recognizes  $\beta$ -1,3-linked Gal but can accommodate  $\beta$ -1,6-bound side chains (6).

Glycoside hydrolases are classified into families based on sequence similarity, whereas they are also divided into two major groups according to their catalytic mechanisms (*i.e.* inverting enzymes and retaining enzymes) (9, 10). Inverting enzymes typically utilize two acidic residues that act as an acid and a base, respectively, and a hydroxyl group connected to anomeric carbon inverts from the glycosidic linkage after the reaction. GH43 enzymes are members of the inverting group and share conserved Glu and Asp as the catalytic acid and base, respectively (8), but GH43\_sub24 enzymes lack the catalytic

This article contains supporting information.

✂ Author's Choice—Final version open access under the terms of the Creative Commons CC-BY license.

\* For correspondence: Kiyohiko Igarashi, [aquarius@mail.ecc.u-tokyo.ac.jp](mailto:aquarius@mail.ecc.u-tokyo.ac.jp).

## Crystal structure of fungal GH43 galactanase

**Table 1**

**Data collection statistics**

Values in parentheses are for the highest-resolution shell.

Data	WT	SeMet				WT Gal3 soaking	E208Q Gal3 co-crystal	E208A Gal3 co-crystal
		Peak	Edge	Low remote	High remote			
Space group	<i>P</i> 1	<i>P</i> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub>	<i>P</i> 3 <sub>2</sub> 2 <sub>1</sub>
<b>Unit-cell parameters</b>								
<i>a</i> , <i>b</i> , <i>c</i> (Å)	40.5, 66.3, 74.0	66.4, 50.5, 75.8				50.8, 66.6, 106.4	66.1, 50.4, 75.7	156.7, 156.7, 147.7
$\alpha$ , $\beta$ , $\gamma$ (degrees)	72.0, 84.7, 82.1	90.0, 111.9, 90.0				90.0, 90.0, 90.0	90.0, 111.3, 90.0	90.0, 120.0, 90.0
Beam line	PF BL-5	PF BL-6A	PF BL-6A	PF BL-6A	PF BL-6A	PF-AR NW12	PF-AR NE3	PF-AR NE3
Detector	ADSC Q315	ADSC Q4R				ADSC Q210	ADSC Q270	ADSC Q270
Wavelength (Å)	0.90646	0.97882	0.97950	0.98300	0.96400	1.0000	1.0000	1.0000
Resolution (Å)	50–1.40	50.0–1.80	50.0–2.00	50.0–2.00	50.0–2.00	100.0–1.50	50.0–2.50	100.0–2.30
	(1.45–1.40)	(1.86–1.80)	(2.07–2.00)	(2.07–2.00)	(2.07–2.00)	(1.55–1.50)	(2.54–2.50)	(2.38–2.30)
<i>R</i> <sub>sym</sub>	0.054 (0.370)	0.079 (0.672)	0.061 (0.307)	0.060 (0.303)	0.062 (0.307)	0.046 (0.109)	0.143 (0.399)	0.167 (0.627)
Completeness (%)	95.6 (89.0)	100.0 (99.9)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	97.5 (94.9)	96.2 (83.0)	99.1 (92.0)
Multiplicity	3.8 (3.1)	14.0 (12.6)	7.2 (6.9)	7.2 (6.9)	7.2 (7.0)	9.2 (8.9)	4.4 (3.0)	9.7 (5.1)
Average <i>I</i> / $\sigma$ ( <i>I</i> )	24.4 (2.8)	36.6 (4.7)	30.9 (8.3)	30.8 (8.2)	31.3 (8.2)	48.9 (21.0)	13.5 (2.7)	17.9 (2.7)
Unique reflections	136,692 (12,747)	43,643 (4,353)	31,744 (3,139)	31,760 (3,144)	31,780 (3,146)	57,278 (5,493)	16,007 (702)	92,497 (8,510)
Observed reflections	520,085	613,162	227,158	228,381	228,595	524,957	69,939	900,469
<i>Z</i>	2	1				1	1	4

base Asp (8, 11, 12). In *Ct1,3Gal43A* (from *Clostridium thermocellum*), Glu<sup>112</sup> was thought to be the catalytic base (13), but in BT3683 (from *Bacteroides thetaiotamicron*), Glu<sup>367</sup> (corresponding to Glu<sup>112</sup> of *Ct1,3Gal43A*) was found not to act as a base but to be involved in recognition of the C-4 hydroxyl group of the nonreducing terminal Gal, and instead, Gln<sup>577</sup> is predicted to be the catalytic base in the form of an unusual tautomerized imidic acid (12). An example of GH lacking a catalytic base, endoglucanase V from *P. chrysosporium* (*PcCel45A*), is already known, and based on the mechanism proposed for this enzyme, it is possible that tautomerized Gln functions as a base in GH43\_sub24 or that this Gln stabilizes nucleophilic water. *PcCel45A* lacks the catalytic base Asp that is conserved in other GH45 subfamilies (14), but it uses the tautomerized imidic acid of Asn as the base, as indicated by neutron crystallography (15). However, it is difficult to understand the situation in GH43\_sub24, because no holo structure with a ligand at the catalytic center has yet been solved in this family. Moreover, no structure of eukaryotic GH43\_sub24 has yet been reported.

The CBM35 module is composed of ~140 amino acids. This family includes modules with various binding characteristics and decorated with xylans, mannans,  $\beta$ -1,3-galactans, and glucans (16–21). The family members are divided into four clusters based on their sequences and binding specificities (17). The structures of CBM35s binding with xylan, mannan, and glucan have already been solved (16–21), but no structure of  $\beta$ -1,3-galactan-binding CBM35 has yet been reported.

In the present work, we solved the apo and liganded structures of *Pc1,3Gal43A*. Based on the results, we discuss the catalytic mechanism and the mode of ligand binding to CBM35 in the two-domain structure.

## Results

### Overall structure of *Pc1,3Gal43A*

The crystal structure of the SeMet derivative of *Pc1,3Gal43A* was first determined by means of the multiwavelength anomalous dispersion method, and this was followed by structure determination of the ligand-free WT, the WT bound with Gal (WT\_Gal), the E208Q mutant co-crystallized with  $\beta$ -1,3-galac-

totriose (Gal3; E208Q\_Gal3), and the E208A mutant co-crystallized with Gal3 (E208A\_Gal3). Data collection statistics and structural refinement statistics are summarized in Tables 1 and 2, respectively.

The recombinant *Pc1,3Gal43A* molecule is composed of a single polypeptide chain of 428 amino acids (Gln<sup>21</sup>–Tyr<sup>448</sup>) with two extra amino acids, Glu<sup>19</sup> and Phe<sup>20</sup>, derived from the restriction enzyme cleavage site, which are disordered and thus were not observed. The protein is decorated with *N*-glycans because it was expressed in *Pichia* yeast. Up to three sugar chains are attached at Asn<sup>79</sup>, Asn<sup>194</sup>, and Asn<sup>389</sup>; the attached chains vary in position and structure, and most contain one or two GlcNAc moieties.

*Pc1,3Gal43A* is composed of two domains, and ligands introduced by soaking or co-crystallization are located in a subsite of the catalytic domain or the binding site of CBM35 (Fig. 1). The N-terminal catalytic domain consists of a five-bladed  $\beta$ -propeller (Gln<sup>21</sup>–Gly<sup>325</sup>), as in other GH clan-F enzymes, and the C-terminal domain (*PcCBM35*) takes a  $\beta$ -jellyroll fold (Thr<sup>326</sup>–Tyr<sup>448</sup>) structure, as in previously reported CBM35s (16–25). *PcCBM35* contains one calcium ion near the end of the first  $\beta$ -strand on a different domain surface from the plane to which the ligand binds (Fig. 1). The structure of *PcCBM35* is similar to those of other known CBM35s. The interface area is 686 Å<sup>2</sup> and includes many water molecules. The PDBEPIA server (RRID:SCR015749) indicates that the enzyme forms a complex in the crystal, but this is an effect of crystallization, and the enzyme exists as a monomer in solution (data not shown).

### Sugar-binding structure of the *Pc1,3Gal43A* catalytic domain

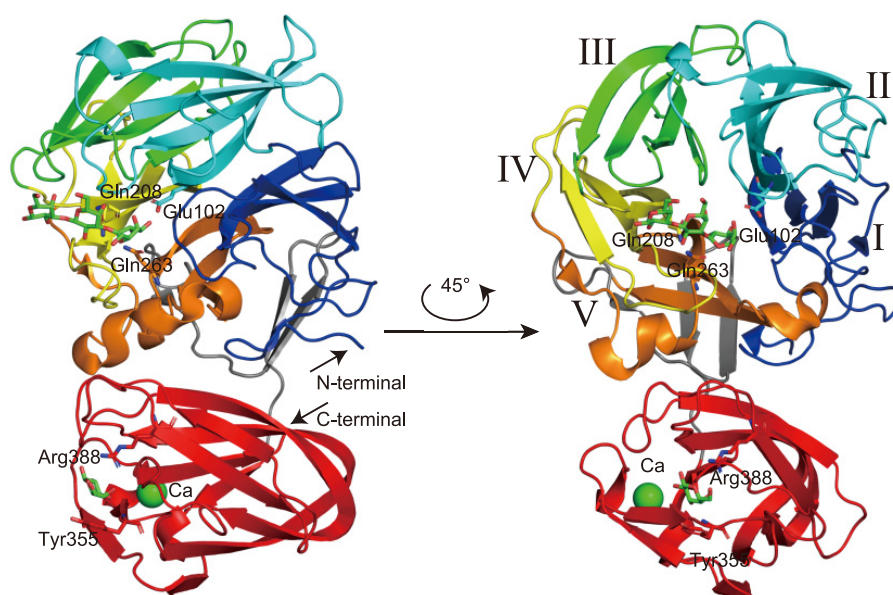
The five-bladed  $\beta$ -propeller exhibits an almost spherical structure, and two central cavities are located at the ends of the pseudo-5-fold axis (Fig. 1). One of them contains the catalytic site and it is common in almost all GH43 enzymes. The catalytic site is located in the center of the five-bladed  $\beta$ -propeller, whose blades are formed by Gln<sup>21</sup> or Asn<sup>22</sup>–Leu<sup>87</sup> (*I* in Fig. 1), Ser<sup>88</sup>–Asp<sup>155</sup> (*II* in Fig. 1), Ser<sup>156</sup>–Gly<sup>204</sup> (*III* in Fig. 1), Ala<sup>205</sup>–Ser<sup>247</sup> (*IV* in Fig. 1), and Ala<sup>248</sup>–Asp<sup>297</sup> (*V* in Fig. 1).



**Table 2**  
Refinement statistics

Values in parentheses are for the highest-resolution shell.

Data	WT	WT_Gal	E208Q_Gal3	E208A_Gal3
Resolution range	7.997–1.398 (1.448–1.398)	41.56–1.500 (1.554–1.500)	29.79–2.499 (2.588–2.499)	30.66–2.300 (2.382–2.300)
Completeness (%)	95.46 (87.82)	97.51 (94.80)	96.41 (85.67)	98.78 (92.17)
Wilson <i>B</i> -factor	12.76	10.11	29.91	30.40
Reflections used in refinement	136,655 (12,497)	57,105 (5,474)	15,762 (1,381)	92,011 (8,507)
Reflections used for <i>R</i> -free	6,862 (630)	2,884 (272)	799 (64)	4,568 (441)
<i>R</i> -work (%)	15.47 (22.50)	13.43 (12.71)	16.62 (25.54)	16.10 (22.39)
<i>R</i> -free (%)	18.56 (26.28)	16.00 (17.93)	24.39 (42.53)	21.43 (28.28)
<b>No. of nonhydrogen atoms</b>	7,966	3,923	3,576	14,570
Macromolecules	6,615	3,290	3,235	12,886
Ligands	109	121	114	678
Solvent	1,242	512	227	1,006
Protein residues	2,106	427	428	1,708
r.m.s. (bonds)	0.008	0.006	0.008	0.011
r.m.s. (angles)	1.22	0.87	0.94	1.05
Ramachandran favored (%)	97.29	97.41	94.13	95.76
Ramachandran allowed (%)	2.71	2.59	5.87	4.24
Ramachandran outliers (%)	0	0	0	0
Rotamer outliers (%)	0.81	0.55	0.29	0.36
Clash score	2.06	1.95	6.94	3.50
<b>Average <i>B</i>-factor (Å<sup>2</sup>)</b>	17.21	12.45	30.48	32.98
Macromolecules	14.97	10.57	29.77	31.60
Ligands	29.38	23.33	52.26	56.11
Solvent	28.09	22.02	29.74	35.03
<b>PDB code</b>	7BYS	7BYT	7BYV	7BYX

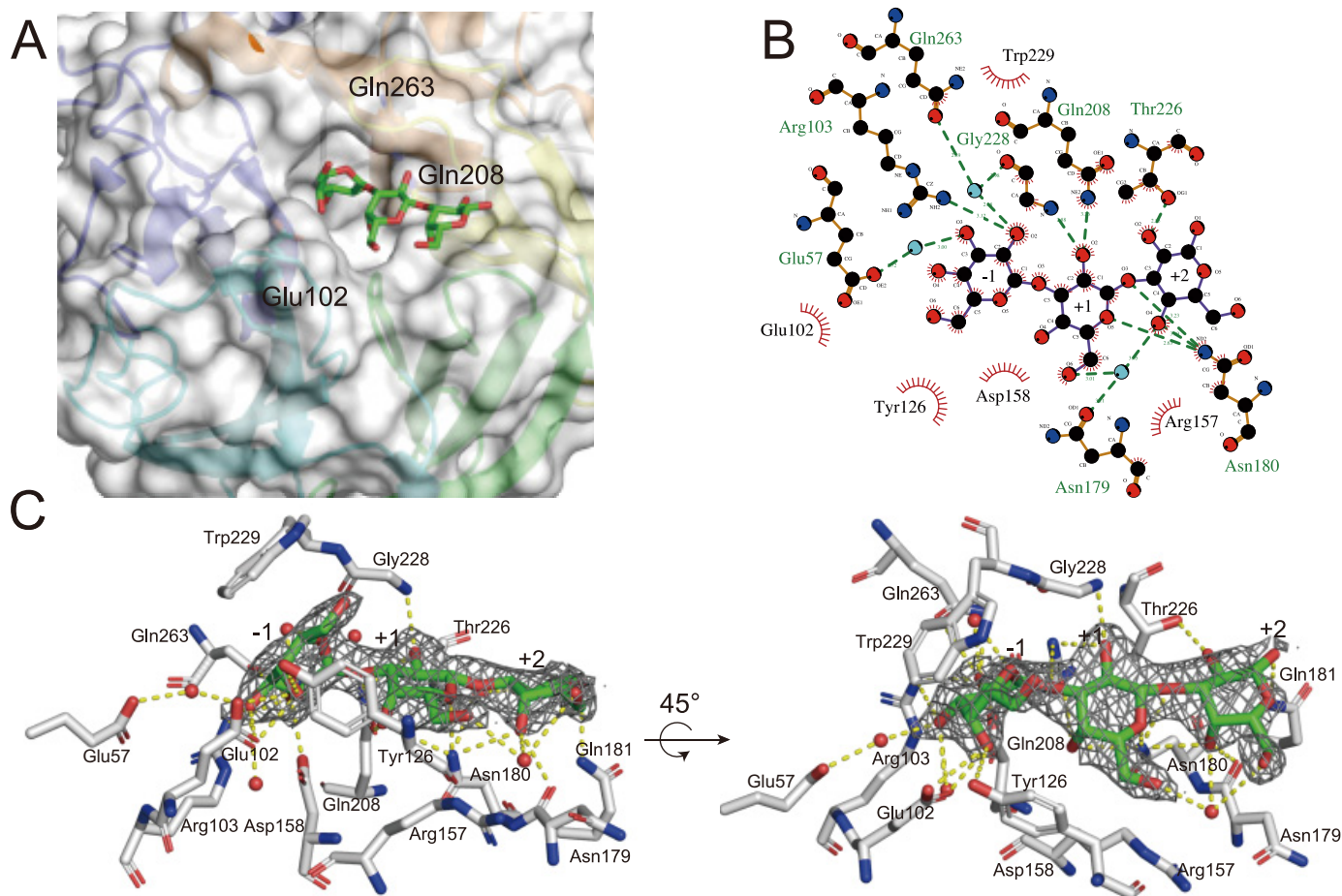
**Figure 1. Overall structure of Pc1,3Gal43A.** In the three-dimensional structure of Pc1,3Gal43A, the five blades of the catalytic domain are shown in *blue* (Gln<sup>21</sup>–Leu<sup>87</sup>), *cyan* (Ser<sup>88</sup>–Asp<sup>155</sup>), *green* (Ser<sup>156</sup>–Gly<sup>204</sup>), *yellow* (Ala<sup>205</sup>–Ser<sup>247</sup>), and *orange* (Ala<sup>248</sup>–Asp<sup>297</sup>) with successive *roman numerals*. The CBM (The<sup>326</sup>–Val<sup>448</sup>) is shown in *orange*. The linker connecting the two domains (Phe<sup>298</sup>–Gly<sup>325</sup>) is shown in *gray*.

As shown in Fig. 2, the Gal3 molecule co-crystallized with the E208Q mutant occupies subsites  $-1$ ,  $+1$ , and  $+2$  of the catalytic site, from the nonreducing end to the reducing end. Gal <sub>$-1$</sub>  is located at the bottom of the catalytic cavity, and Gal <sub>$+1$</sub>  and Gal <sub>$+2$</sub>  extend linearly outwards. Gal <sub>$+1$</sub>  is half-buried in the cavity, whereas Gal <sub>$+2$</sub>  is exposed at the surface (Fig. 2A).

Gal <sub>$-1$</sub>  adopts a <sup>1</sup>S<sub>3</sub> skew boat conformation and interacts with many residues via hydrogen bonds and hydrophobic interactions. As shown in Fig. 2 (B and C), the C-2 hydroxyl group of Gal <sub>$-1$</sub>  forms hydrogen bonds with NH<sub>2</sub> of Arg<sup>103</sup> and with OE1 of Gln<sup>263</sup> via water. In addition, this water molecule is bound

with O of Gly<sup>228</sup>. The C-3 hydroxyl group of Gal <sub>$-1$</sub>  also forms a hydrogen bond with OE2 of Glu<sup>57</sup> via water. Glu<sup>102</sup>, Tyr<sup>126</sup>, Asp<sup>158</sup>, Gln<sup>208</sup>, Thr<sup>226</sup>, Trp<sup>229</sup>, and Gln<sup>263</sup> interact with Gal3 through hydrophobic interactions. Notably, Trp<sup>229</sup> supports the flat C3-C4-C5-C6 structure of Gal <sub>$-1$</sub> , and Tyr<sup>126</sup> recognizes the C-6 methylol and C-4 hydroxyl groups, whereas Glu<sup>102</sup> recognizes the C-3 hydroxyl and C-4 hydroxyl groups. In Gal <sub>$+1$</sub>  (as shown in Fig. 2, B and C), the C-2 hydroxyl group forms a hydrogen bond with NE2 of Gln<sup>208</sup> and N of Gly<sup>228</sup>, whereas O5 forms a hydrogen bond with ND2 of Asn<sup>180</sup>, and C-6 hydroxyl group forms a hydrogen bond with OD1 of Asn<sup>179</sup> via

## Crystal structure of fungal GH43 galactanase



**Figure 2. Gal3-binding mode at the catalytic site.** *A*, surface structure of the catalytic center. Gal3 is represented as green (carbon) and red (oxygen) sticks. *B*, schematic diagram showing the interaction mode at the catalytic center. Black, red, and blue, carbon, oxygen, and nitrogen, respectively. Red lines indicate the hydrophobically interacting residues. This diagram was drawn with LigPlot+ (version 1.4.5). *C*, the  $2F_o - F_c$  omit map is drawn as a blue mesh ( $0.8\sigma$ ). Residues are shown in white (carbon), red (oxygen), and blue (nitrogen). Gal3 is shown in green (carbon) and red (oxygen). Yellow dots, hydrogen bonds and/or hydrophobic interaction; red spheres, water molecules interacting with ligands or residues.

water. Tyr<sup>126</sup>, Arg<sup>157</sup>, Asn<sup>180</sup>, and Gln<sup>208</sup> interact hydrophobically with Gal. In Gal<sub>+2</sub> (Fig. 2, *B* and *C*), the C-2 and C-4 hydroxyl groups form hydrogen bonds with OG1 of Thr<sup>226</sup> and ND2 of Asn<sup>180</sup>, respectively. In addition, Thr<sup>226</sup> interacts with Gal<sub>+2</sub> through hydrophobic interaction. Furthermore, the glycosidic oxygen between Gal<sub>+1</sub> and Gal<sub>+2</sub> interacts with ND2 of Asn<sup>180</sup> through a hydrogen bond.

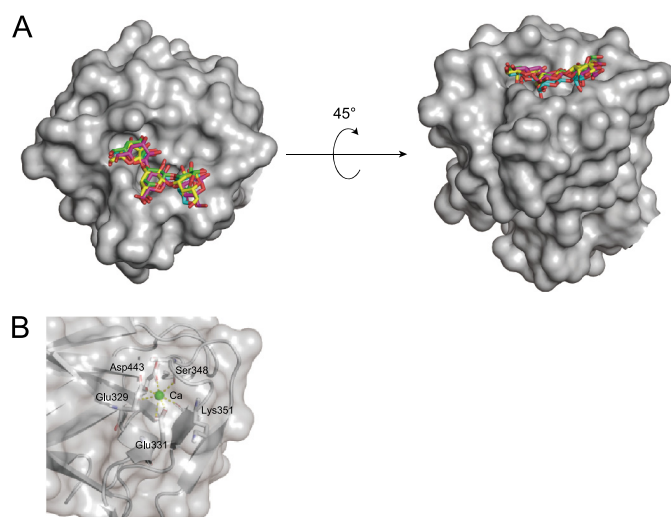
In the structure of WT\_Gal, one Gal was found at subsite -1, taking a <sup>4</sup>C<sub>1</sub> chair conformation with  $\alpha$ -anomeric conformation of the C-1 hydroxyl group (data not shown). The binding mode of Gal<sub>-1</sub> is almost the same as that in E208Q\_Gal3, but the C-1 hydroxyl group in the axial position forms hydrogen bonds with Gly<sup>228</sup> and Gln<sup>263</sup>. No Gal3 molecule was observed at the catalytic domain in the structure of the Gal3 co-crystallized E208A mutant.

To identify the catalytic residues, we examined the relative activity of WT and the six mutants toward  $\beta$ -1,3-galactobiose (Gal2) and Gal3. WT showed  $5.58 \pm 0.35$  and  $11.15 \pm 0.39$  units of activity ( $\mu\text{mol}$  of Gal/min/nmol of enzyme) toward Gal2 and Gal3, respectively, whereas the six mutants showed no detectable activity (Fig. S1), suggesting that these residues are all essential for the catalysis.

### Sugar-binding structure of CBM35 in Pc1,3Gal43A

Pc1,3Gal43A has one CBM35 domain at the C terminus. We previously reported that this enzyme has a CBM6-like domain (6), but it has been reclassified into the CBM35 family (7). The  $\beta$ -jellyroll fold domain is accompanied by a single calcium ion-binding site on a domain surface different from the surface to which the ligand at the end of the first  $\beta$  chain binds, and this corresponds to a conserved calcium ion-binding site in CBM35s. Some CBM35 modules bind another calcium ion at a site at the top of domain (16), but PcCBM35 lacks this second calcium ion-binding site (Fig. 1).

In E208A\_Gal3, electron density of Gal3 was observed in the ligand-binding site of PcCBM35. As illustrated in Fig. 3A and Fig. S2,  $2F_o - F_c$  omit maps showed that the binding mode of PcCBM35 with ligands is “exo-type,” corresponding to type-C CBM (26). The asymmetric unit of E208A\_Gal3 contained four Pc1,3Gal43A molecules, and each molecule binds to the non-reducing end of Gal3 (called Gal<sub>site</sub> 1), as in other CBM35 modules. However, the middle Gal (Gal<sub>site</sub> 2) and the reducing end Gal (Gal<sub>site</sub> 3) are found in two main locations (Fig. 3), although residues involved in the interactions with the ligand in each molecule were mostly shared. The Gal<sub>site</sub> 1 forms



**Figure 3. Surface structures of the CBM.** A, substrate-binding mode at CBM35. Green, cyan, magenta, and yellow, carbons of chains A, B, C, and D of E208A\_Gal3, respectively; red, oxygen. Left, nonreducing end of Gal3; right, reducing end. B, calcium ion-binding mode at CBM35. Calcium ion is represented as green spheres, and interacting residues are shown as stick models. Yellow dots, interaction.

hydrogen bonds with Tyr<sup>355</sup> and Arg<sup>388</sup> and interacts hydrophobically with Leu<sup>342</sup>, Gly<sup>354</sup>, Tyr<sup>438</sup>, and Asp<sup>441</sup>. The Gal<sub>site</sub> 2 interacts hydrophobically with Gly<sup>383</sup> and Asp<sup>384</sup>. The main ligand interaction in the Gal<sub>site</sub> 3 involves Gly<sup>409</sup> and Gly<sup>410</sup>, but in addition to these residues, Asn<sup>411</sup> is also involved in ligand recognition in chain C (Fig. 4).

### Ensemble refinement

To understand the fluctuation of ligands, ensemble refinements were performed with the refined models. This method produces ensemble models by employing a combination of X-ray structure refinement and molecular dynamics. These models can simultaneously account for anisotropic and anharmonic distributions (27). Four different pTLS values of 0.6, 0.8, 0.9, and 1.0% were set for each model. Table 3 shows the statistical scores of the refinement with the most appropriate pTLS value for each model. Focused views of the catalytic site in the catalytic domain and the ligand-binding site of the CBM are shown in Figs. 5 and 6, respectively. Note that structures containing multiple molecules in the asymmetric unit (WT and E208A\_Gal3) are found for all molecules in this paper.

In the catalytic site, the vibration levels of some residues were significantly different between the apo and holo forms. As shown in Fig. 5, Tyr<sup>126</sup>, Arg<sup>157</sup>, Asp<sup>158</sup>, Asn<sup>179</sup>, Asn<sup>180</sup>, Gln<sup>181</sup>, Trp<sup>229</sup>, and Gln<sup>263</sup> in the liganded structures (Fig. 5, B, C, F, G, J, and K) showed smaller vibrations than in the apo structures (Fig. 5, A, D, E, H, I, and L). These results indicate that side-chain fluctuations converge upon ligand binding. Comparison of the Gal<sub>bond</sub> structure (*i.e.* WT\_Gal; Fig. 5, B, F, and J) with the Gal<sub>3-bond</sub> structure (*i.e.* E208Q\_Gal3; Fig. 5, C, G, and K) showed that the fluctuations of Glu(Gln)<sup>208</sup>, Asn<sup>179</sup>, and Thr<sup>226</sup> of E208Q\_Gal3 were smaller than these in WT\_Gal. Therefore, it can be inferred that these residues recognize the ligands at the plus subsites. The catalytic acid, Glu<sup>208</sup>, has two major conformations in WT and WT\_Gal. These two conformations

were also reported in the BT3683 structure (12). Thus, the movement of this residue appears to be important for catalysis. Gln<sup>263</sup> shows one conformation (Fig. 5, A–D) that is identical to the result of the ensemble refinement of Asn<sup>92</sup>, known as imidic acid in *PcCel45A* (Fig. S3). Glu<sup>102</sup> may distinguish non-reducing terminal Gal, because it interacts with the axial C-4 hydroxyl group of Gal<sub>-1</sub> (12). The vibration degree of Glu<sup>102</sup> was different between WTs and mutants, so its conformation does not depend on the ligand localization, but reflects interaction with Glu<sup>208</sup>, which serves as a general acid. Asp<sup>158</sup> of WT and E208A\_Gal3 show greater vibration than WT\_Gal and E208Q\_Gal3. The role of Asp<sup>158</sup> is thought to be a pK<sub>a</sub> modulator; therefore, its function and conformational stability might be related. Focusing on Fig. 5 (I–L), there were large differences in the fluctuation level of Trp<sup>229</sup>. E208Q\_Gal3 (Fig. 5K) showed small movements of Trp<sup>229</sup>, but other structures showed much larger fluctuations (Fig. 5, I, J, and L). These results suggest that this Trp is normally flipped and forms a  $\pi$ - $\pi$  interaction to anchor the ligand in the proper position upon arrival. A histogram of the dihedral angle is shown in Fig. S4.

As regards the ligand-binding site of the CBM, a comparison of each chain of the E208A\_Gal3 asymmetric unit showed no significant difference in the vibration levels of each residue involved in ligand binding (Fig. 6). However, ensemble refinement revealed that Gal<sub>site</sub> 1 and Gal<sub>site</sub> 2 do not show huge fluctuations, whereas Gal<sub>site</sub> 3 has many conformations. They include the same conformation of each Gal chain in X-ray crystallography. Interestingly, a spatial difference in fluctuations was observed between ligands bound to the catalytic site and to the ligand-binding site of CBM35 (Fig. 7). At the catalytic site, Gal<sub>-1</sub> is anchored in the appropriate position, and Gal<sub>+2</sub> appears to fluctuate in a planar fashion as it interacts with the surrounding residues. In the CBM, it was inferred that Gal<sub>site</sub> 1 is fixed and Gal<sub>site</sub> 3 is adsorbed at the appropriate location at the binding site while fluctuating in three dimensions.

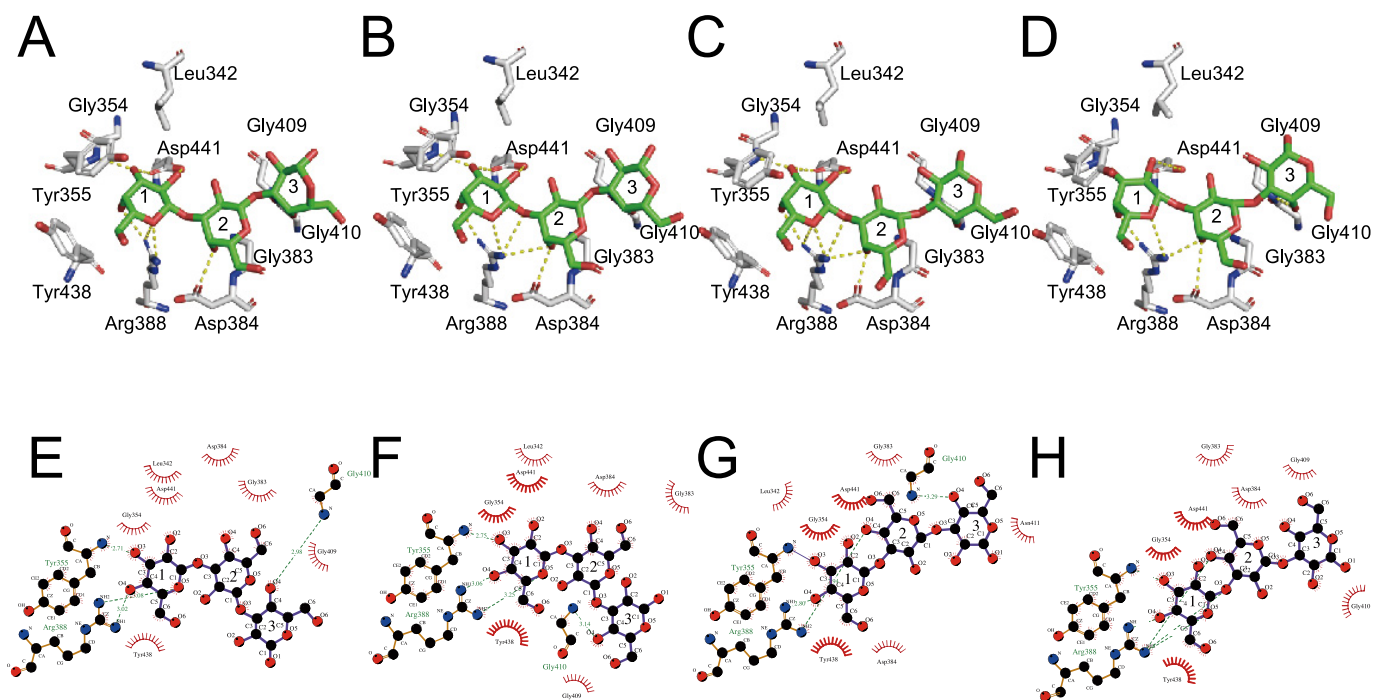
### Discussion

Most exo- $\beta$ -1,3-galactanases belonging to GH43\_sub24 possess CBMs that can be classified into CBM35 or CBM13 (8). In this study, we elucidated the structure of a  $\beta$ -1,3-galactan-binding module for the first time by solving the structure of a GH43\_sub24 containing CBM35 and obtained the ligand-bound structures of both the catalytic and sugar-binding domains of *Pc1,3Gal43A*. This is also the first study to reveal the structure of a eukaryotic exo- $\beta$ -1,3-galactanase. This information will be useful to understand how the CBM35 module interacts with  $\beta$ -1,3-galactan in combination with the GH43\_sub24 catalytic module.

### How does *Pc1,3Gal43A* hydrolyze $\beta$ -1,3-galactan?

Although catalytic residues such as Glu and Asp are conserved in GH43 as a catalytic acid and base, respectively, GH43\_sub24 lacks such a base residue. Cartmell *et al.* (12) suggested that GH43\_sub24 may use Gln in the base role via conversion to imidic acid or use an exogenous base or utilize the Grotthuss mechanism of catalysis (8, 12). In this study, we measured the enzyme activity of six variants of the three

## Crystal structure of fungal GH43 galactanase



**Figure 4. Ligand interaction mode at CBM35.** A and E, B and F, C and G, and D and H, chains A, B, C, and D of E208A, respectively. A–D, interaction modes between ligand and CBM35 residues. Atoms are indicated in the same colors as in Fig. 2. E–G, schematic diagram showing the interaction mode at CBM35. Atoms are indicated in the same colors as in Fig. 2. Sugar-binding sites are named Gal\_site 1, Gal\_site 2, and Gal\_site 3 from the nonreducing end of the sugar, and in this figure, they are labeled 1, 2, and 3, respectively.

**Table 3**

### Refinement statistics of ensemble refinement

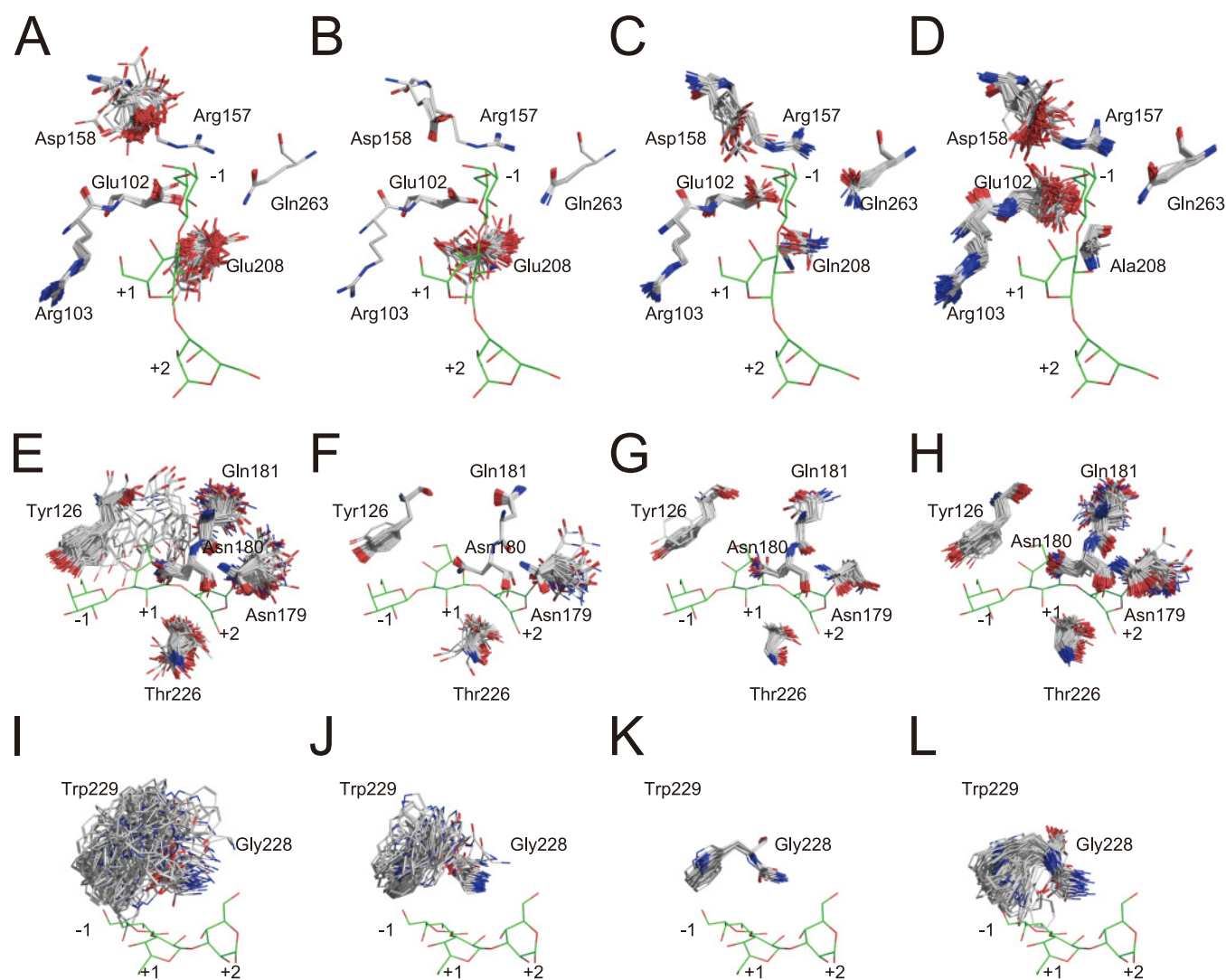
Values in parentheses are for the highest-resolution shell.

Data	WT	WT_Gal	E208Q_Gal3	E208A_Gal3
Resolution range	7.997–1.398 (1.448–1.398)	41.56–1.500 (1.554–1.500)	29.79–2.499 (2.588–2.499)	30.66–2.300 (2.382–2.300)
Completeness (%)	95.97 (82)	97.52 (95)	96.47 (88)	98.93 (87)
pTLS (%)	0.9	0.9	0.9	1.0
Tx	1.0	0.9	0.3	0.4
Wilson <i>B</i> -factor	12.8	10.1	29.9	30.4
Reflections used in refinement	136,649	57,112	15,759	91,994
Reflections used for <i>R</i> -free	6,862	2,885	799	4,569
<i>R</i> -work	13.81 (24.36)	12.08 (10.68)	17.82 (24.73)	15.92 (22.56)
<i>R</i> -free	17.08 (26.30)	15.29 (17.10)	23.33 (32.75)	20.71 (28.84)
r.m.s. (bonds)	0.008	0.010	0.007	0.008
r.m.s. (angles)	1.171	1.312	1.078	1.090
Ramachandran favored (%)	94.06	95.39	88.98	92.62
Ramachandran allowed (%)	5.08	4.03	9.19	7.24
Ramachandran outliers (%)	0.86	0.58	1.83	0.74
Rotamer outliers (%)	7.45	7.00	11.05	7.85
Clash score	0	0	0	0
<b>Average <i>B</i>-factor (Å<sup>2</sup>)</b>	<b>13.65</b>	<b>9.55</b>	<b>28.32</b>	<b>32.83</b>
Macromolecules	13.63	9.54	28.30	32.66
Ligands	14.97	9.82	28.98	36.05
Molprobity score	1.56	1.45	1.87	1.64
Model number	100	103	20	34

residues speculated to be involved in the catalytic reaction. As shown in Fig. S1, production of Gal by the mutants was not detected by means of HPLC analysis, suggesting that all three residues are essential for the catalytic activity of *Pc1,3Gal43A*. Glu<sup>102</sup>, Glu<sup>208</sup>, and Gln<sup>263</sup> are speculated to serve in C-4 hydroxyl group recognition, as a catalytic acid, and as a catalytic base, respectively. These residues are well-conserved in GH43\_sub24, as shown in Fig. S5.

In GH43\_sub24, only bacterial enzyme structures have been solved so far ([http://www.cazy.org/GH43\\_24.html](http://www.cazy.org/GH43_24.html)). To under-

stand the catalytic mechanism of *Pc1,3Gal43A*, we compared its structure with those of BT3683 and *Ct1,3Gal43A* (Fig. 8). Most of the residues that interact with ligands are conserved in these three enzymes. In subsite –1, all residues, Glu<sup>57</sup>, Glu<sup>102</sup>, Arg<sup>103</sup>, Tyr<sup>126</sup>, Asp<sup>158</sup>, Glu<sup>208</sup>, Trp<sup>229</sup>, and Gln<sup>263</sup>, of *Pc1,3Gal43A* are conserved, indicating that the binding mode at subsite –1 is fully conserved in GH43\_sub24. Based on the results of ensemble refinement, Trp<sup>229</sup> showed huge fluctuation, especially in the apo structure (Fig. 5, *I–L*). Trp<sup>541</sup> of BT3683, which corresponds to Trp<sup>229</sup> of *Pc1,3Gal43A*, has a



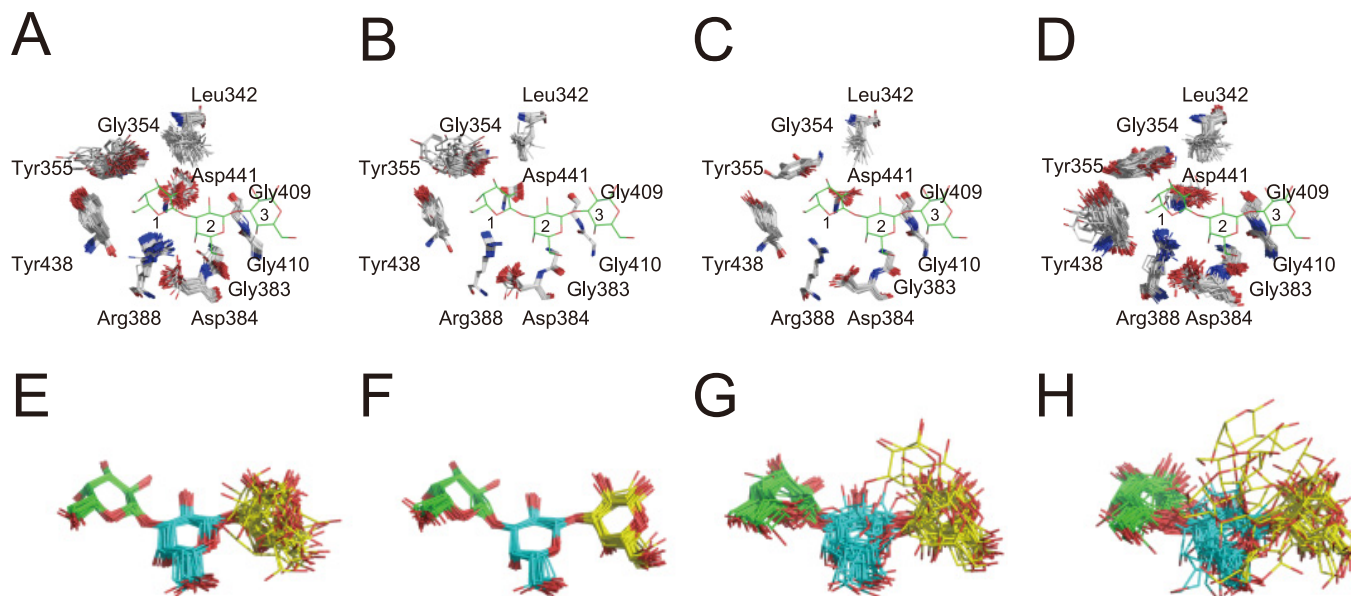
**Figure 5. Results of ensemble refinement at the catalytic site.** Each model is divided into three parts for clarity. *A* (*E* and *I*), *B* (*F* and *J*), *C* (*G* and *K*), and *D* (*H* and *L*) show WT, WT\_Gal, E208Q\_Gal3, and E208A\_Gal3, respectively. Although WT and E208A\_Gal3 contained multiple molecules in an asymmetric unit, the results obtained with multiple molecules were considered as an ensemble of one molecule in the present study. Letters indicate the chain names. Atoms are indicated in the same colors as in Fig. 2. Gal3 of the structure of E208Q\_Gal3 obtained by X-ray crystallography is arranged in each figure to maximize ease of comparison.

polar interaction with Gal (12). Trp<sup>229</sup> fluctuates in solution and plays a role in holding the substrate at the catalytic site through polar interactions. On the other hand, Asn<sup>179</sup> and Thr<sup>226</sup> of *Pc1,3Gal43A* are replaced by Asp<sup>490</sup> and Cys<sup>538</sup> in BT3683 and by Glu<sup>199</sup> and Cys<sup>247</sup> in *Ct1,3Gal43A*. Because all of these enzymes can accommodate a  $\beta$ -1,6–branched side chain (6, 12, 28), we considered that these residues are not related to the mechanism of side-chain accommodation.

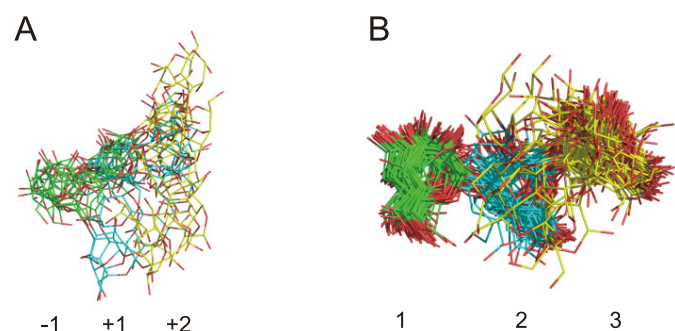
The bypass mechanism of *Pc1,3Gal43A*, which enables accommodation of the  $\beta$ -1,6-galactan side chain so that the  $\beta$ -1,3-galactan main chain can be cleaved, appears to depend on the orientation of the C-6 methylol group of Gal<sub>-1</sub> at each subsite. The C-6 methylol group of Gal<sub>-1</sub> is exposed to the solvent, so that the side chain can be accommodated externally. The C-6 methylol groups of Gal<sub>+1</sub> and Gal<sub>+2</sub> are also exposed to the solvent, so that the enzyme should be able to cleave the

$\beta$ -1,3 linkage of continuously  $\beta$ -1,6–substituted galactan, and a similar situation has been reported for BT3683 (12). Moreover, there are spaces near the nonreducing terminal Gal in these enzymes (12). This enables the enzymes to degrade the main chain, even if the side chain contains multiple carbohydrates. Similarly,  $\beta$ -1,3-galactanases belonging to GH55 also bypass the  $\beta$ -1,6-galactan side chain and degrade  $\beta$ -1,3-galactan from the nonreducing end (29, 30). Comparing the surface structure of the catalytic site of *Pc1,3Gal43A* with that of these GH55 exo- $\beta$ -1,3-galactanases from *P. chrysosporium* (*PcLam55A*), we see that *Pc1,3Gal43A* has a small pocket-like space capable of accepting the C-6 side chain of Gal at subsite -1 (Fig. 9, *A* and *B*). In addition, the C-6 methylol group of Gals, located at the positive subsites of *Pc1,3Gal43A*, are exposed to solvent in a similar manner to that reported for *SacteLam55A*, GH55 exo- $\beta$ -1,3-galactanase from *Streptomyces* sp. SirexAA-E (Fig. 9, *A*

## Crystal structure of fungal GH43 galactanase



**Figure 6. Results of ensemble refinement at the CBM ligand-binding site.** A, B, C, and D, residues related to ligand interaction. In this figure, Gal3 of chain A of refined E208A\_Gal3 is drawn for comparison. E, F, G, and H, the ligands of each chain. Green, cyan, and yellow are used in order from the nonreducing terminal Gal. A and E, B and F, C and G, and D and H represent chains A, B, C, and D of E208A\_Gal3, respectively. Atoms are indicated in the same colors as in Fig. 2.



**Figure 7. Ligand conformation of ensemble refinement at a glance.** A, ligand conformation of E208Q\_Gal3 ensemble model. B, ligand conformation of E208A\_Gal3 ensemble models with four chains aligned. Green, cyan, and yellow are used in order from the nonreducing terminal Gal.

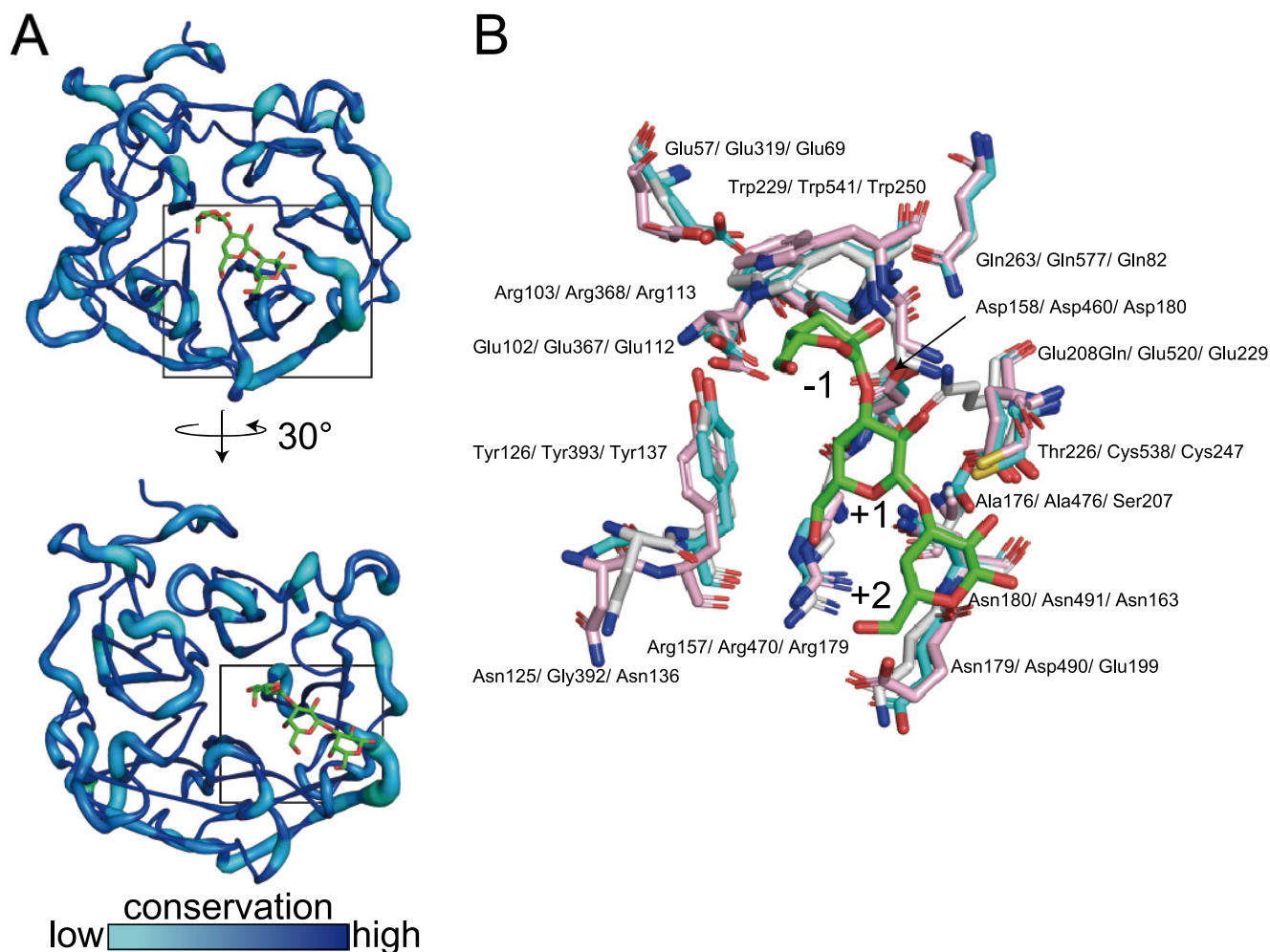
and C). Structures capable of accepting nonreducing terminal Gal with  $\beta$ -1,6-linked Gal are conserved among GH43\_sub24 of known structure (Fig. 8 and Fig. S5). In the nonbypassing GH3 *Hypocrea jecorina*  $\beta$ -glucosidase (*HjCel3A*), the C-6 hydroxyl group of nonreducing glucose is oriented toward the enzyme, introducing steric hindrance (Fig. 9D) (31). In other words, enzymes bypassing side chains have a space adjacent to C-6 of the nonreducing terminal sugar, and the positive subsites are particularly wide, allowing side chains of the substrate to be accommodated. In contrast, enzymes unable to bypass the side chain have no space next to the  $-1$  subsite and have a narrow entrance to the catalytic site, so that they are unable to accommodate side chains (Fig. 9D).

Although the electron density of Gal3 was observed in the present study, *Pc1,3Gal43A* is proposed to have four subsites ranging from  $-1$  to  $+3$ , based on biochemical experiments (6). As mentioned above, although *Pc1,3Gal43A* has a structure capable of accepting the C-6 side chain, its degradation activity toward  $\beta$ -1,3/1,6-galactan is only approximately one-fifth that

of the linear  $\beta$ -1,3-galactan (6). This difference in reactivity may be due to the structure of the sugar. The  $\beta$ -1,3-galactan in solution has a right-handed triple helical structure with 6–8 Gal residues per turn (32, 33), with the C-6 methylol group pointing outward to avoid collisions between the  $\beta$ -1,6-bonded Gal side chains (32). However, as shown in Fig. S6, Gal3 bound to the catalytic site of *Pc1,3Gal43A* is anchored to the enzyme, so that the helix of the glycans differs from the usual state in solution. Therefore, the reason why the hydrolytic activity of *Pc1,3Gal43A* toward  $\beta$ -1,3/1,6-galactan is lower than that toward  $\beta$ -1,3-galactan may be interference between the  $\beta$ -1,6-Gal side chains as a result of changes in the helical state of the main chain.

### How does *PcCBM35* recognize $\beta$ -1,3-galactan?

Although the amino acid sequence similarity of CBM35s is not so high, important residues involved in ligand binding are well-conserved (17). The modules belonging to CBM35 can be divided into four clades according to the mode of ligand binding, and the diversity in ligand binding and in the calcium ion-coordinating residue account for the various ligand-binding specificities (17) (Fig. 10A). Moreover, the residues involved in ligand binding of *PcCBM35* differ from those of CBM35, which binds to  $\alpha$ -Gal of galactomannan. This CBM is one part of a protein predicted to be the  $\beta$ -xylosidase of *C. thermocellum* celulosomal protein (Cte\_2137; Fig. 10), which belongs to the same cluster as *PcCBM35* (17). There are some differences between the residues interacting with  $\alpha$ -Gal of Cte\_2137 and those interacting with  $\beta$ -Gal of *PcCBM35*. For instance, the regions of Ala<sup>352</sup>–Tyr<sup>355</sup> and Tyr<sup>438</sup>–Asp<sup>441</sup> of *PcCBM35* correspond to Val<sup>39</sup>–Gly<sup>42</sup> and Ser<sup>136</sup>–Asn<sup>140</sup> of Cte\_2137, which are related to ligand specificity (Fig. 10). Especially, Asn<sup>140</sup> of Cte\_2137 is not conserved but replaced by Asp<sup>441</sup> in *PcCBM35* and is located at the bottom of the ligand-binding site.



**Figure 8. Catalytic domain structure comparison.** *A*, visualization of the degree of preservation of GH43\_sub24. The degree of conservation of amino acid residues in the catalytic domain of GH43\_sub24 was visualized using the ConSurf server (RRID:SCR002320), the query for BLAST was set to *Pc1,3Gal43A*, and the conservation degree was analyzed based on 150 amino acid sequences in the ConSurf server (47–51). The conservation degrees are shown in a gradient with cyan for the lowest degree of preservation and blue for the highest. *B*, catalytic domain comparison of *Pc1,3Gal43A* and two GH43\_sub24 galactanases. Shown are the catalytic centers of E208Q\_Gal3 of *Pc1,3Gal43A* (white, PDB code 7BYV), BT3683 (cyan, PDB code 6EUI), and Ct1,3Gal43A (pink, PDB code 3VSF). Red, blue, and yellow, oxygen atoms, nitrogen atoms, and sulfur atoms, respectively. Residue names are shown for *Pc1,3Gal43A*/BT3683/Ct1,3Gal43A.

Furthermore, Trp<sup>108</sup> of Cte\_2137 plays a key role in sacking the pyranose ring (17), whereas in CBM35 of *Pc1,3Gal43A*, this Trp residue is replaced with Gly (Fig. 10B). In other words, although *PcCBM35* and Cte\_2137 are in the same cluster, the residues involved in ligand recognition are different, and this difference affects the discrimination between  $\beta$ -Gal and  $\alpha$ -Gal and between galactan and galactomannan. It is still unclear how CBM35s acquire such variation of binding specificity within a similar binding architecture. However, a detailed understanding of the molecular mechanisms of polysaccharide recognition by CBM35 will be essential for efficient utilization of various types of biomass.

In conclusion, we have determined the crystal structure of the catalytic and binding domains of *Pc1,3Gal43A* with the aim of reaching a detailed understanding of the mechanism of substrate accommodation by side chain-bypassing galactanase. *Pc1,3Gal43A* uses Glu as the catalytic acid and Gln as the catalytic base and has a structure in which the side chain of the substrate does not interfere with the catalytic reaction, thus mak-

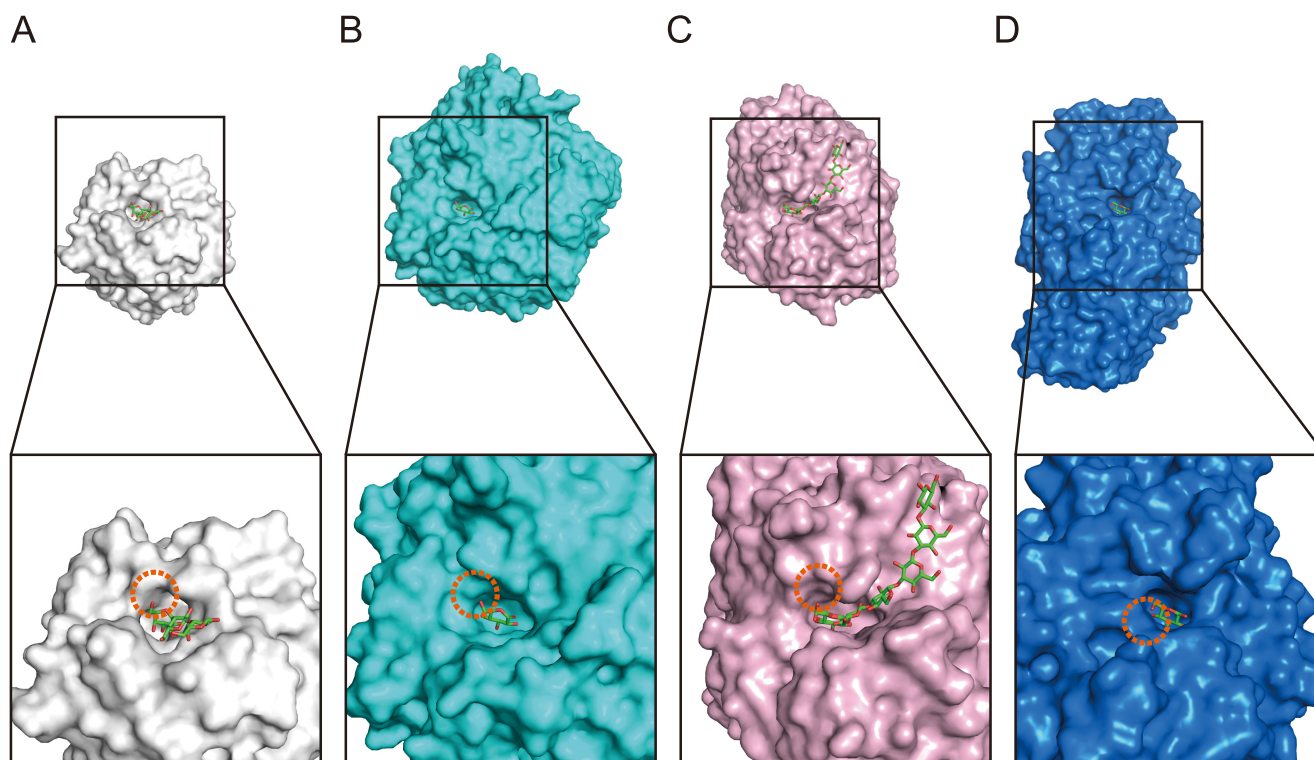
ing it possible to degrade the  $\beta$ -1,3-galactan main chain of AGPs despite the presence of the  $\beta$ -1,6-galactan side chain. Thus, although polysaccharides have a variety of molecular decorations, it appears that the structures of the degrading enzymes enable them to recognize specific features of the substrate while accommodating the variations. The introduction of mutations in substrate recognition residues to create enzymes with altered substrate recognition properties is expected to be helpful in the structural analysis of AGP glycans and also for the preparation of useful oligosaccharides.

## Experimental procedures

### Expression of *Pc1,3Gal43A* and its mutants

The E208Q, E208A, E102Q, E102A, Q263E, and Q263A mutants were constructed by inverse PCR using PrimeSTAR MAX (Takara, Tokyo, Japan). For crystallization, *Pc1,3Gal43A* WT, E208Q, and E208A from *P. chrysosporium* were expressed in *P. pastoris* and purified as reported previously (7). For a reactivity

## Crystal structure of fungal GH43 galactanase



**Figure 9.** Structure comparison of the catalytic sites of *Pc1,3Gal43A* (A), GH55 exo- $\beta$ -1,3-glucanase from *P. chrysosporium* (B; *PcLam55A*; PDB code 3EQO), GH55 exo- $\beta$ -1,3-glucanase from *Streptomyces* sp. *SirexAA-E* (C; *SactLam55A*; PDB code 4PF0), and GH3  $\beta$ -glucosidase from *H. jecorina* (D, *HjCel3A*; PDB code 3ZYZ). A, B, and C hydrolyze the main chain of  $\beta$ -1,3-galactan or  $\beta$ -1,3-glucan, bypassing  $\beta$ -1,6-branched side chains (6, 29, 30). D hydrolyzes four types of  $\beta$ -bonds, and it does not bypass side chains (31, 52). The top panel shows the overall surface structure, and the bottom panel shows an enlarged view of the catalytic region. Orange dashed circles, space near the C-6 position of Gal or glucose at the nonreducing end.

assay, WT and mutants were purified by using SkillPak TOYOPEARL Phenyl-650M (c.v. = 5 ml, Tosoh, Tokyo, Japan) equilibrated with 20 mM sodium acetate buffer, pH 4.0, containing 1 M ammonium sulfate, and the enzymes were eluted with 20 mM sodium acetate buffer, pH 4.0, containing 0.7 M ammonium sulfate. SeMet-labeled *Pc1,3Gal43A* was expressed as reported previously (7).

### Preparation for $\beta$ -1,3-galactooligosaccharides and crystallization of *Pc1,3Gal43A*

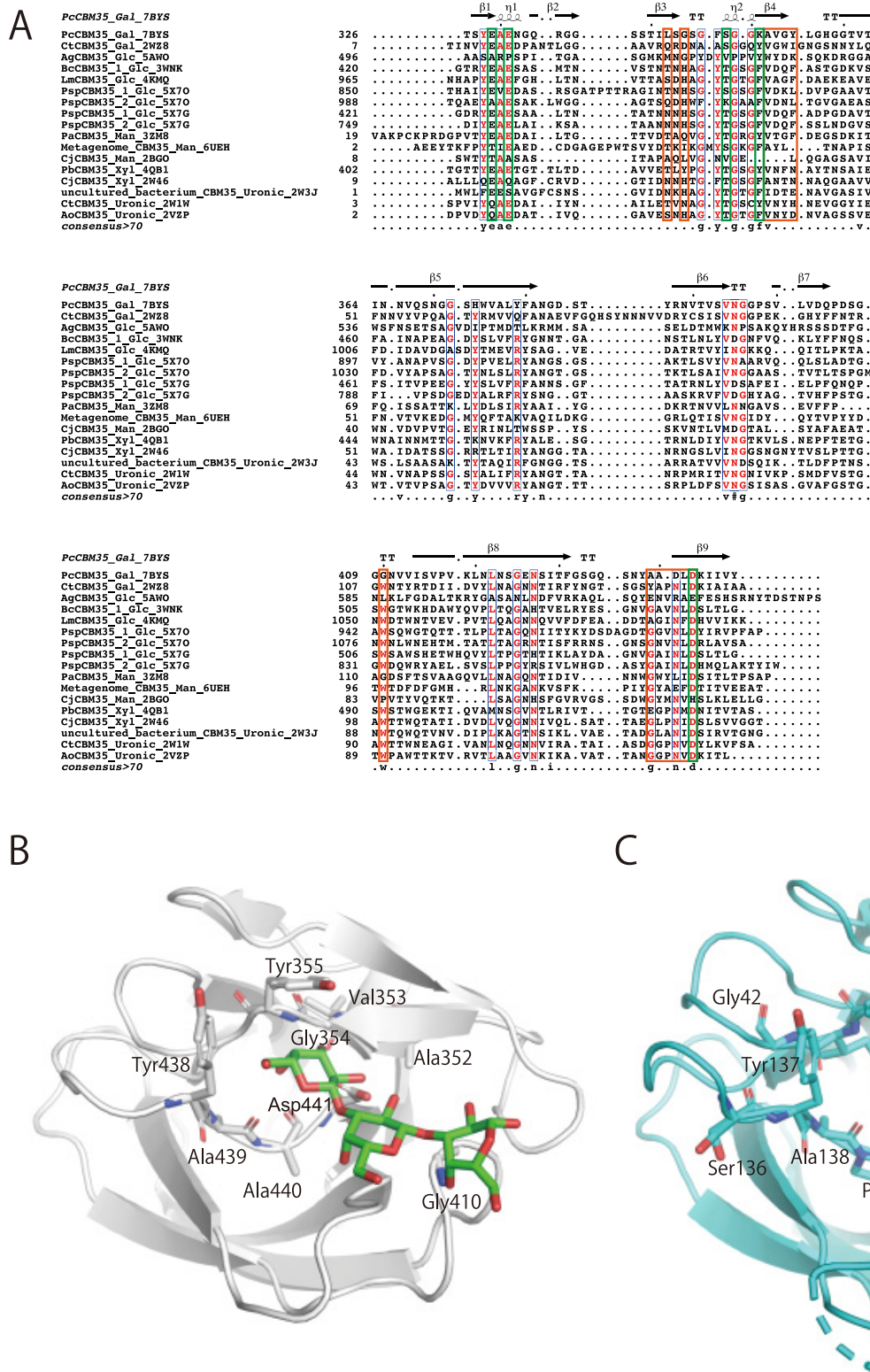
Gal2 and Gal3 were prepared as reported previously (6). The protein solution was concentrated to 4.9–6.9 mg/ml and used for the crystallization setup. The WT plate crystal used for data collection was obtained from a reservoir of 2.1 M ammonium sulfate, 0.1 M citrate buffer, pH 5.5. Other WT crystals were obtained from solutions in 16% (w/v) PEG 10000, 0.1 M ammonium sulfate, 0.1 M bis-tris, pH 5.5, and 5.0% (v/v) glycerol. SeMet crystals were obtained from 16% (w/v) PEG 10000, 95 mM ammonium sulfate, 95 mM bis-tris, pH 5.5, and 4.8% (v/v) glycerol. Two types of crystals, thin plate crystals (space group  $P2_1$ ) and rod crystals ( $P2_12_12_1$ ), appeared under the same condition. Cocrystallization of the E208Q mutant with 10 mM Gal3 in 16% (w/v) PEG 10000, 95 mM ammonium sulfate, 95 mM bis-tris, pH 5.5, and 4.8% (v/v) glycerol afforded thin plate crystals. The E208A mutant was cocrystallized with 10 mM Gal3 in 0.2 M potassium nitrate, 15% (w/v) PEG 6000, 20 mM sodium citrate, pH 4.5, and 5% glycerol to afford bipyramidal crystals.

### Data collection and structure determination

Diffraction experiments for *Pc1,3Gal43A* crystals were conducted at the beamlines of the Photon Factory (PF) or Photon Factory Advanced Ring (PF-AR), High Energy Accelerator Research Organization, Tsukuba, Japan (Table 1). Diffraction data were collected using CCD detectors (Area Detector Systems Corp., Poway, CA, USA). Crystals were cryocooled in a nitrogen gas stream to 95 K. For data collection of the WT enzyme complexed with Gal3, *Pc1,3Gal43A* crystals were soaked in a drop containing 1% (w/v) Gal3 for 10 min before the diffraction experiment. The data were integrated and scaled using the programs DENZO and SCALEPACK in the HKL2000 program suite (34).

Crystal structure was determined by means of the multiwavelength anomalous dispersion method using a SeMet-labeled crystal (7). Initial phases were calculated using the SOLVE/RESOLVE program (35) from five selenium atom positions. The resultant coordinates were subjected to the automodeling ARP/wARP program (36) in the CCP4 program suite (37), and manual model building and molecular refinement were performed using Coot (version 0.8.9, University of Oxford, Oxfordshire, UK) (38), REFMAC5 (version 7.0.063, Science and Technology Facilities Council, Swindon, UK) (39), phenix.refine (40), and phenix.ensemble\_refinement (27, 41, 42) in the Phenix suite of programs (version 1.13-2998-000, Lawrence Berkeley National Laboratory, Berkeley, CA, USA) (43). The refinement statistics are summarized in Table 2.





**Figure 10. Sequence alignment of known CBM35s (A) and structure comparison between CBM35s of Pc1,3Gal43A (B) and Cte\_2137 (C).** A, taxon names are shown as scientific names, ligand specificity, and PDB code only for brevity. When the same enzyme contains two CBM35 domains, the taxon name is indicated with 1 on the N terminus and 2 on the C terminus. *Gal*, *Glc*, *Man*, *Xyl*, and *Uronic*, ligand specificities for Gal, glucose, mannose, xylose, and glucuronic acid and/or galacturonic acid, respectively. Among these, 3ZM8, 6UEH, and 2BGO, which bind to Man, are type B CBMs, which show endo-type binding, whereas the other 14 are all type C CBMs, which show exo-type binding. The alignment was built by using MUSCLE on MEGAX: Molecular Evolutionary Genetics Analysis (53, 54), and the figure was generated with ESPrint 3.0 (RRID:SCR006587) (55). Orange and green boxes represent ligand-binding and calcium ion-binding residues, respectively. B and C, ligand-binding residues of Pc1,3Gal43A (chain A of E208A\_Gal3) and Cte\_2137 (PDB code 2WZ8). Red and blue, oxygen and nitrogen, respectively. The green stick model represents Gal3.

## Crystal structure of fungal GH43 galactanase

For the analyses of WT and ligand-bound structures, structural determination was conducted by the molecular replacement method with the MOLREP program (44) in the CCP4 program suite using the SeMet or ligand-free structure as the starting model. Bound sugars, water molecules, and crystallization agents were modeled into the observed electron density difference maps. Calcium ion was modeled based on the electron density map and the coordination distances. Three *N*-glycans were observed, and the identified sugars were modeled. The stereochemistry of the models was analyzed with LigPlot + (version 1.4.5) (45, 46), and structural drawings were prepared using PyMOL (version 2.2.3, Schrödinger, LLC, New York).

### Enzymatic activity assay of Pc1,3Gal43A and its mutants

To evaluate the reactivity toward Gal2 and Gal3 of WT and each mutant, 20 nM enzyme was incubated with 0.263 or 0.266 mM galactooligosaccharides in 20 mM sodium acetate, pH 5.0, for 30 min at 30 °C, respectively. The reaction was stopped by heating at 95 °C for 5 min. The supernatant was separated with 75% (v/v) acetonitrile on a Shodex Asahipak NH2P-50 4E column (Showa Denko, Tokyo, Japan), and the amount of released Gal was determined by HPLC (LC-2000 series; Jasco, Tokyo, Japan) with a Corona charged aerosol detector (ESA Biosciences, now Thermo Fisher Scientific). One unit of enzyme activity was defined as the amount of enzyme that releases 1 μmol of Gal/1 min/1 nmol of enzyme under our experimental conditions.

### Data availability

The structures presented in this paper have all been deposited in the Protein Data Bank (PDB) with the following codes: 7BYS, 7BYT, 7BYV, and 7BYX. All remaining data are contained within the article.

**Acknowledgments**—We thank Dr. Takuya Ishida (Japan Aerospace Exploration Agency) for helping with the crystallization and structure refinement. We also thank the staff of Photon Factory for X-ray data collection.

**Author contributions**—K. M., K. I., and S. K. conceptualization; K. M., N. K., Z. F., N. S., and K. I. data curation; K. M., N. K., Z. F., and K. I. formal analysis; M. S., K. I., and S. K. supervision; K. I. funding acquisition; K. M., N. K., Z. F., and K. I. validation; K. M., K. I., and S. K. writing-original draft; K. M., N. K., and K. I. project administration; K. M., K. I., and S. K. writing-review and editing; N. S. methodology; T. K. and Y.T. resources.

**Funding and additional information**—This work was supported in part by Grant-in-Aid for Scientific Research (B) 19H03013 (to K. I.) from the Japan Society for the Promotion of Science (JSPS) and a Grant-in-Aid for Innovative Areas 18H05494 from the Japanese Ministry of Education, Culture, Sports, and Technology (MEXT) (to K. I.). In addition, K.I. was supported by Business Finland (BF, formerly the Finnish Funding Agency for Innovation (TEKES)) via the Finland Distinguished Professor (FiDiPro) Program “Advanced approaches for enzymatic biomass utilization and modification (BioAD).”

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: AGP, arabinogalactan protein; Gal, D-galactose; Pc1,3Gal43A, exo-β-1,3-galactanase from *Phanerochaete chrysosporium*; GH, glycoside hydrolase; GH43\_sub24, GH family 43 subfamily 24; CBM, carbohydrate-binding module; Ct1,3Gal43A, exo-β-1,3-galactanase from *Clostridium thermocellum*; BT3683, β-1,3-galactosidase from *Bacteroides thetaiotaomicron* VPI-5482; PcCel45A, endoglucanase V from *P. chrysosporium*; SeMet, selenomethionine; Gal3, β-1,3-galactotriose; WT\_Gal, WT bound with Gal; E208Q\_Gal3, E208Q bound with Gal3; E208A\_Gal3, E208A bound with Gal3; PcCBM35, CBM35 domain of Pc1,3Gal43A; Gal<sub>-1</sub>, Gal residue-occupied subsite -1; Gal<sub>+1</sub>, Gal residue-occupied subsite +1; Gal<sub>+2</sub>, Gal residue-occupied subsite +2; Gal2, β-1,3-galactobiose; Gal\_site 1, the nonreducing terminal Gal residue of Gal3 bound to PcCBM35; Gal\_site 2, middle Gal residue of Gal3 bound to PcCBM35; Gal\_site 3, reducing terminal Gal residue of Gal3 bound to PcCBM35; PcLam55A, exo-β-1,3-glucanase from *P. chrysosporium*; SactLam55A, GH55 exo-β-1,3-glucanase from *Streptomyces* sp.; Cte\_2137, CBM35 of *C. thermocellum* cellulosomal protein; r.m.s., root mean square; bis-tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; PF, Photon Factory; PF-AR, Photon Factory Advanced Ring; PDB, Protein Data Bank.

### References

1. Tsumuraya, Y., Hashimoto, Y., Yamamoto, S., and Shibuya, N. (1984) Structure of L-arabino-D-galactan-containing glycoproteins from radish leaves. *Carbohydr. Res.* **134**, 215–228 [CrossRef](#)
2. Majewska-Sawka, A., and Nothnagel, E. A. (2000) The multiple roles of arabinogalactan proteins in plant development. *Plant Physiol.* **122**, 3–9 [CrossRef](#) [Medline](#)
3. Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010) Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol.* **153**, 403–419 [CrossRef](#) [Medline](#)
4. Tsumuraya, Y., Mochizuki, N., Hashimoto, Y., and Kováč, P. (1990) Purification of an Exo-β-(1-3)-D-galactanase of *Irpex lacteus* (*Polyporus tulipiferae*) and its action on arabinogalactan-proteins. *J. Biol. Chem.* **265**, 7207–7215 [Medline](#)
5. Pellerin, P., and Brillouet, J. M. (1994) Purification and properties of an exo-(1→3)-β-D-galactanase from *Aspergillus niger*. *Carbohydr. Res.* **264**, 281–291 [CrossRef](#) [Medline](#)
6. Ichinose, H., Yoshida, M., Kotake, T., Kuno, A., Igarashi, K., Tsumuraya, Y., Samejima, M., Hirabayashi, J., Kobayashi, H., and Kaneko, S. (2005) An exo-β-1,3-galactanase having a novel β-1,3-galactan-binding module from *Phanerochaete chrysosporium*. *J. Biol. Chem.* **280**, 25820–25829 [CrossRef](#) [Medline](#)
7. Ishida, T., Fujimoto, Z., Ichinose, H., Igarashi, K., Kaneko, S., and Samejima, M. (2009) Crystallization of selenomethionyl exo-β-1,3-galactanase from the basidiomycete *Phanerochaete chrysosporium*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **65**, 1274–1276 [CrossRef](#) [Medline](#)
8. Mewis, K., Lenfant, N., Lombard, V., and Henrissat, B. (2016) Dividing the large glycoside hydrolase family 43 into subfamilies: a motivation for detailed enzyme characterization. *Appl. Environ. Microbiol.* **82**, 1686–1692 [CrossRef](#) [Medline](#)
9. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859 [CrossRef](#) [Medline](#)
10. Rye, C. S., and Withers, S. G. (2000) Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* **4**, 573–580 [CrossRef](#) [Medline](#)
11. Cartmell, A., McKee, L. S., Peña, M. J., Larsbrink, J., Brumer, H., Kaneko, S., Ichinose, H., Lewis, R. J., Viksø-Nielsen, A., Gilbert, H. J., and Marles-Wright, J. (2011) The structure and function of an arabinan-specific α-1,2-arabinofuranosidase identified from screening the activities of bacterial

- GH43 glycoside hydrolases. *J. Biol. Chem.* **286**, 15483–15495 [CrossRef Medline](#)
12. Cartmell, A., Muñoz-Muñoz, J., Briggs, J. A., Ndeh, D. A., Lowe, E. C., Baslé, A., Terrapon, N., Stott, K., Heunis, T., Gray, J., Yu, L., Dupree, P., Fernandes, P. Z., Shah, S., Williams, S. J., *et al.* (2018) A surface endogalactanase in *Bacteroides thetaiotaomicron* confers keystone status for arabinogalactan degradation. *Nat. Microbiol.* **3**, 1314–1326 [CrossRef Medline](#)
  13. Jiang, D., Fan, J., Wang, X., Zhao, Y., Huang, B., Liu, J., and Zhang, X. C. (2012) Crystal structure of 1,3Gal43A, an exo- $\beta$ -1,3-galactanase from *Clostridium thermocellum*. *J. Struct. Biol.* **180**, 447–457 [CrossRef Medline](#)
  14. Igarashi, K., Ishida, T., Hori, C., and Samejima, M. (2008) Characterization of an endoglucanase belonging to a new subfamily of glycoside hydrolase family 45 of the basidiomycete *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* **74**, 5628–5634 [CrossRef Medline](#)
  15. Nakamura, A., Ishida, T., Kusaka, K., Yamada, T., Fushinobu, S., Tanaka, I., Kaneko, S., Ohta, K., Tanaka, H., Inaka, K., Higuchi, Y., Niimura, N., Samejima, M., and Igarashi, K. (2015) “Newton’s cradle” proton relay with amide–imidic acid tautomerization in inverting cellulase visualized by neutron crystallography. *Sci. Adv.* **1**, e1500263 [CrossRef Medline](#)
  16. Montanier, C., Van Bueren, A. L., Dumon, C., Flint, J. E., Correia, M. A., Prates, J. A., Firbank, S. J., Lewis, R. J., Grondin, G. G., Ghinet, M. G., Gloster, T. M., Herve, C., Knox, J. P., Talbot, B. G., Turkenburg, J. P., *et al.* (2009) Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3065–3070 [CrossRef Medline](#)
  17. Correia, M. A. S., Abbott, D. W., Gloster, T. M., Fernandes, V. O., Prates, J. A. M., Montanier, C., Dumon, C., Williamson, M. P., Tunnicliffe, R. B., Liu, Z., Flint, J. E., Davies, G. J., Henrissat, B., Coutinho, P. M., Fontes, C. M. G. A., *et al.* (2010) Signature active site architectures illuminate the molecular basis for ligand specificity in family 35 carbohydrate binding module. *Biochemistry* **49**, 6193–6205 [CrossRef Medline](#)
  18. Couturier, M., Roussel, A., Rosengren, A., Leone, P., Stålbrand, H., and Berrin, J. G. (2013) Structural and biochemical analyses of glycoside hydrolase families 5 and 26  $\beta$ -(1,4)-mannanases from *Podospira anserina* reveal differences upon manno-oligosaccharide catalysis. *J. Biol. Chem.* **288**, 14624–14635 [CrossRef Medline](#)
  19. Okazawa, Y., Miyazaki, T., Yokoi, G., Ishizaki, Y., Nishikawa, A., and Tonozuka, T. (2015) Crystal structure and mutational analysis of isomalto-dextranase, a member of glycoside hydrolase family 27. *J. Biol. Chem.* **290**, 26339–26349 [CrossRef Medline](#)
  20. Fujimoto, Z., Kishine, N., Suzuki, N., Suzuki, R., Mizushima, D., Momma, M., Kimura, K., and Funane, K. (2017) Isomaltooligosaccharide-binding structure of *Paenibacillus* sp. 598K cycloisomaltooligosaccharide glucanotransferase. *Biosci. Rep.* **37**, BSR20170253 [CrossRef Medline](#)
  21. Suzuki, N., Fujimoto, Z., Kim, Y. M., Momma, M., Kishine, N., Suzuki, R., Suzuki, S., Kitamura, S., Kobayashi, M., Kimura, A., and Funane, K. (2014) Structural elucidation of the cyclization mechanism of  $\alpha$ -1,6-glucan by *Bacillus circulans* T-3040 cycloisomaltooligosaccharide glucanotransferase. *J. Biol. Chem.* **289**, 12040–12051 [CrossRef Medline](#)
  22. Light, S. H., Cahoon, L. A., Halavaty, A. S., and Freitag, N. E., Anderson, W. F. (2016) Structure to function of an  $\alpha$ -glucan metabolic pathway that promotes *Listeria monocytogenes* pathogenesis. *Nat. Microbiol.* **2**, 16202 [CrossRef Medline](#)
  23. Fujimoto, Z., Suzuki, N., Kishine, N., Ichinose, H., Momma, M., Kimura, A., and Funane, K. (2017) Carbohydrate-binding architecture of the multi-modular  $\alpha$ -1,6-glucosyltransferase from *Paenibacillus* sp. 598K, which produces  $\alpha$ -1,6-glucosyl- $\alpha$ -glucosaccharides from starch. *Biochem. J.* **474**, 2763–2778 [CrossRef Medline](#)
  24. Ji, S., Dix, S. R., Aziz, A. A., Sedelnikova, S. E., Baker, P. J., Rafferty, J. B., Bullough, P. A., Tzokov, S. B., Agirre, J., Li, F. L., and Rice, D. W. (2019) The molecular basis of endolytic activity of a multidomain alginate lyase from *Defluviitalea phaphyphila*, a representative of a new lyase family, PL39. *J. Biol. Chem.* **294**, 18077–18091 [CrossRef Medline](#)
  25. Mandelli, F., De Morais, M. A. B., De Lima, E. A., Oliveira, L., Persinoti, G. F., and Murakami, M. T. (2020) Spatially remote motifs cooperatively affect substrate preference of a ruminal GH26-type endo- $\beta$ -1,4-mannanase. *J. Biol. Chem.* **295**, 5012–5021 [CrossRef Medline](#)
  26. Guillén, D., Sánchez, S., and Rodríguez-Sanoja, R. (2010) Carbohydrate-binding domains: multiplicity of biological roles. *Appl. Microbiol. Biotechnol.* **85**, 1241–1249 [CrossRef Medline](#)
  27. Burnley, B. T., Afonine, P. V., Adams, P. D., and Gros, P. (2012) Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* **2012**, e00311 [CrossRef Medline](#)
  28. Ichinose, H., Kuno, A., Kotake, T., Yoshida, M., Sakka, K., Hirabayashi, J., Tsumuraya, Y., and Kaneko, S. (2006) Characterization of an exo- $\beta$ -1,3-galactanase from *Clostridium thermocellum*. *Appl. Environ. Microbiol.* **72**, 3515–3523 [CrossRef Medline](#)
  29. Bianchetti, C. M., Takasuka, T. E., Deutsch, S., Udell, H. S., Yik, E. J., Bergeman, L. F., and Fox, B. G. (2015) Active site and laminarin binding in glycoside hydrolase family 55. *J. Biol. Chem.* **290**, 11819–11832 [CrossRef Medline](#)
  30. Ishida, T., Fushinobu, S., Kawai, R., Kitaoka, M., Igarashi, K., and Samejima, M. (2009) Crystal structure of glycoside hydrolase family 55  $\beta$ -1,3-glucanase from the basidiomycete. *J. Biol. Chem.* **284**, 10100–10109 [CrossRef Medline](#)
  31. Karkehabadi, S., Helmich, K. E., Kaper, T., Hansson, H., Mikkelsen, N. E., Gudmundsson, M., Piens, K., Furdala, M., Banerjee, G., Scott-Craig, J. S., Walton, J. D., Phillips, G. N., and Sandgren, M. (2014) Biochemical characterization and crystal structures of a fungal family 3  $\beta$ -glucosidase, Cel3A from *Hypocrea jecorina*. *J. Biol. Chem.* **289**, 31624–31637 [CrossRef Medline](#)
  32. Chandrasekaran, R., and Janaswamy, S. (2002) Morphology of Western larch arabinogalactan. *Carbohydr. Res.* **337**, 2211–2222 [CrossRef Medline](#)
  33. Kitazawa, K., Tryfona, T., Yoshimi, Y., Hayashi, Y., Kawachi, S., Antonov, L., Tanaka, H., Takahashi, T., Kaneko, S., Dupree, P., Tsumuraya, Y., and Kotake, T. (2013)  $\beta$ -Galactosyl Yariv reagent binds to the  $\beta$ -1,3-galactan of arabinogalactan proteins. *Plant Physiol.* **161**, 1117–1126 [CrossRef Medline](#)
  34. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 [CrossRef Medline](#)
  35. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 849–861 [CrossRef Medline](#)
  36. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6**, 458–463 [CrossRef Medline](#)
  37. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 [CrossRef Medline](#)
  38. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 [CrossRef Medline](#)
  39. 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 [CrossRef Medline](#)
  40. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix. refine research papers. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 [CrossRef Medline](#)
  41. Burnley, B. T., and Gros, P. (2013) phenix.ensemble\_refinement: a test study of apo and holo BACE1. *Comput. Crystallogr. Newsl.* **4**, 51–58
  42. Forneris, F., Burnley, B. T., and Gros, P. (2014) Ensemble refinement shows conformational flexibility in crystal structures of human complement factor D. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 733–743 [CrossRef Medline](#)
  43. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons:

## Crystal structure of fungal GH43 galactanase

- Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 [CrossRef Medline](#)
44. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 [CrossRef Medline](#)
  45. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134 [CrossRef Medline](#)
  46. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786 [CrossRef Medline](#)
  47. Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, 529–533 [CrossRef Medline](#)
  48. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **33**, W299–W302 [CrossRef Medline](#)
  49. Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T., and Ben-Tal, N. (2013) ConSurf: using evolutionary data to raise testable hypotheses about protein function. *Isr. J. Chem.* **53**, 199–206 [CrossRef](#)
  50. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 [CrossRef Medline](#)
  51. Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003) ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* **19**, 163–164 [CrossRef Medline](#)
  52. Korotkova, O. G., Semenova, M. V., Morozova, V. V., Zorov, I. N., Sokolova, L. M., Bubnova, T. M., Okunev, O. N., and Sinityn, A. P. (2009) Isolation and properties of fungal  $\beta$ -glucosidases. *Biochemistry* **74**, 569–577 [CrossRef Medline](#)
  53. Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. E* **35**, 1547–1549 [CrossRef Medline](#)
  54. Stecher, G., Tamura, K., and Kumar, S. (2020) Molecular evolutionary genetics analysis (MEGA) for macOS. *Mol. Biol. Evol.* **37**, 1237–1239 [CrossRef Medline](#)
  55. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDScript server. *Nucleic Acids Res.* **42**, W320–W324 [CrossRef Medline](#)
  56. Bohne, A., Lang, E., and von der Lieth, C. W. (1998) W3-SWEET: carbohydrate modeling by internet. *J. Mol. Model.* **4**, 33–43 [CrossRef](#)
  57. Bohne, A., Lang, E., and von der Lieth, C. W. (1999) SWEET—WWW-based rapid 3D construction of oligo- and polysaccharides. *Bioinformatics* **15**, 767–768 [CrossRef Medline](#)