


## RESEARCH ARTICLE

# Chiral recognition mechanism of cellobiohydrolase Cel7A for ligands based on the $\beta$ -blocker propranolol: The effect of explicit water molecules on binding and selectivities

Alexandra Fagerström<sup>1</sup> | Tommy Liljefors<sup>2,†</sup> | Mats Sandgren<sup>3</sup> | Roland Isaksson<sup>4</sup> | Jerry Ståhlberg<sup>3</sup> | Ulf Berg<sup>1</sup> 

<sup>1</sup>Department of Chemistry, Lund University, Lund, Sweden

<sup>2</sup>Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>3</sup>Department of Molecular Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

<sup>4</sup>Department of Biomedical Sciences, Linné university, Kalmar, Sweden

## Correspondence

Ulf Berg, Department of Chemistry, Lund University, P.O. Box 124, Lund SE-221 00, Sweden.

Email: [ulf.berg@chem.lu.se](mailto:ulf.berg@chem.lu.se)

## Funding information

Research School in Pharmaceutical Sciences, Lund University; Linné University, Kalmar

## Abstract

Proteins are useful chiral selectors. In order to understand the recognition mechanism and chiral discrimination, the binding of the (*R*)- and (*S*)-enantiomers of a series of designed amino alcohol inhibitors based on propranolol to cellobiohydrolase Cel7A (*Trichoderma reesei*) has been studied more closely. X-ray crystal structures were determined of the protein complex with the (*R*)- and (*S*)-enantiomers of the strongest binding propranolol analogue. The combination of the structural data, thermodynamic data from capillary electrophoresis and microcalorimetry experiments and computational modelling give a clearer insight into the origin of the enantioselectivity and its opposite thermodynamic signature. The new crystal structures were used in computational molecular flexible dockings of the propranolol analogues using the program Glide. The results indicated that several water molecules in the active site were essential for the docking of the (*R*)-enantiomers but not for the (*S*)-enantiomers. The results are discussed in relation to the enantiomeric discrimination of the enzyme. Both dissociation constants ( $K_d$  values) and thermodynamical data are included to show the effects of the structural modifications in the ligand on enthalpy and entropy in relation to the enantioselectivity.

## KEYWORDS

Cel7A, computational chemistry, enantioselectivity, inhibitor binding, propranolol, thermodynamics

## INTRODUCTION

Access to enantiomerically pure material is of great importance in many scientific areas, particularly in biology and medicine. This can be accomplished either through asymmetric synthesis or by separation of the enantiomers from the racemic mixture. Proteins have been

successfully used as chiral selectors in liquid chromatographic and capillary electrophoretic techniques to determine the enantiomeric composition of various drugs.<sup>1</sup> However, in most cases, it is difficult to foresee affinity and selectivity in the interplay between the enantiomers and the protein (selector), as well as which kind of interactions that are involved. This requires determination of the influence of different functionalities of the ligands, which amino acid residues of the protein are involved in the chiral discrimination and how flexible the

<sup>†</sup>Deceased author.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Natural Sciences* published by Wiley-VCH GmbH.

protein–ligand interactions are. The binding of a ligand to an enzyme can be studied by enzyme activity and inhibition, thermodynamical data and kinetic studies. Together with experimental affinity and selectivity, such studies give useful information regarding the interactions between ligand and protein. Another very attractive way to gain knowledge of a chiral recognition mechanism is to use the X-ray crystal structure of the ligand–protein complex. Such a structure can provide accurate information about positioning and consequent interactions and can be further studied and evaluated by computational methods. This approach is also a powerful tool for structure-based design of new ligands.

A protein that has been successfully exploited as a chiral selector is cellobiohydrolase Cel7A (also known as CBH I or CBH 1) from the filamentous fungus *Trichoderma reesei* (a clonal derivative of *Hypocrea jecorina*). Cel7A plays a key role in the fungal cellulolytic enzyme system and it is the major secreted enzyme when the fungus grows on cellulose. The enzyme can thus be produced in large quantities at a reasonable cost,<sup>2</sup> an important factor for making Cel7A attractive as a chiral selector, and commercial columns are available.<sup>3</sup>

Cel7A has a bimodular organisation with two independently folded domains, a large catalytically active module of ~430 amino acids and a small cellulose-binding module (CBM1) of ~35 amino acids, connected by a flexible and highly glycosylated linker peptide of ~30 amino acids. Three-dimensional structures have been determined for both the catalytic<sup>4</sup> and cellulose-binding<sup>5</sup> modules. Although some enantioselectivity for *rac*-propranolol was found with the isolated CBM of Cel7A,<sup>6</sup> the major selective site is within the active site of the catalytic module. An intriguing structural aspect of the catalytic module is that the active site is buried within a 50 Å-long cellulose-binding tunnel.<sup>7</sup> This seems to be an important feature for enantioselectivity since the homologous endoglucanase Cel7B in *T. reesei*, which has very similar structure but an open cleft rather than a tunnel, showed very poor enantioselectivity.<sup>8</sup> The substrate binding site can harbour 10 glucose residues of a cellulose chain, with seven glucosyl-binding subsites, numbered –7 to –1, from the non-reducing end of the chain, and three subsites, numbered +1 to +3, towards the reducing end of the chain, from the catalytic centre that catalyses hydrolysis of the glycosidic bond between subsites –1 and +1.<sup>9</sup>

Previously, Cel7A has been used as a chiral selector to resolve numerous pharmaceuticals.<sup>10</sup> The enzyme has proven to be particularly useful for separation of enantiomers of adrenergic  $\beta$ -blockers, which is also the most extensively studied application. The experimental conditions in both high-performance liquid chromatography and capillary electrophoresis (CE) regarding pH,<sup>11</sup> type of buffer,<sup>12</sup> temperature<sup>13</sup> and ionic strength<sup>14</sup> have systematically been varied, not only to optimise the separations but also to obtain information about the physical chemical parameters relating to selectivity and affinity. Furthermore, propranolol has been shown to have an inhibitory effect,<sup>8</sup> and together with alprenolol, binding energies have been determined in thermodynamical studies.<sup>13,15,16</sup>

Important insight into the structural basis for ligand binding and chiral discrimination has been provided by the X-ray crystal structure of Cel7A in complex with (*S*)-propranolol bound at the active

site<sup>14</sup> and the structure of the homologous enzyme Cel7D from *Phanerochaete chrysosporium* in complex with (*R*)-propranolol.<sup>17</sup> An overlay of these two structures revealed that the enantiomers were binding in similar locations in glucosyl-binding subsites –1 and +1 with the secondary amine interacting with the catalytic carboxylic acid residues (Glu and Asp), and the naphthyl moiety stacking with the corresponding tryptophan residues (normally the glucosyl-binding platform in site +1). However, in this comparison, the amino alcohol side chains adopted different conformations. Under the assumption that the enantiomers would bind similarly in the two enzymes, it was suggested that the major reasons why (*S*)-propranolol is more tightly bound than (*R*)-propranolol are two favourable hydrogen bonds between the hydrolytic group and the enzyme and further van der Waals interactions of the slightly deeper positioned side chain of the (*S*)-enantiomer.<sup>14</sup> In light of the complete loss of enantiomeric separation upon mutation of either of the catalytic glutamates,<sup>17</sup> it was also concluded that the major driving force behind binding and enantioselectivity of these compounds is the bidentate interaction between their positively charged secondary ammonium and the two negatively charged glutamate residues, while the other interactions contributed less. It should be pointed out that even if the catalytic domains of Cel7A and Cel7D have similar sequences (56% identity) and highly conserved catalytic sites, there are distinct structural differences around the propranolol-binding site, particularly near the naphthyl moiety, which affects the binding differently in the two enzymes. Unfortunately, no direct comparison of the enantiomers could be made since X-ray structures could not be obtained of Cel7A with (*R*)-propranolol or of Cel7D with (*S*)-propranolol despite many trials. The lack of truly comparable structures, that is, with both enantiomers in the same enzyme, has also hampered the use of computational methods for studies of the chiral recognition mechanism.

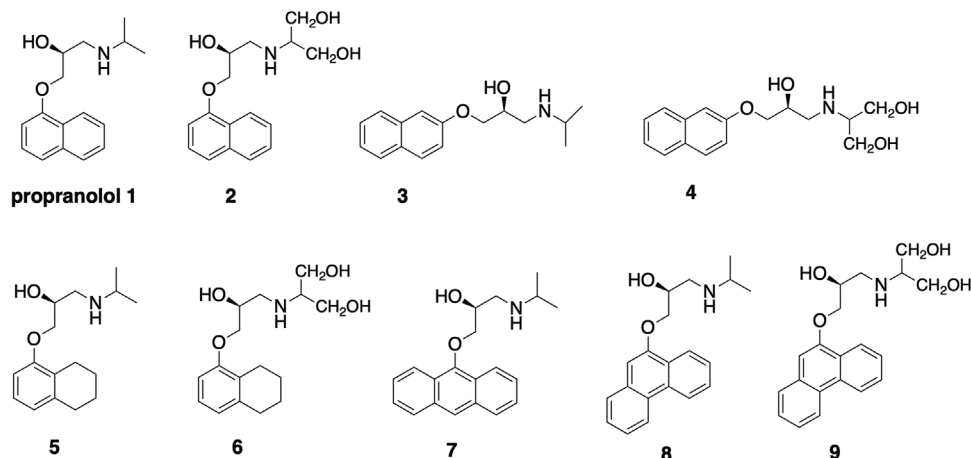
In a previous study, computational modelling based on the Cel7A/(*S*)-propranolol structure was used to design propranolol analogues and to elucidate the importance of different ligand motifs. The aim was to study both weakening and reinforcement of the binding and the relation to selectivity, as well as using the system as a model for rational design of more powerful inhibitors. Several ligands with higher affinity were indeed obtained.<sup>18</sup>

In this work the binding of both the (*S*)- and the (*R*)-enantiomer to Cel7A is studied by X-ray crystallography, CE, microcalorimetry and computational modelling. Scheme 1 shows the compounds involved in the study. This is the first time crystal structures of complexes of both enantiomers have been available, facilitating a comparable in-depth study of the enantioselectivity.

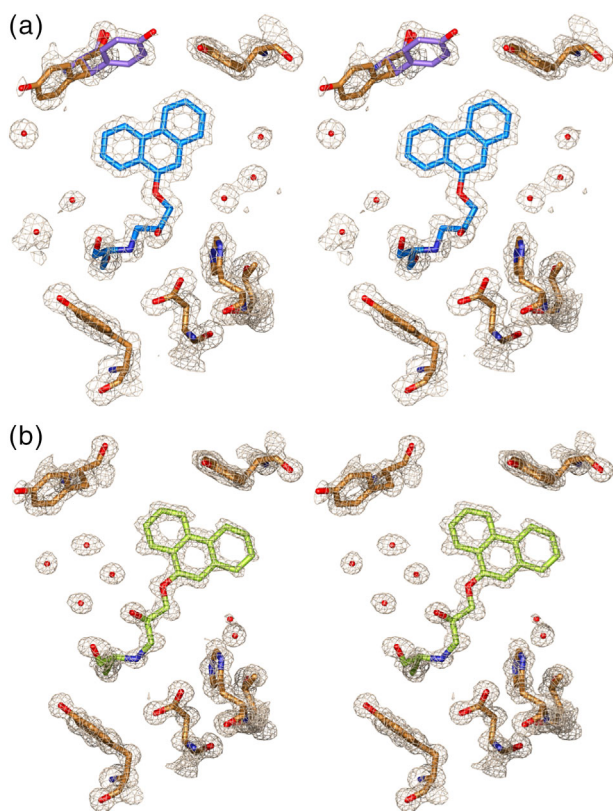
## RESULTS AND DISCUSSION

### Protein structures

Crystals of the catalytic module of *T. reesei* Cel7A were soaked with the (*S*)- and (*R*)-enantiomers of **9**. The crystals were isomorphous with previous crystal structures,<sup>19</sup> and complete diffraction data sets to 1.60



**SCHEME 1** Ligands based on propranolol designed for binding to Cel7A. The *S* enantiomers are shown.



**FIGURE 1** Electron density for (a) (S)-9 and (b) (R)-9 in the crystal structures in complex with Cel7A, shown in divergent stereo. The density maps were contoured at 1.0 sigma around the ligands. The orientation of view is the same for both panels. All the structure figures were made as follows: molecular objects were created with O,<sup>24</sup> and the images were rendered with POV-Ray<sup>25</sup> utilising the Molray web interface.<sup>26</sup>

and 1.05 Å resolution for (S)-9 and (R)-9, respectively (Figure 1), were obtained using synchrotron radiation. The ligands were clearly visible in the electron density prior to inclusion in the structure models. Statistics of crystallographic data and the refined structures are summarised

in Table 1. The structure of the Cel7A protein is almost identical in the two complexes, with a root mean square deviation (rmsd) of 0.13 Å, and there is also a high similarity of the protein structures when compared to the previous complex with (S)-1 (Protein Data Bank [PDB] accession code 1DY4)<sup>19</sup> with rmsd values of 0.12 and 0.16 Å for the (S)-9 and the (R)-9 complexes, respectively.

The electron density of the (S)-9 complex shows that the side chain of Tyr371, which covers the ligand-binding site, is present in two discrete positions related by ~110° rotation around chi1 (Figure 1a) and 9.2 Å shift of the O(H) atom, as was observed in previous structures.<sup>19</sup> Tyr371 was predominantly positioned above glucosyl-binding subsite +1 in the (S)-1 complex, and above subsite -1 in structures with bound carbohydrates. In the current (S)-9 complex, there seems to be a preference for the position above site -1. Flexibility is also evident for Thr246 and Tyr247 at the tip of the so-called “exo-loop”<sup>20</sup> that meets Tyr371 above the ligand-binding site. For these residues, however, the positional shifts are less than 1 Å for corresponding atoms and it is not clear if they take on discrete conformations in conjunction with the flip of the Tyr371 side chain. All other residues with alternate conformations (listed in Section 4) are distant from the ligand-binding site.

The density for the ligand (S)-9 is as good as for surrounding protein atoms (Figure 1a), indicating that it is binding at full occupancy. Around the catalytic centre, the position of (S)-9 is practically identical to that of (S)-1, but the larger phenanthryl group seems to cause a slight shift “upwards” of the ligand atoms closest to the aromatic moiety (Figure 2). The ether oxygen is shifted 0.64 Å, and the largest distance is 1.67 Å between the atoms in the phenanthryl of (S)-9 and the corresponding atoms in the naphthyl of (S)-1. This displacement is within the plane of the aromatic system, which remains at similar distance to the hydrophobic Trp376-binding platform (3.5–4.0 Å). The *B* factors (temperature factors) are slightly elevated for the phenanthrene atoms, indicating some flexibility.

The found interactions of (S)-1 and (S)-9 with the protein were very similar in the two complexes. In addition, the hydroxyl groups introduced in 9 fit well into their respective sites and make favourable

**TABLE 1** Summary of statistics for X-ray diffraction data and the refined Cel7A complex structures.

Parameter	(S)-9 complex, 2V3R	(R)-9 complex, 2V3I
Data collection <sup>a</sup>		
X-ray source, beamline	MAX Laboratory, I711	ESRF, ID23-1
Wavelength (Å)	1.082	0.925
Cell dimensions (Å, °)	82.9, 82.9, 110.6; 90°, 90°, 90°	83.0, 82.9, 110.7; 90°, 90°, 90°
Resolution (Å) <sup>b</sup>	25–1.60 (1.69–1.60)	30–1.05 (1.07–1.05)
Unique reflections	50,461	174,883
Average multiplicity <sup>b</sup>	7.3 (7.2)	9.6 (6.2)
Completeness (%) <sup>b</sup>	100.0 (100.0)	99.2 (90.0)
$R_{\text{merge}}$ (%) <sup>b,c</sup>	7.7 (36.0)	10.8 (35.6)
$\langle I/\sigma I \rangle$ <sup>b</sup>	20.9 (8.0)	18.1 (5.3)
Structure refinement <sup>d</sup>		
R-factor/R-free (%)	13.6 (17.9)	12.0 (14.0)
No. of protein atoms (average B, Å <sup>2</sup> )	3400 (12.9)	3594 (12.5)
No. of water molecules (average B, Å <sup>2</sup> )	688 (26.5)	700 (26.7)
No. of GlcNAc atoms (average B, Å <sup>2</sup> )	28 (35.4)	14 (17.5)
No. of ligand atoms (average B, Å <sup>2</sup> )	25 (11.6)	25 (17.0)
rms bond length (Å)	0.012	0.016
rms bond angle (°)	1.41	1.77
No. of Ramachandran plot outliers (%) <sup>e</sup>	2 (0.5)	4 (1.0)

Note: Protein crystals were of space group I222 with one protein molecule per asymmetric unit.

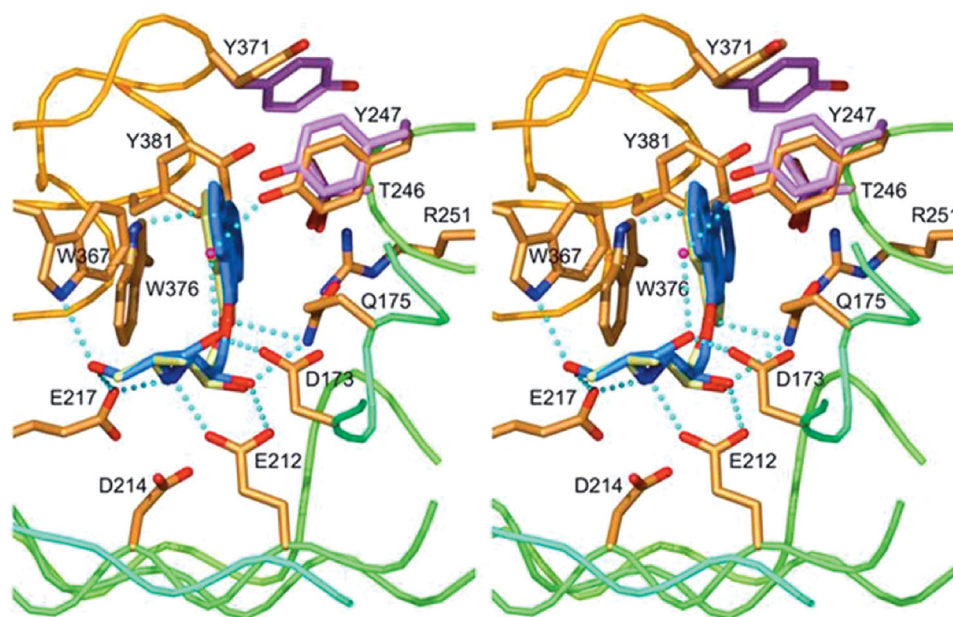
<sup>a</sup>Taken from TRUNCATE.<sup>21</sup>

<sup>b</sup>Values in the highest resolution shell are given in parentheses.

<sup>c</sup> $R_{\text{merge}} = \sum_h \sum_i |I_{h,i} - \langle I_h \rangle| / \sum_h \sum_i I_{h,i}$ .

<sup>d</sup>Calculated using MOLEMAN2.<sup>22</sup>

<sup>e</sup>A stringent-boundary Ramachandran plot was used.<sup>23</sup>



**FIGURE 2** Binding of (S)-9 (atomic colours, with blue carbons) at the active site of Cel7A and comparison with (S)-1 (pale green; Protein Data Bank [PDB] accession code 1DY4),<sup>19</sup> in divergent stereo. The side chains of a number of binding-site residues are shown with carbons in gold, and alternate conformations for Thr246, Tyr247 and Tyr371 are shown in lilac. Hydrogen bonds with the ligand and the closest distance to Trp376 are indicated with cyan bubbled lines. An interacting water molecule is presented as a red sphere. Selected segments of the main chain are shown, rainbow coloured blue to red from N- to C-terminus.

**TABLE 2** Ligand–protein interaction distances in Å in the crystal structures of Cel7A in complex with (*S*)-9, (*R*)-9, (*S*)-1 and *Phanerochaete chrysosporium* Cel7D with (*R*)-1

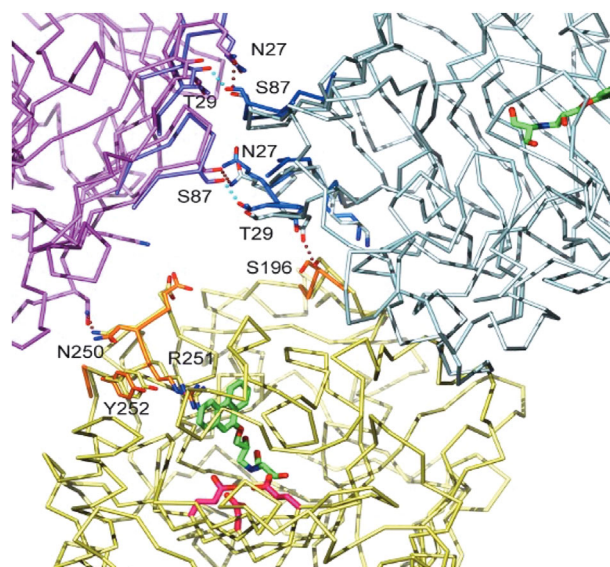
Ligand atom–protein residue	( <i>S</i> )-9	( <i>S</i> )-1	( <i>R</i> )-9	( <i>R</i> )-1 (Cel7D)
PDB accession code	2V3R	1DY4	2V3I	1H46
Ring system–Trp376 (closest atom distance)	3.36	3.62	3.53	3.24 (Trp373)
Amine–Glu212 (nucleophile)	2.75	2.81	2.81	2.53 (Glu207)
Amine–Glu217 (acid/base)	2.79	2.83	2.86	3.10 (Asp209)
Ether O–Gln175	3.02	2.98	–	–
( <i>S</i> )-OH–Glu212	2.87	2.98	–	–
( <i>S</i> )-OH–Gln175	2.90	2.76	–	–
( <i>R</i> )-OH–water	–	–	2.78	3.00
3OH–Trp367 NE	2.91	–	2.95	–
3OH–Glu217	2.61	–	2.64	–
4OH–Asp173	2.62	–	2.74	–
4OH–water	3.03	–	2.93	–

Abbreviation: PDB, Protein Data Bank.

hydrogen bonds with the protein (Table 2). The ligand O3 hydroxyl occupies the site dedicated for binding the 6-hydroxyl of a glucose residue in subsite –1, whereas O4 utilises the site for the 3-hydroxyl of the same glucosyl residue.

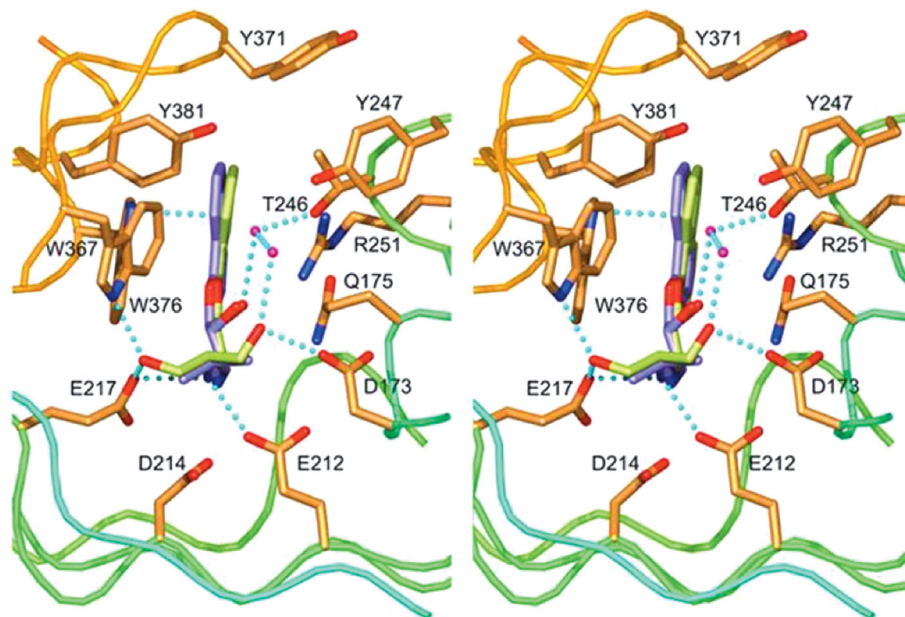
In the (*R*)-9 complex, only the conformation of Tyr371 situated above subsite –1 is visible. On the other hand, alternate conformations are evident for Arg251 and Asp259, which are in van der Waals contact with the phenanthrene of (*R*)-9 on the side opposite to the Trp376 platform. The Arg251 side chain extends to the active site from a surface loop where a stretch of six residues, Asp249–Gly254, shows elongated density for both side chain and main chain. Interestingly, Asp249 and Asn250 in the loop are located at a crystal contact, suggesting that binding events at the active site may be transmitted via Arg251 to the crystal contact surface and perhaps affect protein crystallisation, and may have contributed to previous failures in producing crystals containing the (*R*)-enantiomers. On the same side of the protein, there was yet another loop close to a crystal contact, Pro194–Asn197, with clearly elongated density where alternate conformations were modelled. We were also surprised to find a disordered region on the opposite side of the protein. Alternate conformations were included for 22 residues here, nos. 13–20, 25–30 and 85–92. This region forms the crystal contact with two neighbouring protein molecules, and the C $\alpha$  positional shifts are largest for Thr26 and Asn27 (~1.2 Å), which interact with Ser87 (1.0 Å shift of Ca) of a symmetry-related protein.

The hydrogen bonding pattern between the molecules is different for the two modelled conformations, indicating that it is a matter of a transition between two discrete conformations rather than continuous disorder (Figure 3). Similar disorder was not observed in the structures of Cel7A in complex with (*S*)-1 or (*S*)-9. It thus seems that the binding of the (*R*)-enantiomer of  $\beta$ -blockers tends to induce conformational changes in the protein, and one may speculate that this is the reason that we could not obtain a crystal structure of Cel7A in complex with (*R*)-1.



**FIGURE 3** Alternate conformations in the structure of Cel7A in complex with (*R*)-9. The hydrogen bonding patterns at the crystal contact interface differ depending on which conformations occur at the contact surface. The C $\alpha$  trace of three symmetry-related Cel7A proteins in the crystal is shown with selected residue side chains numbered, the catalytic residues in magenta and the (*R*)-9 molecules in green.

The bound (*R*)-9 molecule shows slightly weaker density (Figure 1b) and higher *B* factors than surrounding protein atoms, indicating some flexibility. Superposition of the Cel7A/(*R*)-9 complex with that of *P. chrysosporium* Cel7D with (*R*)-1 bound (PDB accession code 1H46)<sup>17</sup> showed a nearly perfect overlap of (*R*)-9 and (*R*)-1 with less than 0.5 Å between corresponding atoms (Figure 4). There are small differences in the position of the amino nitrogen in the ligands (0.43 Å) and of the catalytic residue side chains, which alter the interaction pattern



**FIGURE 4** Binding of (*R*)-**9** (atomic colours, with light green carbons) at the active site of Cel7A and comparison with (*R*)-**1** in *P. chrysosporium* Cel7D (violet, Protein Data Bank [PDB] accession code 1H46),<sup>17</sup> in divergent stereo. The side chain of a number of binding-site residues are shown with carbon in gold. Hydrogen bonds with the ligand and the closest distance to Trp376 are indicated with cyan bubbled lines. Interacting water molecules are represented as red spheres. Selected segments of the main chain are shown, rainbow coloured blue to red from N- to C-terminus.

slightly. The interaction between the nitrogen and the catalytic nucleophile (Glu212 in Cel7A, Glu207 in Cel7D) is similar, but in Cel7A/(*R*)-**9**, the amino nitrogen is closer to the catalytic acid/base Glu217 than to assisting aspartate residue Asp214, while the opposite is true for Cel7D/(*R*)-**1** (Table 2). As with (*S*)-**9**, the two additional hydroxyl groups on (*R*)-**9** provide further hydrogen bonds to the protein.

The (*S*)-**9** and the (*R*)-**9** enantiomers bind very similarly at the end of the ligand side chain with less than 0.5 Å atom deviations, and the two hydroxyl groups at the end and the secondary amino group form similar interactions in the enantiomer complexes (Figure 5 and Table 2). Proceeding along the side chain to the stereogenic centre the enantiomers exhibit different conformations and interactions with the protein. The hydroxyl at the stereogenic centre of (*S*)-**9** interacts directly with the protein through hydrogen bonds with Gln175 and Glu212, while the corresponding hydroxyl of (*R*)-**9** interacts indirectly via a hydrogen bond mediated by a water molecule. The phenanthryl groups are at about the same distance from Trp376 platform (~3.5–4.5 Å) but are slightly tilted relative to each other (~10°), and in the case of (*R*)-**9**, the ring system is also buried deeper in the cleft. The distance between corresponding phenanthrene atoms ranges from 2.4 to 3 Å. Schematic illustrations of protein–ligand interactions (LigPlots) are provided in Figure S1.

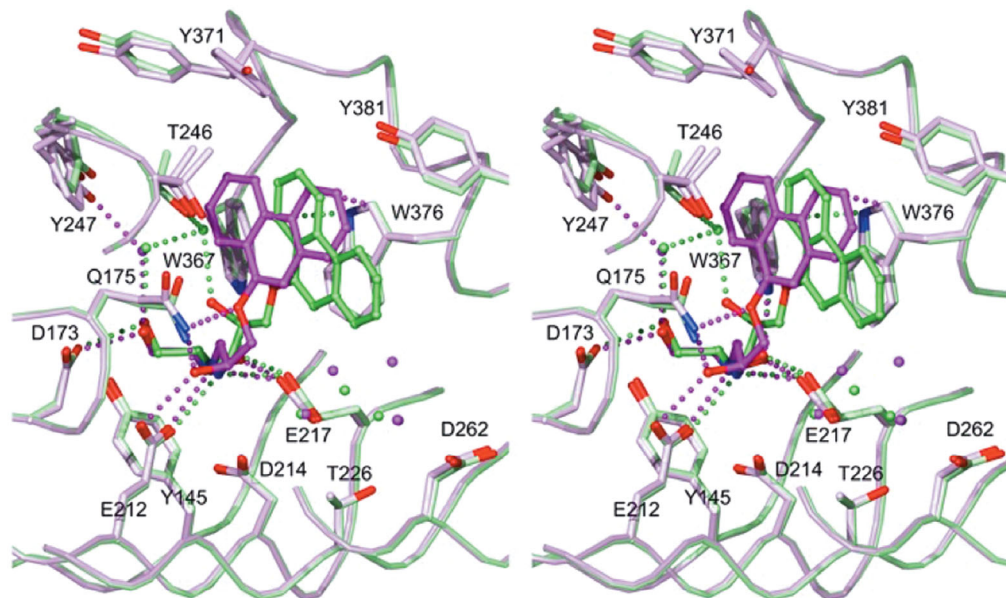
### Calculation of affinity and selectivity

The crystal structures of Cel7A in complex with (*S*)-**9** and (*R*)-**9**, denoted Cel7A/*S***9** and Cel7A/*R***9**, respectively, were initially prepared for molecular flexible docking with Glide.<sup>27,28</sup> Glide is a computational

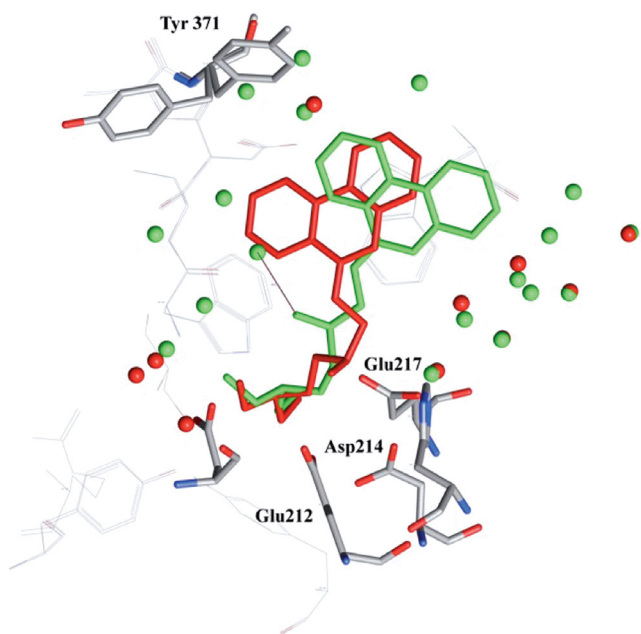
program for finding the correct binding modes for a large set of test cases with high docking accuracy across a diverse range of receptor types. Initial computations were performed without explicit water molecules present. To evaluate the docking performance, the (*R*)- and (*S*)-enantiomers of **9** were docked back into empty protein structures of Cel7A/*R***9** and Cel7A/*S***9**, respectively. For (*S*)-**9**, the resulting poses were very similar to the ligand conformation in the crystal structure. However, in the attempts to dock (*R*)-**9** into the Cel7A/*R***9** protein structure, the ligand did not obtain positions or conformations similar to the one in the crystal structure. In general, the phenanthrene moiety had shifted quite some distance towards the tunnel entrance, and the hydroxyl at the stereogenic centre was positioned in a fashion as in the case of (*S*)-**9**.

A comparison of the original complexes of the two enantiomers of **9** disclosed that the water molecules in the active site were quite differently positioned (Figure 6), and in the case of (*R*)-**9**, it was very likely that the water molecules restricted the flexibility of the ligand through bridging interactions between the ligand and the protein. This was not the case for the (*S*)-enantiomer.

This modelling indicated that the presence of water was crucial for the structure of the (*R*)-enantiomer complex. Through a careful elimination of water molecules and repeated docking, we could identify 14 water molecules in the vicinity of the ligand, which influenced the docking result. Five of these water molecules, between Glu217 and Arg267, formed a cluster via hydrogen bonds to each other and the surrounding amino acid residues. This cluster seemed important to position the phenanthrene moiety in the cavity. Furthermore, a water molecule close to the stereogenic centre of (*R*)-**9** in the crystal structure was at hydrogen bonding distance to both the hydroxyl



**FIGURE 5** Comparison of the binding of (*S*)-**9** (violet) and (*R*)-**9** (green) at the active site of Cel7A in divergent stereo. Selected side chains, main chain, water molecules and ligand–protein interactions are shown in associated colours.

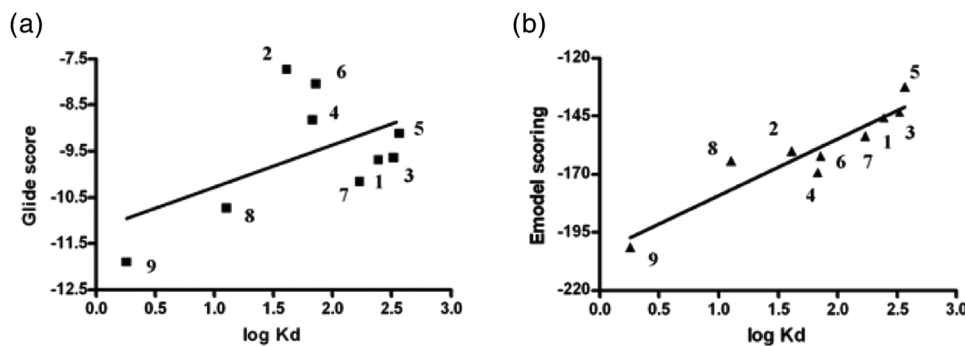


**FIGURE 6** Superposition of crystal structures of complexes with (*R*)-**9** (green) and (*S*)-**9** (red) showing water molecules present in the active site. The hydrogen bond between the hydroxyl at the stereogenic centre of (*R*)-**9** and water is indicated.

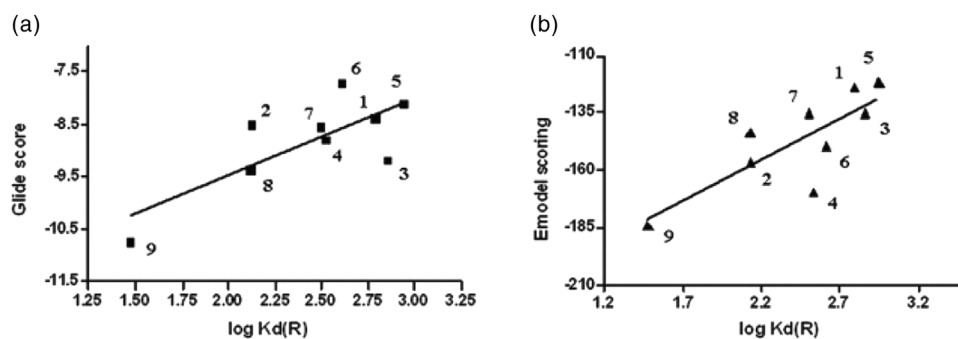
of the ligand and a second water molecule. To favour this directionality of the hydroxyl at the stereogenic centre, both water molecules were required in the dockings. In Cel7A/R9, a glycerol molecule was present close to Asp173 and at hydrogen bonding distance to the dihydroxy functionality of (*R*)-**9**, whereas a water molecule was observed in the same position in the Cel7A/S9 structure. This indicated an important hydrogen bond, and for simplicity the hydroxyl of the glycerol in

Cel7A/R9 was replaced with a water molecule for docking purposes. Another five water molecules, surrounding the remaining part of the phenanthrene ring system were kept in the Cel7A/R9 structure even if these were found to be of minor importance. It should be noted that these were located where Tyr371 was positioned in the (*S*)-**9** complex, shielding the ligand from the bulk water molecules of the solvent. In order to have two comparable protein structures, 14 water molecules in similar positions were kept in the Cel7A/S9 structure. This initial docking evaluation gave the first clear indication of the importance of the water molecules in the ligand–protein interaction.

The protein structures were prepared for the docking through energy minimisation of the water molecules in the structures, keeping all other atoms fixed. An excess of water molecules was present in these minimisations to restrict the freedom to move for the relevant water molecules, therefore keeping them in essentially the same place. The excess water molecules were removed before protein preparation, and grid construction was performed prior to the Glide dockings. The enantiomers of all ligands shown in Scheme 1 were docked into their respective protein structures. This provided us with poses of both enantiomers of **9** that correlated well with the experimental ligand conformations found in the respective crystal structures and confirmed the set-up. When (*S*)-**1** was docked into the Cel7A/S9 structure, the ligand conformation and positioning showed very good agreement with the previously published crystal structure (PDB accession code 1DY4).<sup>19</sup> To evaluate the difference between the two enantiomeric crystal structures, (*R*)-**9** was docked into Cel7A/S9 and (*S*)-**9** into Cel7A/R9. However, the resulting ligand poses did not correlate with their respective crystal structures. Considering the differences in positioning of the enantiomers in the crystal structure, as described above, this was not too surprising. As a result, all (*S*)-enantiomers fitted in the ligand-free Cel7A/S9 and (*R*)-enantiomers in Cel7A/R9.



**FIGURE 7** Calculations for the (S)-enantiomers. (a) GlideScore versus  $\log K_d$  for all ligands ( $r^2 = 0.28$ ). (b) Emodel scoring versus  $\log K_d$  for all ligands ( $r^2 = 0.87$ ).



**FIGURE 8** Calculations for the (R)-enantiomers. (a) GlideScore versus  $\log K_d$  for all ligands ( $r^2 = 0.58$ ). (b) Emodel scoring versus  $\log K_d$  for all ligands ( $r^2 = 0.64$ ).

To compare the docking results to the experimentally obtained dissociation constants ( $K_d$ ), both GlideScore and Emodel scorings for all ligands were plotted as a function of  $\log K_d$ . For both enantiomers, it was shown that the Emodel scorings correlated better with the experimental data than those of GlideScore (Figures 7 and 8). At a closer examination of the resulting poses, it was noted that compounds containing the dihydroxyl functionality had less favoured ligand conformations, with one of the hydroxyl groups at the end of the amino alcohol side chain in an eclipsed conformation. These conformations give a conformational penalty, which affects the GlideScore more than the Emodel scoring. The conformations of Cel7A/1 and Cel7A/2 in solution were recently studied by various nuclear magnetic resonance techniques, and the results concur with our results in the solid state.<sup>29</sup>

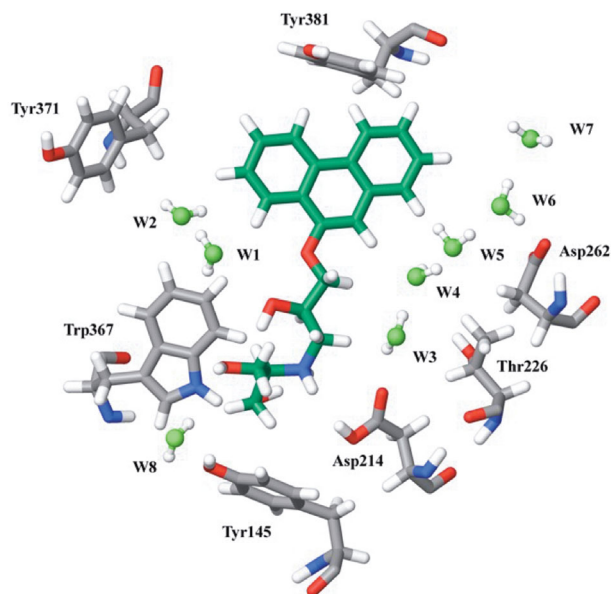
The results from docking and scoring showed an unexpectedly good correlation with the experimental affinities, in particular the Emodel scoring function, considering that it was concluded in a recently reported assessment of docking programs and scoring functions (including Glide and GlideScore) that none of the programs or scoring functions made a useful prediction of ligand binding affinity.<sup>30</sup> Hence, the docking results' ability to distinguish between compounds close in affinity exceeded our expectations. Comparing the results from Emodel scoring function, it can be seen that the more strongly bound compounds are predicted to show a greater selectivity, which is in agreement with experimental data.<sup>31</sup>

## Water molecules in the active site

As described above, several of the water molecules present in the crystal structure of Cel7A/(R)-9 were important to obtain a good working model, that is, a model where the (R)-9 ligand molecule adopts the same position and conformation as in the crystal structure. To determine the interaction energies for the water molecules, the GRID program was used with water as a probe (OH2) to identify positions of strongly bound water molecules.<sup>32,33</sup> When the receptor was used without ligand or water molecules present, the only energy density of interest shown was at  $-10 \text{ kcal mol}^{-1}$  and overlapped with the ammonium group of the ligand. As noted before, several of the water molecules included in the docking experiments were involved in additional hydrogen bonds to neighbouring water molecules and/or to the ligand. Therefore, we also performed calculations with the ligand included in the protein and then with different combinations of the water molecules from the crystal structure to obtain values from an environment more closely related to what was found in the crystal structure (Figure 9).

In the presence of the ligand, water molecules W3, W4 and W8 were calculated to have interaction energies between  $-13$  and  $-10 \text{ kcal mol}^{-1}$ . The binding cavity was explored with a water probe while keeping one or several alternating water molecules in the positions given by the crystal structure. The results showed that both W1





**FIGURE 9** Water molecules W1–W8 (light green oxygens), which were important for the docking results of (*R*)-**9** (dark green carbons), studied with respect to displacement energies using GRID.

and W2 have interaction energies of about  $-11 \text{ kcal mol}^{-1}$ . The clustered water molecules W3–W7 were suggested to restrict the ligand movement and were calculated to have interaction energies between  $-9.8$  and  $-13.6 \text{ kcal mol}^{-1}$ , where W3 and W4 were most strongly bound. This confirms that when the water molecules are clustered, as seen in the crystal structure of Cel7A/(*R*)-**9**, they most probably do impose a limitation on the flexibility of the ligand. Furthermore, the hydroxyl from glycerol, in the vicinity of Asp173 in Cel7A/*R*-**9** was also calculated to be tightly bound, which justifies the use of a water molecule in this position in the dockings. Thus, the GRID computations support the importance of the water molecules W1–W8 obtained in the crystallographic study.

### Thermodynamical data

Enthalpies and entropies have been determined for weaker binding ligands (*S*)-**1**, (*R*)-**1**, (*R*)-**2**, (*S*)-**3** and (*R*)-**3** using CE at pH 5.2 (Table 3). The results show that the selectivity is entropy driven for all the amino alcohols in the study and that the additional dihydroxy functionality makes the enthalpy term more negative, in accordance with stronger electrostatic interactions.

Attempts to determine binding energies of **9** were made by means of microcalorimetry using the same buffer and pH as in the CE experiments (Table 3).<sup>34</sup> However, only data from (*S*)-**9** were obtained under these conditions and show the same importance of the entropy factor as, for example, (*S*)-propranolol (**1**).

Unfortunately, the free energies,  $\Delta G^\circ$ , from microcalorimetry and CE experiments differ in absolute values. Hence, the two experimentally different methods do not seem to be comparable with respect to

the absolute numerical values even though the relative importance of entropy versus enthalpy is maintained. Regardless of the experimental conditions, both data sets show the importance of the entropy contribution. Concerning (*R*)-**9**, no reliable data were obtained at pH 5.2 due to a too low signal-to-noise ratio. Similar behaviour has previously been observed in measurements of alprenolol, which would be expected to have similar interactions and selectivity mechanisms as propranolol in this system.<sup>35</sup> In addition, since the buffer and pH have a great impact on the absolute numbers, a comparison is not possible between the previously obtained  $K_d$  values<sup>20</sup> and those calculated from the thermodynamical experiments presented here. However, a similar common trend would be expected.

### CONCLUSIONS

In conclusion, we present the crystal structure of the complexes of Cel7A with the (*S*)- and the (*R*)-enantiomers of the designed propranolol derivative **9**. Both enantiomers bind to the active site of the enzyme but show some significant conformational differences and altered interactions with the protein and water molecules in the active site. A microcalorimetric study confirms earlier results, which shows that the binding and enantioselectivity are entropy driven. A computational analysis of the docking of the ligands to the protein indicates that a likely origin of the entropic force behind the chiral discrimination is due to the presence of important water clusters in the (*R*)-enantiomeric complexes. The water molecules impose a limitation on the flexibility of the ligand and the intra-molecular interactions reduced the free movement of several water molecules instead of expelling them from the site, which is further supported by the thermodynamical data. This is new and important information, not only in further studies of Cel7A and the binding and selectivity of different ligands, but also to provide a model system for further computational calculations and other studies of the effect of water on binding and selectivities.

### EXPERIMENTAL

#### Protein crystallisation and structure determination

The catalytic domain of *T. reesei* Cel7A was prepared and crystallised as previously described<sup>19</sup> using the hanging-drop vapour-diffusion method<sup>36</sup> with 3 or 6 mg mL<sup>-1</sup> Cel7A protein in 10 mM sodium acetate, pH 5.0, 18 or 19.4% (w/v) monomethyl ether polyethylene glycol 5000, 14% (v/v) glycerol and 11 mM CoCl<sub>2</sub>. Crystals appeared within 1–5 days at room temperature. A few grains of dry powder of the inhibitors, (*R*)-**9** and (*S*)-**9**, respectively, were added to the surface of the drops and allowed to diffuse into the crystals for 24 h, after which individual crystals were picked with 0.1–0.5 mm loops and flash-frozen in liquid nitrogen.

Synchrotron X-ray diffraction data were recorded at 100 K. Data for the Cel7A/(*S*)-**9** complex were collected at beamline I711, MAX Laboratory, Lund, Sweden, and indexed, integrated and scaled with Mosflm

**TABLE 3** Thermodynamical data of selected ligands

Compound	$\Delta H$ (kJ mol <sup>-1</sup> )	$\Delta S$ (J mol <sup>-1</sup> K <sup>-1</sup> )	$\Delta G$ (kJ mol <sup>-1</sup> )	$K_D$ ( $\mu$ M)
(R)-1 <sup>a</sup>	8.3	93	-19.9	620
(S)-1 <sup>a</sup>	14.2	128	-24.3	245
(R)-2 <sup>a</sup>	-4.4	69	-24.9	134
(S)-2 <sup>a</sup>	-	-	-29.6	41
(R)-3 <sup>a</sup>	3.8	75	-18.6	720
(S)-3 <sup>a</sup>	13.2	124	-23.6	330
(S)-9 <sup>b</sup>	13.5	132	-26.1	1.8

<sup>a</sup>Thermodynamical data obtained using capillary electrophoresis.<sup>34</sup>

<sup>b</sup>Data obtained from calorimetric titrations.

and Scala.<sup>37,38</sup> Data for the Cel7A/(R)-9 complex were collected at beamline ID23-1, ESRF, Grenoble, France, and processed and scaled using Denzo and Scalepack.<sup>39</sup> The crystals belonged to space group I222 with one protein molecule per asymmetric unit.

Initial phases were obtained from the refined protein coordinates for Cel7A in complex with (S)-propranolol (PDB accession code 1DY4).<sup>19</sup> The structures were refined with altering cycles of model building using O<sup>24</sup> and maximum likelihood refinement using Refmac 5.0 through the ccp4i interface, with isotropic temperature factor refinement for the (S)-9 complex and anisotropic for the (R)-9 complex.<sup>38,40</sup> Water molecules were added with the program ARP/wARP<sup>38</sup> during refinement and manually selected or discarded by visual inspection. The models contain the complete catalytic domain of Cel7A (residues 1–434), one *N*-acetyl glucosamine residue covalently linked to Asn270, two Co<sup>2+</sup> ions, one molecule of 9, and 688 and 700 water molecules in the (S)-9 and (R)-9 complexes, respectively. The (R)-9 complex also contains one glycerol molecule, and the (S)-9 complex contains a second *N*-acetyl glucosamine residue linked to Asn284. Statistics from diffraction data processing and structure refinement are summarised in Table 1.

Alternate conformations were modelled at 50% occupancy in the electron density for 23 residues in the (S)-9 complex: Gln7, Pro13, Ser47, Thr55, Gln98, Met111, Asp114, Glu119, Pro159, Pro194, Ser195, Ile215, Thr246, Tyr247, Asp249, Gly253, Gly254, Asp328, Ser342, Ser357, Tyr371, Ser388 and Val407; and for 50 residues in the (R)-9 complex: Gln7, Pro13–Ser20, Cys25–Gly30, Ser47, Thr55, Ser80, Thr85–Ser92, Met111, Asp114, Leu140, Tyr158, Pro194–Asn197, Ile215, Asp249–Gly254, Asp259, Tyr266, Ser272, Gln293, Asp328, Ser342, Ser357, Val407 and Ser409.

The (S)-9 and (R)-9 complex structures were superimposed with each other and with the previously published structures of Cel7A in complex with (S)-1 (PDB accession code 1DY4)<sup>19</sup> and *P. chrysosporium* Cel7D in complex with (R)-1 (PDB accession code 1H46)<sup>17</sup> using the least-squares fitting algorithm of the O program.<sup>24</sup>

The atomic coordinates and experimental structure factor amplitudes have been deposited with the PDB. PDB accession codes are 2V3R and 2V3I for the (S)-9 and (R)-9 complexes, respectively.

## Microcalorimetry

Thermodynamical quantities were obtained on a Thermometric 2277 Thermal Activity Monitor (Järfälla, Sweden) and all data were analysed using Digitam 4.1 (Thermometric AB, Järfälla Sweden) for determination of enthalpy,  $\Delta H$ , and the equilibrium constant,  $K$ . The measurements were performed at 25°C in 50 mM sodium acetate buffer, pH 5.2, and all entities were dissolved in the same buffer. The buffer system of Cel7A was exchanged using PD-10 desalting column (Amersham Biosciences, now Cytiva AB, Uppsala, Sweden) prior to the experiments, and the concentration was determined on a Shimadzu UV-2401 PC (Kyoto, Japan) spectrophotometer at 280 nm,  $\lambda = 78,800 \text{ M}^{-1} \text{ cm}^{-1}$ . The titrant volume of 700  $\mu$ L of 1.43 mM Cel7A was titrated with 15 injections of 20  $\mu$ L ligand (7 mM) in quadruplicate. Dilution experiments were performed as a triplicate using 700  $\mu$ L of buffer solution, injecting 20  $\mu$ L of 7 mM ligand solution 15 times. The mean value for each injection was subtracted from the corresponding injection value of the titration before calculation of  $\Delta H$  and  $K$  based on a 1:1 complex model.

## Computational methods

The preparation of the protein structures and docking studies were performed using the MacroModel 9.1 software<sup>41,42</sup> with the Maestro interface (version 7.5).<sup>43</sup>

## Preparation of ligand

All ligands were built in Maestro and minimised using multiple minimisations with water as a solvent. Merck molecular force field (static) (MMFFs)<sup>44</sup> was used with truncated Newton conjugate gradient (TNCG) as minimisation method. Next, a conformational search<sup>45,46</sup> (MMFFs, TNCG, water as solvent) was performed and the global minimum structure of each ligand was used for docking.

## Protein preparation of docking studies to the empty protein

The structures of Cel7A/R9 and Cel7A/S9 were retrieved from the PDB files of the refined crystal structures, hydrogens were added, and cobalt atoms were removed. The ligand was corrected with respect to bond orders and charge and amino acid residues were adjusted accordingly: His228 was protonated and positively charged, Asp214 was rotated 180°, protonated and the negative charge was removed. All water molecules were removed, as well as the glycerol present in Cel7A/R9 for docking to the empty protein. The ligands (R)-9 and (S)-9 were docked into the structures prepared through protein preparation (refinement only) and grid generation using Grid Receptor as implemented in Glide 4.0.<sup>28</sup>

## Protein preparation for docking studies to predict affinity and selectivity

The PDB file for Cel7A/R9 was treated as described above with two exceptions: the hydroxyl on the glycerol present close to Asp214 was replaced with a water molecule and a total of 34 water molecules in the part of the enzyme close to the active site were kept in the structure. Keeping the protein frozen, a Monte Carlo multiple minimisation<sup>46,47</sup> using MMFFs with water as a solvent and TNCG as the minimisation method was performed to obtain correct conformations of the water molecules. The number of water molecules present in the structure was reduced through a stepwise removal of water molecules followed by redocking of (R)-9 in an iterative process to determine the importance of the removed water molecules with respect to the resulting poses of the ligand.

The 73 water molecules in Cel7A/S9 were minimised after the same principles as for Cel7A/R9. Only 14 water molecules were included in each of the final structures. The water molecules in Cel7A/S9 were selected to resemble the positioning of the necessary water molecules in the structure of Cel7A/R9.

## Docking studies of complexes Cel7A/R9 and Cel7A/S9

The protein structures containing the selected water molecules were prepared using protein preparation (refinement only) and receptor grid generation as implemented in Glide (version 4.0). The ligands were prepared as described above and docked using both Standard-Precision Glide and Extra-Precision Glide.

## Interaction energies of water in the active site

Six sets of the Cel7A/R9 receptor were used in the GRID<sup>33</sup> calculations, all derived from the X-ray structure data files used in the docking experiments, and new PDB files were made from the following combinations: protein without water or ligand present, protein with (R)-9

present, protein with (R)-9 and W1 present, protein with (R)-9 and W2 present, protein with (R)-9 and W3, W5 and W7 present alternatively with W4 and W6 present. Hydrogens were removed, and all heteroatoms (HETATM) were placed last in the PDB file. Water was used as a probe (OH2) with the grid spacing of 0.5 Å (number of grid planes per Ångstrom, NPLA = 2) and the grid dimensions in Å were set to  $X_{\max}/X_{\min}$ , 10/46;  $Y_{\max}/Y_{\min}$ , 37/68;  $Z_{\max}/Z_{\min}$ , 69/105, centred on Glu175. To analyse the results, PyMOL (version 0.99rc6)<sup>47</sup> was used for visualisation.

## AUTHOR CONTRIBUTIONS

*Conceptualisation (equal), data curation (equal), formal analysis (equal), investigation (equal), methodology (equal), validation (equal) and writing—original draft (equal):* Alexandra Fagerström. *Conceptualisation (equal), formal analysis (equal), investigation (equal), methodology (equal), resources (equal), software (equal), supervision (equal), validation (equal), visualisation (equal) and writing—original draft (equal):* Tommy Liljefors. *Conceptualisation (supporting), data curation (supporting), formal analysis (equal), funding acquisition (equal), investigation (supporting), methodology (supporting), resources (equal), supervision (supporting), validation (equal), visualisation (supporting) and writing—review and editing (supporting):* Mats Sandgren. *Conceptualisation (equal), investigation (equal), methodology (equal), resources (equal), supervision (equal), validation (equal) and writing—original draft (supporting):* Roland Isaksson. *Conceptualisation (equal), data curation (equal), formal analysis (equal), funding acquisition (equal), investigation (equal), methodology (equal), project administration (supporting), resources (equal), supervision (equal), validation (equal), visualisation (equal), writing—original draft (equal) and writing—review and editing (equal):* Jerry Ståhlberg. *Conceptualisation (equal), data curation (equal), formal analysis (equal), funding acquisition (equal), investigation (equal), methodology (equal), project administration (lead), resources (equal), supervision (equal), validation (equal), visualisation (equal), writing—original draft (equal) and writing—review and editing (equal):* Ulf Berg.

## ACKNOWLEDGEMENTS

We thank Lars-Erik Briggner from AstraZeneca, Lund, Sweden, for providing us with equipment and guidance concerning all microcalorimetric measurements, Gerd Olofsson, Physical Chemistry, Lund University, Sweden, for kindly sharing their equipment, Mikel Hedeland for initiating the computational calculations on the propranolol/Cel7A complex, the Research School in Pharmaceutical Sciences, Lund University and Kalmar University for funding and Dr. Igor Sabljic, Department of Molecular Sciences, SLU, for help with making LigPlot figures of ligand–protein interactions.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## ETHICS STATEMENT

The authors confirm that they have followed the ethical policies of the journal.

## DATA AVAILABILITY STATEMENT

Data available on request from the authors.

## ORCID

Ulf Berg  <https://orcid.org/0000-0001-9030-9156>

## PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1002/ntls.20220050>

## REFERENCES

- Scriba GKE. Recognition mechanisms of chiral selectors: an overview. *Methods Mol Biol.* 2019;1985:1-33. doi: [10.1007/978-1-4939-9438-0\\_1](https://doi.org/10.1007/978-1-4939-9438-0_1)
- Foreman PK, Brown D, Dankmeyer L, et al. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. *J Biol Chem.* 2003;278(34):31988-31997. doi: [10.1074/jbc.M304750200](https://doi.org/10.1074/jbc.M304750200)
- CHIRALPAK® CBH. Daicel Chiral Technologies. Accessed December 28, 2022. <https://www.daicelchiral.com/en/chiral-columns/protein/>
- Divne C, Ståhlberg J, Reinikainen T, et al. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science.* 1994;265(5171):524-528. doi: [10.1126/science.8036495](https://doi.org/10.1126/science.8036495)
- Kraulis J, Clore GM, Nilges M, et al. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry.* 1989;28(18):7241-7257. doi: [10.1021/bi00444a016](https://doi.org/10.1021/bi00444a016)
- Marle I, Jönsson S, Isaksson R, Pettersson C, Pettersson G. Chiral stationary phases based on intact and fragmented cellobiohydrolase I immobilised on silica. *J Chromatogr A.* 1993;648:333-347. doi: [10.1016/0021-9673\(93\)80415-5](https://doi.org/10.1016/0021-9673(93)80415-5)
- Divne C, Ståhlberg J, Teeri TT, Jones TA. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *J Mol Biol.* 1998;275(2):309-325. doi: [10.1006/jmbi.1997.1437](https://doi.org/10.1006/jmbi.1997.1437)
- Henriksson H, Ståhlberg J, Isaksson R, Pettersson G. The active sites of cellulases are involved in chiral recognition: a comparison of cellobiohydrolase 1 and endoglucanase 1. *FEBS Lett.* 1996;390(3):339-344. doi: [10.1016/0014-5793\(96\)00685-0](https://doi.org/10.1016/0014-5793(96)00685-0)
- Knott BC, Haddad Momeni M, Crowley MF, et al. The mechanism of cellulose hydrolysis by a two-step, retaining cellobiohydrolase elucidated by structural and transition path sampling studies. *J Am Chem Soc.* 2014;136(1):321-329. doi: [10.1021/ja410291u](https://doi.org/10.1021/ja410291u)
- Hedeland M, Isaksson R, Pettersson C. Cellobiohydrolase I as a chiral additive in capillary electrophoresis and liquid chromatography. *J Chromatogr A.* 1998;807(2):297-305. doi: [10.1016/S0021-9673\(98\)00083-1](https://doi.org/10.1016/S0021-9673(98)00083-1)
- Hedeland M, Henriksson H, Isaksson R, Pettersson G. Studies on the enantioselective retention mechanisms of cellobiohydrolase I (CBH I) by covalent modification of the intact and fragmented protein. *Chirality.* 1998;10(8):760-769. doi: [10.1002/\(SICI\)1520-636X\(1998\)10:8<760::AID-CHIR7>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1520-636X(1998)10:8<760::AID-CHIR7>3.0.CO;2-6)
- Harang V, Tysk M, Westerlund D, Isaksson R, Johansson G. A statistical experimental design to study factors affecting enantioseparation of propranolol by capillary electrophoresis with cellobiohydrolase (Cel7A) as chiral selector. *Electrophoresis.* 2002;23(14):2306-2319. doi: [10.1002/1522-2683\(200207\)23:14<2306::AID-ELPS2306>3.0.CO;2-U](https://doi.org/10.1002/1522-2683(200207)23:14<2306::AID-ELPS2306>3.0.CO;2-U)
- Fornstedt T, Sajonz P, Guiochon G. Thermodynamic study of an unusual chiral separation. Propranolol enantiomers on an immobilized cellulase. *J Am Chem Soc.* 1997;119(6):1254-1264. doi: [10.1021/ja9631458](https://doi.org/10.1021/ja9631458)
- Henriksson H, Ståhlberg J, Koivula A, et al. The catalytic amino-acid residues in the active site of cellobiohydrolase 1 are involved in chiral recognition. *J Biotechnol.* 1997;57:115-125.
- Hedeland M, Henriksson H, Bäckman P, Isaksson R, Pettersson G. Microcalorimetric studies on the complex formation between cellobiohydrolase I (CBH I) from *Trichoderma reesei* and the (R)- and (S)-enantiomers of the  $\beta$ -receptor blocking agent alprenolol. *Thermochim Acta.* 2000;356:153-158. doi: [10.1016/S0040-6031\(00\)00473-1](https://doi.org/10.1016/S0040-6031(00)00473-1)
- Jönsson S, Schön A, Isaksson R, Pettersson C, Pettersson G. An unexpected temperature effect obtained on enantiomer separation using CBH I-silica as a chiral stationary phase: increase in retention and enantioselectivity at elevated column temperature: a chromatographic and microcalorimetric study. *Chirality.* 1992;4(8):505-508. doi: [10.1002/chir.530040808](https://doi.org/10.1002/chir.530040808)
- Munoz IG, Mowbray SL, Ståhlberg J. The catalytic module of Cel7D from *Phanerochaete chrysosporium* as a chiral selector: structural studies of its complex with the beta blocker (R)-propranolol. *Acta Crystallogr Sect D: Biol Crystallogr.* 2003;D59:637-643.
- Fagerström A, Nilsson M, Berg U, Isaksson R. New propranolol analogues: binding and chiral discrimination by cellobiohydrolase Cel7A. *Org Biomol Chem.* 2006;4:3067-3076. doi: [10.1039/b605603b](https://doi.org/10.1039/b605603b)
- Ståhlberg J, Henriksson H, Divne C, et al. Structural basis for enantiomer binding and separation of a common  $\beta$ -blocker: crystal structure of cellobiohydrolase Cel7A with bound (S)-propranolol at 1.9 Å resolution. *J Mol Biol.* 2001;305:79-93. doi: [10.1006/jmbi.2000.4237](https://doi.org/10.1006/jmbi.2000.4237)
- von Ossowski I, Ståhlberg J, Koivula A, et al. Engineering the exo-loop of *Trichoderma reesei* cellobiohydrolase, Cel7A. A comparison with *Phanerochaete chrysosporium* Cel7D. *J Mol Biol.* 2003;333:817-829. doi: [10.1016/S0022-2836\(03\)00881-7](https://doi.org/10.1016/S0022-2836(03)00881-7)
- French S, Wilson K. On the treatment of negative intensity observations. *Acta Crystallogr Sect.* 1978;A34:517-525.
- Kleywegt GJ. Validation of protein models from C $\alpha$  coordinates alone. *J Mol Biol.* 1997;273:371-376. doi: [10.1006/jmbi.1997.1309](https://doi.org/10.1006/jmbi.1997.1309)
- Kleywegt GJ, Jones TA. Phi/psi-chology: Ramachandran revisited. *Structure.* 1996;4:1395-1400. doi: [10.1016/s0969-2126\(96\)00147-5](https://doi.org/10.1016/s0969-2126(96)00147-5)
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr Sect A.* 1991;A47:110-119.
- Persistence of Vision Development Team™. <http://www.povray.org>
- Harris M, Jones TA. Molray—a web interface between O and the POV-Ray ray tracer. *Crystallogr Sect D: Biol Crystallogr.* 2001;57:1201-1203.
- Glide, Version 4.0.108. Portland, OR: Schrödinger Inc.; 2006.
- a) Friesner RA, Banks JL, Murphy RB, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem.* 2004;47:1739-1749. doi: [10.1021/jm0306430](https://doi.org/10.1021/jm0306430) b) Friesner RA, Murphy RB, Repasky MP, et al. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem.* 2006;49:6177-6196. doi: [10.1021/jm051256o](https://doi.org/10.1021/jm051256o)
- Hamark C, Pendrill R, Landröm J, et al. Enantioselective binding of propranolol and analogues thereof to cellobiohydrolase Cel7A. *Chem Eur J.* 2018;24:17975-17998. doi: [10.1002/chem.201803104](https://doi.org/10.1002/chem.201803104)
- Fagerström A, Nilsson M, Berg U, Isaksson R. pH dependency of ligand binding to cellobiohydrolase 1 (Cel7A)—affinity, selectivity and inhibition for designed propranolol analogues. *J Chromatogr A.* 2007;1138:276-283. doi: [10.1016/j.chroma.2006.10.098](https://doi.org/10.1016/j.chroma.2006.10.098)
- Warren GL, Andrews CW, Capelli A-M, et al. A critical assessment of docking programs and scoring functions. *J Med Chem.* 2006;49:912-931. doi: [10.1021/jm050362n](https://doi.org/10.1021/jm050362n)
- GRID, Version 2.2a. Pinner, Middlesex: Molecular Discovery Ltd.; 2005.

33. Goodford PJ. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem*. 1985;28:849-857. doi: [10.1021/jm00145a002](https://doi.org/10.1021/jm00145a002)
34. Nilsson M, Fagerström A, Berg U, Isaksson R. Thermodynamical studies of designed ligands binding to Cel7A using partial-filling capillary electrophoresis. *Electrophoresis*. 2008;29:358-362. doi: [10.1002/elps.200700370](https://doi.org/10.1002/elps.200700370)
35. Goetmar G, Ozen C, Serpersu E, Guiochon G. Isothermal microcalorimetric study of the pH dependence of the interactions between a cellulase and a  $\beta$ -blocker. *J Chromatogr A*. 2004;1046:49-53. doi: [10.1016/j.chroma.2004.06.089](https://doi.org/10.1016/j.chroma.2004.06.089)
36. McPherson A. *Preparation and Analysis of Protein Crystals*. New York, NY: John Wiley & Sons; 1982.
37. Leslie AGW. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography. 1992.
38. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr Sect D: Biol Crystallogr*. 1994;50:760-763. doi: [10.1107/S0907444994003112](https://doi.org/10.1107/S0907444994003112)
39. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 1997;276:307-326. doi: [10.1016/S0076-6879\(97\)76066-X](https://doi.org/10.1016/S0076-6879(97)76066-X)
40. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr Sect D: Biol Crystallogr*. 1997;D53:240-255. doi: [10.1107/S0907444996012255](https://doi.org/10.1107/S0907444996012255)
41. *MacroModel, Version 9.1.106*. Portland, OR: Schrödinger Inc.; 2006.
42. Mohamadi F, Richards NGJ, Guida WC, et al. MacroModel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J Comput Chem*. 1990;1440-1467. doi: [10.1002/jcc.540110405](https://doi.org/10.1002/jcc.540110405)
43. *Maestro, Version 7.5.106*. Portland, OR: Schrödinger Inc.; 2006.
44. Halgren TA. MMFF VI. MMFF94s option for energy minimization studies. *J Comput Chem*. 1999;20:720-729. doi: [10.1002/\(SICI\)1096-987X\(199905\)20:7<720::AID-JCC7>3.0.CO;2-X](https://doi.org/10.1002/(SICI)1096-987X(199905)20:7<720::AID-JCC7>3.0.CO;2-X)
45. Chang G, Guida WC, Still WC. An internal coordinate Monte Carlo method for searching conformational space. *J Am Chem Soc*. 1989;111:4379-4386. doi: [10.1021/ja00194a035](https://doi.org/10.1021/ja00194a035)
46. Meng EC, Cieplak P, Caldwell JW, Kollman PA. Accurate solvation free energies of acetate and methylammonium ions calculated with a polarizable water model. *J Am Chem Soc*. 1994;116:12061-12062. doi: [10.1021/ja00105a060](https://doi.org/10.1021/ja00105a060)
47. *The PyMOL Molecular Graphics System, Version 0.99rc6*. Schrödinger, LLC.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Fagerström A, Liljefors T, Sandgren M, Isaksson R, Ståhlberg J, Berg U. Chiral recognition mechanism of cellobiohydrolase Cel7A for ligands based on the  $\beta$ -blocker propranolol: The effect of explicit water molecules on binding and selectivities. *Nat Sci*. 2023;3:e20220050. <https://doi.org/10.1002/ntls.20220050>