The structural basis for catalysis and substrate specificity of a rhomboid protease

Kutti R. Vinothkumar, Kvido Strisovsky, Antonina Andreeva, Yonka Christova, Steven Verhelst and Matthew Freeman

Supplementary Data

Figure legends

Figure S1:

A: Surface representation of GlpG and Gurken tetrapeptide shown as stick representation. GlpG molecule is colour coded as in Figure 4.

B: Interaction of P_2 and P_1' side chains of Gurken substrate with loop 5.

C: Possible substrate-enzyme interaction in GlpG-Gurken model: Gurken tetrapeptide consists of a sequence (M-A-H-I) corresponding to the P_2 to P_2' respectively. In comparison to the TatA peptide (Figure 5), the side chains at P_2 and P_1' are larger and in our present model point up towards the solvent. As with the TatA peptide model in Figure 4D, possible hydrogen bonding interactions between substrate and enzyme (red dashed lines) are shown.

Figure S2:

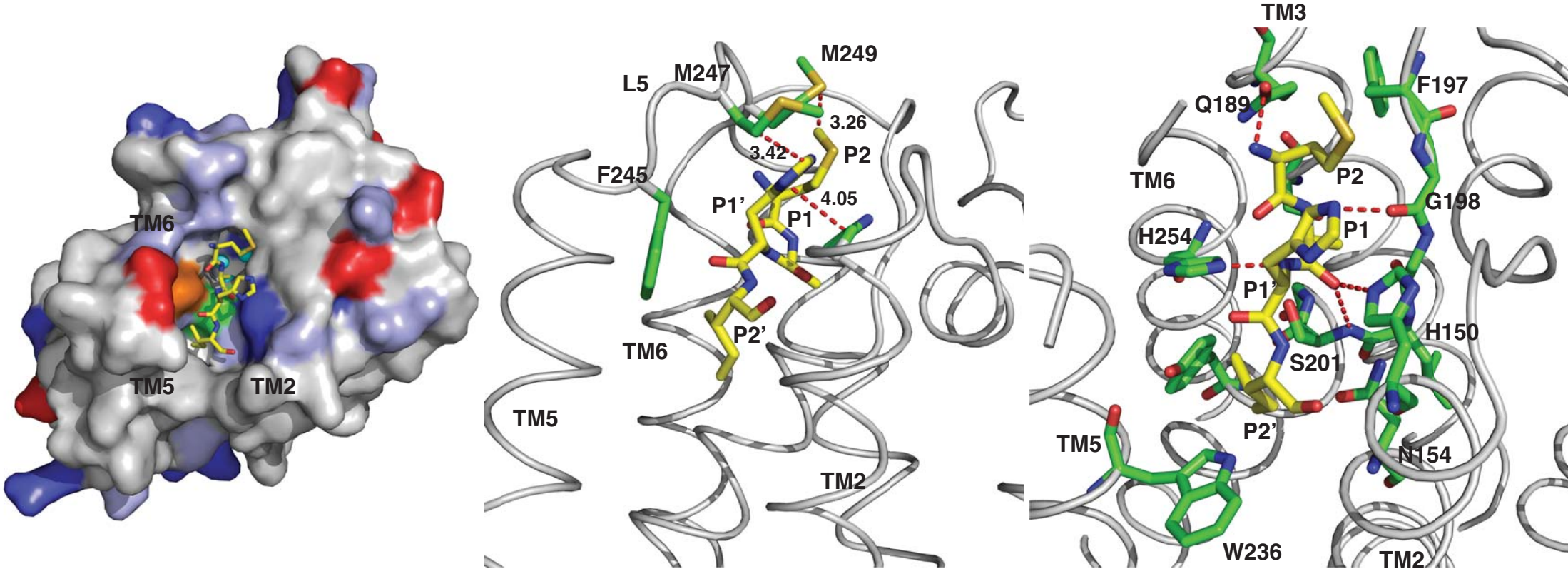
Comparison of isocoumarin binding in porcine pancreatic elastase (PPE) and GlpG. Carbon atoms of inhibitor molecules are shown in white and key residues are shown as green stick representation.

A: PPE in complex with amino-methoxy-isocoumarin (PDB – 1JIM).

B: PPE in complex with bromoethoxy-isocoumarin (PDB – 9EST)

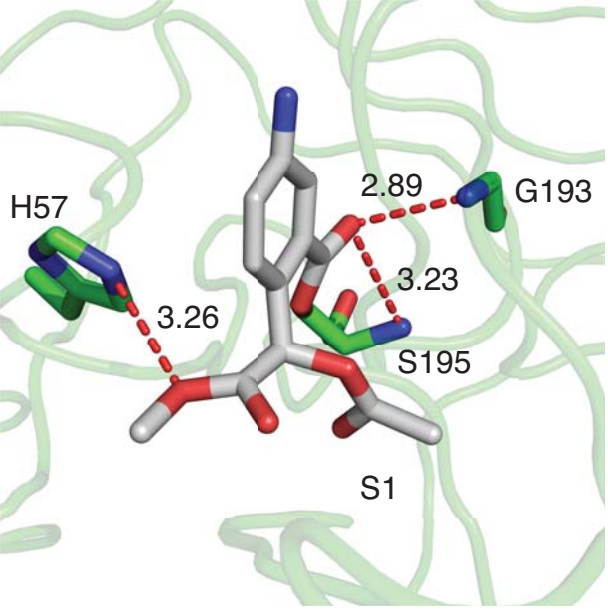
C: GlpG in complex with amino-methoxy-isocoumarin (PDB – 2xow)

Three structures of porcine pancreatic elastase (PPE) in complex with different isocoumarins either as a single or a doubly covalent bonded enzyme have been described. In complex with amino-methoxy-isocoumarin, the compound described in this study, PPE not only reacts with the active site serine but subsequently with an acetate molecule from the buffer. In this PPE structure, the benzoyl carbonyl oxygen of the isocoumarin points towards the oxyanion hole defined by backbone amides of Gly193 and the active site Ser195 (Fig S2A). The acetoxy group points to the S_1 pocket of the enzyme. Two other structures in complex with either 7-amino 3-bromoethoxy isocoumarin or 7-guanidino 3-ethoxy isocoumarin show a doubly covalent bonded or a chloroacylenzyme of PPE respectively (Powers et al, 1990; Vijayalakshmi et al, 1991). In both these structures the benzoyl carbonyl oxygen points away from the oxyanion hole (Fig S2B). The ethoxy group in the chloroacyl enzyme points to the S_1 pocket but the bromoethoxy group in the doubly bonded alkylated acyl enzyme points towards the S_3 subsite of the enzyme. These precedents indicate that a degree of caution is needed in inferring the position of the oxyanion hole from the position of the inhibitor carbonyl oxygen. Nevertheless, as outlined in the main text, we believe that all evidence favours our interpretation of the GlpG oxyanion hole.

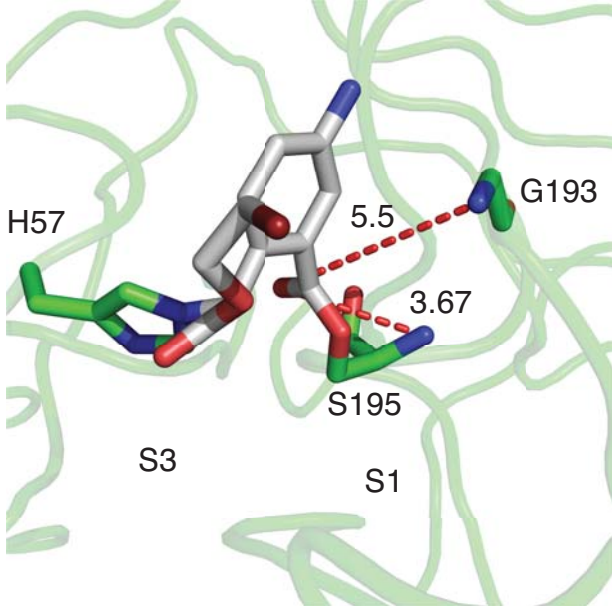


(A) (B) (C)

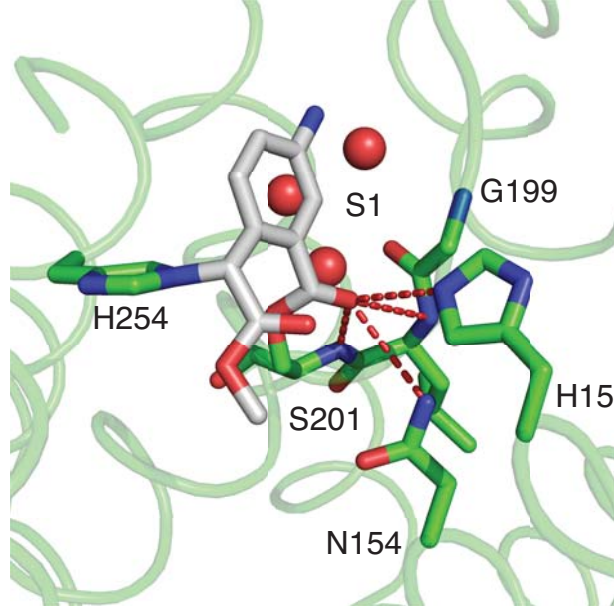
Figure S1



(A)



(B)



(C)

Figure S2

Table SI

Comparison of acyl enzyme structure with models of native GlpG from *E.coli* and *Haemophilus influenzae*

Reference model – Acyl enzyme structure

PDB	RMSD (Å)[a]	Comments
3B45	0.789	High-resolution structure with L5 covering the active site
2O7L	0.846	Soaking with DCI results in disorder of loop5, called the open cap conformation
2NRF – molecule A	1.816	TM5 is tilted 35° and thus called the open conformation of enzyme
2NRF - molecule B	0.759	Structure is much similar to 3B45 but residues 242-250 are disordered
2IRV – molecule A	1.125	Structure determined independently using phases derived from heavy atoms, loop5 is well ordered and adopts a different conformation
2IRV – molecule B	1.406	A lipid molecule is pointing to the active site and loop5 is disordered
2NR9 – *H.influenzae*	1.552	A homolog of GlpG from *E.coli*, shows a distinct TM5 conformation but loop5 still covering the active site

[a] – structural similarity was calculated using CCP4 program Superpose with loop5 (residues 245-250) omitted.