The NRP peptidase ClbP as a target for the inhibition of genotoxicity, cell proliferation and tumorogenesis mediated by *pks*-harboring bacteria

INTRODUCTION

Nougayrède *et al.* recently reported a genomic island designated *pks* that is involved in the production of a genotoxin named colibactin (Nougayrède et al. 2006). *pks* island is observed in both *Escherichia coli*, a versatile commensal bacterial inhabitant of the large intestine of humans (Croxen and Finlay 2009; Kaper, Nataro, and Mobley 2004). *pks*-harboring bacteria (*pks*+) induce DNA double-strand breaks in eukaryotic cells both *in vitro* and *in vivo* (Cuevas-Ramos et al. 2010; Nougayrède et al. 2006). The DNA damage induced by *pks*+ *E. coli* is followed by a transient activation of the DNA damage cellular signaling pathway, megalocytosis, signs of incomplete DNA repair, chromosomal instability and increasing level of mutations in host cells (Cuevas-Ramos et al. 2010); lesions which are frequently observed in cancers (Boffetta et al. 2007).

Genetic and functional analyses of the *pks* island indicate that it encodes non-ribosomal peptide synthetases (NRPS) and polyketide synthetases (PKS), which are required for the production of one or more polyketide - non-ribosomal peptide (PK-NRP) hybrid compounds designated colibactin (Nougayrède et al. 2006). In addition to NRPS and PKS, the *pks* island encodes the peptidase ClbP, which was identified as a crucial enzyme in colibactin maturation (Dubois et al. 2010). In this work, we identified small molecules, which inhibit ClbP, with K_i values in the nanomolar range and determined the crystal structure of ClbP in complex with these compounds. The coordinates and structural factors have been deposited in the RCSB Protein Data Bank (PDB IDs: 4E6W and 4E6X).

RESULTS

The structure of ClbP in complex with two nanomolar inhibitors (compounds 1 and 2) was determined by x-ray diffraction to confirm their ability to bind ClbP active site and understand how compounds having differences in chemical structures can exhibit such closely related inhibiting activity. The structures were refined against diffraction data extending to 2.19 and 2.24 Å resolutions (Table 1).

table 1. Data collection and refinement statistics.

	ClbP / Compound 1	ClbP / Compound 2
Data collection		
Space group	*C121*	*C121*
Cell dimensions		
a, *b*, *c* (Å)	104.5, 149.4, 88.2	105.1, 152.5, 86.9
α, β, γ (°)	90.0, 124.2, 90.0	90.0, 123.4, 90.00
Resolution (Å)	52.2-2.19 (2.31-2.19)	52.6-2.24 (2.36-2.24)
R_{merge}	0.087 (0.311)*	0.069 (0.290)*
$I / \sigma I$	10.4 (3.7)*	10.7 (3.6)*
Completeness (%)	99.7 (99.8)*	79.0 (80.9)*
Redundancy	3.3 (3.2)*	3.0 (2.9)*
Refinement		
Resolution (Å)	52.2 – 2.19	52.6 – 2.24
No. unique reflections	54118	40985
R_{work} / R_{free}	0.194 / 0.238	0.2004 / 0.2492
No. atoms		
Protein	7626	7054
Water	400	172
Ligand	30	20
B-factors		
Protein	27.45	38.66
Water	26.99	32.03
Ligand	23.11	48.19
R.m.s. deviations		
Bond lengths (Å)	0.013	0.014
Bond angles (°)	1.510	1.713
Ramachandran plot		
% residues in favored region	97.6	98.1
% residues in allowed region	2.4	1.9

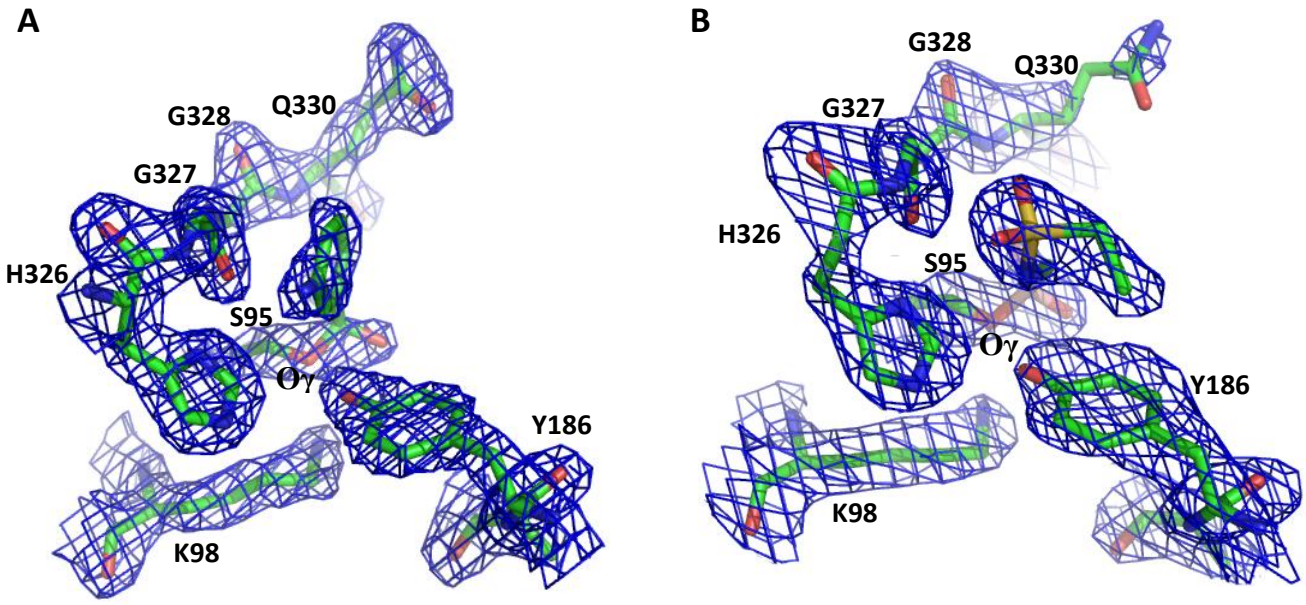
*, values for the highest range of resolution

The electron density map showed well defined density throughout most of the structure. The stereochemical parameters of the model were satisfactory; a Ramachandran plot showed no residues in disallowed regions of Φ/ψ space and 98% of residues were a favoured conformation. The final models included three ClbP monomers in the crystallographic asymmetric unit. The three-dimensional structure of ClbP consists of two structural regions, all-α and α/β. The conserved motifs ^{95}SMSK and ^{186}YAS, which correspond to the catalytic center in S12 enzymes of the MEROPS classification (Rawlings, Barrett, and Bateman 2010), are located in a large groove (Dubois et al. 2010). The root-mean-square deviation (RMSD) of the Cα positions is approximately 0.31 Å between the complex structures and the apoenzyme, showing no dramatic change in geometry resulting from the binding of inhibitors.

The inhibitors bind the same area of the catalytic pocket

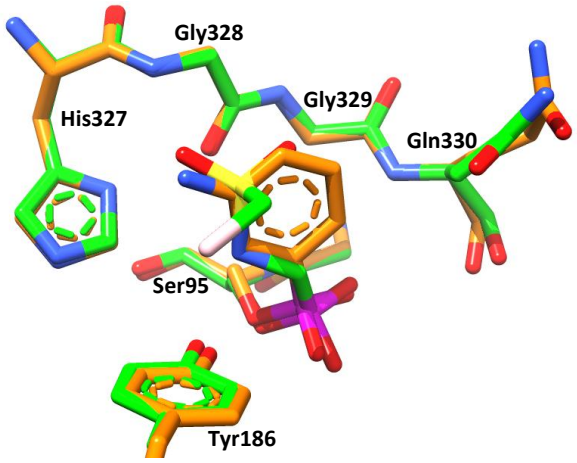
The positions of the inhibitors in the active site were unambiguously identified in the 2Fo-Fc omit density maps (figure 1).

Figure 1. Electron density at the vicinity of the serine active residue (S95) for ClbP in complex with compounds 1 and 2. The position of the inhibitors (A, compound 1; B, compound 2) in the active site is unambiguously identified.



The conformation of the inhibitors was similar in all monomers of the crystallographic asymmetric unit. The compounds 1 and 2 were accommodated in the same zone, in the vicinity of β-strand β11 (residues 324-332) between residues 327-330 and Tyr150 (Figure 1). Intriguingly, most atoms of compounds perfectly overlapped despite major differences in chemical structure (Figure 1).

Figure 1. Superimpositions of ClbP structures. A, superimposition of ClbP active site in complex with inhibitors 1 and 2. The atoms of compounds overlap perfectly and are accommodated in the zone surrounded by the residues 95, 186 and 327-330. Carbon atoms are shown in gold for ClbP in complex with the compound 1 and in green for ClbP in complex the compound2**.** Oxygen atoms are shown in red, nitrogen atoms in blue and sulfurs in yellow. The aromatic systems are indicated by dashed circles.



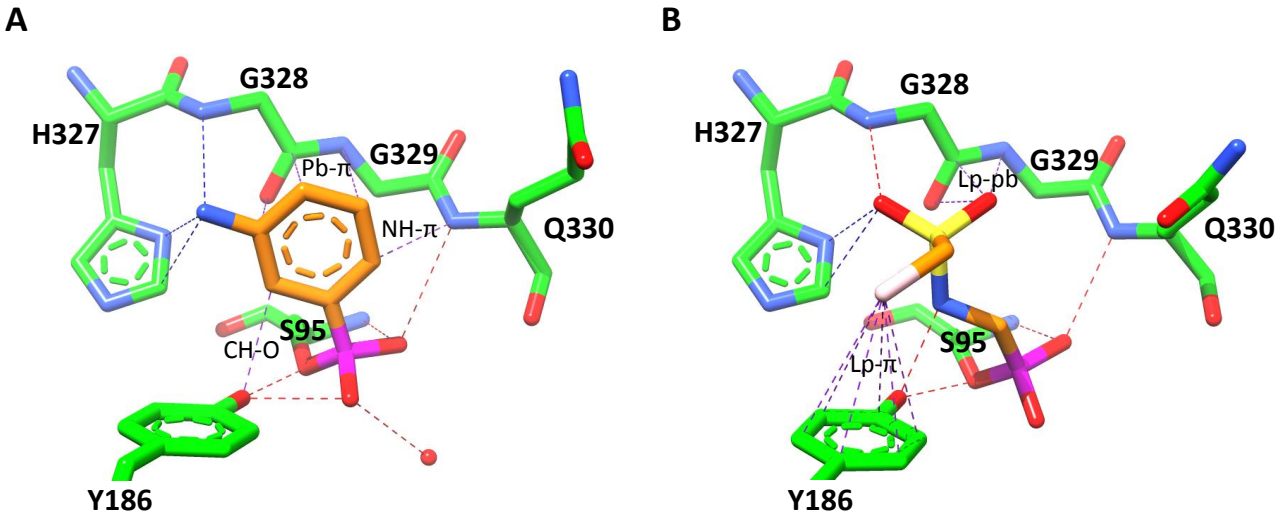
Interactions of inhibitor moieties with ClbP active site

The moieties of compounds 1 and 2 established conventional interactions with ClbP active site (Figure 2). Van der Waals contacts were observed between the nitrogen atom of the compound 1 and residues His237 (nitrogen-Cε1 distance: 3.6 Å and nitrogen-Nδ1 distance: 3.7 Å) and Gly328 (nitrogen-N distance: 3.7 Å). The compound 2 established van der Waals contact with His237 (oxygen-Cε1 distance: 3.5 Å and oxygen-Nδ1 distance: 3.5 Å) and hydrogen bond to the N atom of Gly328 (oxygen-nitrogen distance: 3.1 Å), in a similar manner to the nitrogen atom of compound 1.

Non-canonical interactions were also observed (Figure 2). For compound 1, two C atoms of the aromatic ring were at 2.8 Å and 3.0 Å of the Tyr188 hydroxyl function and the N atom of Gln330 respectively. The geometry of these close contacts corresponds to a NH···π interaction involving Gln330 and a CH···O interaction involving Tyr186 (Cotesta and Stahl 2006; Malone et al. 1997). The plane of the aromatic ring was parallel with β-strand β11 and established a π-stacking-like interaction (Bendova et al. 2007) with the peptide bond between residues Gly328 and Gly329. For compound 2, the chlorine atom established with the aromatic ring of Tyr188 a Cl···π electrostatic interaction (Chlorine-aromatic centroid distance: 3.6 Å), which exhibited a face-on geometry as previously reported (Imai et al. 2008). One oxygen atom of the sulfonamide was positioned in close contact with the C and N atoms of the Gly328-Gly329 peptide bond (3.1 Å and

3.6 Å respectively). Because of the π electron resonance cloud of peptide bonds, this close interaction resembled a Lone pair···π interaction (Egli and Sarkhel 2007; Jain, Ramanathan, and Sankararamakrishnan 2009; Mooibroek, Gamez, and Reedijk 2008).

Figure 2. Structure of ClbP active site in complex with the inhibitors. The compounds (A, compound 1; B, compound 2) establish similar conventional interactions (van der Waals contact, NH…O and OH…O hydrogen bonds), but differ by the non-canonical interactions mediated by lone pair electrons (oxygen and chloride atoms, and π electrons in aromatic systems or the peptide bond G328-G329. Carbon atoms are shown in green for ClbP and in gold for the compounds. Oxygen atoms are shown in red, nitrogen atoms in blue, boron in magenta, sulfur in yellow and chloride in white. The aromatic systems are indicated by dashed circles. The hydrogen bonds are shown in red dashed lines. Van der Waals contacts are indicated by blue dashed lines. Non-canonical interactions are indicated by purple dashed lines (Pb-π, peptide bond-π stacking; Lp-π, lone pair…π interaction; Lp-pb, lone pair…peptide bond interaction; CH-O, CH…O hydrogen bond; NH-π, NH…π interaction).



Conclusion

We determined the first crystal structure of ClbP in complex with ligands. These two ligands are nanomolar inhibitor of ClbP and may be interesting tools in decreasing the production of genotoxic compounds synthesized by *pks* island. We recently identified a third inhibitor of ClbP and are planning to another proposal/experiment at ESRF.

MATERIAL AND METHODS

Protein production and purification. The enzymatic domain of ClbP, was produced in the *E. coli* BL21(DE3) from the pET9a derivative plasmid pClbPpep (Dubois et al. 2010). Bacteria were cultured in terrific-broth II medium (Euromedex, Souffelweyersheim, France) supplemented with sorbitol 400mM (Sigma-Aldrich), betaine 2.5 mM (Sigma-Aldrich), and kanamycin 50 µg/ml (Sigma-Aldrich). The expression of proteins was induced with 0.2 mM isopropyl-ß-D-thiogalactopyranoside (Euromedex). The purification of ClbP was performed as previously described (Dubois et al. 2010). The purification of FmtA-like and ZmaM was carried out by ion-exchange chromatography into a HiTrapTM Q SepharoseTM High Performance column (Healthcare Europe, Velizy-villacoublay, France) equilibrated with 20 mM Tris-HCl pH 7.5 and eluted with a linear NaCl gradient (0 to 500 mM). The elution peak was then purified by ion-exchange chromatography into a HiTrapTM Q SepharoseTM High Performance column (Healthcare Europe) equilibrated with 50 mM MES pH 6.0 and eluted with a linear NaCl gradient (0 to 500 mM). The elution peaks were extensively dialyzed against NaCl 50 mM Tris-HCl 5 mM pH 7.0 and concentrated by ultrafiltration to a final concentration of 10 mg/ml. The purified enzymes were more than 95% homogeneous as determined by Coomassie blue staining after SDS-PAGE (Laemmli 1970).

Crystallization and structure determination. ClbP crystals were grown in hanging drops over a solution of 0.8 M monosodium dipotassium phosphate buffer (pH 7.0) as previously reported (Dubois et al. 2010). The crystals were soaked in the crystallization buffer supplemented with 50 mM boronic acid compounds and cryo-protected with 30% sucrose before flashcooling in liquid nitrogen. Data were collected using a Q315r ADSC-CCD detector on ESRF beamline 23-1 at the European Synchrotron Radiation Facility (Grenoble, France). Reflexions were indexed, integrated, and scaled using CCP4 package (Potterton et al. 2003). The structure was solved by molecular replacement with the program PHASER (McCoy 2007) and ClbP x-ray structure (pdb ID, 3O3V) as a search model. The structure was automatically and manually refined with REFMAC5 and COOT programs, respectively (Emsley et al. 2010; Murshudov, Vagin, and E. J. Dodson 1997; Steiner, Lebedev, and Murshudov 2003). Cross-validation was used throughout and 5 % of the data were used for the Rfree calculation. The stereochemical quality of the models was monitored with the PROCHECK program (Laskowski, Moss, and Thornton 1993). Ramachandran plots were calculated by RAMPAGE (Lovell et al. 2003). Processing and crystallographic refinement statistics are listed in supplemental table 2. The coordinates and structural factors of ClbP have been deposited in the RCSB Protein Data Bank (PDB IDs: 4E6W and 4E6X).