ClbP is a prototype of a peptidase subgroup involved in biosynthesis of non-ribosomal peptides

Introduction

The pks genomic island of Escherichia coli encodes polyketide (PK) and nonribosomal peptide (NRP) synthases that allow assembly of a putative hybrid PK-NRP compound named colibactin that induces DNA double-strand breaks in eukaryotic cells. The pks-encoded machinery harbors an atypical essential protein, ClbP. We determined the crystal structure of the ClbP major domain, its biochemical activity, and we characterized distant ClbP homologous proteins that form a new subgroup of serine-reactive peptidases associated with NRP. The corresponding article (Dubois D., J Biol Chem 2011) has been the subject of comments in the section « Choice's of Editors » of the journal Science in 2012. The structure has been deposited at PDB databank (PDB ID: 3O3V).

Main crystallographic results

To investigate the function and activity of ClbP, the structure of the ClbPpep domain was determined by x-ray diffraction at ESRF, Grenoble, France. The structure was refined against diffraction data extending to 2.4Å resolution. The electron density map (Fig. 1) showed well defined density throughout most of the structure, with the exception of the six N-terminal residues.

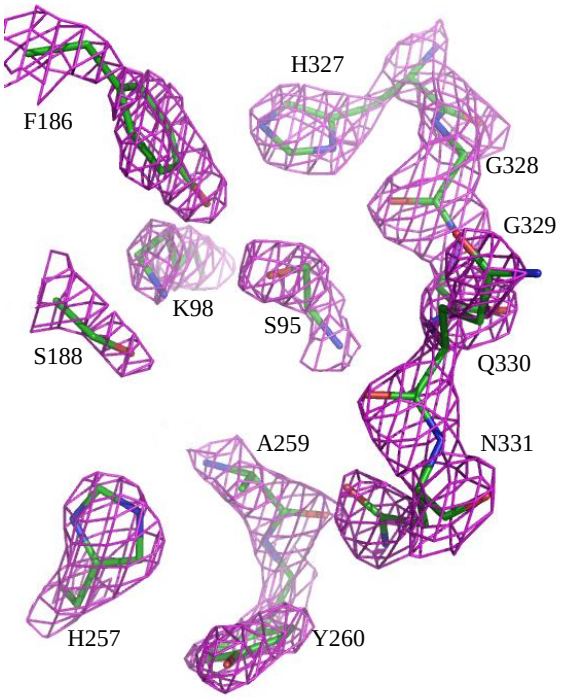

Figure 1. Electron density map of ClbPpep in the vicinity of the active site.
The 2Fo−Fc electron density of the refined model, contoured at 1σ, is shown in mauve. Carbon atoms are in green, oxygen atoms in red, and nitrogen atoms in blue.

The stereochemical parameters of the model were satisfactory (Table 1); a Ramachandran plot showed no residuesin disallowed regions of φ/ϕ space. The final model included three ClbPpep monomers and 277 water molecules.

Table 1. Data collection and structure determination statistics for ClbPpep crystal structure.

Data collection	
Space group	C2
Cell dimensions	
a, b, c (Å)	103.9, 149.9, 87.3
α, β, γ (°)	90.0, 123.9, 90.0
Resolution (Å)	20 - 2.4 (2.5 – 2.4)[a]
R_{merge} (%)[b]	8.6 (28.0) [a]
[I / σ(I)] (°)	11.5 (3.7) [a]
Completness (%)	90.9 (94.3) [a]
Redundancy	3 (3) [a]
Refinement	
No. reflections	39,788
R_{work} / R_{free}	19.85/ 22.47
No. atoms	
Protein	7719
Water	249
B-factors	
Protein	22.6
Water	19.2
R.m.s deviations	
Bond lengths (Å)	0.005
Bond angle (°)	0.919
Ramanchandran plot:	
Favored region (%)	97.2
Allowed region (%)	2.8

[a] Values in parentheses represent the highest resolution shells.

The three-dimensional structure of ClbPpep consists of two structural regions, all-α and α/β (Fig. 2*A*). The most pronounced homology using the DALI program was obtained with the structure of S12 family members, especially *E. coli* AmpC beta-lactamase and Pab87 peptidase. Although their overall fold was clearly similar, there are significant differences affecting the accessibility of the active site from the surrounding solvent.

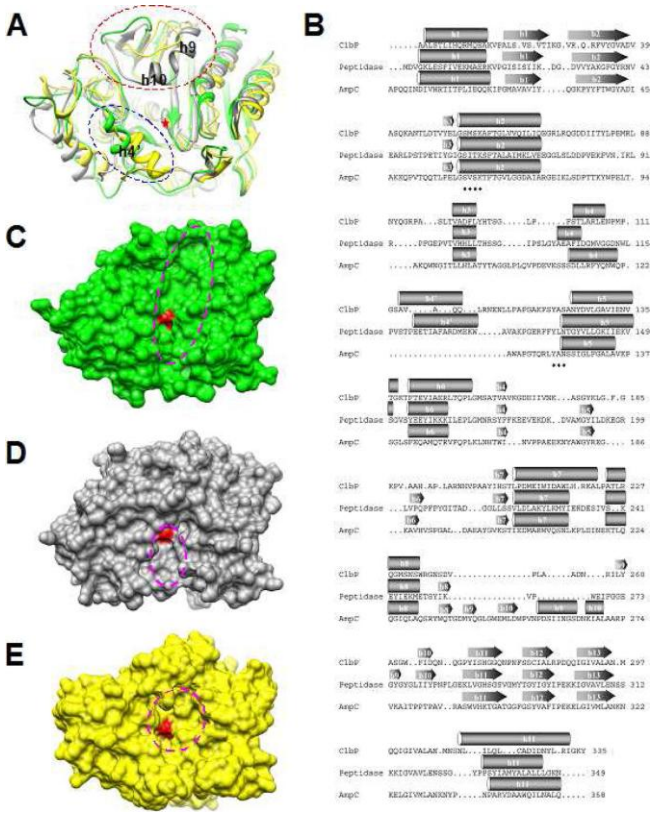

FIGURE 2. Overall structure of ClbPpep, AmpC beta-lactamase (Protein Data Bank code 1KVL) and Pab87 peptidase (Protein Data Bank code 2QMI). A, superimposition of ClbPpep (green), AmpC (gray), and Pab87 peptidase (yellow) presented as ribbons. The loops including helixes h9 and h10 are circled in red. Helixes h4_ are circled in blue. The active serine is indicated by a red star. ClbPpep (B), AmpC (C), and Pab87 peptidase (D) surfaces. The catalytic grooves are circled in pink, and the active serine surfaces are shown in red. E, sequence alignments with secondary structures of ClbPpep with Pab87 peptidase and E. coli AmpC beta-lactamase. The alignment was performed based on the superposed structures. The motifs SxxK and YxN, which defined the MEROPS S12 peptidase family, are indicated by a diamond.

In comparison with AmpC β-lactamases, the upper part of the catalytic pocket was largely open, because of the absence of two helices (h9 and h10) (Fig. 2, *A* and *E*), as observed in Pab87 peptidase and R61 carboxypeptidase. Moreover, the remaining structural element of ClbP in this area, the loop between residues 295–315, b9, and b10 strands (residues 310 to 311 and 316–318, respectively), moved away from the catalytic pocket, resulting in a widening of the catalytic groove, which was unusually larger than the known structures of S12 enzymes. In the bottom part of the catalytic pocket, ClbPpep and Pab87 peptidases harbored the additional helix h4′ (residues 166–170) in contrast with AmpC enzymes (Fig. 2, *A* and *E*). This helix slightly obstructed accessibility to the ClbPpep catalytic pocket. However, the entrance of the binding site was more open in ClbPpep than in S12 peptidases, which harbor one, two, or several bulky helices in this area. Of note, between Cys-337 and Cys-367, ClbP harbored a disulfide bond, which links b12 strand (residues 335–341) with the C-terminal helix h11. To our knowledge, this disulfide bond is not observed in other known structures of S12 enzymes. Overall, the ClbPpep structure was closely similar to that of S12 peptidases. However, ClbPpep had noticeable conformational differences in the active site, especially in the two extremities of the catalytic groove, which was unusually large (Fig. 2,

B–D). This atypical shape suggests an adaptation of ClbP catalytic pocket to a specific substrate.

Structural similarities were observed between the active site of ClbPpep, class C beta-lactamases, and other S12 enzymes such as Pab87 peptidase, in particular, for the motif ^{95}SxxK and the Tyr-186 residue. The positioning of this last residue and the presence of His-327 make the ClbPpep active site very similar to that of Pab87 peptidase (Fig. 3, A and B). In addition, the location and geometry of the residues Ser-95, Lys-98, and Tyr-186 suggest a direct role in the catalytic mechanism mediated by ClbP. To confirm the importance of these residues in ClbP, substitutions S95A, K98T, and Y186G were introduced by site-directed mutagenesis residues into His-tagged ClbP. The resulting mutants were assessed for their ability to restore the cytopathic activity of the pks-positive E. coli ΔclbP isogenic mutant. Infected HeLa cells were analyzed for megalocytosis, G2/M cell cycle arrest, and phosphorylation of histone H2AX. None of these mutants were able to restore the cytopathic activity of the pks-positive E. coli ΔclbP isogenic mutant in contrast with the wild-type ClbP protein (Fig. 3, C and D). These results confirm the identification of the active site and the importance of ClbP main domain for the functionality of pks island. They also suggest that these enzymes and other S12 serine active-site enzymes share a catalytic mechanism.

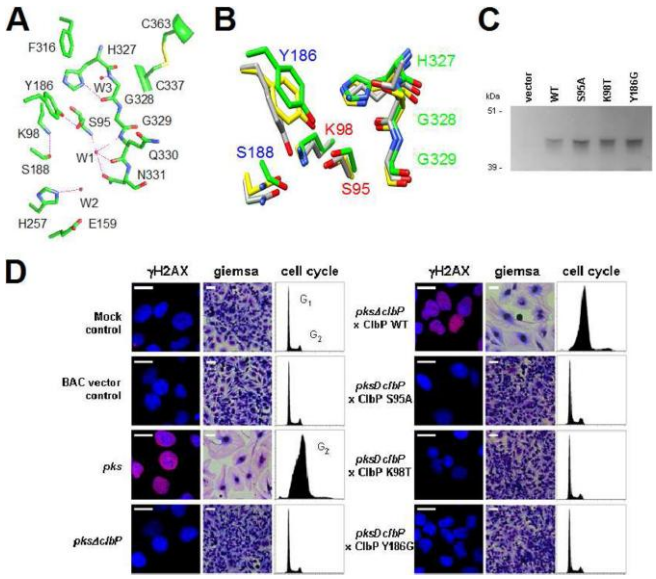

FIGURE 3. Key residues in the ClbP active site. A, residues of ClbP located in close vicinity of the binding site. Hydrogen bonds are indicated by pink dashed lines, and the water molecules are indicated by W. Carbon atoms are colored green, oxygen atoms are colored red, and nitrogen atoms are colored blue. B, overlay of penicillin-recognizing motifs of ClbP, Pab87 peptidase, and AmpC beta-lactamase. Carbon atoms of ClbP, Pab87 peptidase, and AmpC beta-lactamase are in green, yellow, and gray, respectively. The motifs SxxK, Yx(N/S/T), and (K/H)(S/T/G)G are numbered in red, blue, and green, respectively, and residue numbering is according to the ClbP sequence. C and D, phenotypic analysis of pks island activity in E. coli DH10B/pBACpks deleted or not for the clbP gene and trans-complemented with ClbP WT protein or the S95A, K98T, or Y186G active site ClbP mutants. HeLa cells were infected 4 h with 100 E. coli per cell, and then the cells were washed and incubated with gentamicin for 4–72 h. C, anti-5His Western blot analysis of the expression of His-tagged WT and active site ClbP mutants in E. coli DH10B/pBACpksΔclbP after infection. D, host histone H2AX Ser-139 phosphorylation (γH2AX) indicative of DNA double-strand breaks was assayed by confocal immunofluorescence 4 h after infection. DNA and γH2AX are pseudocolored in blue and red, respectively (bars, 20 μm). Cell swelling (megalocytosis) was observed following Giemsa staining 72 h after infection (bars, 50 μm). G2 cell cycle arrest following DNA damage was assessed by flow cytometric DNA content analysis, 48 h after infection.

Conclusion

In conclusion, ClbP is a prototype of a new subfamily of extracytoplasmic serine-reactive peptidases probably acting as maturating enzymes for NRP compounds. The essential role of ClbP in the production of the biological effect make this protein a key target for the control of the bioactivity of the PK-NRP producing pks gene cluster, which may affect commensalism and/or pathogenicity and constitute a predisposing factor for the development of colorectal cancer. The engineering of NRP compounds should take into account these new catalytic units, which could be interesting tools for modifying the structures and bioactivity of NRP and PK-NRP compounds and be a target for identifying bioactive molecules.