J. Mol. Microbiol. Biotechnol. (2000) 2(3): 321-330.

JMMB Research Article

Modeling of *Enterococcus faecalis* D-alanine:D-alanine Ligase: Structure-Based Study of the Active Site in the Wild-Type Enzyme and in Glycopeptide-Dependent Mutants

Martine Prévost,[1*] Daniel Van Belle[1,4], Paul M. Tulkens[2], Patrice Courvalin[3], and Françoise Van Bambeke[2]

[1]Ingénierie Biomoléculaire, Université Libre de Bruxelles, CP 165/64, av. F. Roosevelt, B-1050 Brussels, Belgium
[2]Unité de Pharmacologie Cellulaire et Moléculaire, Université Catholique de Louvain, Brussels, Belgium
[3]Unité des Agents Antibactériens, Institut Pasteur, Paris, France
[4]Present address: Unit of Bioinformatics, Department of Molecular Biology, Université Libre de Bruxelles, B-6041 Gosselies, Belgium.

Abstract

A model for the 3-D structure of *Enterococcus faecalis* D-Ala:D-Ala ligase was produced using the X-ray structure of the *Escherichia coli* enzyme complexed with ADP and the methylphosphinophosphate inhibitor as a template. The model passed critical validation criteria with an accuracy similar to that of the template crystallographic structure and showed that ADP and methylphosphinophosphate were positioned in a large empty pocket at the interface between the central and the C-terminal domains, as in *E. coli*. It evidenced the residues important for substrate binding and catalytic activity in the active site and demonstrated a large body of conserved interactions between the active sites of the *E. faecalis* and the *E. coli* D-Ala:D-Ala ligase, the major differences residing in the balance between the hydrophobic and aromatic environment of the adenine. The model also successfully explained the inactivity of four spontaneous mutants (D295 --> V, which impairs interactions with Mg^{2+} and R293, which are both essential for binding and catalytic activity; S319 --> I, which perturbs recognition of D-Ala2; DAK251-253 --> E, in which the backbone conformation in the vicinity of the deletion remains unaltered but phosphate transfer from ATP is perturbed because of lack of K253; T316 --> I, which causes the loss of a hydrogen bond affecting the positioning of S319 and therefore the binding of D-Ala2). Since D-Ala:D-Ala ligase is an essential enzyme for bacteria, this approach, combining molecular modeling and molecular biology, may help in the design of specific ligands which could inhibit the enzyme and serve as novel antibiotics.

Received January 3, 2000; accepted February 1, 2000. *For correspondence. Email mprevost@ulb.ac.be; Tel. +32-2-650.20.49; Fax. +32-2-650.36.06.

Introduction

World-wide increase in resistance to antibiotics makes essential the discovery of new agents directed against unexploited bacterial targets (Allsop, 1998). The D-alanine:D-alanine (D-Ala:D-Ala) ligase, which catalyses dimerization of D-alanine before its incorporation in late peptidoglycan precursors, is essential for bacteria and may therefore constitute a useful target since its inactivation prevents bacterial growth (Baptista *et al.*, 1997; Van Bambeke *et al.*, 1999). The reaction catalysed by the D-Ala:D-Ala ligase involves the formation of D-alanylphosphate as a tetrahedral intermediate, followed by the nucleophilic attack of this D-alanylphosphate by the deprotonated amine of a second D-Ala, to generate D-Ala-D-Ala by elimination of inorganic phosphate (Walsh, 1989; Fan *et al.*, 1994; Carlson *et al.*, 1999). The enzyme is strongly inhibited by methylphosphinophosphate (Parsons *et al.*, 1988) and weakly by D-cycloserine (Zawadzke *et al.*, 1991), but none of these inhibitors fulfil the requirements for becoming useful drugs. Knowledge of the protein at the molecular level could help for the rational design of improved agents as suggested for many other potential microbial targets (see Moche *et al.*, 1999 and Schluckebier *et al.*, 1999 for recent examples). In this paper, we present a computer-generated structural model of the D-Ala:D-Ala ligase of *Enterococcus faecalis* describing the three-dimensional (3-D) structure of the enzymatic cavity. We used the comparative modeling methodology based on sequence homology with the D-Ala:D-Ala ligase of *E. coli*, the X-ray structure of which is available (Fan *et al.*, 1994). The model was then validated by comparison with the *E. coli* enzyme and by analysis of the interactions of specific residues with ADP and the methylphosphinophosphate. We further validated the model and tested its usefulness by examining the impact of four natural mutations known to inactivate the enzyme in vancomycin-dependent strains of *E. faecalis* (Baptista *et al.*, 1997; Van Bambeke *et al.*, 1999).

Results

Molecular Modeling

E. coli D-Ala:D-Ala ligase (isoform from the *ddlB* gene) was the only template with a known X-ray structure (PDB code 2DLN) to exhibit a significant sequence identity (35 %) with the *E. faecalis* enzyme. Three alignments were used to construct a 3-D model of the target protein (Figure 1). As shown in alignment 1, the main difference between the *E. faecalis* and the *E. coli* sequence is a large insertion of approximately 40 residues spanning positions 40 to 80 (*E. faecalis* numbering). This insertion in alignment 1 is split into smaller insertions of more than 10 residues in

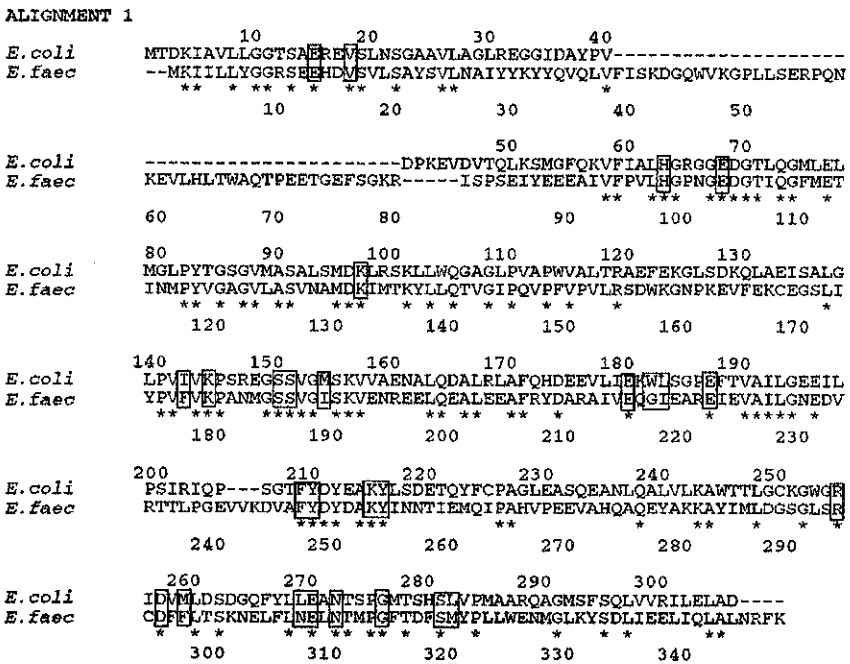

Figure 1. Alignments of the *E. faecalis* D-Ala:D-Ala ligase target sequence with that of the corresponding *E. coli* enzyme. Alignments 1 and 2 were based on amino acid sequences only. Gap initiation and extension penalties were modified in alignment 2 to optimize the number of conserved amino acids. In alignment 3, the gap penalty was based on the *E. coli* enzyme 3-D structure and such that insertions and deletions were less preferred within helices and sheets, buried regions, and straight segments. Identical residues are indicated by asterisks. Residues in contact with ADP and methylphosphinophosphate are boxed. Sequence numbering is given at the top for *E. coli* and at the bottom for *E. faecalis*.

Table 1. Stereochemical Quality of the Three Constructed Models from Alignments 1 to 3 (see Figure 1) and Typical Values Expected for a Crystal Structure at 2.3 Å Resolution (Determined by the Procheck Programme; Laskowski *et al.*, 1993).

Parameter	Model 1	Model 2	Model 3	Typical values at 2.3 Å resolution
% residues in most favoured regions	69.7	73	74	79.8 ± 10
Bad contact per 100 residues	16.1	13.2	11.5	7.6 ± 10

alignments 2 and 3. As for the *E. coli* enzyme, the models based on each alignment and on the template structure (Figure 2) display three domains, each consisting of a ß-sheet surrounded by helices and loops, and a pocket formed by the central and C-terminal domains. The pocket can accommodate ADP and the inhibitor. The root mean square deviation, calculated after superposition of the C_α coordinates of *E. faecalis* residues that are aligned with those of *E. coli*, amounts to the rather low values of 0.25, 0.67, and 0.47 Å for the three models, respectively. The regions most likely to contain errors in the *E. faecalis* model are those for which the accuracy of sequence alignment is critical. They are located in the 40-residue insertion stretch and in the small insertion around residue 240, since there are no equivalent residues in the template structure. These portions had therefore to be modeled *ab initio*, a procedure which is known to yield relatively inaccurate results for insertions larger than seven to eight residues (Sanchez and Sali, 1997).

Evaluation of the Model

Stereochemistry

The various models passed successfully all criteria implemented in Procheck and Modeller nearly to the same degree as the template X-ray structure which is known to a resolution of 2.3 Å (Fan *et al.*, 1994). These criteria

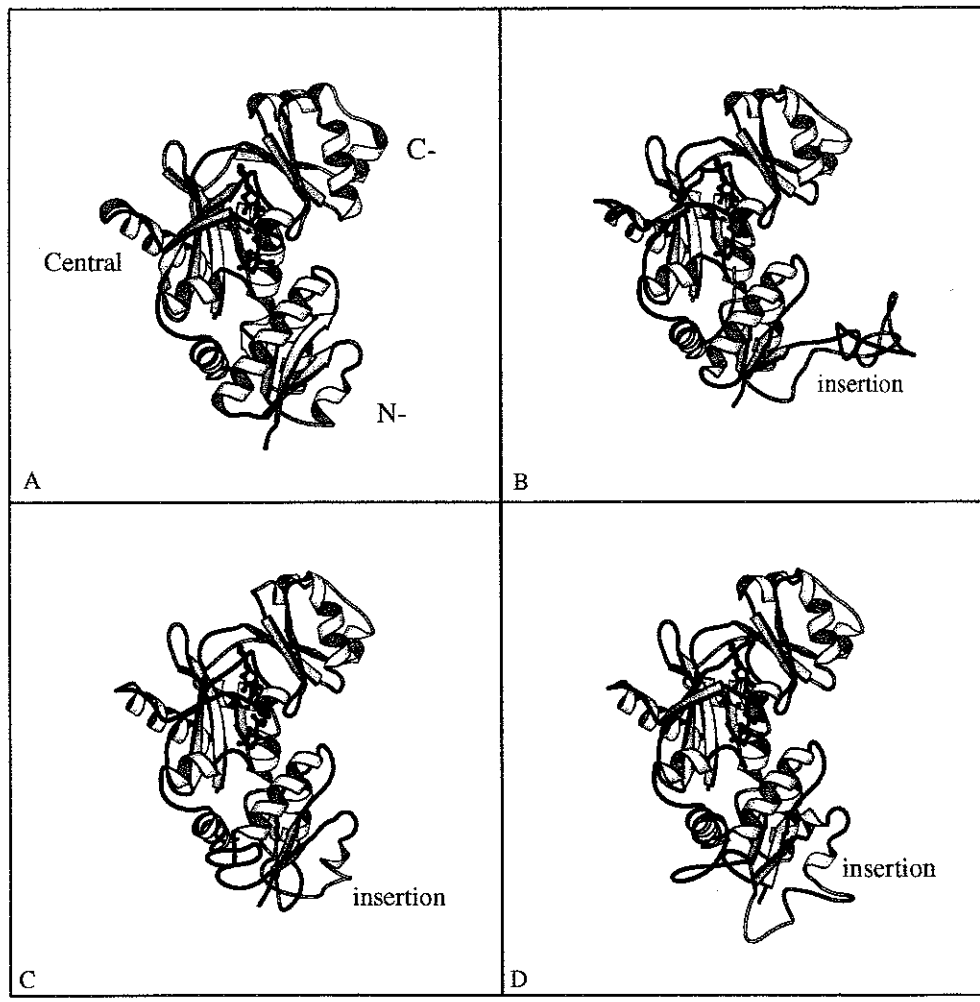

Figure 2. Structural comparison of *E. coli* and *E. faecalis* D-Ala:D-Ala ligases complexed with ADP and the phosphinophosphate inhibitor. A: schematic representation of the X-ray structure of *E. coli* D-Ala:D-Ala ligase determined at 2.3 Å resolution. B-D: three-dimensional models computed with the Modeller software (Sali and Blundell, 1993) from the alignments 1 to 3 in Figure 1. Each structure is composed of three domains featuring helices packed onto a sheet and exhibits a pocket housing ADP and phosphinophosphate. The latter two molecules are depicted as ball-and-stick in black. The location of the large insertion in the *E. faecalis* enzyme is indicated.

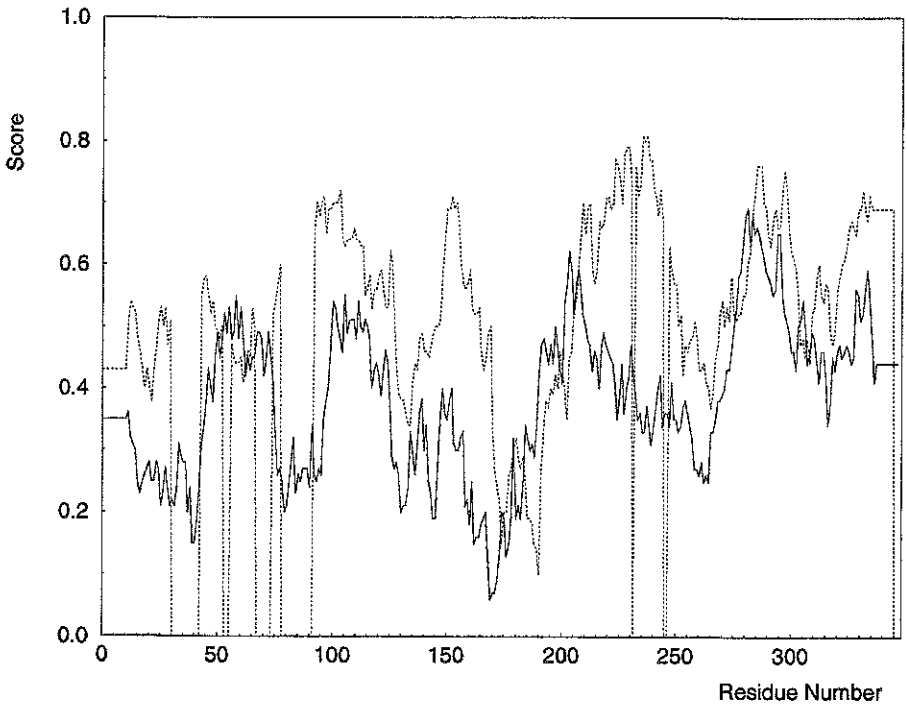

Figure 3. Compatibility scores between the *E. coli* D-Ala:D-Ala ligase X-ray structure (dotted line) and the *E. faecalis* D-Ala:D-Ala ligase 3-D model structure (solid line) obtained from the third alignment in Figure 1. Scores higher than 0 are deemed acceptable for the 3-D profile program (Lüthy *et al*, 1992). The numbering is that of the *E. faecalis* enzyme. Regions in the *E. coli* enzyme with a zero score correspond to insertions in the *E. faecalis* sequence.

(Ramachandran plot, peptide bonds planarity, non-bonded interactions, hydrogen bonds, closeness of side chain dihedral angles to ideal values) were consistent with crystal structures determined at a resolution of 2 to 2.5 Å. For example, the percentage of non-glycine and non-proline residues in most favored regions of the Ramachandran plot reaches 88 % for the *E. coli* X-ray structure and 77 % for the third model. This lower value is due, as expected, to the residues in the insertion located between positions 40 and 80 and also to a few glycine residues in the *E. coli* enzyme which are aligned with non-glycine residues in the *E. faecalis* enzyme. The model generated from the third alignment had better stereochemistry than the two other models (Table 1) and was therefore selected for further analysis.

3D Profiles

A more stringent test of the overall fold and side chain packing of the model was provided by the 3-D profile programs. These methods plot, in a window moving along the sequence, a score that correlates with the accuracy of the model in the corresponding region and that depends only on the sequence and structure of the model. The computed 3-D profiles of the *E. coli* experimental structure and of the *E. faecalis* model are shown in Figure 3. Improperly built segments can be revealed by low overall profile scores. None of the 3-D profile values were lower than zero. Both curves had roughly similar profiles, although the X-ray structure altogether scored better than the model. In particular, two regions spanning residues 130-170 and 210-250 displayed a lower score than their *E. coli* counterparts. These fragments contained less than 20 % identical residues aligned with those of the template in the third alignment (Figure 1). This low percentage value is mainly due to lack of identity in the 155-170 region. Three other portions spanning residues 40-80, 175-210, and 270-320 exhibited score values as high as those of the X-ray *E. coli* structure. These close values were expected for the latter two segments since they encompass approximately 30 % of identical residues in the two enzymes. This result is, however, more surprising for the first stretch that poorly aligns with the *E. coli* sequence.

Structural Analysis of the Binding of ADP and of Methylphosphinophosphate

The region of the active site is well defined in the 3-D model. This allowed us to compare its topology with that of the active site in the *E. coli* enzyme and to identify residues establishing strong interactions with ADP or with the methylphosphinophosphate inhibitor (Figure 4). The computed contact area reveals the contacts that prevail between the enzyme and ADP or the inhibitor without discriminating the contacts promoted only by non-specific Lennard-Jones interactions from those arising, for instance, from the more specific H bond or charge-charge interactions. Table 2 lists the residues in the two enzymes that are in contact with ADP and the phosphinophosphate inhibitor, through their main and/or side chain, in comparison with those identified previously in *E. coli* based on analysis of the X-ray structure (Fan *et al.*, 1994). The main observation is that there is a large body of conserved interactions between the two active sites (Figure 4 a-c). Detailed examination of the model constructed for the *E. faecalis* enzyme indicated that, like in *E. coli*, the adenine ring lies in a pocket mainly composed of hydrophobic and aromatic amino acids. Interestingly, however, a few interactions in the *E. faecalis* model differed from those established in the *E. coli* X-ray structure. The main differences resided in the vicinity of ADP, where two residues in *E. coli*, namely Trp182 and Leu269, are

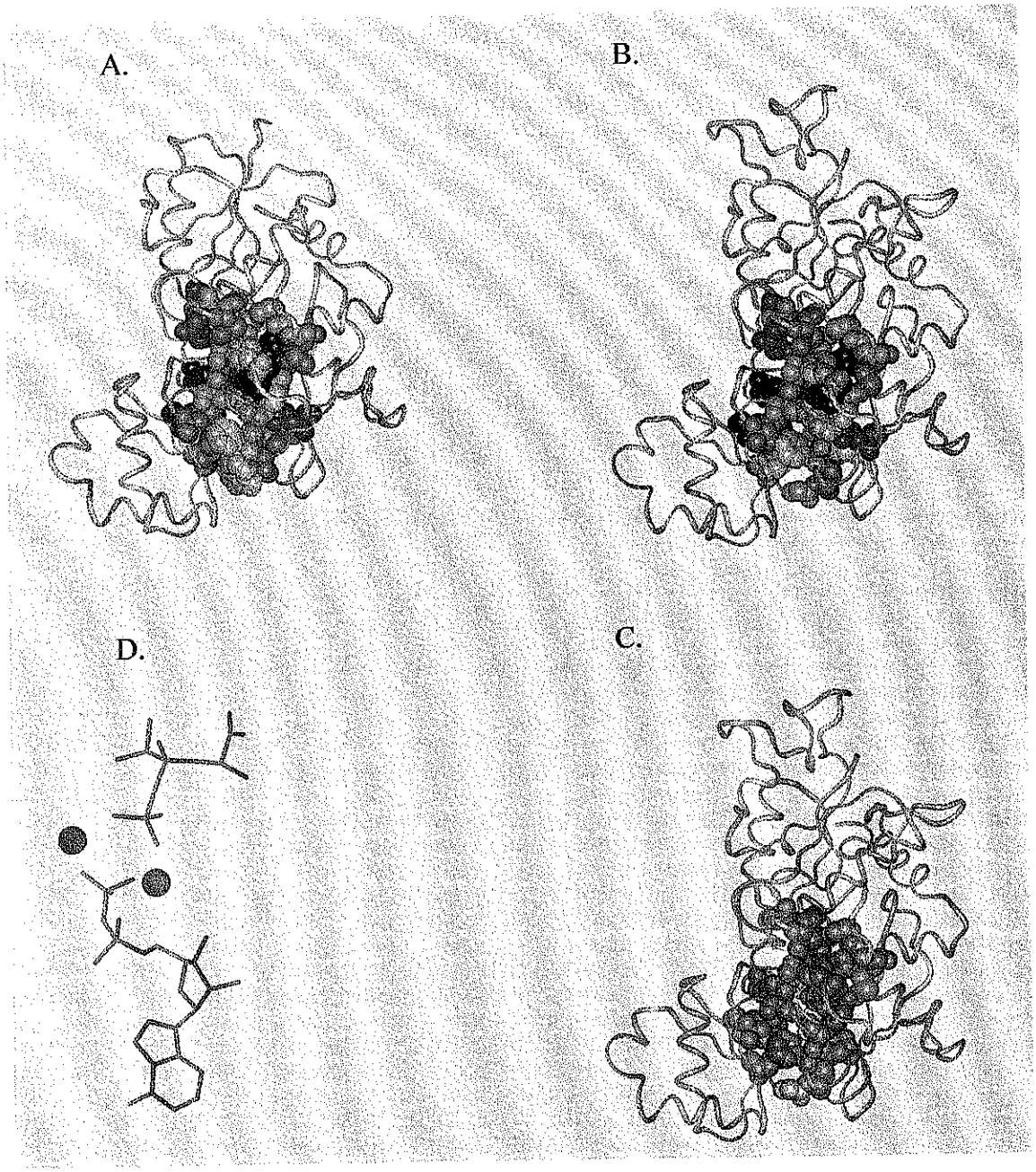

Figure 4. Upper panel: Tube diagrams of *E. coli* (A) and *E. faecalis* (B) D-Ala:D-Ala ligases highlighting the residues (depicted as balls) found in contact with ADP and phosphinophosphate. Color code of residues: green, hydrophobic; pink, aromatic, dark blue, basic; light blue, polar; red, acidic. Lower panel (C): tube diagram of *E. faecalis* D-Ala:D-Ala ligase highlting the conserved residues in the active site (depicted as balls). Color code: white, homologous residues; yellow, non-homologous residues; orange, conserved residues. (D) ADP and methylphosphinophosphate inhibitor are depicted as sticks, roughly in their active site orientation relative to A, B, and C; the two Mg^{2+} ions are shown as spherical balls.

substituted by Gly217 and Asn307 in *E. faecalis* (Table 2). The loss of the bulky side chain of Trp is counterbalanced by the presence of two additional aromatic side chains arising from the substitutions Ile142 to Phe177 and Met259 to Phe297 (Table 2). The latter substitutions are responsible for the relative paucity in aliphatic amino acids of the *E. faecalis* active site. The change of Leu269 to Asn307 should entail the formation of one additional H bond with the adenine since the latter makes two more H bonds with the conserved Glu215 and Lys179 side chains. The ribose, the beta- and gamma-phosphate groups, as well as the two Mg^{2+} cations, interact with almost identical residues in *E. faecalis* and *E. coli* (Table 2). The ribose makes H bonds with Glu222 side chain and Tyr254 main chain NH group. Each ligase possesses three positively charged amino acids, Lys132, Lys179, and Lys253, which form electrostatic links with the phosphate groups of ADP. The two Mg^{2+} cations, that are essential for phosphate transfer and bridge ADP and the inhibitor phosphate group, interact with the conserved Glu308, Asn310, and Asp295 (Table 2). Thus, clusters of negatively and positively charged residues provide the electrostatic environment to

counterbalance the phosphoryl groups and the two Mg^{2+}. In both structures, the ammonium group of the inhibitor interacts with two Glu, 13 and 103, and Ser185. However, the contact area with Glu103 is smaller than that provided by the other two residues. The carboxylic group is maintained at the other end of the inhibitor by one hydrogen bond to Ser319 side chain. In the center, the P-O bond and the phosphate group interact with identical residues (Lys253, Arg293, and Tyr254) in both enzymes. The two side chains of the inhibitor make contact with hydrophobic groups that are conserved in the active site of both enzymes. Careful examination of the interactions summarised in Table 2 also evidenced in the *E. coli* structure a few residues that interact with ADP and the inhibitor, in addition to those identified previously (Fan *et al.*, 1994). Among those, Met154, Leu183, and, as already mentioned, Leu269 (*E. coli* numbering) are not conserved in *E. faecalis*, suggesting that the interactions established may be less specific than anticipated and that substitutions of these amino acids are partly counterbalanced by simultaneous changes of neighbouring residues.

Structure-Function Relationships

Lack of D-Ala:D-Ala ligase activity in the four natural mutants of *E. faecalis* (D295 --> V, S319 --> I, DAK251-253 --> E and T316 --> I; see Figure 5) could easily be explained by the model. The Asp295 residue is invariant in ligases (Evers *et al.*, 1996), and the model shows that it occupies a crucial location in the active site. It interacts directly with one of the two Mg^{2+} ions which assist the phosphate group transfer from ATP. Furthermore it forms a H bond with Arg293, which probably stabilises the carboxylate group of D-Ala1 substrate since it makes contact with the inhibitor P-O group (Figure 6b). The corresponding Asp residue in *E. coli* was shown to have an important catalytic role. The substitution of Asp295 to Val should therefore grossly perturb the phosphate group transfer. Ser319 is also a conserved residue and is likely to play a key role in the recognition of the D-Ala2 substrate since it holds the terminal carboxylic acid group of the phosphinophosphate inhibitor through a hydrogen bond (Figure 6c). It is therefore not surprising that substitution of Ser319 by the large hydrophobic Ile amino acid should have a detrimental effect on the binding of D-Ala2 and abolish the enzymatic activity. The model shows that Lys 253, another conserved residue, is likely to be involved in the transfer of the phosphate group of ATP (Figure 6a-b). Its loss in the D251-253E mutant is therefore also expected to totally impair activity. The concomitant loss of amino acids at positions 251 and 252 raises, however, the question as to whether the loss of enzymatic activity arises solely from the absence of Lys253 or from a more drastic rearrangement of the polypeptide chain due to the deletion. The mutant structure was therefore modeled following the protocol used to produce the structure of the wild-type enzyme, which exhibits a helical conformation in the 250-254 region (Figure 6b). The local rearrangement resulting from the loss of Ala252 and Lys253 caused the corresponding segment to adopt an extended conformation allowing the side chain of the new Glu residue to take a position very similar to that of the Asp251 in the wild type enzyme (Figure 6a). The geometry of the active site remained therefore essentially unchanged. Finally, the model revealed that the side chain of the conserved Thr316 forms a H bond with Ser319 (Figure 6c). The loss of this H bond is likely to alter the orientation and position of Ser319, affecting thereby the binding of the D-Ala2 substrate. The model therefore explains why mutation of Thr316 to Ile may result in enzyme inactivation.

Discussion

This study presents a model for the 3-D structure of the D-Ala:D-Ala ligase of *E. faecalis*, that describes the organisation of the enzymatic cavity and the role played by critical residues in the activity of the enzyme, an information that may be useful in designing new inhibitors. Those presently available, namely D-cycloserine and phosphinates, are indeed unsuitable as drugs. The former shows only marginal activity because of lack of strong anchoring in the active site (Zawadzke *et al.*, 1991) and the latter, although very potent and only slowly reversible, are inactive against bacteria because they are too polar to cross the pericellular membrane (Parsons *et al.*, 1988).

Constructed by the comparative modeling approach using the X-ray structure of the homologous *E. coli* enzyme as a template, the accuracy of the proposed model was checked against sequence, structural and biological data. Correct sequence alignments are crucial for the production of reliable 3-D models constructed by comparative modeling, and careful analyses must be undertaken to detect regions of weak homology. This is illustrated by the iterative changes in the alignment made during calculation of the 3-D model by maximising the degree of identity and valuing the information contained in the three-dimensional structure of the template. The corrections introduced in the alignment, for instance in the region spanning residues 40 to 80 which presented the lowest homology with the *E. coli* sequence, allowed a progressive shift from a totally disorganized loop in the first model to a local rearrangement better adjusted to the *E. coli* structure in the second model and embodying a ß-sheet and two small α-helices in the third model. Detailed evaluation of the stereochemistry and track of the improperly built segments were also highly valuable to validate the models. The degree of accuracy in all the stereochemical features of the three models was comparable to that of the X-ray template structure, with a slightly better performance for the third model generated with an improved alignment. The regions of the protein with the poorest stereochemistry corresponded to low or non homologous portions in the alignment. The large 40-residue insertion in *E. faecalis* for which there is no real counterpart in *E. coli* sequence was located in the N-terminal domain. It was not expected to have strong influence on the structural accuracy of remote parts of the 3-D model and did not directly participate in the active site. In contrast, the insertion of 3 (or 2) residues near position 240 should be examined very critically since it is close to the active site pocket. None of the *E. faecalis* mutations occurred in the vicinity of this insertion so that the model could not be tested experimentally. Yet, improperly built segments, if they are in small numbers, can be detected by low scoring regions in profile calculations of the 3-D environment of each residue. The third model, which performed best with the stereochemical features, had a 3-D profile in which 80 % of the residues involved in key interactions within the active site were located in regions which fitted well with those of the template. In particular,

Table 2. Amino acids in the active site of the D-Ala:D-Ala ligase interacting with ADP, Mg^{2+}, or phosphinophosphate, as determined by X-ray (Fan *et al.*, 1994) or by comparative modeling

Interaction with:		D-Ala:D-Ala ligase		
		E. coli (Fan *et al.*, 1994)	*E. coli* (this study)	*E. faecalis* (3-D model)
ADP	ADENINE	Ile 142	Ile 142	Phe 177 [a]
		Trp 182	Trp 182	Gly 217
		Phe 209	Phe 209	Phe 247 [b]
		Met 259	Met 259	Phe 297 [a]
			Leu 269	Asn 307
			Met 154	Ile 189 [a]
		Lys 144	Lys 144	Lys 179 [b]
		Glu 180	Glu 180	Glu 215 [b]
			Leu 183	Ile 218 [a]
	RIBOSE	Glu 187	Glu 187	Glu 222 [b]
		Tyr 210	Tyr 210	Tyr 248 [b]
			Met 154	Ile 189 [a]
		Lys 215	Lys 215	Lys 253 [b]
	α-P	Lys 97	Lys 97	Lys 132 [b]
		Lys 144	Lys 144	Lys 179 [b]
	ß-P	Lys 97	Lys 97	Lys 132 [b]
		Ser 151	Ser 151	Ser 186 [b]
		Lys 215	Lys 215	Lys 253 [b]
Mg^{2+}		Glu 270	Glu 270	Glu 308 [b]
		Asn 272	Asn 272	Asn 310 [b]
Mg^{2+}		Asp 257	Asp 257	Asp 295 [b]
METHYLPHOSPHINOPHOSPHATE	NH_3 +	Glu 15	Glu 15	Glu 13 [b]
			Glu 68	Glu 103 [b]
			Ser 150	Ser 185 [b]
	PO_4 -	Lys 215	Lys 215	Lys 253 [b]
			Tyr 216	Tyr 254 [b]
			Arg 255	Arg 293 [b]
	CH_3 (1)	His 63	His 63	His 98 [b]
			Val 18	Val 16 [b]
	CH_3 (2)		Leu 282	Met 320 [a]
			Tyr 210	Tyr 248 [b]
			Lys 215	Lys 253 [b]
	P = O	Arg 255	Arg 255	Arg 293 [b]
		Gly 276	Gly 276	Gly 314 [b]
	COO -	Ser 281	Ser 281	Ser 319 [b]
		Leu 282	Leu 282	Met 320 [a]

[a] interactions found to be of similar nature in *E. coli* and in *E. faecalis* by comparative modeling
[b] interactions found to be identical in *E. coli* and in *E. faecalis* by comparative modeling

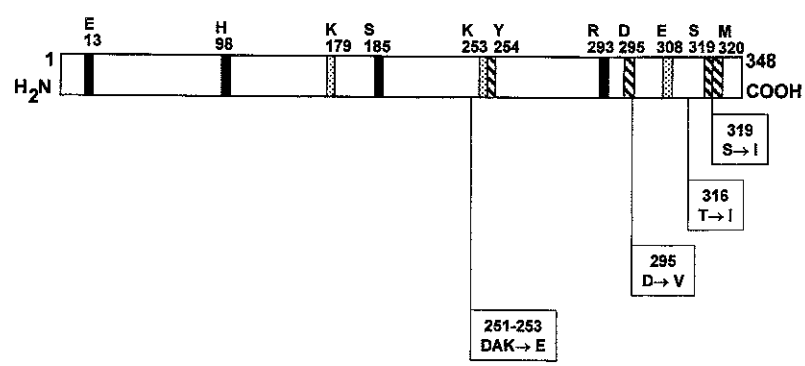

Figure 5. Schematic representation of the *E. faecalis* D-Ala:D-Ala ligase. The positions of the amino acids implicated in the binding of D-Ala1, D-Ala2, and of ATP, as defined by Shi and Walsh (1995) from the X-ray structure, are indicated in black, hatched, and dotted bars, respectively. Substitutions relative to the wild-type enzyme in mutants are based on the data of Baptista *et al.*, 1997 and Van Bambeke *et al.*, 1999.

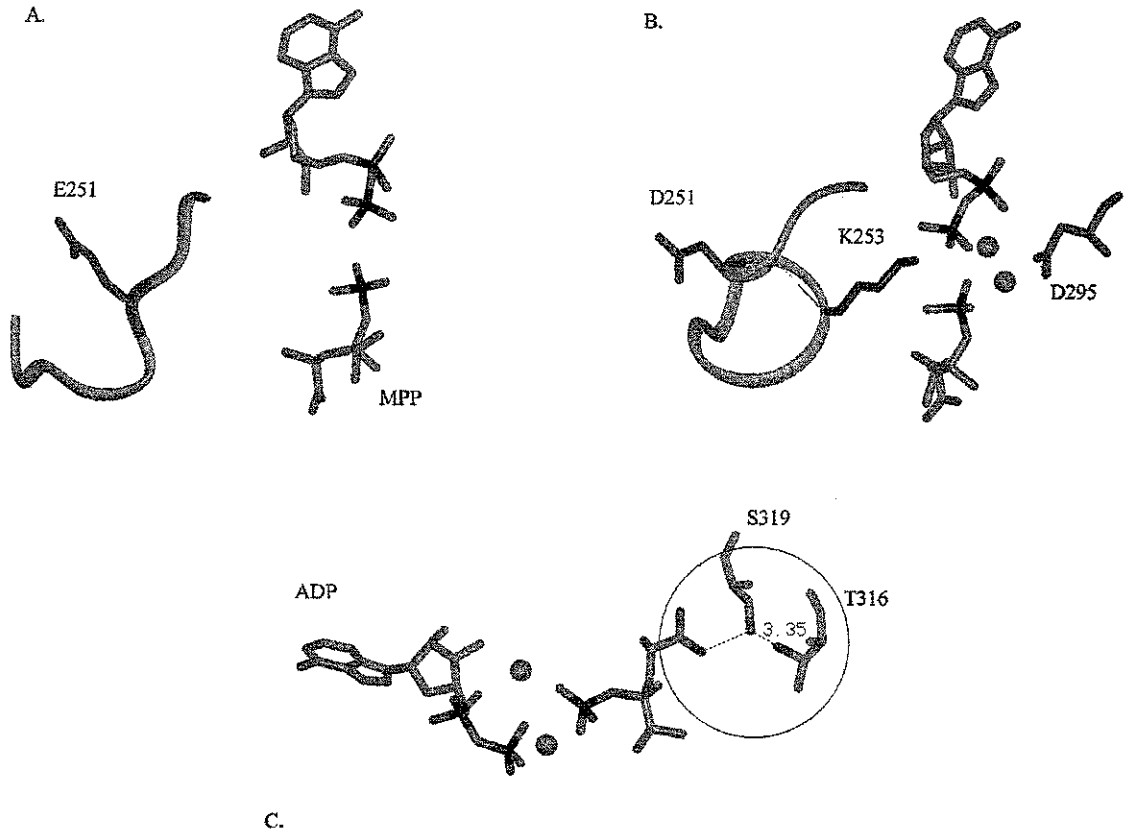

Figure 6. Predicted location of the residues in the binding site of *E. faecalis* D-Ala:D-Ala ligase for which substitution was experimentally shown to impair enzymatic activity (see Figure 5). A . The DAK251-253 --> E mutant. The ribbon represents the modeled conformation of the backbone 248-256 region, nearby the mutation. The E251 side chain is represented. B. The wild-type active site structure is shown with the DAK251-253 segment for sake of comparison with panel A, and with D295, substituted into valine in one mutant (see Figure 5). The ribbon represents the backbone conformation of the 248-256 region. K253 and D251 side chains are represented. The two Mg^{2+} ions near D295 side chain are depicted as spherical balls. C. S319 and T316, substituted into isoleucine in distinct mutants (see Figure 5). The circled region encompasses the carboxylate group of methylphosphinophosphate, S319, and T316. H bonds are represented by broken lines. ADP and methylphosphinophosphate (MPP) are represented as sticks. The phosphorus atoms of the phosphate groups are shown in black.

approximately 50 % of the residues that were in contact with ADP or the inhibitor (Table 2) belonged to the 175-210 and 270-320 regions for which the score of the 3-D profile equals that of the X-ray structure.

Analysis of the residues forming the binding site for ADP and the methylphosphinophosphate inhibitor showed that not only those residues but also their spatial arrangement within the active site (Figure 4) were highly conserved in the *E. faecalis* and the *E. coli* ligases. This suggests that the molecular mechanism of the enzymatic reaction is likely to be similar. This finding, made possible by the computer-based approach, is important for the development of inhibitors with a broad spectrum of activity. One should keep in mind however that the two active sites differ at a few positions, which may hinder the binding of too specific inhibitors. The model also evidences structural similarities between the D-Ala:D-Ala ligases and the enzymes of the ATPgrasp family. These enzymes differ in their substrate specificity but share a mechanism of action involving ATP-dependent ligation of the carboxyl group of a substrate to the amino group of another substrate, yielding an acylphosphate intermediate (Galperin and Koonin, 1997). Our conformational analysis identified the role of several conserved amino acids present in all ATPgrasp proteins [Galperin and Koonin, 1997], namely Lys132, Lys179, Glu215, Gly217, Ile218, Glu222, Lys253, Arg293, Asp295, Glu308, and Asn310 (*E. faecalis* numbering). Interestingly, the model also highlighted the positions recognised as critical for the activity of carbamoyl synthetase (positions 98, 132, 179, 215, 222, 293, 295, 308, and 310 [Javid-Majd *et al.*, 1996]), and the GSH synthetase (positions 13, 132, 179, 185, 215, 217, 218, 253, 295, 308, and 310 [Fan *et al.*, 1995]). It also underscored all of the positions, but one, involved in the recognition of adenine in cAMP-dependent kinases (positions 189, 215, 217, 218, 297, 307 [Kobayashi and Go, 1997]; position 216 was not highlighted in our model but was not proposed as being critical in the *E. coli* enzyme). Thus, except for the replacement of an aromatic residue by a Gly at position 217, the active site of the *E. faecalis* enzyme presented a high degree of homology with that of other enzymes displaying a similar mode of action. This information is critical for the design of inhibitors acting as novel antibiotics since these should not affect eukaryotic enzymes.

Our model also allows a better understanding of the structure-function relationship based on the analysis of mutations, and, could lead to improved engineering of potentially useful inhibitors. Most of man-engineered mutants have indeed been directed at substitutions of residues suspected to be involved in enzyme activity (Shi and Walsh, 1995) or substrate specificity (Park *et al.*, 1996; Fan *et al.*, 1997; Haely *et al.*, 1998; Lessard *et al.*, 1999). By contrast, spontaneous impaired mutants, that are sometimes difficult to explain by simple visual inspection of the X-ray structure (especially if this structure is that of another bacterial species), may point to unexpected residues. The substitutions observed in three of our four *E. faecalis* mutants (D295 --> V, S319 --> I, and DAK251-253 --> E) were expected to result in impairment of activity on the basis of identity with residues known to be involved in activity in the *E. coli* enzyme. Such an impairment could not be anticipated for the T316 --> I mutation. The model therefore rationalised the experimental observation and allowed to rank T316 among the residues which could usefully be targeted for the design of new inhibitors.

In conclusion, the mathematical models and methodology which have been applied to the D-Ala:D-Ala ligase of *E. faecalis* yielded data which correlated fairly well with those generated by experimental studies and with the known general properties of this enzyme. This implies that computer models may usefully complement experimentally-determined three-dimensional structures.

Experimental Procedures

The 3-D structure of the *E. faecalis* D-Ala:D-Ala ligase was calculated using the Modeller4 software (Sali and Blundell, 1993). The calculation involved four steps: (i) identification of proteins with a known 3D structure homologous to the sequence to be modeled (target sequence); (ii) alignment of the target sequence with those of homologous proteins and the selection of sequences that will serve as templates for the building of the model; (iii) building of the model; and (iv) evaluation of the model against criteria of intrinsic coherence and experimental data.

1. Identification of Homologs with a Known 3-D Structure

Proteins similar in sequence to the target sequence and with a known 3-D structure were identified in a set of approximately 1000 non-homologous sequences representative of the Brookhaven protein databank (PDB). These sequences, and their homologs exhibiting an identity score higher than 30%, were considered as potential templates.

2. Alignment of the Target Sequence with Those of Homologous Proteins

First, an alignment of the target sequence with those of homologs was performed taking into account the amino acid sequences only. Second, the gap initiation and extension penalties were varied to obtain a maximal number of conserved amino acids between the two alignments. Third, the gap penalty was allowed to depend also on the 3-D structure of the template proteins. The gap penalty was then chosen such that insertions and deletions were less preferred within helices and sheets, buried regions, and straight segments. These alignment studies identified the *E. coli* D-Ala:D-Ala ligase (isoform from the *ddlB* gene) as the only structural template with an identity score higher than 30 %.

3. Building of the Model

The alignments were used to calculate 3-D models containing all the heavy atoms of the *E. faecalis* enzyme. For each alignment, five models were derived based on the satisfaction of spatial restraints extracted from the alignment of the target sequence with the template structure. The representative model was selected on the basis of the lowest value of the objective function that describes the spatial restraints. The 3-D models were produced for the entire *E. faecalis* amino acid sequence and with both ADP and the methylphosphinophosphate, the initial positions of which were taken from the *E. coli* X-ray structure. These positions were then optimised by taking into account the non bonded intermolecular interactions specific to the *E. faecalis* active site.

4. Evaluation of the Model

The validity of the model was evaluated in three respects. First, its stereochemical validity was assessed using the Modeller (Sali and Blundell, 1993) and Procheck (Laskowski *et al*, 1993) programs. Additional evaluation was done by '3-D profile' program (Lüthy *et al*, 1992) which relies on statistical preferences of each of the 20 amino acids in a particular environment. The environmental properties include information about the water accessibility of the side chains, the fraction of the side chain exposed to polar atoms, and the local secondary structure. In a second step, a careful structural comparison of the active sites of the *E. coli* and *E. faecalis* enzymes was performed. The residues lying in close contact with ADP and methylphosphinophosphate were identified by computing the contact area, which was evaluated as the sum of the areas of the polyhedra faces that atoms of a given residue (or molecule) pair have in common (Alard, 1990). In a last step, we modeled four protein mutants (D295 --> V; S319 --> I; T316 --> I; DAK251 --> 253E) which are inactive (Baptista *et al.*, 1997; Van Bambeke *et al.*, 1999). Strains harboring these mutations are vancomycin-dependent since they can only synthesize peptidoglycan with late precursors terminating in the depsipeptide D-Ala:D-Lac. The latter results from the activity of a D-Ala:D-Lac ligase induced by the presence of vancomycin in the medium.

Acknowledgements

M.P. and F.V.B. are *Chercheur Qualifié* and *Chargé de Recherches*, respectively, of the Belgian *Fonds National de la Recherche Scientifique.*

References

Alard, P. 1990. Computations of surface areas and energies in the fold of macromolecules. PhD Thesis, Université Libre de Bruxelles, Brussels.

Allsop, A.E. 1998. New antibiotic discovery, novel screens, novel targets and impact of microbial genomics. Curr. Opin. Microbiol. 1: 530-534.

Baptista, M., Depardieu, F., Reynolds, P.E., Courvalin, P., and Arthur, M. 1997. Mutations leading to increased levels of resistance to glycopeptide antibiotics in VanB-type enterococci. Mol. Microbiol. 25: 93-105.

Carlson, H.A., Briggs, J.M., and McCammon, J.A. 1999. Calculation of the pKa values for the ligands and side chains of *Escherichia coli* D-alanine:D-alanine ligase. J. Med. Chem. 42: 109-117.

Evers, S., Casadewall, B., Charles, M., Dutka-Malen, S., Galimand, M., and Courvalin, P. 1996. Evolution of structure and substrate specificity in D-alanine:D-alanine ligases and related enzymes. J. Mol. Evol. 42: 706-712.

Fan, C., Moews, P.C., Shi Y., Walsh, C.T., and Knox, J.R. 1995. A common fold for peptide synthetases cleaving ATP to ADP: glutathione synthetase and D-alanine:D-alanine ligase of *Escherichia coli*. Proc. Natl. Acad. Sci. USA 92:1172-1176.

Fan, C., Moews, P.C., Walsh, C.T., and Knox, J.R. 1994. Vancomycin-resistance: structure of D-alanine:D-alanine ligase at 2.3 Å resolution. Science (Wash.) 266: 439-443.

Fan, C., Park, I. S., Walsh, C. T., and Knox, J.R. 1997. D-alanine: D-alanine ligase: phosphonate and phosphinate intermediates with wild-type and the Y216F mutant. Biochemistry (US) 36: 2531-2538.

Galperin, M.Y., and Koonin, E.V. 1997. A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol activity. Protein Sci. 6:2639-2643.

Haely, V.L., Park, S.I., and Walsh, C.T. 1998. Active-site mutants of the VanC2 D-alanyl-D-serine ligase, characteristic of one vancomycin-resistant bacterial phenotype, revert towards wild-type D-alanyl-D-alanine ligases. Chem. Biol. 5: 197-207.

Javid-Majd, F., Stapleton, M.A., Harmon, M.F., Hanks, B.A., Mullins, L.S., and Raushel, F.M. 1996. Comparison of the functional differences for the homologous residues within the carboxy phosphate and carbamate domains of carbamoyl phosphate synthetase. Biochemistry (US) 35: 14362-14369.

Kobayashi, N., and Go, N. 1997. A method to search for similar protein-protein local structures at ligand binding sites and its application to adenine recognition. Eur. J. Biophys. 26: 135-144.

Laskowski, R.A., McArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26: 283-291.

Lessard, I.A.D., Haely, V.L., Park, I.S., and Walsh, C.T. 1999. Determinants for differential effects on D-Ala-D-Lactate vs D-Ala-D-Ala formation by the VanA ligase from vancomycin-resistant enterococci. Biochemistry (US) 38: 14006-14022.

Lüthy, R., Bowie, J.U., and Eisenberg, D. 1992. Assessment of protein models with three-dimentional profiles. Nature (London) 356: 83-85.

Moche, M., Schneider, G., Edwards, P., Dehesh, K., and Lindqvist, Y. 1999. Structure of the complex between the antibiotic cerulin and its target,

beta-ketoacyl-acyl carrier protein synthase. J. Biol. Chem. 274: 6031-6034.

Park, I. S., Lin, C. H., and Walsh, C. T. 1996. Gain of D-alanyl:D-lactate or D-lactyl:D-alanine synthetase activities in three active site mutants of the *Escherichia coli* D-alanyl:D-alanine ligase B. Biochemistry (US) 35: 10464-10471.

Parsons, W.H., Patchett, A.A., Bull, H.G., Schoen, W.R., Taub, D., Davidson, J., and Busch, R.D. 1988. Phosphinic acid inhibitors of D-alanyl:D-alanine ligase. J. Med. Chem. 31: 1772-1778.

Sali, A., and Blundell, T.L. 1993. Comparative protein modeling by satisfaction of spacial restraints. J. Mol. Biol. 234: 779-815.

Sanchez, R., and Sali A. 1997. Advances in comparative protein-structure modeling. Curr. Op. Struct. Biol. 7: 206-214.

Schuckebier, G. Zhong, P., Steward, K.D., Kavanaugh, T.J., and Abad-Zapatera, C. 1999. The 2.2 Å structure of the rRNA methyltransferase ErmC' and its complexes with cofactor and cofactor analogs: implications for the reaction mechanism. J. Mol. Biol. 289: 277-291.

Shi, Y., and Walsh C.T. 1995. Active site mapping of *Escherichia coli* D-Ala:D-Ala ligase by structure-based mutagenesis. Biochemistry (US) 34: 2768-2776.

Van Bambeke, F., Chauvel, M., Reynolds, P.E., Fraimow, H.S., and Courvalin, P. 1999. Vancomycin-dependent *Enterococcus faecalis* clinical isolates and revertant mutants. Antimicrob. Agents Chemother. 43: 41-47.

Walsh, C.T. 1989. Enzymes in the D-Alanine branch of bacterial cell wall peptidoglycan assembly. J. Biol. Chem. 264: 393-2396.

Zawadzke, L.E., Bugg, T.D.H. and Walsh, C.T. 1991. Existence of two D-alanine:D-alanine ligases in *Escherichia coli*: cloning and sequencing of the *ddlA* gene and purification and characterization of the DdlA and DdlB enzymes. Biochemistry (US) 30: 1673-1682.