doi:10.1006/jmbi.2000.4215 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2000) 304, 435–445

JMB

Crystal Structure of *Escherichia coli* UDPMurNAc-tripeptide D-alanyl-D-alanine-adding Enzyme (MurF) at 2.3 Å Resolution

Youwei Yan[1]*, Sanjeev Munshi[1], Barbara Leiting[2], Matt S. Anderson[2] John Chrzas[3] and Zhongguo Chen[1]*

[1]*Department of Structural Biology, Merck Research Laboratories, West Point PA 19486, USA*

[2]*Department of Endocrinology and Chemical Biology, Merck Research Laboratories, Rahway NJ 07065, USA*

[3]*Biological, Chemical, and Physical Sciences Department Illinois Institute of Technology Chicago, IL 60616, USA*

**Corresponding authors*

MurF is required to catalyze the final step in the synthesis of the cytoplasmic precursor of the bacterial cell wall peptidoglycan, rendering it an attractive target for antibacterial drug development. The crystal structure of the MurF apo-enzyme has been determined using the multiwavelength anomalous dispersion method and refined to 2.3 Å resolution. It contains three consecutive open α/β-sheet domains. In comparison with the complex crystal structures of MurD and its substrates, The topology of the N-terminal domain of MurF is unique, while its central and C-terminal domains exhibit similar mononucleotide and dinucleotide-binding folds, respectively. The apo-enzyme of MurF crystal structure reveals an open conformation with the three domains juxtaposed in a crescent-like arrangement creating a wide-open space where substrates are expected to bind. As such, catalysis is not feasible and significant domain closure is expected upon substrate binding.

Keywords: MurF; peptidoglycan; bacterial cell wall synthesis; antibacterial target; crystal structure

Introduction

Disrupting the assembly of the bacterial cell wall peptidoglycan has been successfully exploited in developing many of the currently available antibiotics (Bugg *et al*., 1992). However, the emergence of drug resistant bacteria has underlined the need to explore new avenues for developing efficacious antibacterials (Walsh, 1993; Spratt, 1994). The monomeric precursor of the peptidoglycan, UDP-*N*-acetylmuramyl-pentapeptide, is synthesized in the cytoplasm or at the inner surface of the cytoplasmic membrane and the enzymes involved in its synthesis are attractive targets. Synthesis of UDP-*N*-acetylmuramyl-pentapeptide involves successive addition of L-Ala, D-Glu, *meso*-diaminopimelate (*m*-DAP) or L-Lys and D-Ala-D-Ala dipeptide to UDP-*N*-acetylmuramic acid (Rogers *et al*., 1980; van Heijenoort *et al*., 1994). These reactions are catalyzed by the ATP-dependent cytoplasmic enzymes MurC, MurD, MurE and MurF, respectively. Their high specificity, uniqueness, and occurrence only in eubacteria render them interesting therapeutic targets. The roles of MurC, MurD and MurE may be substituted by the *mpl* gene product that adds the tripeptide, L-Ala-D-Glu-*m*-DAP, onto the UDP-MurNAc in a single step in the recycling process of peptidoglycan (Goodell, 1985; Mengin-Lecreulx *et al*., 1996). However, MurF remains as the sole D-Ala-D-Ala adding enzyme.

The MurF protein from *Escherichia coli* consists of 452 amino acid residues and shares 10-20 % sequence identity with MurC, MurD and MurE. MurF catalyzes the addition of D-Ala-D-Ala to UDPMurNAc-tripeptide according to the following reaction:

$$\text{UDPMurNAc-L-Ala-}\gamma\text{-D-Glu-}m\text{-DAP} + \text{D-Ala-D-Ala} + \text{ATP} \rightleftharpoons$$

$$\text{UDPMurNAc-L-Ala-}\gamma\text{-D-Glu-}m\text{-DAP-D-Ala-D-Ala} + \text{ADP} + P_i$$

All four of these enzymes utilize ATP while incorporating peptides sequentially onto the C terminus

Abbreviations used: *m*-DAP, *meso*-diaminopimelate; MAD, multiwave length anomalous dispersion.

E-mail addresses of the corresponding authors: youwei_yan@merck.com; zhongguo_chen@merck.com

0022-2836/00/030435–11 $35.00/0

of the peptide chain in a non-ribosomal fashion and have been suggested to share a similar reaction mechanism. The C-terminal carboxylate of UDPMurNAc-peptide is phosphorylated by the γ-phosphate of ATP to form an acyl phosphate intermediate. This is followed by nucleophilic attack by the amide group of the incoming amino acid resulting in the extension of the UDPMurNAc-peptide precursor, releasing an ADP and inorganic phosphate (Falk *et al*., 1996; Tanner *et al*., 1996; Vaganay *et al*., 1996). Kinetic studies suggest that MurF binds substrates in an ordered manner, starting with ATP, followed by UDPMurNAc-L-Ala-γ-D-Glu-*m*-DAP and then D-Ala-D-Ala (Anderson *et al*., 1996). However, details of MurF-substrate interactions are unclear.

Recently, crystal structures of MurD protein bound to its substrates have become available (Bertrand *et al*., 1997, 1999). The MurD polypeptide chain comprises three domains with topologies reminiscent of nucleotide binding folds. These structures reveal the binding site of substrates and suggest a reaction mechanism that is consistent with the general mechanism proposed for these enzymes.

Towards our goal to provide a structural framework for the rational design of novel antibacterials, we have determined the crystal structure of the apo form of MurF protein from *E. coli* to 2.3 Å resolution. The MurF polypeptide folds into three open α/β-sheet domains. The N-terminal domain reveals a novel topology , in comparison with the MurD structures, while the central and the C-terminal domains remain similar to the classic nucleotide binding folds. It is interesting that these two domains of MurF are also homologous to the another available ligase structure, Folylglutamate synthetase (Sun *et al*. 1998). It is composed of only two domains. The three domains of MurF juxtapose in a crescent-like arrangement creating a wide-open solvent accessible space where substrates are expected to bind. Substrate diffusion experiments indicate that potential domain rearrangement may be induced upon ATP binding. The structure of the substrate-free MurF is suggestive of an open conformation.

Results and Discussion

Structure determination

The native and selenomethionyl MurF were expressed, purified and crystallized as reported (Anderson *et al*., 1996; Pryor *et al*., 1997; Yan *et al*., 1999). The crystals belong to space group $P6_1$ with the cell dimensions of $a = b = 74.2$ Å, $c = 429.3$ Å and $z = 2$. The structure was determined using the multiwavelength anomalous dispersion (MAD) method (Hendrickson *et al*., 1997). The statistics of the data collection and reduction as well as the structure refinement are shown in Table 1. The protein geometry analyzed by PROCHECK (Laskowski *et al*., 1993) shows 89 % of the residues in the most favored regions and none of the non-glycine residues in disallowed regions. One loop of residues 186 to 193 and five residues at the C terminus of both molecules in an asymmetric unit have no visible electron density and thus have not been included in the model. Figure 1 shows a portion of the experimental and refined maps superimposed with the refined MurF model.

Overall structure

There are two MurF molecules in an asymmetric unit in this hexagonal crystal form related by a local 2-fold symmetry. Each molecule consists of three open α/β-sheet domains formed from contiguous segments in the amino acid sequence as shown in Figure 2. The N-terminal domain comprises residues 1 to 81 and consists of a five stranded β-sheet (β2, β4-β7) surrounded by three helices, two in α-helix (α1 and α3) and one in 3_{10}-helix conformation (α2). In addition, there is a very short two stranded antiparallel β-sheet (β1, β3) perpendicular to the main sheet as shown in Figure 3(a). Data base search using program DALI(Holm *et al*., 1993) revealed no known protein structure homologs to this domain. This N-terminal domain of MurF shows an unique topology and does not contain the so called ''topological switch point'' commonly serving as the substrate binding site seen in known open α/β-sheet domains (Branden *et al*., 1991).

The central domain comprises residues 82 to 310 and consists of a six-stranded parallel central β-sheet (β8-β13) surrounded by eight α-helices (α4-α11). In addition, there is a small three stranded antiparallel β-sheet (β14-β16), as shown in Figure 3(b). The helices not only interact with the central β-sheet but also form two three-helix bundles at each side of it, composed of helices α4, α6 and α7 and helices α5, α10 and α11, respectively. The fold of the central domain is similar to the classic ''mononucleotide-binding fold'' as found in many ATP-binding proteins (Denessiouk *et al*., 1998; Denessiouk & Johnson, 2000).

The C-terminal domain comprises residues 311- to 447 and consists of a six stranded β-sheet (β17-β22) surrounded by five α-helices (α12-α16). As illustrated in Figure 3(c), this domain contains the dinucleotide-binding fold, also known as the Rossmann fold (Rossmann *et al*., 1975; Schulz, 1992) as seen in dehydrogenases, although the signature nucleotide binding sequence is not highly conserved in comparison with the classic dinucleotide-binding domain.

The three domains of the unliganded MurF polypeptide juxtapose in a crescent-like arrangement (Figure 2). Since MurF in solution functions as a monomer (Anderson *et al*., 1996), the observed dimer in this crystal form is hence merely a result of crystal packing. The substrate binding sites are likely at the concave site opposite to the dimerization face. In the quaternary arrangement of the monomer, the N and C-terminal domains are

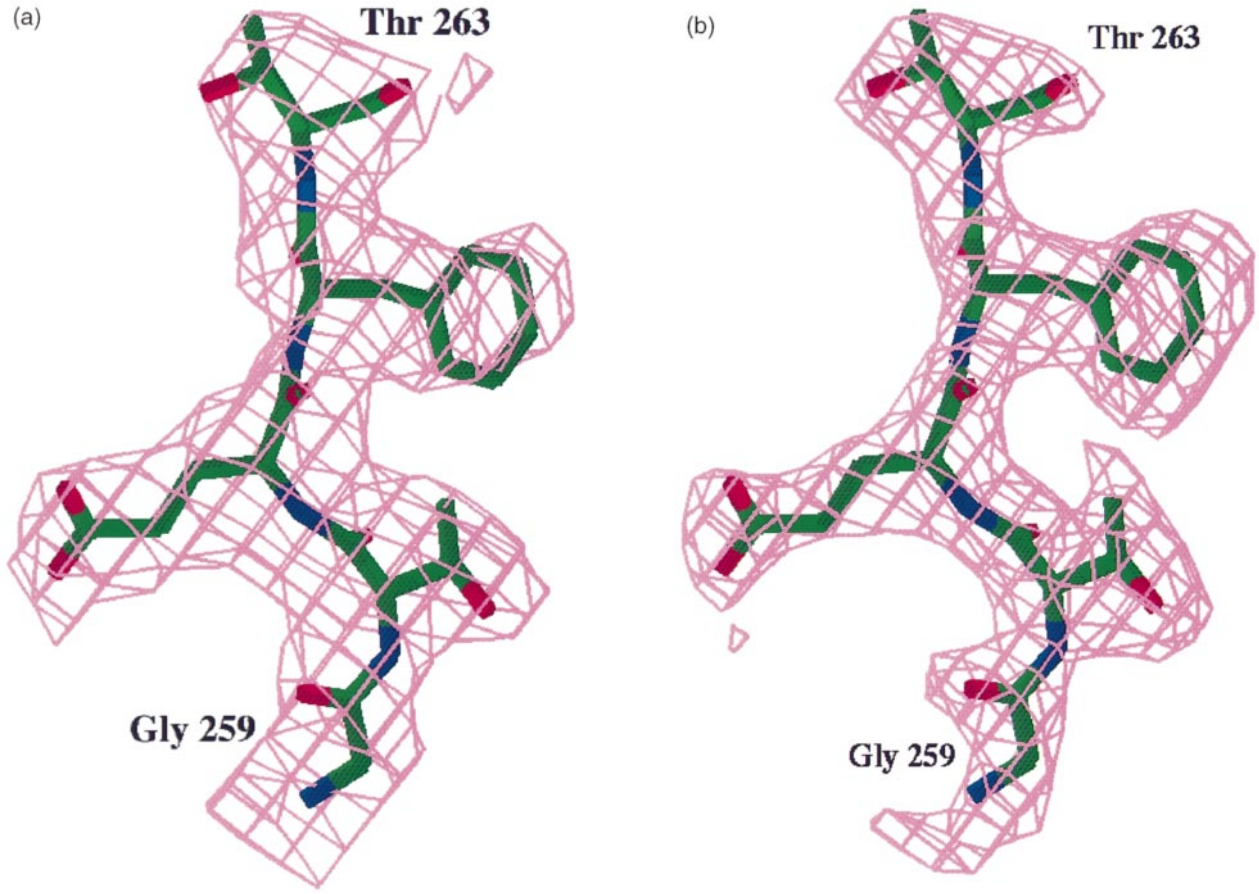

Figure 1. Electron density map for residues 259 to 263 in the central domain. (a) Experimental density after solvent flattening and non-crystallographic averaging at 2.8 Å resolution; and (b) final refined electron density at 2.3 Å resolution. Final refined model is depicted. Both maps were contoured at 1σ. Figure was prepared with program O.

sufficiently distant from each other, creating a wide-open solvent accessible surface; as such it is unfavorable for substrate binding. Based on this apo enzyme structure, a significant domain rearrangement is likely upon substrate binding, bringing the domains into a close proximity of each other. This hypothesis is supported by the substrate diffusion data, by kinetic evidence and by comparison with the ligand-bound forms of MurD (Bertrand *et al.*, 1997, 1999), as discussed below.

Substrate-binding sites and a bound ligand

In an attempt to obtain the substrate bound information of MurF, we soaked the apo MurF crystals in the cryo-protectant solution containing substrates (see Materials and Methods). It was not possible to obtain information regarding the ATP-binding site by the soaking method. When ATP was diffused into crystals of apo MurF, the crystals were immediately destroyed. Since ATP is the first substrate to bind in the ordered binding mechanism, this finding is consistent with a substrate-induced conformation change of the enzyme for an ordered binding mechanism to prepare the enzyme for the productive binding of the second and third substrates. Nonetheless, the secondary structure of MurF indicates its central domain has the characteristic mononucleotide-binding fold. It is likely that the conspicuous sequence pattern located within a large loop between β8 and α5 of this domain is involved in ATP binding. This loop has the most conserved sequence among 20 Mur synthetases from various bacterial organisms (Bouhss *et al.*, 1997) and is similar to the ATP binding regions of known mononucleotide-binding proteins (Schulz *et al.*, 1974; Denessieuk & Johnson, 2000).

However, using crystals soaked with UDPMurNAc-tripeptide and D-ala-D-ala dipeptide substrates, we collected a data set to 3.0 Å resolution. The statistics of the diffraction data are shown in Table 1. As shown in Figure 4, the difference Fourier map phased with the apo protein reveals positive density extending from the N-terminal domain to the central domain. The uracyl ring could be modeled in close vicinity of a protruding loop β5-α3 within the N-terminal domain; in this model, the phosphate is fixed by the β9-α6 loop in the central domain. But the detailed interaction between the substrate and protein is not clear due to the low resolution and the poor quality of the map. A lack of density for the remaining part of

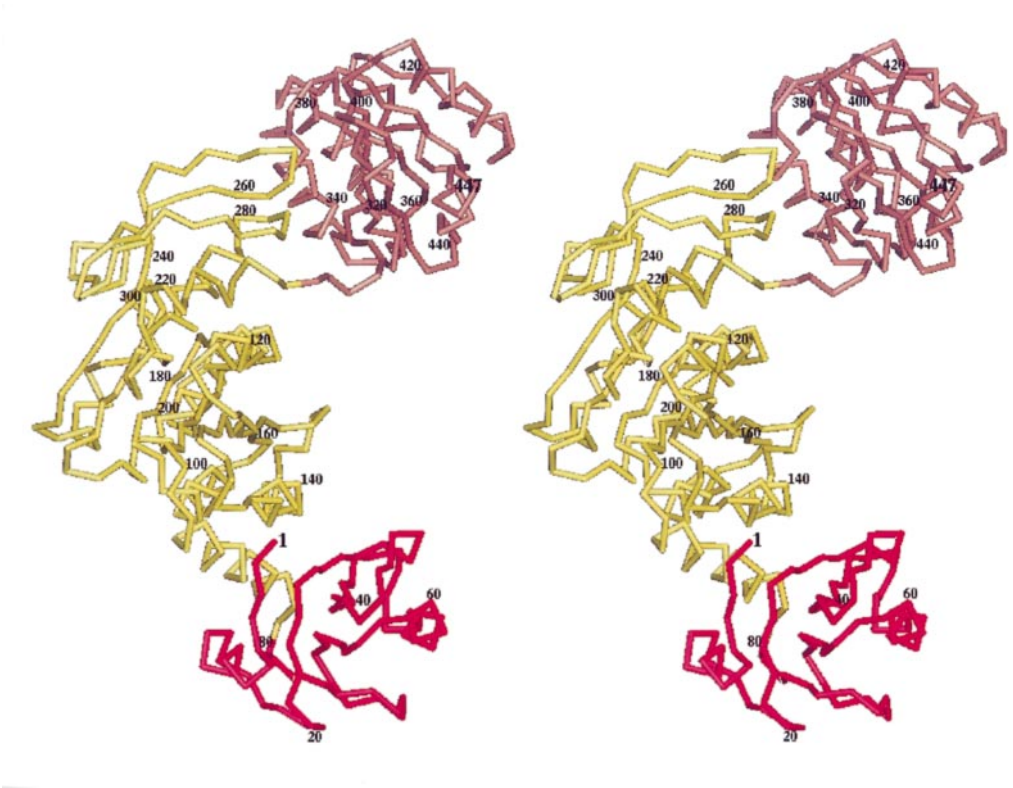

Figure 2. C^{α} atom tracing of MurF. Every 20 residues are labeled. The MurF protein is composed of three open α/β-sheet domains with the N-terminal domain in red, the central domain in yellow and the C-terminal domain in brown. Figure was prepared with program QUANTA (MSI, San Diego).

UDPMurNAc-tripeptide and D-Ala- D-Ala dipeptide also indicates that an optimized substrate binding environment is unavailable in the absence of the first substrate ATP and its induced protein conformational changes. The partial inhibitor density was only seen in one molecule in the asymmetric unit. The binding site of the another molecule is blocked by symmetry related atoms.

Structural comparison with MurD, substrate binding and domain rearrangement

The central and C-terminal domains of MurF show the highest structural similarity to the corresponding domains in MurD (Protein Data Bank code 1UGA) compared using program DALI. The RMS deviations are 2.4 Å for either 181 C^{α} atoms in the central domains or 126 C^{α} atoms in the C-terminal domains of MurF and MurD. Even though the sequence identity of the central and the C-terminal domains between the two proteins are only 22 % and 13 %, respectively. Figure 5 shows the sequence alignment between MurF and MurD based on the tertiary structure comparison and by visual inspection. All seven most conserved residues at the active site in MurD orthologs (Bouhss *et al.*, 1999) are also conserved in MurF. As mentioned before, the N-terminal domains have no structural similarity, and hence there is no alignment. It is worth to note, another member of the ligase superfamily with known structure, folylpolyglutamate synthetase from Lactobacillus casei (Protein Data Bank code 1FGS), consisted of only two domains, are homologous to the central and the C-terminal domains of MurF with second highest scores, respectively.

The domain-domain interactions in MurF are completely different from those seen for MurD. The substrate free MurF looks like an open form of the substrate bound MurD, as shown in Figure 6. While superpositioning these two proteins by their homologous central domains, the center of the N-terminal domain of MurF moves outward by 10 Å regarding to that of MurD, but the residues (residues 43 to 45 in MurF and 34 to 36 in MurD) that are involved in the uracyl ring fixation in both proteins remain in close proximity to each other (the mean shift is 4.0 Å), despite the fact that the orientation and connectivity of the secondary structural elements are different. Asp45 of MurF may play similar role as that of Asp35 which is one of the most conserved residues at the active site in MurD orthologs. At the other end, the C-terminal domain of MurF, although similar in fold to the corresponding domain in MurD, swings away from the central domain by about 47° and interacts with the central domain differently from that in MurD. The β9-α8 loop (residues 176 to 190) of the central domain in MurD is well ordered and interacts with the C-terminal domain. Whereas the corresponding loop in MurF, loop β11-α8 of residues

Table 1. Data Collection and refinement statistics

Data sets[a]				
MAD data at 3.5 Å	Front (1.000 Å)	Edge (0.9796 Å)	Peak (0.9794 Å)	Remote (0.9611 Å)
Total observation	138,300	142,967	142,964	146,623
Unique reflections	20,285	20,228	20,260	20,245
Completeness (%)	99 (88)[b]	99 (95)	99 (94)	100 (100)
$R_{sym\ (\%)}$	5.1 (18.9)	7.0 (18.6)	6.3 (18.2)	6.8 (15.7)
MAD data at 2.8 Å				
Total observation	401,175	206,573	412,036	
Unique reflections	31,628	31,294	31,863	
Completeness (%)	96 (77)	95 (71)	97 (83)	
$R_{sym\ (\%)}$	5.1 (22.0)	7.0 (19.7)	6.8 (20.4)	
Native data at 2.3 Å[c]				
Total observation	545,211			
Unique reflections	47,573			
Completeness (%)	82 (54)			
$R^{d}_{sym\ (\%)}$	6.5 (33.8)			
Substrate bound data at 3.0 Å[e]				
Total observation	249,583			
Unique reflections	22,978			
Completeness (%)	85 (20)			
$R_{sym\ (\%)}$	7.8 (28.1)			

Overall figure of merit before density modification at 3.5 Å resolution: 0.82
Z-value: 47.6; the standard deviation of the local rms. 0.274; the correlation of rms density in neighboring boxes: 0.401.
Overall figure of merit before density modification at 2.8 Å resolution: 0.61
Reflections used in refinement at 2.3 Å resolution: 46,843
Number of non-hydrogen protein atoms: 6484
Number of solvent molecules: 307
R_{work} [f] = 0.203
R_{free} = 0.282
RMSD from ideal
Bond lengths: 0.010 Å
Angles: 1.86°

[a] All data were collected from selenomethionyl MurF crystals.
[b] Values within parentheses are for the highest resolution shell.
[c] Data collected from a selenomethionyl MurF protein crystal and used as a native in structure refinement.
[d] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I = observed intensity, $\langle I\rangle$ = average intensity of symmetry relation reflections.
[e] Selenomethionyl MurF crystal soaked with the substrates of UDP-MurNAc-L-Ala-γ-D-Glu-*m*-DAP and D-Ala-D-Ala
[f] R_{work} and $R_{free} = \Sigma\ ||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are observed and calculated amplitudes, respectively. R_{free} was calculated using 10% of data excluded from refinement.

186 to 193 is disordered and is unlikely to be involved in domain-domain interactions. The interface in MurF moves to the flanking three-stranded β-sheet (β14-β16) area. The loop of β14-β15 becomes fixed in MurF whereas its counterpart in MurD is disordered. The residues 223 to 230 in this area form an α-helix (α9) but the corresponding residues in MurD form a random coil along with four missing residues. It has been noticed that the Lys198 of MurD is carbamylated (Bertrand *et al*., 1999). There is no convincing extra electron density at the side chain of the corresponding residue Lys202 of MurF. Therefore, it has been modeled as a regular lysine residue in this substrate free MurF structure. In the substrate bound MurD structure, the N and C-terminal domains bend toward each other, create a crevice where the substrates bind. In contrast, the N and C-terminal domains in the substrate free MurF structure are opened up, leaving a much bigger solvent accessible area (Figure 6). The tip to tip distance between the N-terminal and C-terminal domains in MurD is 15 Å but is 34 Å in MurF. The domain rotation pivot seems to be the peptide bond between the residues 310 and 311 in the loop connecting the central and C-terminal domains in MurF. It is conceivable that the residues around 311 at the connecting loop represents a conformationally flexible region of the protein, permitting the possible domain rearrangement required for the enzymatic reaction.

The MurD-substrate complex structures (Bertrand *et al*., 1999) reveal that the ADP is sandwiched between the central and the C-terminal domains. The corresponding domains of MurF are closely homologous to them. In the MurF protein, the β8-α5 moiety consists of residues 104 to 112 with the sequence of Ala-Leu-Thr-Gly-Ser-Ser-Gly-Lys-Thr, that is conserved with the ATP-binding P-loop in MurD. Three more residues, Asn271, Arg302 and Asp317, that are in direct contact with ADP in MurD, are also strictly conserved both in primary and secondary structural context in MurF (see Figure 5). They are Asn285 in the central domain, Arg316 and Asp331 in the C-terminal domain. All these similarities strongly suggest that the ATP-binding site may be conserved between these two enzymes. However the C-terminal domain in the unliganded MurF is significantly tilted away from the putative nucleotide-binding site. Given the fact that these two domains are chemically connected by a single stretch of amino

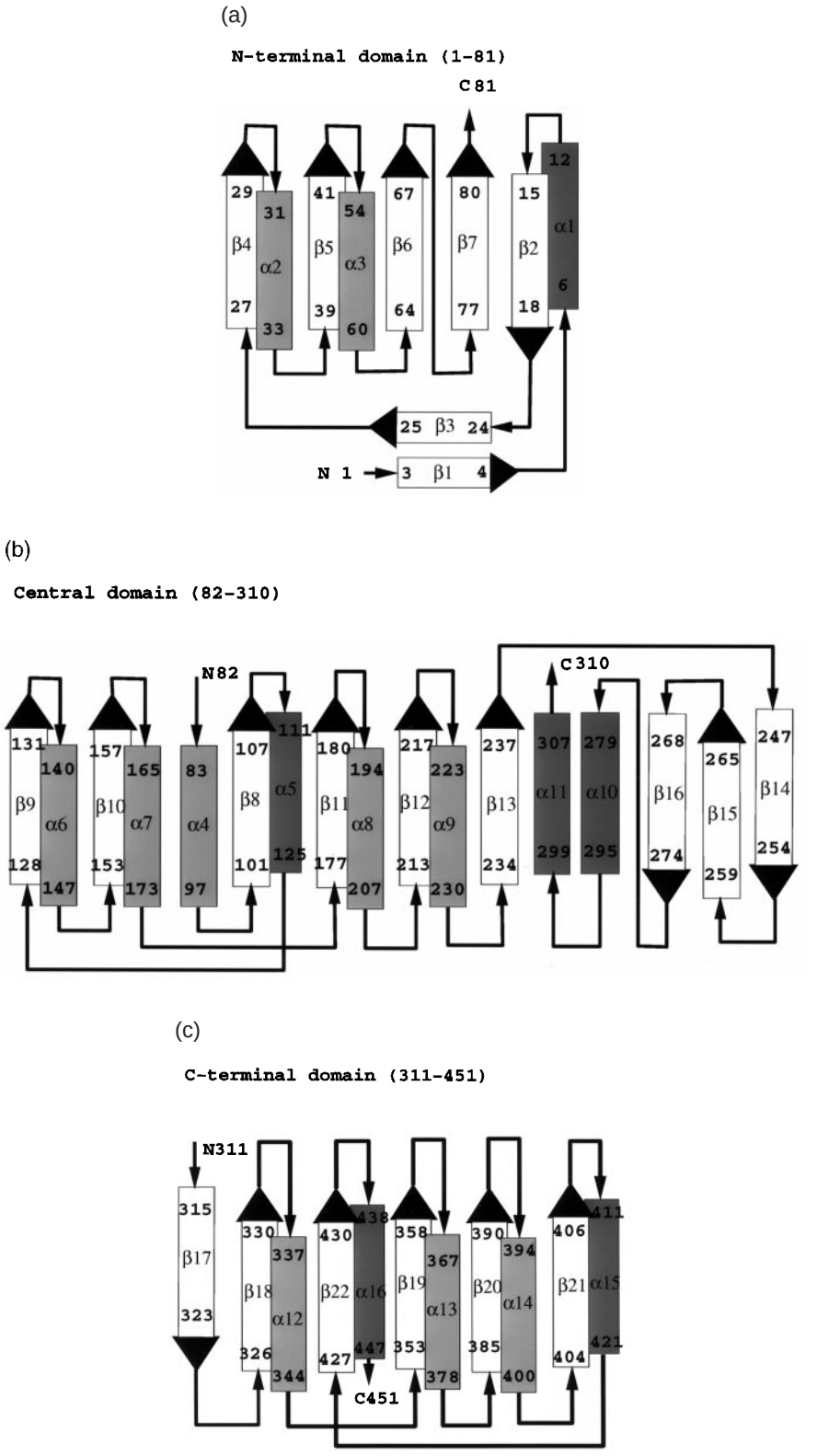

Figure 3. Schematic diagram of the polypeptide topology of the three domains of MurF, (a), (b), and (c) for the N-terminal, central and C-terminal domains, respectively. The β-strands are depicted as arrows, with the arrowheads indicating the direction of the chain. α-helices above and below the β-sheet are represented by light and dark shaded rectangles, respectively.

acid residues and have a mere 500 Å^2 inter-domain contact area, domain rearrangement would be possible for ATP binding. This hypothesis is consistent with the enzyme kinetic study that suggests MurF binds to its substrates in a specific order with ATP binding first (Anderson *et al.*,

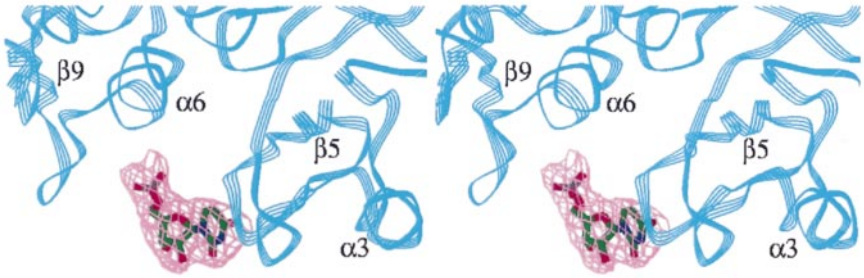

Figure 4. Stereo view of MurF in complex with its partial substrate. The difference Fourier map, in pink, was calculated with the amplitudes from the MurF crystals soaked with two substrates, UDPMurNAc-tripeptide and D-Ala-D-Ala dipeptide and the phases from the apo MurF model. The map was contoured at 2.5σ and superimposed with the manually built partial substrate, the uridine-ribose moiety and one phosphate. The protein moieties, β5-α3 loop in the N-terminal domain and β9-α6 loop in the central domain, of MurF at close vicinity of the substrate are shown as ribbon. Figure was prepared with program O.

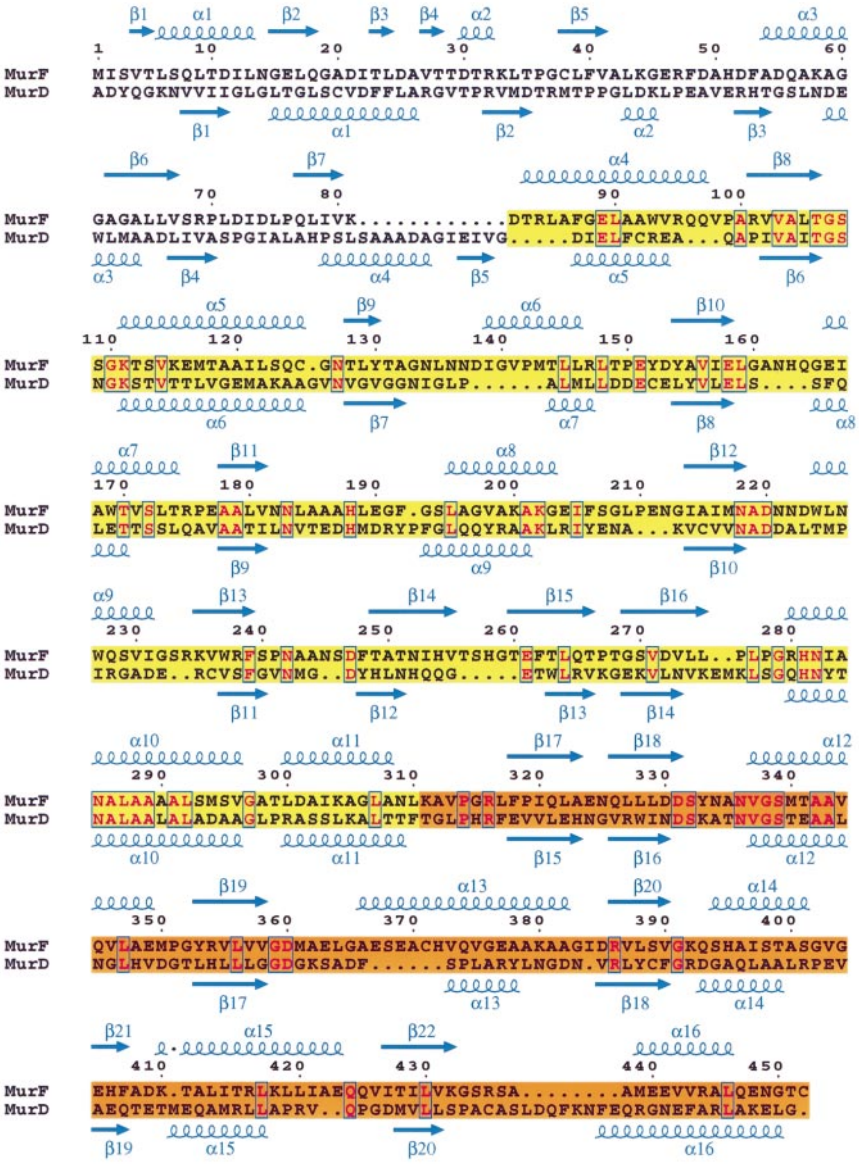

Figure 5. A structure-based sequence alignment of MurF and MurD from *E. coli*. The N-terminal domains are not aligned. The cental and C-terminal domains in yellow and brown background, respectively, were aligned. The locations of secondary structure elements in MurF and MurD are indicated over and below the aligned sequences, respectively. The identical residues are boxed in red. This Figure was prepared using ESPrint (Gouet *et al.*, 1999).

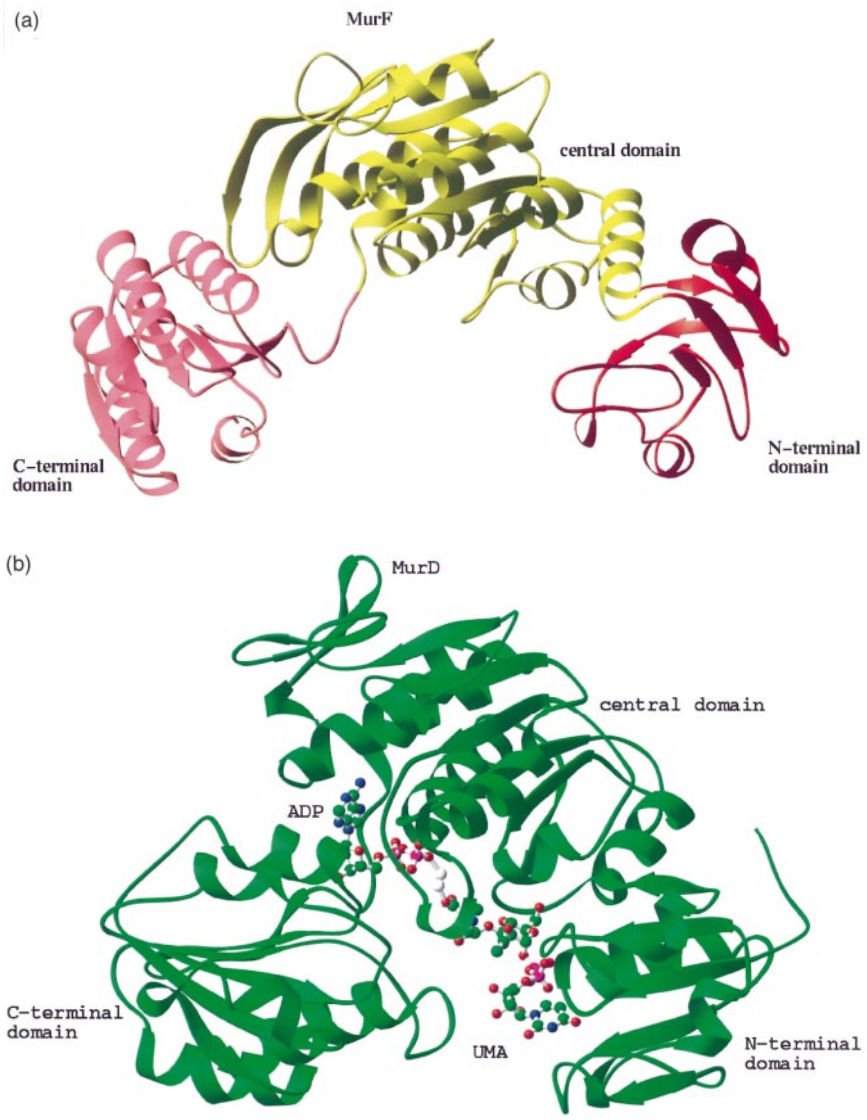

Figure 6 (*legend opposite*)

1996). A similar sequential binding mode is also observed for the MurC protein (Bouhss *et al.*, 1997). In fact, domain rearrangement is commonly seen in many protein kinases (Bossemeyer *et al.*, 1993; Goldsmith *et al.*, 1994). Consistent with this hypothesis is also the fact that soaking of MurF crystals in ATP containing solution destroys the crystals, perhaps due to ensuing structural change upon ATP incorporation.

Conclusion

The MurF protein consists of three open α/β-sheet domains. Two of them are very similar to their correspondents in MurD, which are responsible for ATP binding, the only common substrate between the two enzymes. The topology of the N-terminal domain in MurF is different from that in MurD. However, the uracyl ring of UDP still binds to this domain at a similar position. The N-terminal domain in MurF does not have the characteristic nucleotide-binding fold and is not involved in the fixation of the pyrophosphate group. Instead, part of the central domain, the β9-α6 loop, extends out towards the N-terminal domain and is involved in pyrophosphate binding. The substrate-free MurF structure reveals that its substrate-binding site is highly solvent exposed. Significant domain rearrangement is required for forming a functional enzyme-substrate complex. Binding of ATP, the first substrate to bind, may be responsible for inducing or stabilizing the proper inter-domain interface.

Materials and Methods.

Purification and Crystallization

The selenomethionyl MurF protein was expressed and purified in a similar way as of the native MurF protein and was concentrated to 13.0 mg/ml in

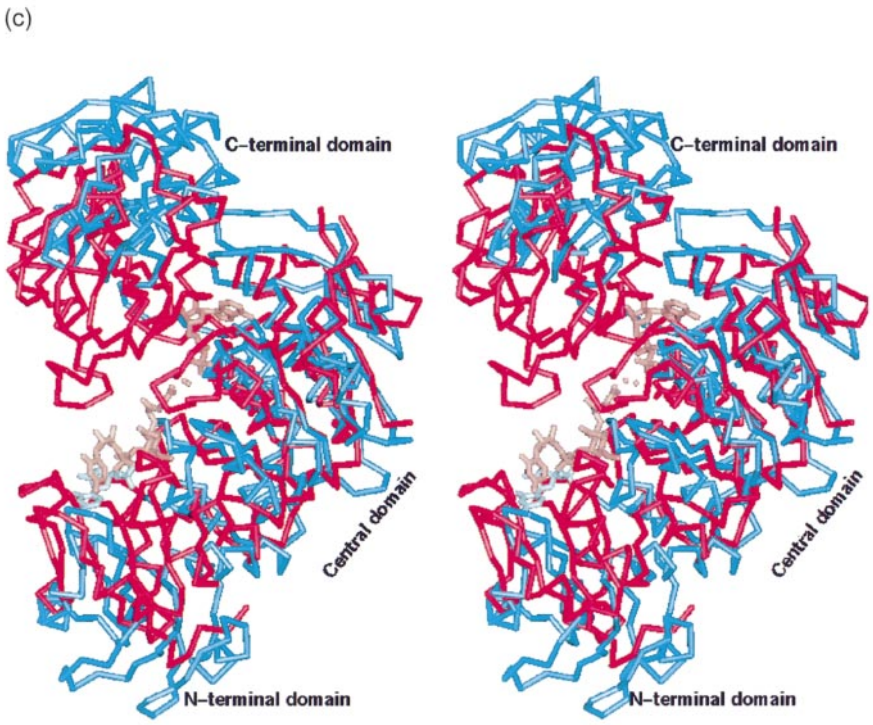

Figure 6. The structural comparison between MurF and MurD. (a) Ribbon diagram of the substrate free MurF, and (b) the substrate bound MurD, protein in ribbon and substrates in ball and stick. (c) MurF in blue and its partial substrate in light blue was superpositioned with MurD in red and its substrates in brown-based on the conserved central domains. Figures were prepared with program RIBBON (Carson, 1991) for (a) and (b) and QUANTA for (c). The coordinates utilized for generating the MurD figures were retrieved from the Protein Data Bank under accession code 2UAG.

10 mM Tris-HCl buffer at pH 7.4 in the presence of 10 mM DTT. The best crystals were obtained when the well solution contained 0.1 M Bis-Tris propane buffer at pH 9.4, 16% (w/v) PEG 8 K, 0.12 M $MgSO_4$ and 10% (w/v) glycerol at 4°C. The hanging-droplets contained an equal volume of the protein solution and the reservoir solution of 5 µl each and micro-seeds from the previous crystallization trays. The crystals would reach the maximum size, 0.2 mm × 0.2 mm × 0.4 mm, in about three weeks. The substrate bound crystals were obtained by soaking the selenomethionyl MurF crystals in the cryo-protectant solution in the presence of 10 mM UDPMurNac-tripeptide and D-Ala-D-Ala dipeptide (SIGMA) for one day.

Data Collection

MAD data sets at 3.5 Å and 2.8 Å and the high resolution data set at 2.3 Å were collected at 100 K using a single selenomethionyl MurF crystal for each data set. Prior to data collection, the crystals were soaked in the mother liquor mixture containing a cryo-protective agent, 20% (v/v) glycerol for one to two hours at 4°C and flash-frozen in liquid nitrogen. The data sets were collected at synchrotron beamline 17-ID of the Advanced Photon Source† at Argonne National Laboratory equipped with the MAR CCD detector (MAR Research, Hamburg) at the crystal to detector distance of 280 mm with 1° oscillation per frame. In order to record the high resolution data at 2.8 Å and 2.3 Å, the detector center was shifted away from the x-ray beam center by withdrawing one and two supporting blocks, respectively. The C axis of the crystals was always manually oriented roughly parallel to the camera oscillation spindle axis in order to reduce the overlap of the diffraction spots due to the long cell dimension. The diffraction intensities were collected using the inverse beam technique over a 120° rotation in ω, and were indexed, processed and scaled with program HKL (Otwinowski *et al.*, 1997). The data set from the substrate soaked crystals was recorded to 3.0 Å resolution. (Data collection and statistics seen in Table 1).

Phase determination and model building

There are nine selenomethionine residues per molecule and 18 in an asymmetric unit. The selenomethionine residue at the N terminus was considered most likely to be flexible in the crystal, only 16 selenium atoms were aimed to search and all of them were

† The facilities of the Industrial Macromolecular Crystallography Association Collaborative Access Team (IMCA-CAT) at the Advanced Photon source. These facilities are supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with Illinois Institute of Technology (IIT), executed through the IIT's Center for Synchrotron Radiation Research and Instrumentation.

found with the use of program SOLVE (Terwilliger *et al*., 1999) based on the 3.5 Å resolution MAD data. The selenium positions and *B*-factors were refined and initial phases at the low resolution were derived with SOLVE program package. The phases were further refined at 2.8 Å resolution with MLPHARE in CCP4 program package (CCP4, 1994). The experimental map quality was improved by the solvent flattening and electron density modification procedure performed with program DM (Cowtan, 1994; CCP4, 1994). As the protein boundary, the domain structure and the major secondary structure elements became clearer, the local two fold axis and more accurate molecular masks were obtained with program O (Jones *et al*., 1991). The map quality was further improved by the local two-fold averaging combined with the solvent flattening procedures (figure of merit = 0.74). The complete poly-peptide chains were traced unambiguously in the solvent flattened and two-fold averaged map at 2.8 Å resolution using the selenium sites as reference markers, with program O. The initial model gave an *R*-factor of 0.44. For the substrate bound structure, the phases were derived from the refined substrate free MurF model.

Structure Refinement

The structure refinement was performed using simulated annealing and energy minimization as well as the *B*-factor refinement as implemented in X-PLOR (Brünger, 1992a), at 2.8 Å resolution with tight non-crystallographic symmetry (NCS) constrains. The phases were extended to 2.3 Å resolution in two steps coupled with manual adjustments of the model. Ten per cent of the data were set aside as a reference for the free *R*-factor calculation (Brünger, 1992b). The model was further refined with NCS restrain. A total of 307 water molecules were included in the model at the later refinement stage. The NCS restrain was released at the last cycle of the refinement. The final model gives a R-factor of 0.20 and free *R*-factor of 0.28 for all data between 8.0 to 2.3 Å resolution(Table 1).

Protein Data Bank accession codes

The coordinates of the MurF structure have been deposited with the RCSB Protein Data Bank under the accession code 1GG4.

Acknowledgments

We thank Dr Lawrence C. Kuo for prompting this project, for his encouragement throughout this work, and for his critical reading of the manuscript. We acknowledge K.A.D. Pryor and F. Marsilio for the protein expression and purification. We would also like to thank Dr Thomas Terwilliger at Los Alamos National Laboratory for helping in locating the positions of the selenium atoms at 3.5 Å resolution using the program SOLVE. The use of the Advanced Photon Source is supported by the US Department of Energy, Basic Energy Sciences, Office of Energy Research, under the Contract No. W-31-109-Eng-38. We would like to thank the staff at beam line 17-ID and MAR-USA for their help during the data collection.

References

Anderson, M. S., Eveland, S. S., Onishi, H. R. & Pompliano, D. L. (1996). Kinetic mechanism of the *Escherichia coli* UDPMurNAc-tripeptide D-alanyl-D-alanine-adding enzyme: use of a glutathione S-transferase fusion. *Biochemistry,* **35**, 16264-16269.

Bertrand, J. A., Auger, G., Fanchon, E., Martin, L., Blanot, D., van Heijenoort, J. & Dideberg, O. (1997). Crystal structure of UDP-*N*-acetylmuramoyl-L-alanine: D-glutamate ligase from *Escherichia coli*. *EMBO J.* **16**, 3416-3425.

Bertrand, J. A., Auger, G., Martin, L., Fanchon, E., Blanot, D. & Le, Beller, *et al*. (1999). Determination of the MurD mechanism through crystallographix analysis of enzyme complexes. *J. Mol. Biol.* **289**, 579-590.

Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H. & Huber, R. (1993). Phosphotransferase and substrate binding mechanism of the cAMP-dependent proteinkinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with Mn+ adenylyl imidodiphosphate and inhibitor peptide PKI(5-24). *EMBO J.* **12**, 849-859.

Bouhss, A., Mengin-Lecreulx, D., Blanot, D., van Heijenoort, J. & Parquet, C. (1997). Invariant amino acids in the Mur peptide synthetases of bacterial peptidoglycan synthesis and their modification by site-directed mutagenesis in the UDP-MurNAc: L-alanine ligase from *E. coli*. *Biochemistry,* **36**, 11556-11563.

Bouhss, A., Dementin, D., Parquet, C., Mengin-Lecreulx, D., Bertrand, J. A. & Le, Beller D., *et al*. (1999). Role of the ortholog and paralog amino acid invariants in the active site of the UDP-MurNAc-L-alanine:D-glutamate ligase (MurD). *Biochemistry,* **38**, 12240-12247.

Branden, C. & Tooze, J. (1991). *Introduction to Protein Structure*, Garland Publishing, Inc., New York.

Brünger, A. T. (1992a). *X-PLOR, Version 3.1, A System for Crystallography and NMR*, Yale University Press, New Haven.

Brünger, A. T. (1992b). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structure. *Nature,* **355**, 472-475.

Bugg, T. D. H. & Walsh, C. T. (1992). Intracellular steps of bacterial cell wall peptidoglycan biosyntheses: enzymology, antibiotics, and antibiotic resistance. *Nature Prod. Rep.* **9**, 199-215.

Carson, M. (1991). Ribbons 2.0. *J. Appl. Crystallog.* **24**, 958-961.

Collaborative Computational Project No. 4 (CCP4) (1994). *Acta. Crystallog. sect D*, **50**, 760-763.

Cowtan, K. (1994). An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM News. Protein Crystallog.* **31**, 34-38.

Denessiouk, K. A. & Johnson, M. S. (2000). When fold is not important: a common structural framework for adenine and AMP binding in 12 unrelated protein families. *Proteins: Struct. Funct. Genet.* **38**, 310-326.

Denessiouk, K. A., Lehtonen, J. V. & Johnson, M. S. (1998). Enzyme-mononucleotide interactions: three different folds share common structural elements for ATP recognition. *Protein Sci.* **7**, 1768-1771.

Falk, P. J., Ervin, K. M., Volk, K. S. & Ho, H. T. (1996). Biochemical Evidence for the Formation of a covalent acyl-phosphate linkage between UDP-*N*-acetylmuramate and ATP in the *E. coli* UDP-*N*-acet-

ylmuramate: L-alanine ligase - catalyzed reaction. *Biochemistry*, **35**, 1417-1422.
Goldsmith, E. J. & Cobb, M. H. (1994). Protein kinases. *Curr. Opin. Struct. Biol.* **4**, 833-840.
Goodell, E. W. (1985). Recycling of Murein by *Escherichia coli*. *J. Bacteriology*, **163**, 305-310.
Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). ESPrint: multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305-308.
Hendrickson, W. A. & Ogata, C. M. (1997). Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol*, **276**, 494-523.
Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. D*, **47**, 110-119.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thomton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.
Mengin-Lecreulx, D., van Heijenoort, J. & Park, J. T. (1996). Identification of the *mpl* Ggene encoding UDP-*N*-acetylmuramate: L-alanyl-γ-D-glutamyl-meso-diaminopimelate ligase in *Escherichia coli* and its role in recycling of cell wall peptidoglycan. *J. Bacteriol.* **178**, 5347-5352.
Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
Pryor, K. A. D. & Leiting, B. (1997). High-level expression of soluble protein in *Escherichia coli* using a His_6-Tag and maltose-binding-protein double-affinity fusion system. *Protein Expr. Purif.* **10**, 309-319.
Rogers, H. T., Perkins, H. R. & Ward, J. B. (1980). *Microbial Cell wWalls and Membranes*, Chapman & Hall Ltd., London.
Rossmann, M. G., Liljas, A., Branden, C. I. & Banaszak, L. J. (1975). Evolutionary and structural relationships among dehydrogenases. In *The Enzymes* (Boyer, P. D., ed.), pp. 61-102, Academic Press, New York..
Schulz, G. E. (1992). Binding of nucleotides by proteins. *Curr. Opin. Struct. Biol.* **2**, 61-67.
Schulz, G. E., Eizinga, M., Mark, F. & Schirmer, R. H. (1974). Three-dimensional structure of adenylate kinase. *Nature*, **250**, 120-123.
Spratt, B. G. (1994). Resistance to antibiotics mediated by target alterations. *Science*, **264**, 388-393.
Sun, X., Bognar, A. L., Baker, E. N. & Smith, C. A. (1998). Structural homologies with ATP- and folate-binding enzymes in the crystal structure of folylpolyglutamate synthetase. *Proc. Natl Acad. Sci. USA*, **95**, 6647-6652.
Tanner, M. E., Vagnanay, S., van Heijenoort, J. & Blanot, D. (1996). Phosphinate inhibitors of the D-glutamic acid-adding enzyme of peptidoglycan biosynthesis. *J. Org. Chem.* **61**, 1756-1760.
Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849-861.
Vaganay, S., Tanner, M. E., van Heijenoort, J. & Blanot, D. (1996). Study of the reaction mechanism of the D-glutamic acid-adding enzyme from *Escherichia coli*. *Microbial Drug Res.* **2**, 51-54.
van Heijenoort, J. (1994). Biosynthesis of the bacterial peptidoglycan unit. In *Bacterial Cell Wall* (Ghuysen, J. M. & Hakenbeck, R., eds), pp. 39-54, Elsevier Science B. V., Amsterdam.
Walsh, C. T. (1993). Vancomycin resistance: decoding the molecular logic. *Science*, **261**, 308-309.
Yan, Y., Munshi, S., Li, Y., Pryor, K. A. D., Marsilio, F. & Leiting, B. (1999). Crystallization and preliminary X-ray analysis of the Echerichia coli UDPMurNAc-tripeptide D-alanyl-D-alanine-adding enzyme MurF. *Acta Crystallog. sect. D*, **55**, 2033-2034.

Edited by I. Wilson

(Received 20 July 2000; received in revised form 6 October 2000; accepted 8 October 2000)