Three-Dimensional Structure of Escherichia coli Asparagine Synthetase B: A Short Journey from Substrate to Product

ARTICLE *in* BIOCHEMISTRY · JUNE 2000

Impact Factor: 3.02 · DOI: 10.1021/bi9915768 · Source: PubMed

7 AUTHORS, INCLUDING:

Susan K Boehlein
University of Florida
38 PUBLICATIONS **706** CITATIONS
SEE PROFILE

James B Thoden
University of Wisconsin–Madison
131 PUBLICATIONS **5,159** CITATIONS
SEE PROFILE

Ivan Rayment
University of Wisconsin–Madison
233 PUBLICATIONS **14,681** CITATIONS
SEE PROFILE

Three-Dimensional Structure of *Escherichia coli* Asparagine Synthetase B: A Short Journey from Substrate to Product[†,‡]

Todd M. Larsen,[§] Susan K. Boehlein,[||] Sheldon M. Schuster,*[,||] Nigel G. J. Richards,*[,||] James B. Thoden,[§] Hazel M. Holden,[§] and Ivan Rayment*[,§]

Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Gainesville, Florida 32611, and Department of Chemistry and Biotechnology Program, University of Florida, Gainesville, Florida 32611

Received July 8, 1999; Revised Manuscript Received August 23, 1999

ABSTRACT: Asparagine synthetase B catalyzes the assembly of asparagine from aspartate, Mg^{2+}ATP, and glutamine. Here, we describe the three-dimensional structure of the enzyme from *Escherichia coli* determined and refined to 2.0 Å resolution. Protein employed for this study was that of a site-directed mutant protein, Cys1Ala. Large crystals were grown in the presence of both glutamine and AMP. Each subunit of the dimeric protein folds into two distinct domains. The N-terminal region contains two layers of antiparallel β-sheet with each layer containing six strands. Wedged between these layers of sheet is the active site responsible for the hydrolysis of glutamine. Key side chains employed for positioning the glutamine substrate within the binding pocket include Arg 49, Asn 74, Glu 76, and Asp 98. The C-terminal domain, responsible for the binding of both Mg^{2+}ATP and aspartate, is dominated by a five-stranded parallel β-sheet flanked on either side by α-helices. The AMP moiety is anchored to the protein via hydrogen bonds with O^{γ} of Ser 346 and the backbone carbonyl and amide groups of Val 272, Leu 232, and Gly 347. As observed for other amidotransferases, the two active sites are connected by a tunnel lined primarily with backbone atoms and hydrophobic and nonpolar amino acid residues. Strikingly, the three-dimensional architecture of the N-terminal domain of asparagine synthetase B is similar to that observed for glutamine phosphoribosylpyrophosphate amidotransferase while the molecular motif of the C-domain is reminiscent to that observed for GMP synthetase.

Asparagine synthetase, isolated from both prokaryotes and eukaryotes, catalyzes the ATP-dependent conversion of aspartic acid to asparagine, employing either glutamine or ammonia as the nitrogen source (*1*). In *Escherichia coli*, there are two unlinked genes, *asnA* and *asnB*, that code for mechanistically distinct enzymes *(2*−*4*). Asparagine synthetase A is strictly ammonia-dependent, whereas the preferred source of nitrogen for asparagine synthetase B is glutamine. According to the presently available experimental data, the overall transformation of aspartic acid to asparagine occurs via three separate reactions as indicated in Scheme 1: the activation of aspartate by the formation of a β-aspartyl-AMP intermediate, the hydrolysis of glutamine to glutamate and ammonia, and the final breakdown of the β-aspartyl-AMP intermediate via a nucleophilic attack by the ammonia (*5*, and references therein).

On the basis of amino acid sequence analyses, asparagine synthetase B has been shown to belong to a larger family of enzymes referred to as the glutamine amidotransferases and elegantly reviewed in refs *6* and *7*. In all of these enzymes, glutamine serves as the preferred source of nitrogen whereby an acceptor molecule is subsequently aminated. Two different categories of glutamine amidotransferases have been identified thus far, namely the Triad (formerly referred to as *TrpG* or class I) and the Ntn (formerly referred to as the *PurF* or class II) types. The Triad enzymes, such as GMP synthetase and carbamoyl phosphate synthetase, employ a histidine−cysteine couple for catalytic activity (*8*−*10*). In the Ntn enzymes, of which asparagine synthetase is a member, the active-site cysteine is invariably located at the N-terminus. Proteins belonging to the Ntn family do not contain a catalytic triad *(6, 11)*.

The glutamine amidotransferases are of significant research interest, in part, because their biochemical transformations occur via two or more distinct reactions. Strikingly, these reactions have been shown to take place at separate and quite spatially distinct active sites. Indeed, in glutamine phosphoribosylpyrophosphate amidotransferase (GPATase),[11] an Ntn class enzyme, the two active sites are separated by 20 Å (*12*), while in carbamoyl phosphate synthetase, a Triad class

[†] This research was supported in part by NIH grants (AR35186 to I.R., GM55513 to H.M.H., CA28725 to S.M.S., and GM39082 to G.H.R.).

[‡] The X-ray coordinates have been deposited in the Protein Data Bank (1CT9).

* To whom correspondence should be addressed. Phone: (608) 262-0529. Fax: (608) 265-2904. E-mail: Ivan_Rayment@biochem.wisc.edu.

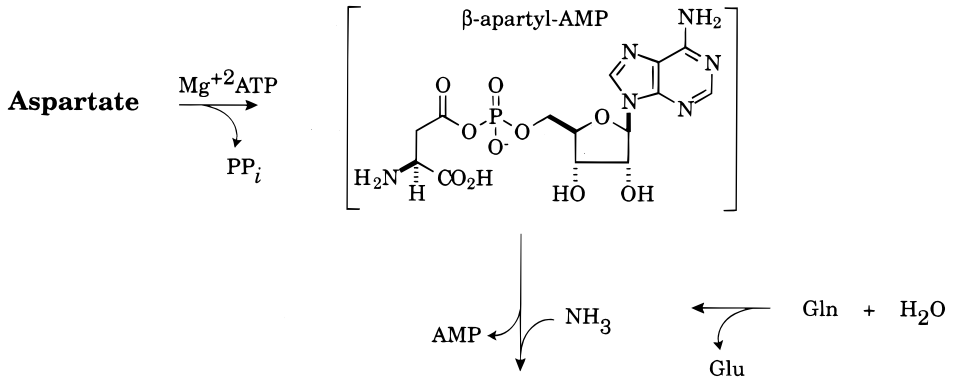
[§] University of Wisconsin.

[||] Department of Biochemistry and Molecular Biology and Department of Chemistry and Biotechnology Program, University of Florida.

[1] Abbreviations: GPATase, glutamine phosphoribosylpyrophosphate amidotransferase; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-(ethanesulfonic acid); MAD, multiple wavelength anomalous dispersion; DON, 6-diazo-5-oxonorleucine; PPi, pyrophosphate.

10.1021/bi9915768 CCC: $18.00
Published on Web 11/12/1999

Scheme 1



protein, the three active sites are separated by nearly 100 Å (*13*). Molecular tunnels connecting the various active sites have been observed in both of the above-mentioned enzymes (*12*, *13*). The manner in which the active sites in these proteins are coordinated and regulated remains an area of intense investigation. Recent results have shown that, for GPATase, the binding of phosphoribosylpyrophosphate to the C-terminal domain results in three structural changes occurring in the protein: a kinking of the C-terminal helix (Asp 471−Gln 492), an ordering of a flexible loop from Val 325−Arg 354, and a restructuring of the loop defined by residues Arg 73 to Ser 79, which allows the guanidinium group of Arg 73 to interact with the glutamine substrate in the N-terminal domain (*12*). These structural changes clearly play a pivotal role in signaling between the two active sites of GPATase.

In an effort to both complement and extend our current understanding of asparagine synthetase, and of the Ntn class amidotransferases in general, an X-ray crystallographic analysis of asparagine synthetase B from *E. coli* was initiated. Here we report the molecular architecture of the catalytically inactive Cys1Ala mutant of the enzyme complexed with L-glutamine and AMP and determined to a nominal resolution of 2.0 Å. This study was conducted in order to address several key structural issues concerning asparagine synthetase. For example, what is the spatial relationship between the active sites utilized for the hydrolysis of glutamine and for the formation of the β-aspartyl AMP intermediate? How similar is the molecular motif of asparagine synthetase to the structures of other Ntn class enzymes? What residues are important for the binding of glutamine? How is ATP accommodated within the active site and what amino acids are critical for the production and stabilization of the β-aspartyl AMP intermediate? How is the ammonia generated from the hydrolysis of glutamine transported to the active site responsible for the final production of asparagine? Does a molecular tunnel exist in asparagine synthetase, and if so, how similar is it to that observed in GPATase? The model described here begins to address several of these key structural/functional questions.

MATERIALS AND METHODS

Purification and Crystallization Procedures. Protein employed for this investigation was purified as described elsewhere (*14*) and stored in 2.0 M ammonium sulfate. For crystallization experiments, the protein was dialyzed against 20 mM Bis-Tris (pH 6.5), 150 mM NaCl, 0.5 mM EDTA, and 2 mM DTT and either used immediately or frozen in liquid nitrogen in 30 μL droplets. Crystals were grown by the hanging drop method of vapor diffusion at 4 °C. For these experiments, the protein concentration was adjusted to 6.5 mg/mL and contained 10 mM $MgCl_2$, 5 mM glutamine, and 10 mM AMP. Typically 10 μL of the protein solution was mixed with 10 μL of a precipitant solution containing 14% (w/v) poly(ethylene glycol) 8000 and 50 mM Hepes (pH 7.0). These droplets were then suspended against wells containing 15% poly(ethylene glycol) 8000 and 50 mM Hepes (pH 7.0). Small crystals appeared within approximately 3 days and grew as rectangles with well-defined faces. Large crystals were subsequently obtained by the technique of macroseeding (*15*, *16*). For these experiments, small crystals were washed in 8% poly(ethylene glycol) 8000 and 50 mM Hepes (pH 7.0) and placed in a fresh 20 μL droplet of the protein solution as described before but with a slightly lower poly(ethylene glycol) 8000 concentration. These 20 μL droplets were subsequently equilibrated against 15% poly(ethylene glycol) 8000 and 50 mM Hepes (pH 7.0). This procedure was repeated an additional time to produce crystals with overall dimensions of 0.3 mm × 0.5 mm × 1.0 mm. The crystals belonged to the space group $P2_12_12_1$ with unit cell dimensions of $a = 101.0$ Å, $b = 127.8$ Å, and $c = 205.6$ Å and contained two dimers per asymmetric unit.

X-ray Data Collection and Processing. Initial X-ray data were recorded at −3 °C with a Bruker HiSTAR area detector positioned at a crystal-to-detector distance of 25 cm. For heavy-atom derivative searches, crystals were transferred to a synthetic mother liquor containing 14% poly(ethylene glycol) 8000, 50 mM Hepes (pH 7.0), 100 mM NaCl, 100 mM tetramethylammonium sulfate, 10 mM $MgCl_2$, 8 mM glutamine, and 10 mM AMP and allowed to equilibrate for at least 24 h. Over 40 heavy metal salts were surveyed. Only two compounds, namely K_2PtCl_4 and uranyl oxalate, produced interpretable isomorphous heavy-atom derivatives. The uranyl oxalate derivative was subsequently utilized to determine the structure of asparagine synthetase by the technique of multiple wavelength anomalous dispersion (MAD).

Table 1: MAD X-ray Data Collection Statistics

resolution	edge	peak	high	low	high-resolution data set
wavelength (Å)	0.7219	0.7216	0.7012	0.7433	0.7433
resolution (Å)	2.4	2.4	2.4	2.4	1.9
total observations	2 339 392	2 341 855	2 303 772	2 321 221	1 085 969
unique reflections	105 296	105 284	104 341	104 251	200 469
redundancy[a]	11.1 (4.1)	11.1 (4.2)	11.0 (3.7)	11.1 (4.0)	5.4 (3.2)
completeness	99.9 (99.8)	99.9 (99.8)	99.8 (99.1)	99.9 (99.8)	98.0 (93.6)
R_{merge} (%)	8.8 (31.4)	8.7 (29.7)	9.4 (33.1)	8.6 (31.8)	6.3 (33.1)
R	5.0		6.96	7.1	

[a] Numbers in parentheses are related to the highest resolution shell, which is 2.49−2.40 Å for the MAD X-ray data and 1.97−1.90 Å for monochromatic X-ray data.

For MAD X-ray data collection, the crystals were soaked in 2 mM uranyl oxalate for 45 h and then serially transferred in six steps to 14% poly(ethylene glycol) 8000, 50 mM Hepes (pH 7.0), 280 mM NaCl, 100 mM tetramethylammonium sulfate, 10 mM $MgCl_2$, 8 mM glutamine, and 10 mM AMP containing 24% glycerol and 2 mM uranyl oxalate over a period of 1 day. The crystals were captured in a loop of surgical suture and flash frozen in a stream of nitrogen gas at −160 °C.

X-ray data were collected on beam line 19ID of the Structural Biology Center at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). All X-ray data were recorded from a single frozen crystal. Diffraction patterns were recorded with a 3 × 3 mosaic CCD area detector (*17*) and reduced with the HKL2000 package (*18*). The MAD data were collected at four wavelengths with 1° frames at a crystal-to-detector distance of 30 cm and an exposure time of 2 s. Although the crystal exhibited diffraction to a nominal resolution of 1.9 Å resolution, radiation damage limited the MAD X-ray data to 2.6 Å. Upon completion of the MAD X-ray data collection runs, the crystal was translated and a high-resolution X-ray data set to 1.9 Å was collected with 0.25° frames, 4 s per exposure, and a crystal-to-detector distance of 23 cm. X-ray data collection statistics are summarized in Table 1.

Structural Determination. The initial heavy-atom positions were determined by visual inspection of Patterson maps calculated to 4.0 Å resolution. Originally, four heavy-atom binding sites were located. Upon closer examination of difference Fourier maps, however, it became apparent that these major sites were actually composed of three "subsites" each and separated from one another by ∼4 Å. This arrangement led to a total of 12 uranium atoms per asymmetric unit. The heavy-atom positions were refined and protein phases calculated to 2.6 Å resolution with the software package SOLVE (*19*, *20*). Since there were four copies of asparagine synthetase in the asymmetric unit, it was possible to improve the quality of the electron density map by the techniques of molecular averaging and solvent-flattening as implemented in the software package DM (*21*). The relationships between the four subunits in the asymmetric unit were determined with the program MUNCHKINS (*22*). Residues Ala 1−Phe 196, Glu 210−Lys 247, Leu 267−Tyr 357, His 370−Ile 415, and Val 456−Val 514 were built into the averaged electron density map. This model was subsequently expanded into the unit cell and subjected to least squares refinement with the program TNT (*23*) at 2.0 Å. The X-ray data from 2.0− 1.9 Å were weak and were not included in the refinement.

Table 2: Least-Squares Refinement Statistics

resolution limits (Å)	30.0−2.0
$R_{overall}$ (%)[a]	19.7
R_{free} (%)	29.7
no. of reflections used	172 857
no. of protein atoms	15 780
no. of water molecules[b]	1048
weighted root-mean-square deviations from ideality	
bond length (Å)	0.012
bond angle (deg)	2.4
planarity (trigonal) (Å)	0.006
planarity (other planes) (Å)	0.012
torsional angle (deg)[c]	18.3

[a] R-factor $= \Sigma|F_o - F_c|/\Sigma|F_o|$ where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude. $R_{overall}$ includes all measured X-ray data. R_{free} is based on 10% of the X-ray data. [b] In addition to the water molecules, the X-ray model also included 16 uranium ions, four chloride ions, four AMP moieties, and four glutamines. [c] The torsional angles were not restrained during the refinement.

Within the unit cell, it was possible to locate many of the missing residues. In each subunit, however, approximately 40 of the C-terminal residues were disordered. During the early stages of least-squares refinement, 10% of the X-ray data were excluded for the required calculation of R_{free}. In that all X-ray data are important for the Fourier synthesis, however, these data were ultimately included in the final stages of the refinement and model building. Alternate cycles of refinement and model building reduced R_{free} to 29.7% and $R_{overall}$ to 19.7% for all measured X-ray data from 30 to 2.0 Å. Note that the high-resolution native X-ray set employed for the least-squares refinement was collected from a crystal soaked in uranyl oxalate. Consequently, in addition to water molecules, four uranium sites per subunit were included in the refinement. Each subunit in the asymmetric unit also contained a bound glutamine and an AMP moiety. The following regions were disordered in the electron density map: Ala 251 to Gln 266, Gly 423 to Gly 425, and Gly 517 to Lys 553 (subunit I), Ala 251 to Trp 263, Gly 423 to Gly 425, Lys 449 to Ser 453, and Gly 517 to Lys 553 (subunit II), Ala 251 to Trp 263, Gly 423 to Gly 425, Lys 449 to Ser 453, and Gly 517 to Lys 553 (subunit III), and Arg 252 to Pro 265, Asn 424 to Gly 425, Lys 449 to Ser 453, and Gly 517 to Lys 553 (subunit IV). Relevant refinement statistics can be found in Table 2. Portions of the electron density map corresponding to the bound ligands are displayed in Figure 1. For the sake of simplicity, the monomer described in the Results refers to subunit I in the X-ray coordinate set while the dimer presented refers to subunits I and IV.

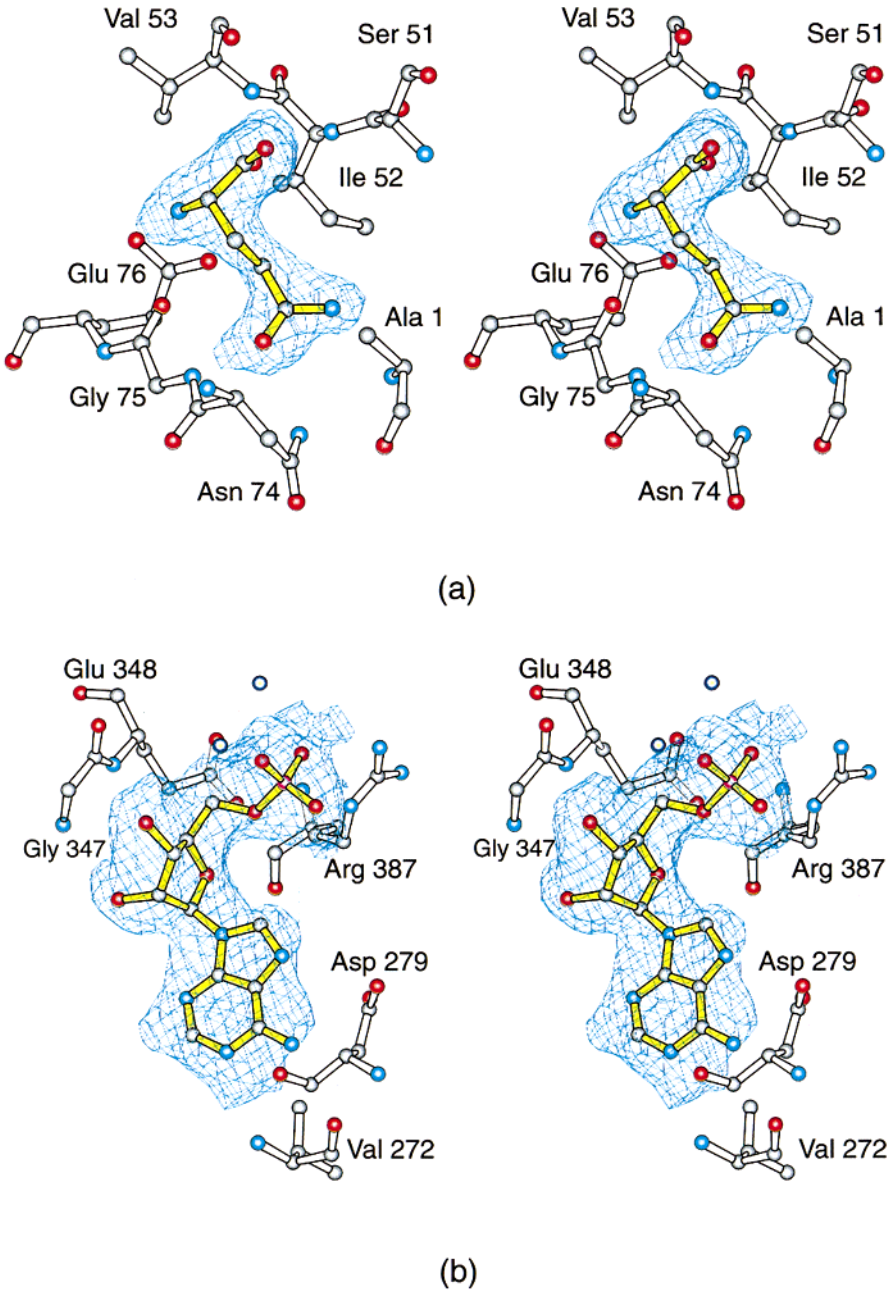


FIGURE 1: Observed electron density for the bound ligands. (a) Shown is a portion of the electron density corresponding to the bound glutamine. The map was contoured at 2σ and calculated with coefficients of the form ($F_o - F_c$), where F_o was the native structure factor amplitude and F_c was the calculated structure factor amplitude. Electron density for the AMP moiety is displayed in panel b. In this case, the map was contoured at 1.5σ. In both panels a and b, the ligands were omitted from the least-squares refinement and the phase calculations.

RESULTS

Tertiary and Quaternary Structure. Asparagine synthetase contains 553 amino acid residues (excluding the N-terminal methionine) and functions as a dimer (*4*). For the numbering employed here, residue 1 corresponds to the N-terminal alanine which in the native protein is a cysteine residue. A ribbon representation of one asparagine synthetase subunit is displayed in Figure 2a. As can be seen, the monomer consists of two structural motifs formed by Ala 1 to Asp 194 and Trp 195 to Gly 516 and referred to as the N- and C-terminal domains, respectively. In addition to the surface loops between Ala 250−Leu 267 and Cys 422−Ala 426, the last 37 residues of the C-terminus are also disordered in the present structure (subunit I). Most likely these C-terminal residues play a role in binding the aspartate substrate.

The N-terminal domain, responsible for the hydrolysis of glutamine, consists of four α-helical regions that flank the outer surfaces of two layers of antiparallel β-sheet. Each layer of sheet contains six β-strands that range in length from three to nine residues. The helical regions are formed from four to 12 residues with the α-helix defined by Met 161 to Pro 166 being decidedly distorted. These secondary structural elements are connected to one another by ten type I, one type II, one type II′, and four type III turns. Nearly 80% of the amino acid residues in the N-terminal domain occur in well-defined secondary structural elements. In addition, Pro 61 adopts the cis-conformation. As expected from the known biochemical features of asparagine synthetase, the glutamine

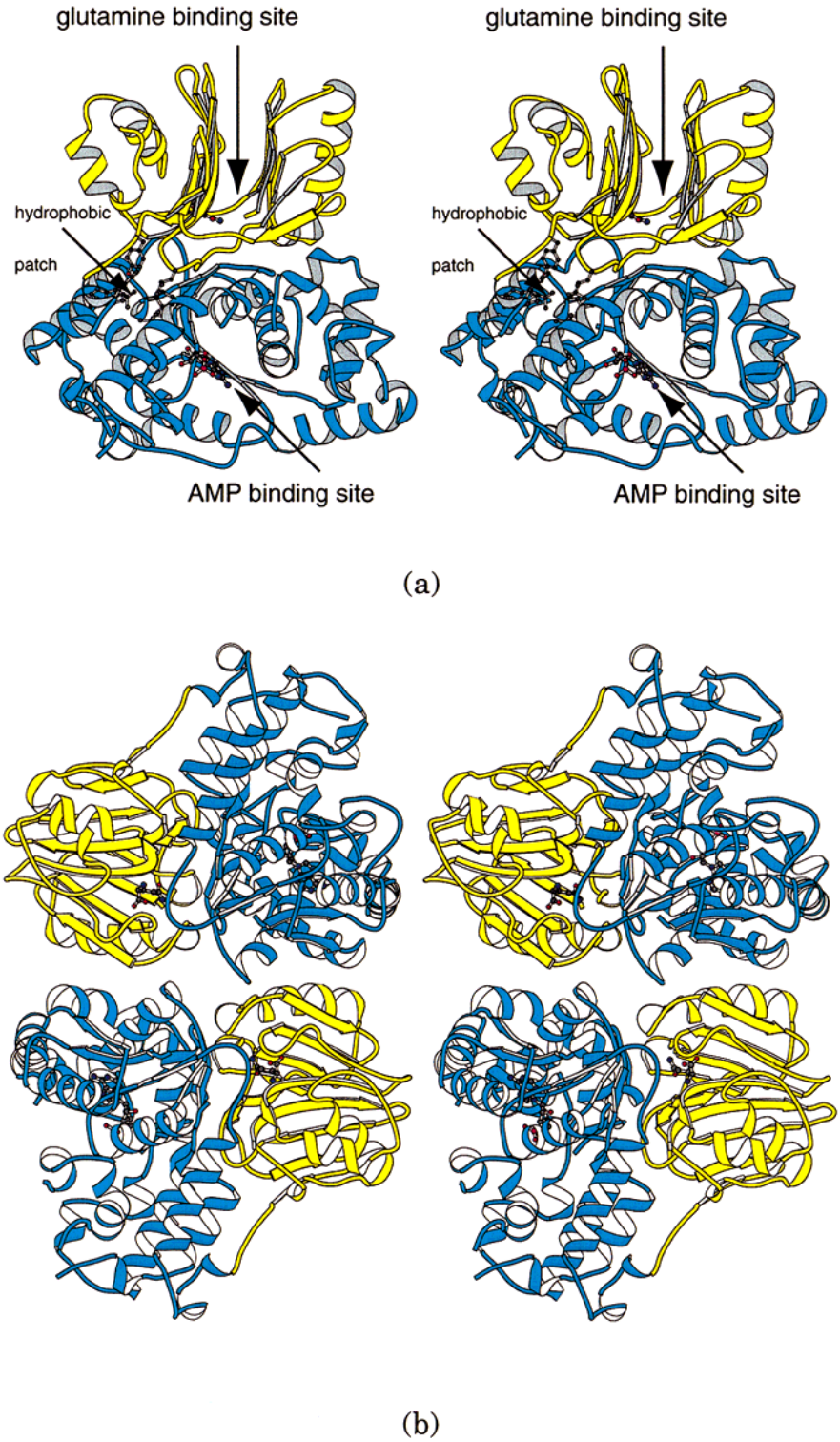


FIGURE 2: Ribbon representation of asparagine synthetase. An individual subunit of asparagine synthetase is shown in panel a while the dimeric species is displayed in panel b. The N-terminal domain, depicted in yellow, is responsible for the hydrolysis of glutamine while the C-terminal motif, shown in blue, is required for the production of the aspartyl-AMP intermediate.

ligand is situated near the N-terminal residue. The location of this glutamine is indicated by the ball-and-stick representation in Figure 2a and as can be seen the carboxamide group of the ligand projects into the interface formed by the two layers of β-sheet.

While the N-terminal domain is composed mostly of antiparallel β-sheet, the larger C-terminal domain is characterized by a five-stranded parallel β-sheet surrounded by a total of 16 α-helices. The strands of β-sheet range in length from three to eight residues. Many of the helices are quite substantial with one ordering up to 27 residues (Ala 366 to Met 392). This particular helix is distorted as a result of Tyr 383, which adopts dihedral angles of $\phi = -114°$ and $\psi = -102°$. The next residue in the polypeptide chain, Asp 384, serves as a ligand to one of the bound uranium ions and in the absence of such metals, this region of polypeptide chain may possibly adopt a less strained conformation. Note that Tyr 383 is the only amino acid residue in asparagine

synthetase whose dihedral angles lie well outside of the allowed regions of the Ramachandran plot. In addition to the α-helices and β-strands, there are seven type I, two type II, and two type III turns also present in the C-terminal domain. Again, most of the residues forming the C-terminal domain are situated in well-defined secondary structural elements with only ~24% lying in random coil regions. The AMP moiety observed in the electron density map lies across the C-terminal edge of the β-sheet as indicated by the ball-and-stick representation in Figure 2a. The glutamine and AMP ligands are separated by approximately 19 Å.

From Figure 2a, it is immediately obvious that the N- and C-domains of asparagine synthetase are quite distinct biochemical units with no commonly shared secondary structural elements. The molecular interface between these motifs is extensive with a buried surface area of ~2500 Å^2 as calculated according to the algorithm of GRASP (*24*). Specifically, this domain:domain interface is built primarily from the regions of polypeptide delineated by His 29 to Asp 33, Leu 50 to Asp 54, Glu 76 to Tyr 78, Asp 115 to Gln 118, His 139 to Tyr 146, and Glu 160 to Glu 173 in the N-terminal domain and by Ser 221 to Asp 226, Tyr 313 to Glu 316, Lys 342, Lys 367 to His 381, and Asn 389 to Arg 414 in the C-terminal domain. While the intramolecular interface contributed by the N-terminal domain is formed primarily by reverse turns and random coil regions, the molecular surface donated by the C-terminal domain is composed mostly of α-helices and one β-strand. Both hydrophobic and electrostatic interactions occur within these regions. As an example of the former, the molecular interface formed by Leu 50 to Asp 54 is reasonably hydrophobic with the side-chain groups of Leu 50 and Ile 52 pointing toward the interior. As examples of the latter, there are salt bridges formed between the functional groups of Asp 33 and Lys 342, Asp 115 and Lys 406, and Lys 163 and Glu 316. A striking hydrophobic patch exists at the transitional region between the two domains as indicated in Figure 2a. In this area there is a crown of leucine residues (140, 377, 378, 380, 404, 409) surrounded by Tyr 191, Trp 195, and Phe 196. Interestingly, the distorted α-helix that includes Tyr 383 discussed above forms part of the intramolecular interface thus providing a direct link between the active site of the C-terminal domain and the surface of the N-terminal motif.

As depicted in Figure 2b, the asparagine synthetase dimer is fairly symmetrical with overall dimensions of approximately 98 Å × 103 Å × 74 Å and a buried surface area of ~2220 Å^2. The two bound glutamine ligands within the dimer are 27 Å apart while the two nucleotide monophosphate moieties are separated by 39 Å. A distance of 25 Å separates the glutamine/nucleotide pairs in the neighboring subunits of the physiological dimer. The dimeric interface is formed predominately by four regions of polypeptide chain contributed by each subunit and delineated by Arg 17 to Glu 48 (an α-helix-turn-β-sheet-turn-β-sheet motif), Gln 302 to His 314 (an α-helix), Leu 331 to Lys 342 (an α-helix), and Ser 393 to Gly 396 (a type I turn). Along with the numerous water molecules that line the subunit:subunit interface, there are various salt bridges that occur including those between Arg 17 and Asp 306*, Glu 48 and Arg 334*, Asp 306 and Arg 17*, and Arg 334 and Glu 48* (where those residues indicated by the asterisks belong to subunit II). Additionally, the side-chain groups of Arg 25 and Arg 28 (subunit I) project into the dimeric interface such that both of their guanidinium groups lie within ~3.3 Å of the carboxylate group of Asp 310 (subunit II). The symmetry-related ion pairs between Arg 25 and Arg 28 of subunit II and Asp 310 (subunit I) are also present in the dimeric species. In addition to numerous electrostatic interactions, there are several hydrophobic patches that contribute to the subunit:subunit interface of asparagine synthetase, including the region surrounding the side chain of Trp 395, which is located in the ~Type I turn listed above.

Active Site of the N-Terminal Domain. Since the structure of asparagine synthetase reported here is that of a mutant protein whereby the active nucleophile Cys 1 was changed to an alanine residue, it was possible to crystallize the enzyme in the presence of its natural glutamine substrate. A close-up view of the glutaminase active site in the N-terminal domain is displayed in Figure 3a, and a cartoon of potential hydrogen bonding interactions between the ligand and the protein is given in Figure 3b. For the sake of clarity, only those residues that lie within approximately 4 Å of the ligand are shown. There are only two bound water molecules observed in the active site. One of these solvents interacts with the α-carboxylate group of the substrate while the second forms a hydrogen bond with the carbonyl oxygen of side-chain carboxamide group. As indicated by the dashed lines in Figure 3b, the side-chain functional groups of Arg 49, Asn 74, Glu 76, and Asp 98 form hydrogen bonds with the glutamine ligand. These hydrogen bonds range in length from 2.6 to 3.1 Å. In addition to these side chains, the backbone amide groups of Ile 52, Val 53, and Gly 75 and the carbonyl oxygens of Leu 50 and Gly 75 also play key roles in the proper positioning of the ligand within the active-site pocket. The observed interactions between the ligand and Arg 49, Asn 74, and Asp 98 were, in fact, previously predicted from model building studies of asparagine synthetase (*5*) based on the structural analyses of the *E. coli* GPATase (*25*) and an N-terminal proteolytic fragment of the *E. coli* glutamine fructose-6-phosphate amidotransferase (*26*).

Note that C$^\beta$ of Ala 1 is situated at 3.9 Å from the carbonyl carbon of the substrate that would be attacked by the N-terminal cysteine residue in the native protein. With the assumption that the glutamine ligand binds to this Cys1Ala mutant protein in an identical manner to that expected for the native protein, it can be speculated that the nucleophilic attack by the thiolate of Cys 1 occurs at the *si*-face of the glutamine substrate, yielding a tetrahedral intermediate with an S-configuration. The oxyanion of the tetrahedral intermediate presumably is stabilized via hydrogen bonds donated by the backbone amide group of Gly 75 and N$^{\delta 2}$ of Asn 74. In the structure reported here, the distance between the side-chain carbonyl oxygen of the ligand and the backbone amide group of Gly 75 is 2.5 Å. Similarly, N$^{\delta 2}$ of Asn 74 is positioned within 3.0 Å to this oxygen atom.

Active Site of the C-Terminal Domain. Unlike the active site in the glutaminase portion of asparagine synthetase, the C-terminal active site is less well-defined in the present complex. From Figure 2a, it is immediately apparent that the binding pocket for the observed AMP moiety is quite open and most likely the disordered regions in the present model play critical roles in the binding of Mg^{2+}ATP and aspartate. Even in light of this caveat, the general region for

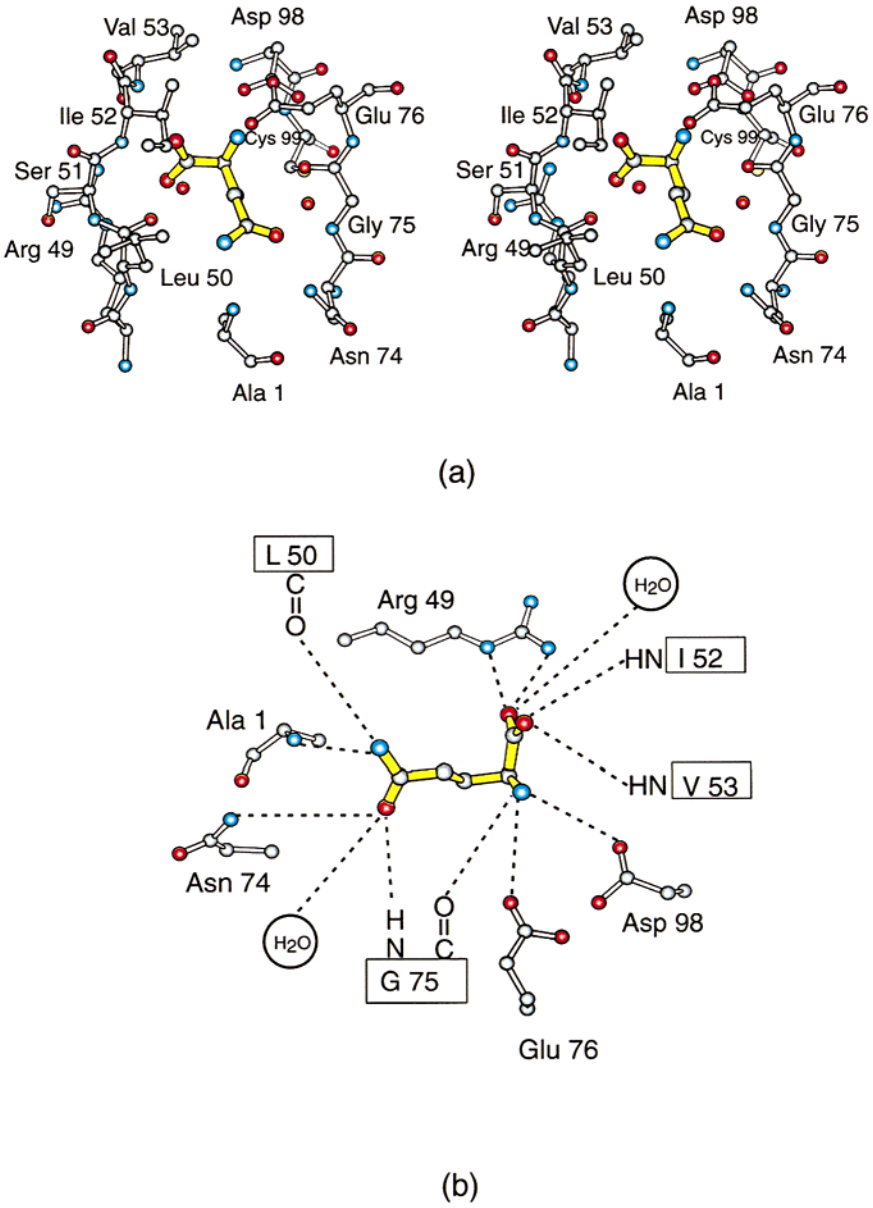


FIGURE 3: Close-up view of the N-terminal domain active site. Those protein atoms located within approximately 4 Å of the glutamine ligand are shown in panel a. A cartoon of the possible electrostatic interactions between the ligand and the protein, within 3.2 Å, is given in panel b.

the C-terminal domain active site has been partially defined in the present structural analysis and is shown in Figure 4a. In addition to the bound nucleotide, three uranium ions were observed in the active site. Most likely some of these metals are mimicking the binding of magnesium ions. The ribose of the nucleotide adopts the $C_{3'}$-endo pucker while the adenine ring lies in the anti-conformation. Two specific electrostatic interactions exist between the purine ring of the nucleotide and the protein: one of these occurs between the backbone carbonyl group of Val 272 and the amino group at position 6 of the adenine ring while the second occurs between the amide group of Val 272 and N1 of the purine moiety. In both cases, the distances between the oxygens and nitrogens in these hydrogen bonds are ~3.1 Å. The 2′-hydroxyl group of the nucleotide ribose is linked to the protein via O of Leu 232, O^{γ} of Ser 346, and N of Glu 347. The only other interactions between the nucleotide and the protein occur via the bound uranium ions. One of the uranium ions sits at 2.6 Å from the 3′-hydroxyl group of the ribose. The other two uranium sites are coordinated by phosphoryl oxygens of the AMP moiety. One of these ions is further ligated to the protein via the carboxylate side chains of Asp 238 and Asp 351 while the second metal is surrounded by the carboxylate side chains of Glu 352 and Asp 384 and O^{η} of Tyr 357. The metal:ligand bond distances range in length from 2.3 to 2.8 Å. The coordination geometry for the metals and the observed protein:nucleotide interactions will most likely change when the structure of asparagine synthetase is solved in the presence of β-aspartyl-AMP mimics and magnesium ions.

DISCUSSION

Significant progress was made on our understanding of the amidotransferases by the informative X-ray crystallographic studies of GMP synthetase from *E. coli* by Davisson and Smith (*9*) and of GPATase, also from *E. coli*, by Zalkin and Smith (*12*, *25*, *27*). These specific enzymes belong to the Triad and Ntn classes, respectively. Subsequent

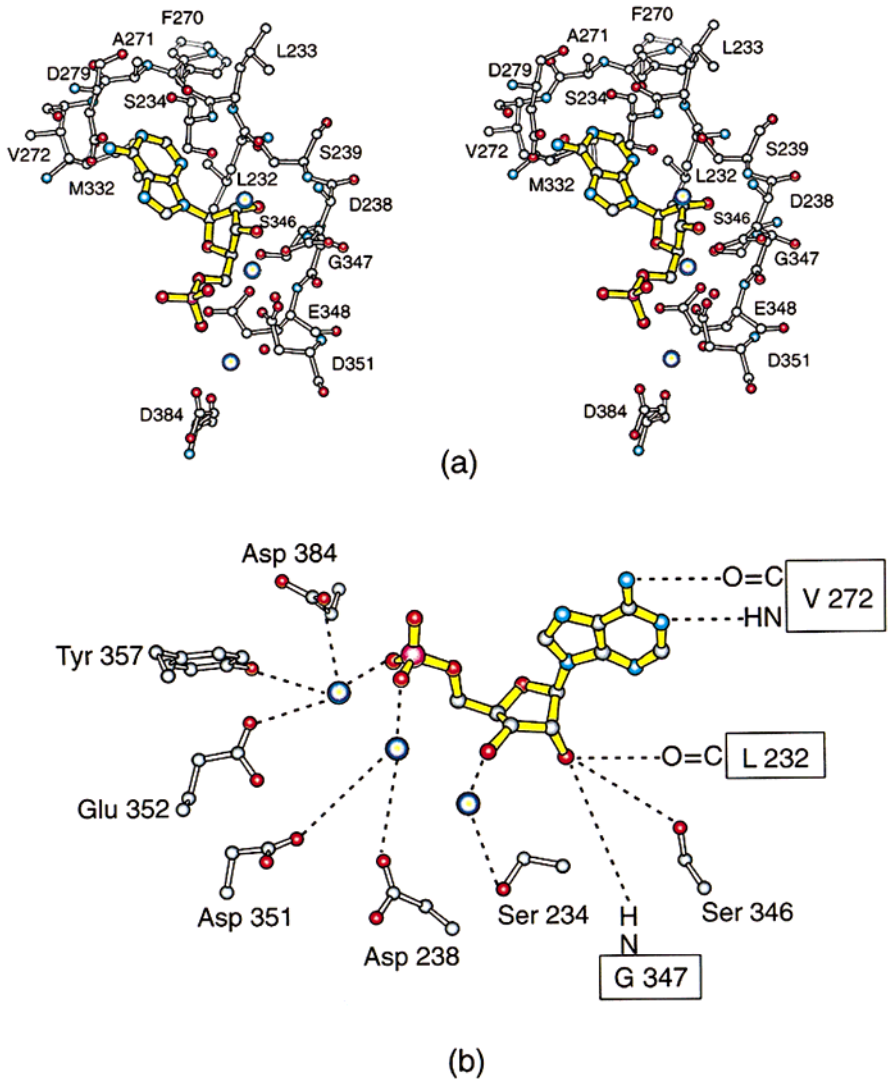


FIGURE 4: Close-up view of the AMP binding region in the C-terminal domain. Those protein atoms located within approximately 4 Å of the nucleotide monophosphate are shown in panel a. A cartoon of the possible electrostatic interactions between the protein and the AMP, within 3.2 Å, is displayed in panel b. The bound uranium ions are indicated by the yellow/gray spheres.

structural analyses of two additional enzymes, namely the N-terminal domain of glucosamine 6-phosphate synthase (*26*) and carbamoyl phosphate synthetase (*10*, *13*, *28*, *29*, *30*) added important new information regarding the Ntn and Triad families, respectively. From a combination of both biochemical and structural data, it is now known that the Ntn amidotransferases, as the name implies, employ an N-terminal cysteine residue as the active-site nucleophile for the hydrolysis of glutamine to glutamate and ammonia. As a consequence of the location of the nucleophile, the glutaminase domains of the Ntn class proteins are always located at the N-terminal portions of the molecules. In the case of the Triad family, however, this restriction does not apply. These enzymes contain a conserved Cys-His couple that is located in the N-terminal domain of GMP synthetase, but in the C-terminal domain of the carbamoyl phosphate synthetase small subunit. While these enzymes are referred to as the Triad family, this may, in fact, be a misnomer. Presently, there is some discrepancy in the literature regarding the third residue in the Triad, namely a conserved glutamate. In GMP synthetase, it has been speculated that this glutamate is part of a catalytic triad (*9*). However, a study on the glutaminase subunit of *p*-aminobenzoate synthetase has demonstrated that replacement of this residue with an alanine or aspartate results in little loss of catalytic activity (*31*). Regardless of the actual role of the conserved glutamate, these enzymes will most likely continue to be referred to as the Triad class in the literature, and, as such, this nomenclature will be retained here.

Prior to the structural analysis of asparagine synthetase described in this report, it had already been predicted that its N-terminal domain would have a similar molecular fold to that observed in GPATase and that its C-terminal domain would be homologous to the C-terminal domain of GMP synthetase (*9*). These predictions of similar structures were based on both amino acid sequence homologies and consideration of common catalytic mechanisms and were exceedingly accurate (*32*). Indeed, asparagine synthetase can be envisioned simply as a chimera of GPATase and GMP synthetase whereby an Ntn class glutaminase domain has been attached to a synthetase domain belonging to a Triad family. This mixing and matching of donor (glutaminase) and acceptor (synthetase) domains is a common theme among the amidotransferases.

Shown in Figure 5a is a superposition of the N-terminal domain of GPATase onto the N-terminal domain of aspar-

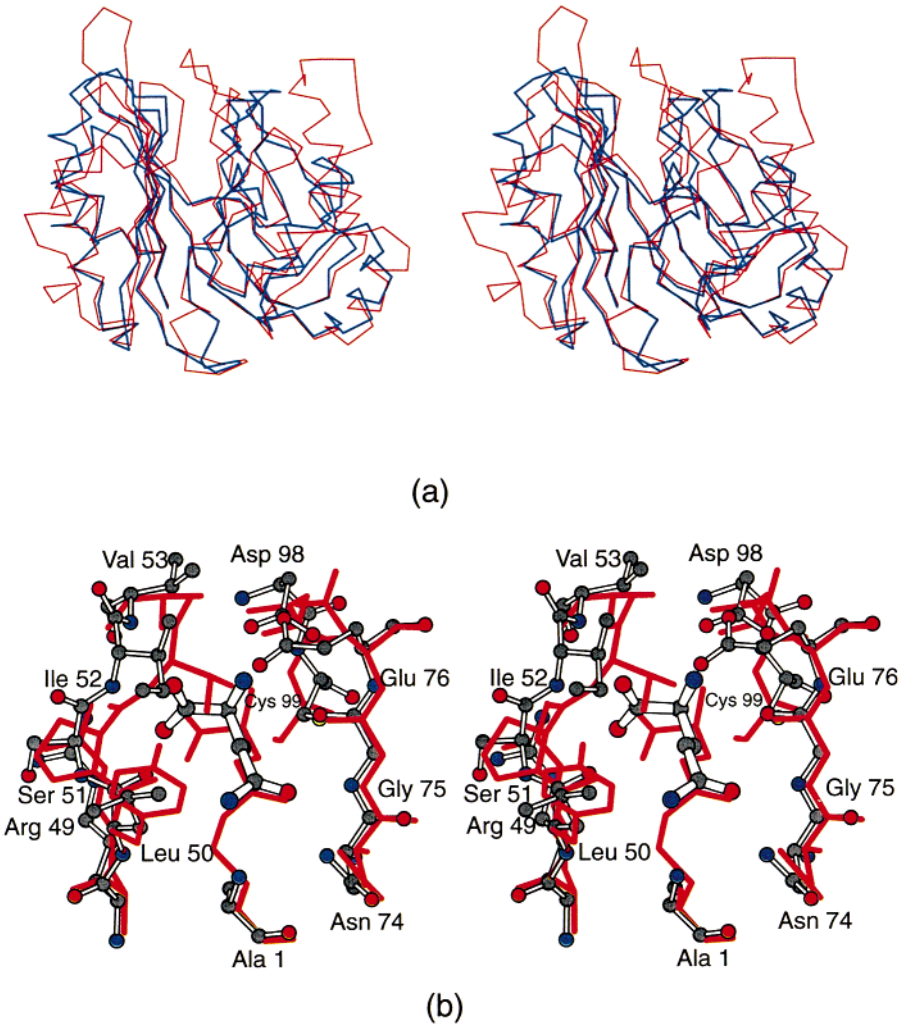


FIGURE 5: Comparison of the glutaminase domains in asparagine synthetase and GPATase. X-ray coordinates for GPATase (*12*) were obtained from the Brookhaven Protein Data Bank (1ECC). A superposition of the α-carbons for the two N-terminal domains in asparagine synthetase (blue) and GPATase (red) is depicted in panel a. A close-up view of the active site regions is shown in panel b. Note that the structure of GPATase was solved in the presence of a glutamine affinity analogue covalently attached to the N-terminal cysteine residue.

agine synthetase. The α-carbons for these two models superimpose with a root-mean-square deviation of 1.4 Å for 176 structurally equivalent residues. It is immediately obvious that the two structures differ only in a few regions. The largest difference occurs at Ser 40 (asparagine synthetase) or Ile 36 (GPATase). Here, there is an insertion in the GPATase model whereby the polypeptide chain wraps around to form an additional β-strand-helix motif. As a consequence, the first β-sheet in GPATase contains seven rather than six β-strands. The two polypeptide chains superimpose again at Ile 44 (asparagine synthetase) or Gly 68 (GPATase). There is a second, rather large insertion, that occurs in the region defined by Glu 108 to Phe 113 in asparagine synthetase or Ser 137 to Thr 159 in GPATase. Finally, there are two smaller insertions that occur between Glu 151 to Gly 153 and Ser 182 to Gly 185 in asparagine synthetase and correspond to Asp 196 to Thr 202 and Tyr 231 to Phe 239 in GPATase.

The structure of GPATase was solved in the presence of a glutamine analogue, namely 6-diazo-5-oxonorleucine and referred to as DON (*12*). Shown in Figure 5b is a superposition of the active sites for the glutaminase domains of GPATase and asparagine synthetase. The chemical environments of the active sites observed in asparagine synthetase and GPATase are somewhat different. For example, Tyr 74, Pro 75, Thr 76, Ala 77, Asn 103, and Ser 128 in GPATase are replaced by Leu 50, Ser 51, Ile 52, Val 53, Glu 76, and Cys 99 in asparagine synthetase. As can be seen, there is a close three-dimensional correspondence between the DON inhibitor bound to GPATase and the actual substrate, glutamine, as observed in asparagine synthetase. In GPATase, the side-chain carbonyl group of the inhibitor is hydrogen bonded to the backbone amide group of Gly 102 and $N^{\delta 2}$ of Asn 101. These interactions are identical to those depicted for asparagine synthetase in Figure 3b. Likewise, as observed in asparagine synthetase, the carboxylate group of the DON inhibitor forms a salt bridge with an arginine residue in GPATase, namely Arg 73. The amino group of the DON inhibitor interacts with GPATase in a slightly different manner to that observed for glutamine bound to asparagine synthetase. Specifically, O^{γ} of Thr 76 and $O^{\delta 1}$ of Asp 127 lie within hydrogen-bonding distance to the amino group. These residues are replaced by Glu 76 and Asp 98 in asparagine synthetase. The structure of asparagine synthetase reported here is that of a site-directed mutant, Cys1Ala. It is important to note the close correspondence between the N-terminal alanine of asparagine synthetase and the N-terminal cysteine of GPATase. This similarity indicates that

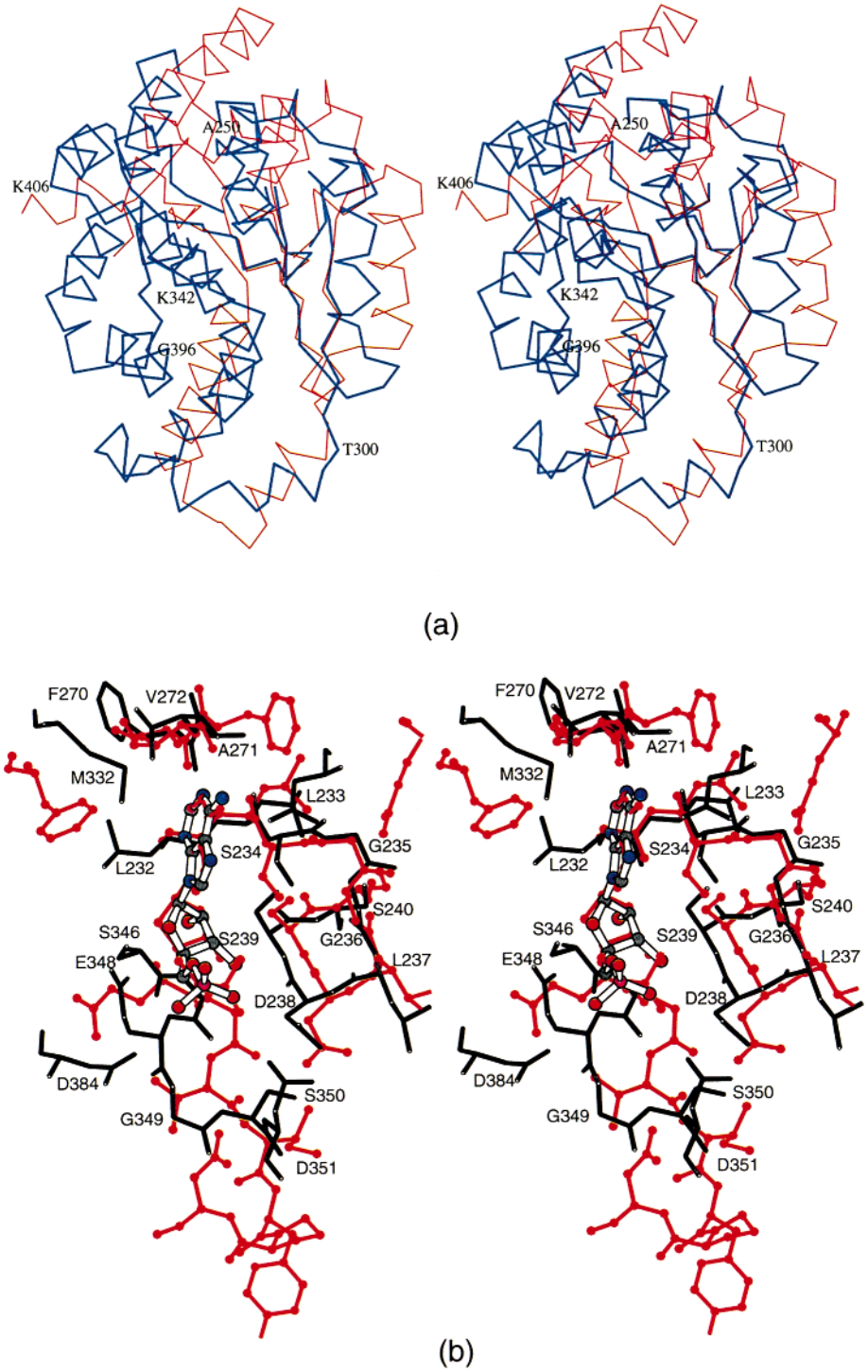


FIGURE 6: Comparison of the acceptor domains in asparagine synthetase and GMP synthetase. X-ray coordinates for GMP synthetase (*9*) were obtained from the Brookhaven Protein Data Bank (1GPM). A superposition of the α-carbons for the C-terminal domains in asparagine synthetase (blue) and GMP synthetase (red) is depicted in panel a. A close-up view of the protein environment surrounding the nucleotides is displayed in panel b. For the sake of clarity, the Mg^{2+} ions and the pyrophosphate moiety in GMP synthetase and the uranium ions in asparagine synthetase were omitted.

the structural results presented here, albeit based on a mutant protein, represent a good mimic for the native enzyme.

The synthetase domain of asparagine synthetase is responsible for the binding of Mg^{2+}ATP and aspartate and the subsequent production of the *β*-aspartyl intermediate and PP_i. These functional properties place asparagine synthetase into the category of ATP pyrophosphatases (*33*). Likewise, the C-terminal domain of GMP synthetase catalyzes the production of the O^2-adenyl-XMP intermediate and PP_i from XMP and ATP. A superposition of these two domains is given in Figure 6a. The α-carbons for these two models superimpose with a root-mean-square deviation of 1.9 Å for 79 structurally equivalent residues. Both of these structures contain bound AMP, and in GMP synthetase, PP_i is also observed. It can be speculated that the structural similarity between these two domains would be even greater if both models contained the full complement of ligands bound to their respective active sites. Indeed, in both GMP synthetase and asparagine

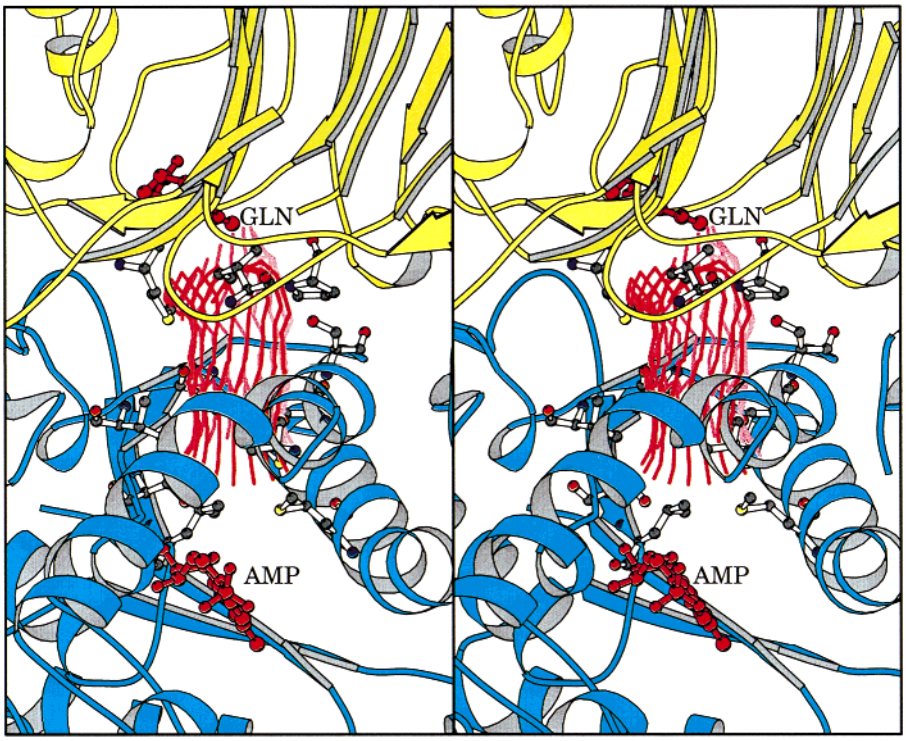


FIGURE 7: Molecular tunnel connecting the two active sites in asparagine synthetase. The location of the tunnel connecting the two active sites is indicated by the wire representation. Amino acid side chains that line the tunnel are included. This pathway was approximated visually on a graphics system as a series of base points. The program Alimentary was employed to refine the position of the tunnel (*13*).

synthetase, there are residues missing from the present models: the loop between Ala 346 and Lys 369 in GMP synthetase and the loops between Ala 250 to Leu 267 and Cys 422 to Ala 426 in asparagine synthetase. Interestingly, the structural correspondence observed between the C-terminal domains of asparagine synthetase B and GMP synthetase does not extend to the molecular motif recently determined by Nakatsu et al., (*34*) for the *E. coli* asparagine synthetase A. Rather, this enzyme closely resembles the catalytic domain of yeast aspartyl-tRNA synthetase (*35*).

There are four major regions where the two polypeptide chains differ significantly between GMP synthetase and asparagine synthetase: (i) a 13 residue insertion in asparagine synthetase beginning at Ala 250, (ii) a four residue insertion in asparagine synthetase beginning at Thr 300 and resulting in a completely different disposition of the surface loops defined by Thr 300 to Ala 325 in asparagine synthetase and Glu 291 to Ile 311 in GMP synthetase, (iii) an extended helix-turn-helix motif delineated by Val 353 to Gly 396 in asparagine synthetase which corresponds to the disordered region in GMP synthetase, and (iv) a 21 residue insertion in asparagine synthetase beginning at Lys 406.

A close-up view of the protein environment surrounding the bound nucleotides in asparagine synthetase and GMP synthetase is given in Figure 6b. Like that observed in asparagine synthetase, both N1 of the adenine ring and the N6 amino group of the base are hydrogen bonded to GMP synthetase via the backbone nitrogen and carbonyl oxygen of Val 260, respectively. The "P-loops", delineated by Ser 234 to Ser 240 in asparagine synthetase and Ser 235 to Ser 241 in GMP synthetase, are virtually identical with the only exception being Leu 237 in asparagine synthetase being replaced with a valine residue in GMP synthetase. As can be seen in Figure 6b, there are significant differences between the two models near the phosphate binding region which, most likely, are a result of the different constituents bound within the active sites (i.e., uranium ions in asparagine synthetase and PP_i in GMP synthetase).

While there are a number of disordered residues in the model of asparagine synthetase B described in this report, it is still absolutely clear that the active sites of the N- and C-terminal domains are separated by approximately 19 Å. This distance argues for some type of molecular tunnel connecting the site of ammonia formation to the site of its utilization. The concept of tunnels or channels is not unique, however, and was first described in the elegant structural analyses of tryptophan synthetase (*36*). A recent review of such tunnels in other enzyme systems, including GPATase and carbamoyl phosphate synthetase, can be found in ref *37*. Visual inspection of the asparagine synthetase model and a computational search with the software GRASP (*24*) has revealed the presence of a tunnel as highlighted in Figure 7. This molecular conduit begins at the apex formed by the two layers of antiparallel β-sheet in the N-terminal domain and leads to the domain:domain interface. In the C-terminal domain, there are two α-helices formed by Val 320 to Ile 336 and Ala 366 to Met 392 that splay outward and away from the five-stranded parallel β-sheet. The tunnel passes between these α-helices and the edge of the sheet. Like that observed in GPATase, the tunnel is formed primarily by polypeptide chain backbone atoms and hydrophobic or nonpolar side chains. There are no aspartates or glutamates lining the channel except for Glu 348 which is positioned at the end of the tunnel and near to the AMP moiety. Glu 348 is absolutely conserved in all 25 amino acid sequences that have been determined thus far for asparagine synthetases B. The side chain of Arg 30 is near the tunnel, but its guanidinium group is hydrogen bonded to the side-chain

carboxamide groups of Asn 74 and Asn 389. Both Arg 30 and Asn 74 are absolutely conserved while Asn 389 is replaced with an aspartate residue in some asparagine synthetases. Those residues with side chains pointing into the tunnel include Met 120, Ile 142, Ile 143, Leu 232, Met 329, Ser 346, Ala 388, Met 392, Ser 393, Ala 399, and Val 401. Leu 232, Met 329, Ser 346, and Ala 388 form a cluster near the C-terminal portion of the tunnel and are absolutely conserved among all 25 asparagine synthetase B amino acid sequences. Most of the other residues listed above are located more toward the N-terminal portion of the tunnel and are not strictly conserved. Both Met 120 and Ile 142, however, are replaced with side chains of comparable chemical reactivities. In the present model of asparagine synthetase with bound glutamine and AMP, there are several water molecules located within the tunnel. The positions of these solvents most likely change upon catalysis.

In summary, the structure of asparagine synthetase presented here provides another entry into the growing database concerning the Ntn-class amidotransferases. From this investigation, the spatial relationships between the two active sites have been determined, and the manners in which glutamine and AMP are accommodated within their respective binding pockets have been defined. In addition, a tentative molecular tunnel connecting the two active sites has been identified. Most likely the amino acid residues that are disordered in the current model play a key role in binding both the aspartate substrate and the β-aspartyl AMP intermediate. Further definition of the C-terminal domain active site and the molecular tunnel must await additional structural analyses of the enzyme in the presence of aspartate analogues and β-aspartyl AMP mimics; however, the current study establishes a solid foundation for these efforts. This work is in progress.

ACKNOWLEDGMENT

We thank Dr. W. W. Cleland for critically reading this manuscript and Dr. G. H. Reed for the financial support of T. M. L. In addition, we thank Dr. Gary Wesenberg for helpful discussions regarding the molecular tunnel. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Energy Research, under Contract W-31-109-ENG-38. We thank Drs. Frank Rotella and Norma Duke for their assistance in setting up the MAD experiment. Finally, we acknowledge the support of the Florida Space Grant Consortium for their summer fellowship to Stephen Goble who was involved in the initial crystallization experiments.

REFERENCES

1. Milman, H. A., and Cooney, D. A. (1979) *Biochem. J. 181*, 51−59.
2. Zalkin, H., and Truit, C. D. (1977) *J. Biol. Chem. 252*, 5341−5346.
3. Nakamura, M., Yamada, M., Hirota, Y., Sugimoto, K., Oka, A., and Takanami, M. (1981) *Nucleic Acids Res. 9*, 4669−4676.
4. Scofield, M. A., Lewis, W. S., and Schuster, S. M. (1990) *J. Biol. Chem. 265*, 12895−12902.
5. Richards, N. G., and Schuster, S. M. (1998) *Adv. Enzymol. 72*, 145−198.
6. Zalkin, H. (1993) *Adv. Enzymol. Relat. Areas Mol. Biol. 66*, 203−309.
7. Zalkin, H., and Smith, J. L. (1998) *Adv. Enzymol. Relat. Areas Mol. Biol. 72*, 87−144.
8. Amuro, N., Paluh, J. L., and Zalkin, H. (1985) *J. Biol. Chem. 260*, 14844−14849.
9. Tesmer, J. J. G., Klem, T. J., Deras, M. L., Davisson, V. J., and Smith, J. L. (1996) *Nat. Struct. Biol. 3*, 74−86.
10. Thoden, J. B., Miran, S. G., Phillips, J. C., Howard, A. J., Raushel, F. M., and Holden, H. M. (1998) *Biochemistry 37*, 8825−8831.
11. Smith, J. L., Zaluzec, E. J., Wery, J. P., Niu, L., Switzer, R. L., Zalkin, H., and Satow, Y. (1994) *Science 264*, 1427−1433.
12. Krahn, J. M., Kim, J. H., Burns, M. R., Parry, R. J., Zalkin, H., and Smith, J. L. (1997) *Biochemistry 36*, 11061−11068.
13. Thoden, J. B., Holden, H. M., Wesenberg, G., Raushel, F. M., and Rayment, I. (1997) *Biochemistry 36*, 6305−6316.
14. Boehlein, S. K., Richards, J. G. J., and Schuster, S M. (1994) *J. Biol. Chem. 269*, 7450−7457.
15. Thaller, C., Weaver, L. H., Eichele, G., Wilson, E., Karlsson, R., and Jansonius, J. N. (1981) *J. Mol. Biol. 147*, 465−469.
16. Thaller, C., Eichele, G., Weaver, L. H., Wilson, E., Karlsson, R., and Jansonius, J. N. (1985) *Methods Enzymol. 114*, 132−135.
17. Westbrook, E. M., and Naday, I. (1997) *Methods Enzymol. 276*, 244−268.
18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
19. Terwilliger, T. C., and Eisenberg, D. (1983) *Acta Crystallogr., Sect. A 39*, 813−817.
20. Terwilliger, T. C. (1997) *Methods Enzymol. 276*, 530−537.
21. Cowtan, K., and Main, P. (1998) *Acta Crystallogr., Sect. D 54*, 487−493.
22. Rypniewski, W. R., Breiter, D. R., Benning, M. M., Wesenberg, G., Oh, B.-H., Markley, J. L., Rayment, I., and Holden, H. M. (1991) *Biochemistry 30*, 4126−4131.
23. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr., Sect. A. 43*, 489−501.
24. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.
25. Kim, J. H., Krahn, J. M., Tomchick, D. R., Smith, J. L., and Zalkin, H. (1996) *J. Biol. Chem. 271*, 15549−15557.
26. Isupov, M. N., Obmolova, G., Butterworth, S., Badet-Denisot, M. A., Badet, B., Polikarpov, I., Littlechild, J. A., and Teplyakov, A. (1996) *Structure 4*, 801−810.
27. Muchmore, C. R., Krahn, J. M., Kim, J. H., Zalkin, H., and Smith, J. L. (1998) *Protein Sci. 7*, 39−51.
28. Thoden, J. B., Raushel, F. M., Benning, M. M., Rayment, I., and Holden, H. M. (1999) *Acta Crystallogr., Sect. D 55*, 8−24.
29. Thoden, J. B., Wesenberg, G., Raushel, F. M., and Holden, H. M. (1999) *Biochemistry 38*, 2347−2357.
30. Thoden, J. B., Raushel, F. M., Wesenberg, G., and Holden, H. M. (1999) *J. Biol. Chem* (in press).
31. Roux, B., and Walsh, C. T. (1993) *Biochemistry 32*, 3763−
32. 3768.
33. Boehlein, S. K., Walworth, E. S., Schuster, S. M., and Richards, N. G. J. (1999) Manuscript in preparation.
34. Bork, P., and Koonin, E. V. (1994) *Proteins: Struct., Funct., Genet. 20*, 347−355.
35. Nakatsu, T., Kato, H., and Oda, J. (1998) *Nat. Struct. Biol. 5*, 15−19.
36. Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., and Moras, D. (1991) *Science 252*, 1682−1689.
37. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., and Davies, D. R. (1988) *J. Biol. Chem. 263*, 17857−17871.
38. Miles, E. W., Rhee, S., and Davies, D. R. (1999) *J. Biol. Chem. 274*, 12193−12196.

BI9915768