# Structure of the Allosteric Regulatory Enzyme of Purine Biosynthesis

Janet L. Smith,* Eugene J. Zaluzec,† Jean-Pierre Wery,‡ Liwen Niu,§ Robert L. Switzer, Howard Zalkin, Yoshinori Satow

**Multiwavelength anomalous diffraction (MAD) has been used to determine the structure of the regulatory enzyme of de novo synthesis of purine nucleotides, glutamine 5-phosphoribosyl-1-pyrophosphate (PRPP) amidotransferase, from *Bacillus subtilis*. This allosteric enzyme, a 200-kilodalton tetramer, is subject to end product regulation by purine nucleotides. The metalloenzyme from *B. subtilis* is a paradigm for the higher eukaryotic enzymes, which have been refractory to isolation in stable form. The two folding domains of the polypeptide are correlated with functional domains for glutamine binding and for transfer of ammonia to the substrate PRPP. Eight molecules of the feedback inhibitor adenosine monophosphate (AMP) are bound to the tetrameric enzyme in two types of binding sites: the PRPP catalytic site of each subunit and an unusual regulatory site that is immediately adjacent to each active site but is between subunits. An oxygen-sensitive [4Fe-4S] cluster in each subunit is proposed to regulate protein turnover in vivo and is distant from the catalytic site. Oxygen sensitivity of the cluster is diminished by AMP, which blocks a channel through the protein to the cluster. The structure is representative of both glutamine amidotransferases and phosphoribosyltransferases.**

Glutamine amidotransferases catalyze the incorporation of reduced nitrogen in the biosynthesis of cellular components. At least 16 different amidotransferases function to transfer the glutamine amide nitrogen to substrates in biosynthetic pathways for nucleotide bases, amino acids, enzyme cofactors, amino sugars, and antibiotics (*1*). Phosphoribosyltransferases (PRTase) function in the biosynthesis and metabolism of nucleotides and are important chemotherapeutic targets (*2*). They catalyze reactions at the 1-position of ribose with PRPP as a substrate. Glutamine PRPP amidotransferase is a member of both enzyme families. We now report the crystal structure of glutamine PRPP amidotransferase from *B. subtilis* and thus provide three-dimensional structural information applicable to other members of the amidotransferase or PRTase families.

Biochemical and sequence studies of glutamine amidotransferases point to several common properties of the enzymes (*1*). Most glutamine amidotransferases can use ammonia ($NH_3$) as an alternative nitrogen source. Sites for glutamine binding and for $NH_3$-dependent synthesis have been localized to separate domains that are in some cases on dissimilar subunits. The site for glutamine utilization is distinguished by a catalytic cysteine residue in a "glutamine" domain and that for substrate (PRPP) binding and $NH_3$-dependent activity resides in a "transferase" domain. Nevertheless, the functions of the glutamine and transferase domains are tightly coupled. Sequences of different amidotransferases have similar glutamine domains but unrelated transferase domains. Two families of glutamine domains have been detected. Glutamine PRPP amidotransferase is the product of the *purF* gene in *Escherichia coli* and is the prototypical member of the *purF* family of glutamine domains. This family is distinguished by a catalytic cysteine at the amino terminus of the mature protein as well as by several other invariant residues.

Glutamine PRPP amidotransferse catalyzes the first committed step in de novo biosynthesis of purine nucleotides as follows:

$$\text{PRPP} + \text{glutamine} + H_2O \rightarrow \text{phosphoribosyl-1-amine} + \text{glutamate} + \text{pyrophosphate}$$

The amidotransferase is the only enzyme in the 10-step pathway to inosine monophosphate (IMP) that is regulated by purine nucleotides. Regulation occurs through feedback inhibition by adenosine and guanosine mono- and diphosphate end products of purine biosynthesis (*3–5*). *Bacillus subtilis* glutamine PRPP amidotransferase is a homotetramer of subunits, each with 465 amino acid residues. Thus the glutamine and transferase domains reside on one polypeptide chain. Amino acid sequences of six glutamine PRPP amidotransferases have been derived from cloned genes (*6, 7*). Pairs of amino acid sequences are 31 to 93 percent identical, and 23 percent of the residues are invariant in the six sequences.

The amidotransferase from *B. subtilis* contains in each subunit of the tetramer a [4Fe-4S] cluster, which has regulatory but not catalytic function (*8*). The structural integrity of the cluster and of the protein are interdependent. It has been proposed that *B. subtilis* regulates de novo purine biosynthesis in part by controlling the oxygen-dependent inactivation of the enzyme, a signal for proteolytic degradation. Oxygen-dependent inactivation leads to cluster decomposition and irreversible enzyme denaturation in vitro (*9*), and stability of the [4Fe-4S] cluster is a key determinant of enzyme half-life in vivo (*10*). Sensitivity to air has precluded the purification and characterization of stable enzyme preparations from vertebrates. However, the *E. coli* enzyme, which is not a metalloprotein, has been purified. Perhaps the strongest evidence that the [4Fe-4S] cluster does not have a cryptic catalytic function is the 35 to 43 percent sequence identity between the *B. subtilis* enzyme and those of other species, regardless of whether or not the proteins have metal centers.

PRPP amidotransferase is distinct from several other nonredox proteins with [4Fe-4S] clusters (*8*), including hydro-lyases such as aconitase, in which three of the Fe atoms are cysteine-liganded and the fourth interacts with substrate (*11*). An intact [4Fe-4S] cluster is required for hydro-lyase activity. Some of the [4Fe-4S] hydro-lyases have been characterized in an inactive form with a stable [3Fe-4S] cluster; in other cases oxygen causes destruction of the cluster. The [4Fe-4S] hydro-lyases have Mössbauer spectra distinct from that of glutamine PRPP amidotransferase (*12*), and evidence for a correlation between hydro-lyase stability and cluster integrity has been lacking. A regulatory role has been established for the [4Fe-4S] cluster in the iron-responsive element (IRE)–binding protein, which is an aconitase when it contains an intact [4Fe-4S] cluster and an RNA-binding regulatory protein without the cluster (*13*). Another class of nonredox [4Fe-4S] proteins is the DNA *N*-glycosylases (*14*), represented by endonuclease III and adenine glycosylase. The native [4Fe-4S] cluster is ligated by four cysteine residues (*15*), is not sensitive to oxy-

J. L. Smith, E. J. Zaluzec, J.-P. Wery, and L. Niu are in the Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA. R. L. Switzer is in the Department of Biochemistry, University of Illinois, Urbana, IL 61801, USA. H. Zalkin is in the Department of Biochemistry, Purdue University, West Lafayette, IN 47907, USA. Y. Satow is with the Faculty of Pharmaceutical Sciences, University of Tokyo, Tokyo, Japan.

*To whom correspondence should be addressed.
†Present address: Department of Biochemistry, Michigan State University, East Lansing, MI 48824, USA.
‡Present address: Eli Lilly & Company, Indianapolis, IN 46285, USA.
§Permanent address: Department of Biology, University of Science and Technology, Heifei, Anhui, China.

gen, and appears to maintain the structure of the protein domain to which it is bound.

The [4Fe-4S] cluster in glutamine PRPP amidotransferase was the basis for our determination of the crystal structure of the enzyme. Experimental phase information was derived solely from MAD by the Fe atoms. The advent of tunable sources of synchrotron radiation has led to the development of MAD as an important new tool in macromolecular crystallography. The method is based on "anomalous" scattering, which is caused by resonance frequencies of the scattering atoms. It exploits differences in scattering between wavelengths near a resonance frequency of a subset of atoms in the crystal, in this case Fe, for direct determination of crystallographic phases. Crystal structures of several small (less than 30 kD) proteins have been solved by MAD (*16*). Our work demonstrates the applicability of MAD to the structure determination of a large (200 kD) protein.

**Structure determination.** Experimental phases for the structure determination came solely from multiwavelength anomalous diffraction of the Fe atoms in the four [4Fe-4S] clusters of the tetramer (Table 1). The MAD phasing signal from Fe is moderate (5 to 6 percent of the structure amplitude $|F|$) when the [4Fe-4S] clusters scatter as single centers but falls to less than 3 percent of $|F|$ for data beyond 5 Å spacings, as the irons begin to scatter as individual atoms. Wavelengths for data collection were carefully selected to maximize this weak signal. The experimental phase set was obtained by (i) determination and refinement of the partial structure of Fe atoms and (ii) computation of phase probabilities and structure amplitudes for the total structure. Normal structure amplitudes for the Fe partial structure were derived by least squares fit of the sets of multiwavelength measurements for each reflection to the MAD observational equation (*17*). The Fe partial structure was solved by Patterson analysis and refined against these amplitudes. Phase information from the least squares fit was not used. Rather, phase probability coefficients (*18*) for every observation in the multiwavelength data set were computed from the observations and the refined partial structure of Fe. Phase probabilities were combined for all observations of each unique reflection from all wavelengths, crystals and asymmetric units to produce the final set of experimental phases to 3.0 Å spacings. Experimental phases were improved by fourfold noncrystallographic symmetry averaging and solvent flattening of electron density maps beginning at 5.5 Å resolution and extending gradually to 3.0 Å. The final phase-refined 3.0 Å map (Fig. 1) had a clear trace of the polypeptide chain throughout. We were thus able to build an accurate and complete atomic model with an initial R factor of 34 percent against all data to 3.0 Å. The model was subsequently refined to an R factor of 19 percent (*19*).

Our structure determination of glutamine PRPP amidotransferase demonstrates that MAD phasing readily produces an accurate structural result for proteins as large as 200 kD from a signal as weak as 5 percent of $|F|$. The error in MAD experimental phases was 55° for data to 5.5 Å spacings, where the [4Fe-4S] clusters scattered predominantly as single centers. Beyond 5.5 Å, the cluster scattering was rapidly lost and phase errors increased as the MAD signal diminished to 2 percent, substantially below the level of experimental error (Table 1). Electron density maps at 5.5 Å resolution cannot be interpreted in terms of a detailed atomic structure, and the phase refinement was critical to produc-

**Table 1.** Crystallographic data. Crystals were grown in an inert atmosphere and diffracted to at least 2.3 Å spacings (*29*). Collection of the MAD data collection was limited to 3.0 Å spacings by constraints of detector geometry and synchrotron beam time. MAD data were recorded at Photon Factory beam line 14A by the oscillation method on an imaging plate detector (*30*). The fluorescence spectrum of an amidotransferase crystal was used to determine scattering factors (*31*) and to select optimal wavelengths for data collection, specifically, the inflection point of the Fe K edge (1.7425 Å), the maximum of the first peak above the edge (1.7390 Å), and a remote point at higher energy (1.5000 Å). Friedel-pair data were measured at each of the three wavelengths in an experiment designed to record all measurements contributing to the phase of each reflection from the same asymmetric unit of the same crystal at nearly the same time. Data from eight crystals were processed with the program DENZO (*32*); only fully recorded reflections were kept. Friedel-pair MAD data were scaled, phased and merged with the MADSYS programs (*31, 33*), which were modified to accommodate the multiple-crystal data set. Because the overlap between adjacent oscillation images (0.15°) was insufficient for the crystal mosaicity (average = 0.48°), a significant proportion of the simultaneously measured, multiwavelength reflection sets was not complete enough for phasing. The data that could be phased were adequate for determination and refinement of the Fe partial structure, initially for cluster positions with data to 5.5 Å and later for positions of individual Fe atoms with data to 3.0 Å. Final experimental amplitudes were the averages of all observations of each unique reflection, and phase distributions were calculated as described. The correct enantiomorph for the structure was chosen by correlation of electron densities about the *D*2 molecular symmetry operator; the correlation coefficient for the initial 5.5 Å MAD map in the correct hand was 0.49 and in the incorrect hand was 0.09.

| Wavelength (Å) | Anomalous scattering factors | | Diffraction ratios* for data to $d_{min}$ = 5 Å at (Å) | | | Unmerged data ($d_{min}$ = 3.0 Å) | Merged data ($d_{min}$ = 3.0 Å) | |
|---|---|---|---|---|---|---|---|---|
| | f′ ($e^-$) | f″ ($e^-$) | 1.7425 | 1.7390 | 1.5000 | Total MAD observations (N) | Unique reflections (N) | Completeness (%) |
| 1.7425 | −7.864 | 2.304 | 0.060 (0.033) | 0.028 (0.011) | 0.053 (0.050) | 125,462 | 36,761 | 80.3 |
| 1.7390 | −6.257 | 4.053 | | 0.071 (0.058) | 0.046 (0.038) | 116,349 | 35,922 | 78.4 |
| 1.5000 | −0.897 | 3.065 | | | 0.053 (0.044) | 109,796 | 34,801 | 76.0 |
| Overall | | | | | | 351,607 (101,564 multi-wavelength sets) | 41,267 | 90.1 |

| $d_{min}$ (Å) | $R_{sym}$† (%) | Maximum MAD signal‡ (%) | MAD ⟨fom⟩§ | Phase refined ⟨fom⟩ | MAD phase error‖ | $R$ factor¶ (%) |
|---|---|---|---|---|---|---|
| 5.5 | 4.7–5.5 | 4.4–5.2 | 0.46 | 0.82 | 55° | 24.9# |
| 4.3 | 4.7–8.0 | 2.2–2.6 | 0.28 | 0.81 | 67° | 15.0 |
| 3.7 | 5.2–7.4 | 1.7–2.0 | 0.19 | 0.79 | 77° | 15.7 |
| 3.3 | 6.2–9.4 | 2.1–2.4 | 0.19 | 0.76 | 82° | 18.4 |
| 3.0 | 9.4–13.6 | 3.3–3.8 | 0.20 | 0.72 | 83° | 22.4 |
| Overall | 5.5–7.6 | 2.7–3.1 | 0.25 | 0.78 | 73° | 18.2 |

*Diffraction ratio = $\langle||F_{\lambda 1}| - |F_{\lambda 2}||\rangle/\langle|^0F|\rangle$, or $\langle|I F_\lambda^+| - |F_\lambda^-|\rangle/\langle|^0F|\rangle$, where $|F_\lambda|$'s are total structure amplitudes including anomolous scattering, and $|^0F|$'s are for normal scattering only. In parentheses are theoretical values calculated for $d = \infty$. †$R_{sym} = (\Sigma_h \Sigma_i ||I_{hi} - \langle I\rangle_h|/\Sigma_h \langle I\rangle_h)$; values include the MAD signal; ranges reflect data from different wavelengths. ‡MAD signal =

$$(1/\sqrt{2})(\text{avg}\,||F'_{Fe,\lambda 1}| - |F'_{Fe,\lambda 2}||/\text{avg}|F_0|), \text{ or } (1/\sqrt{2})[\text{avg}(2|F''_{Fe,\lambda}|)/\text{avg}|F_0|],$$

where $F'_{Fe}$ and $F''_{Fe}$ are calculated from the Fe atomic structure, $|F_0|$ are the experimental structure amplitudes, and wavelengths (λ) are chosen to give the largest signal. §⟨fom⟩ is the mean figure-of-merit (cosine of estimated phase error). ‖Average phase discrepancy relative to the refined model. ¶$R = \Sigma\,||F_0| - |F_c||/\Sigma|F_0|$ for data with $|F_0| > 2\sigma(|F_0|)$, where $F_c$ are calculated from the refined model. #7.0 to 5.5 Å.

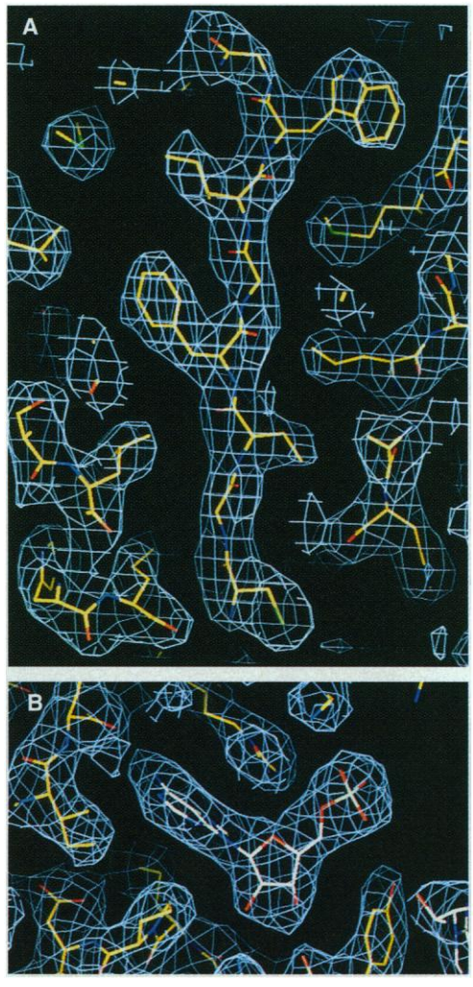


**Fig. 1.** Electron density at 3.0 Å resolution for (**A**) a β strand and for (**B**) one AMP molecule in the final experimental map. Contours are drawn at the rms value of the electron density; no phase information from the atomic model was used in computing this map. A starting MAD electron density map at 5.5 Å resolution was improved with 10 cycles of fourfold averaging and solvent flattening. Phase extension was in 25 steps of one reciprocal lattice unit with five cycles of averaging and solvent flattening in each step (*34*). At each step, experimental MAD phases were combined with calculated phases during the first cycle only, for the newly added data only. The molecular envelope was constructed by an automatic procedure from the 5.5 Å MAD map. The D2 molecular symmetry operator was determined from the refined Fe positions and was refined by least squares fit of rotated and unrotated densities in unaveraged maps. Density is continuous in the phase-refined map for all of the protein backbone with the exception of the glycine-rich loop 76-78 and the adjacent COOH-terminal peptide 461 to 465. Residue 402 was reported to be glycine but has strong side chain density bridging to the side chain of $Arg^{405}$. An aspartate side chain was built into this position.

ing interpretable electron density at 3.0 Å resolution. The quality of the 3.0 Å map (Fig. 1) also suggests that errors in experimental phases from MAD may be less systematic, and thus more easily removed by phase refinement, than are errors due to lack of isomorphism in experimental phases from isomorphous replacement or errors due to model bias in phases from molecular replacement.

The amidotransferase tetramer is a doughnut-shaped molecule of approximate dimensions 105 by 95 by 55 Å (Fig. 2). The four subunits occupy corners of the molecule and are related by mutually perpendicular twofold axes, which we have labeled P, Q, and R. There are no contacts between subunits related by the molecular Q axis, which is perpendicular to the doughnut plane. Contacts along the molecular P axis are much more extensive than are those along the R axis. In solution the amidotransferase equilibrates between dimeric and tetrameric aggregation states (*20*). We infer from the structure that the P-axis dimer is the solution dimer.

Each amidotransferase subunit is organized in two domains of approximately equal size but with distinctive folds (Fig. 3). The $NH_2$-domain (residues 1 to 230) is folded in a four-layer structure with two antiparallel β sheets sandwiched between layers of α helices. The front and back sheets (Fig. 3A) have, respectively, seven and six β strands. This fold has, to our knowledge, not been described; antiparallel β structure is not normally encased in other protein secondary structure. The $NH_2$-domain is reminiscent of the structure of deoxyribonuclease I (DNase I) (*21*), which has two mostly antiparallel sheets surrounded by helices. However, the topologies of the DNase I and amidotransferase domains are unrelated, and the amidotransferase fold has no pseudosymmetry.

The COOH-domain (residues 231 to 465) of the amidotransferase subunit has a smaller secondary structural core and more extensive loops than those of the $NH_2$-domain. The core of the COOH-domain is a five-stranded parallel β sheet packed between α helices. The topology is like that of the familiar six-stranded nucleotide-binding fold except that amidotransferase lacks the third β strand and α helix. A 30-residue insertion, which we call the "flag," inter-

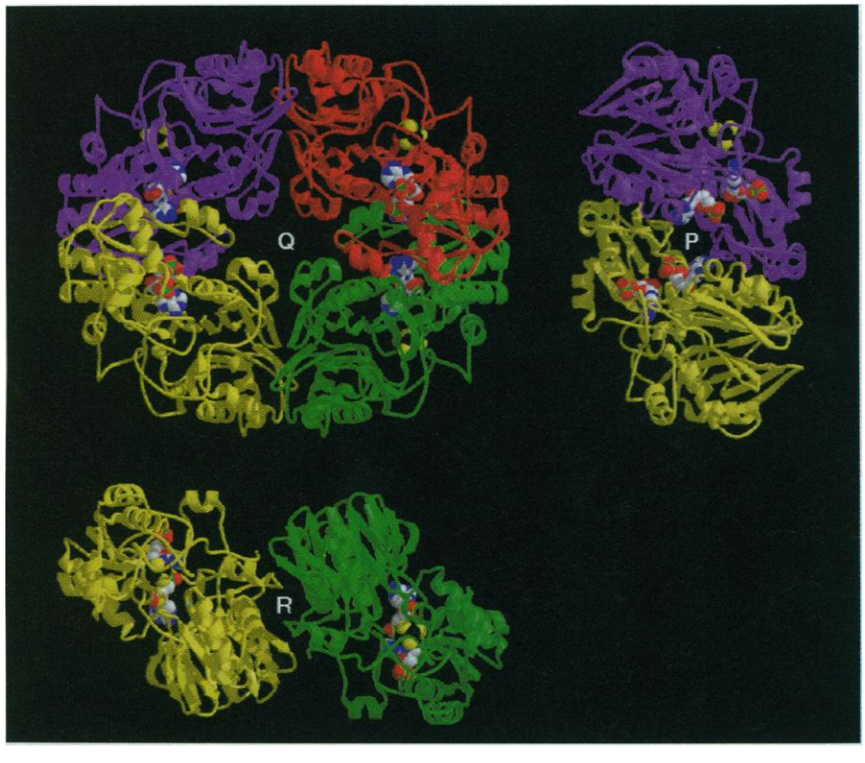


**Fig. 2.** Ribbon diagrams of glutamine PRPP amidotransferase (*35*). The central diagram is of the tetramer along the molecular Q axis. Perpendicular views of dimer association along the molecular P axis and R axis are shown in the upper right and lower left, respectively. Individual subunits are separately colored. Van der Waals spheres are shown for atoms of the [4Fe-4S] clusters and AMP molecules.

rupts the COOH-domain after the second β strand and protrudes from the COOH-terminal edge of the β sheet. The flag staff is a pair of antiparallel β strands connected by the flag itself, which is a long loop with a short α helix at one end. The "flag" makes all of the contacts with the $NH_2$-domain of a neighboring subunit related by the molecular P axis and may be involved in feedback regulation of enzyme activity (see below).

**[4Fe-4S] cluster.** The four [4Fe-4S] clusters are located at the corners of the tetramer between the $NH_2$- and COOH-domains of each subunit (Fig. 2). Its position is consistent with a noncatalytic role for the cluster. The Sγ of catalytic $Cys^1$ is 13 Å from the nearest cluster atom, and C1′ of AMP in the catalytic site is 11 Å from the cluster. The cluster has the expected cubane structure with three sulfide ligands to each Fe. Additional ligation is provided by four cysteine S atoms from the protein (Fig. 4A): $Cys^{236}$, $Cys^{382}$, $Cys^{437}$, and $Cys^{440}$. The protein environment around the cluster is relatively hydrophobic, although there are a typical number of hydrogen bonds from peptide nitrogens to sulfide and cysteine Sγ atoms (*22*).

Ligation by $Cys^{236}$ was not expected; rather $Cys^{434}$ was predicted to be an Fe ligand because the clustering of Cys residues at positions 434, 437, and 440 matches the Cys-X-X-Cys-X-X-Cys pattern for Fe ligands in ferredoxins. $Cys^{434}$, which is not an Fe ligand, occurs in a well ordered loop on the outside of the tetramer. The protein can accommodate substantial structural variability in this loop, evidenced by a mutant with rearranged cysteinyl ligands (*23*) and by the variable length and character of the loop among glutamine PRPP amidotransferases from different species (Fig. 4B). $Cys^1$ and the four cysteinyl ligands to the Fe-S cluster in the *B. subtilis* enzyme are the only conserved cysteine residues among metal-containing glutamine PRPP amidotransferases, leading to the conclusion that the four conserved cysteines in the transferase domain function as ligands to a [4Fe-4S] cluster in the vertebrate enzymes as well. Evidence for an [Fe-S] cluster in avian and mammalian amidotransferase has been summarized (*1*).

The cluster has been proposed to function as an oxygen sensor in a mechanism to detect nutrient limitation, leading to sporulation (*8*). The cluster is shielded from bulk solvent by the protein except for one sulfide, which is accessible through a small channel (7 Å wide and 7 Å deep). The channel seems to be the most probable route to the cluster for a destructive oxygen molecule. In the crystal structure, access to the channel is blocked by the adenine base of the AMP molecule that is bound in the presumed active site of the enzyme (Fig. 4C), consistent with biochemical data on oxygen sensitivity. AMP strongly protects *B. subtilis* amidotransferase from oxygen-dependent degradation (*9*). The substrate PRPP prevents AMP from protecting the enzyme against oxygen but does not itself protect, indicating that the adenine base of AMP blocks access to the cluster. No atom of AMP directly contacts the cluster or any of its ligands. The location of the cluster in the domain-domain boundary and the position of ligand $Cys^{236}$ in the peptide linking the domains suggest that loss of cysteinyl ligation upon oxygen attack may "unhinge" the protein tertiary structure, accounting for the dependence on the cluster for structural integrity of the protein.

The $NH_2$-domain of glutamine PRPP amidotransferase is the glutamine domain. Several residues in the $NH_2$-domain are implicated in glutamine utilization. Foremost among these is the catalytic residue $Cys^1$. Removal of the sulfhydryl function by mutagenesis or chemical modification results in an enzyme that functions to transfer $NH_3$ to PRPP but cannot utilize glutamine as the $NH_3$ source (*24*). Other residues believed to function in glutamine utilization have been identified by comparison of amino acid sequences of *purF*-type glutamine domains (*1*). $Cys^1$ and most of the other conserved residues are located on the lower edge of the $NH_2$-domain, facing the COOH-domain. The side chain of $Cys^1$ is buried in a cavity within the $NH_2$-domain. This cavity has approximate dimensions of 12 by 10 by 9 Å and is large enough to accommodate a glutamine residue, or a glutamyl-ε-thiol ester modification of the $Cys^1$ side chain. The cavity is lined with conserved residues, including $His^{70}$, $Arg^{72}$, $Tyr^{73}$, $Thr^{75}$, $Gln^{85}$, $Ala^{100}$, $His^{101}$, $Asn^{102}$, $Gly^{103}$, $Ser^{126}$, $Asp^{127}$, $Gly^{385}$, and $Ile^{386}$.

**Nucleotide binding sites.** Eight molecules of the feedback inhibitor AMP are bound to the amidotransferase tetramer in two distinct types of binding sites. The sites are designated as C, for catalytic, and R, for regulatory. Each subunit has one C site, which appears to be the transferase active

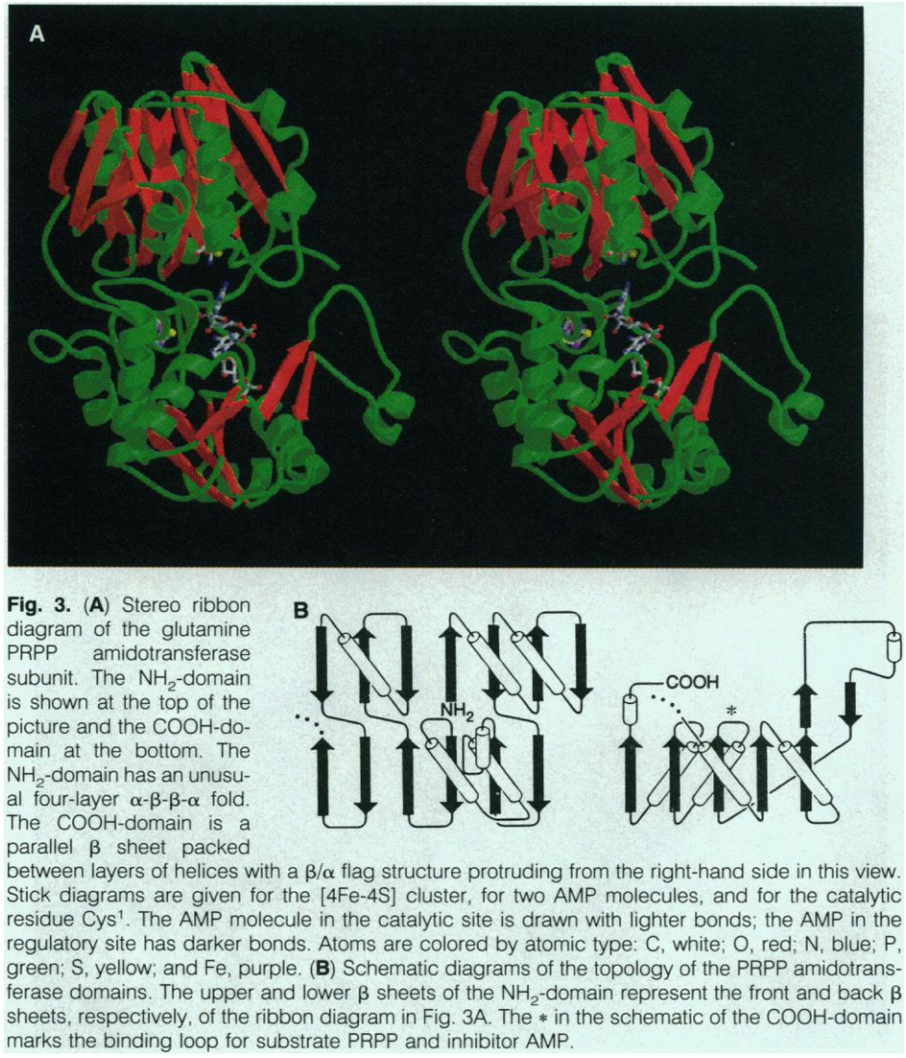


**Fig. 3.** (**A**) Stereo ribbon diagram of the glutamine PRPP amidotransferase subunit. The $NH_2$-domain is shown at the top of the picture and the COOH-domain at the bottom. The $NH_2$-domain has an unusual four-layer α-β-β-α fold. The COOH-domain is a parallel β sheet packed between layers of helices with a β/α flag structure protruding from the right-hand side in this view. Stick diagrams are given for the [4Fe-4S] cluster, for two AMP molecules, and for the catalytic residue $Cys^1$. The AMP molecule in the catalytic site is drawn with lighter bonds; the AMP in the regulatory site has darker bonds. Atoms are colored by atomic type: C, white; O, red; N, blue; P, green; S, yellow; and Fe, purple. (**B**) Schematic diagrams of the topology of the PRPP amidotransferase domains. The upper and lower β sheets of the $NH_2$-domain represent the front and back β sheets, respectively, of the ribbon diagram in Fig. 3A. The * in the schematic of the COOH-domain marks the binding loop for substrate PRPP and inhibitor AMP.

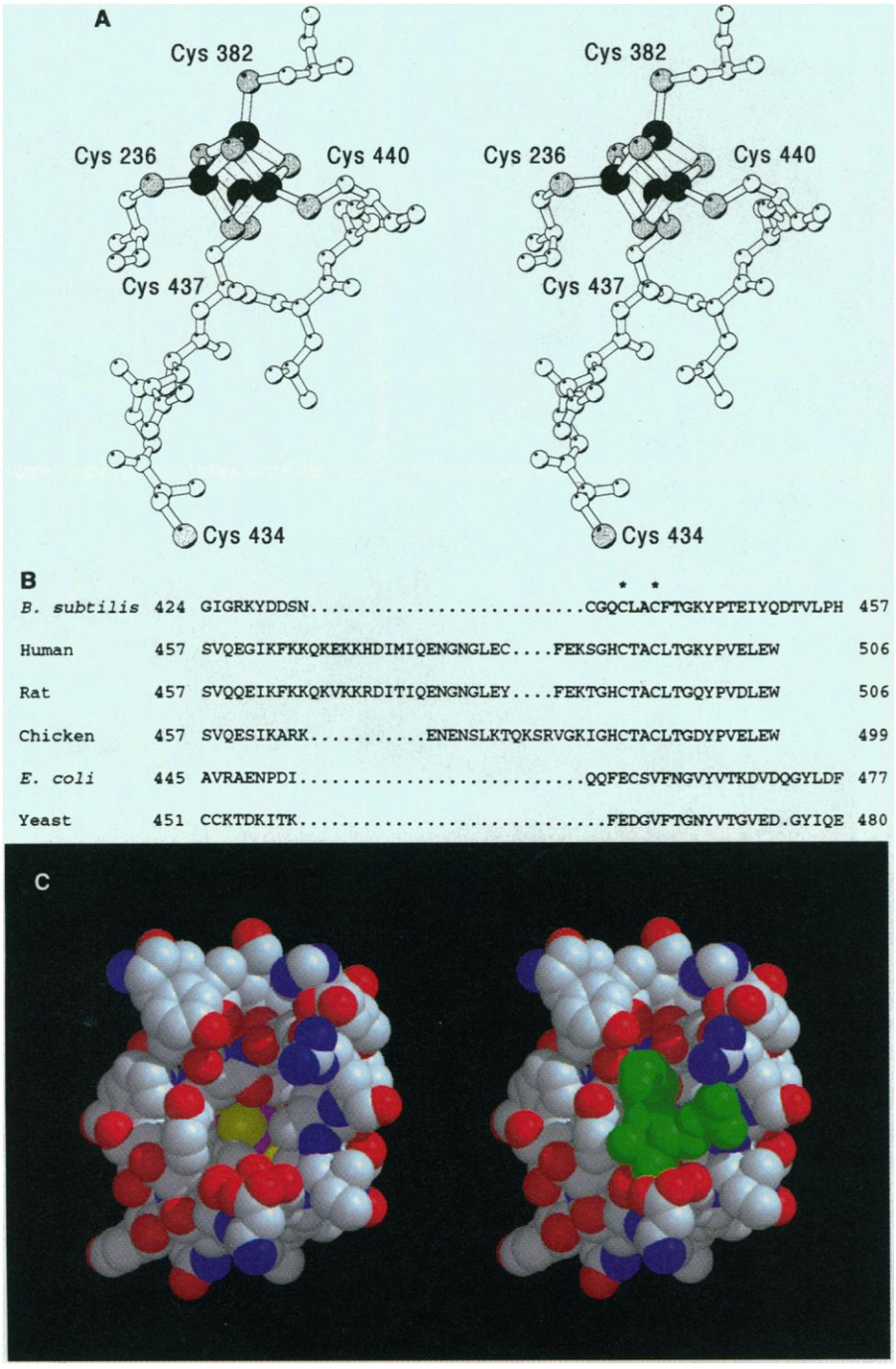


**Fig. 4.** (**A**) Stereo diagram of the [4Fe-4S] cluster in glutamine PRPP amidotransferase showing the four cysteinyl ligands and $Cys^{434}$, which was predicted to be an Fe ligand. (**B**) Alignment of glutamine PRPP amidotransferase sequences near clustered Fe ligands showing conservation of Cys residues in the iron-containing enzymes (*36*). This alignment differs from early published alignments (*7*), which differ from each other in this region. The highly variable peptide between the final α helix (residues 417 to 426) in the structure and the last two cysteinyl Fe ligands (residues 437 and 440) is a well ordered loop on the outside of the *B. subtilis* molecule. Seven of the 35 total differences between the 93 percent identical human and rat sequences occur between residues 460 and 489. Sequence numbers are given for the first and last residues of each peptide; "." indicates a gap in one sequence relative to the others; Cys ligands to Fe atoms in the *B. subtilis* structure are marked with "*". (**C**) Space-filling diagrams of the accessible sulfide channel leading from bulk solvent to the [4Fe-4S] cluster in the amidotransferase. The channel is shown without (left) and with (right) the AMP molecule that blocks access to the cluster in the crystal structure and protects the enzyme from oxygen-dependent inactivation. Atomic van der Waals surfaces of the protein and cluster are colored according to atomic type with protein C, white; N, blue; O, red; S, yellow; and Fe, magenta; the AMP is green.

site. The four R sites are between subunits along the molecular P axis. All amino acid residues in the R and C sites are from the COOH-domain of the polypeptide. Neither the R site nor the C site exactly fits structural paradigms for nucleotide binding to proteins. Amidotransferase does not have a Gly-X-X-Gly-X-Gly sequence motif, which has been found in mononucleotide binding sites of many proteins (*25*), and the nucleotide binding sites in the amidotransferase do not include glycine-rich peptides.

The C site illustrates the likely binding mode of PRPP to the PRTase enzyme family. It consists of a strand-loop-helix structure in the core of the COOH-domain (Fig. 5A). The strand and loop contain a sequence motif (residues 341 to 353) that is common to several PRTases and has been proposed as a PRPP binding site (*26*). Residues in the PRPP fingerprint encircle the ribose 5′-phosphate of AMP. The acidic side chains $Asp^{345}$ and $Asp^{346}$, a remarkable feature of the PRPP fingerprint considering the net negative charge of PRPP, are hydrogen bonded to the ribose O3′-hydroxyl of AMP. The ribose ring is best fit to the electron density with a C3′-endo pucker. The C site could easily accommodate PRPP bound in the same position and orientation as the ribose 5′-phosphate of AMP. The adenine base of AMP in the C site is sandwiched between the side chains of $Tyr^{242}$ and $Val^{349}$ but forms no hydrogen bonds with the protein.

Although the primary function of the C site is binding of the substrate PRPP, it has some features in common with mononucleotide binding sites in other proteins (*25*). The C site is near the COOH-terminal edge of the parallel β sheet, with AMP bound to the loop between a central β strand and an α helix at a crossover point in the β sheet. However, in contrast to other mononucleotide binding sites, the β strand is not the first strand of the COOH-domain (Fig. 3B). The 5′-phosphate sits directly over the $NH_2$-terminus of the α helix. Peptide dipoles from the loop and second residue of the α helix are oriented to make favorable electrostatic interactions with the phosphate, which also contacts the charged side chain of $Arg^{350}$ in the loop. Unlike other mononucleotide binding sites, the ribose O2′ and O3′ directly contact residues in the loop, in line with the function of the site to orient PRPP for catalysis at the ribose C1′ atom.

The flag loop (residues 302 to 332) figures prominently in the nucleotide R site, which is between subunits related by the molecular P axis (Fig. 5B). The AMP molecule in the R site contacts the conserved loop 240 to 246, $Lys^{305}$, and $Lys^{328}$ in the primary subunit and, in the secondary subunit, the side chain of $Arg^{307}$ and

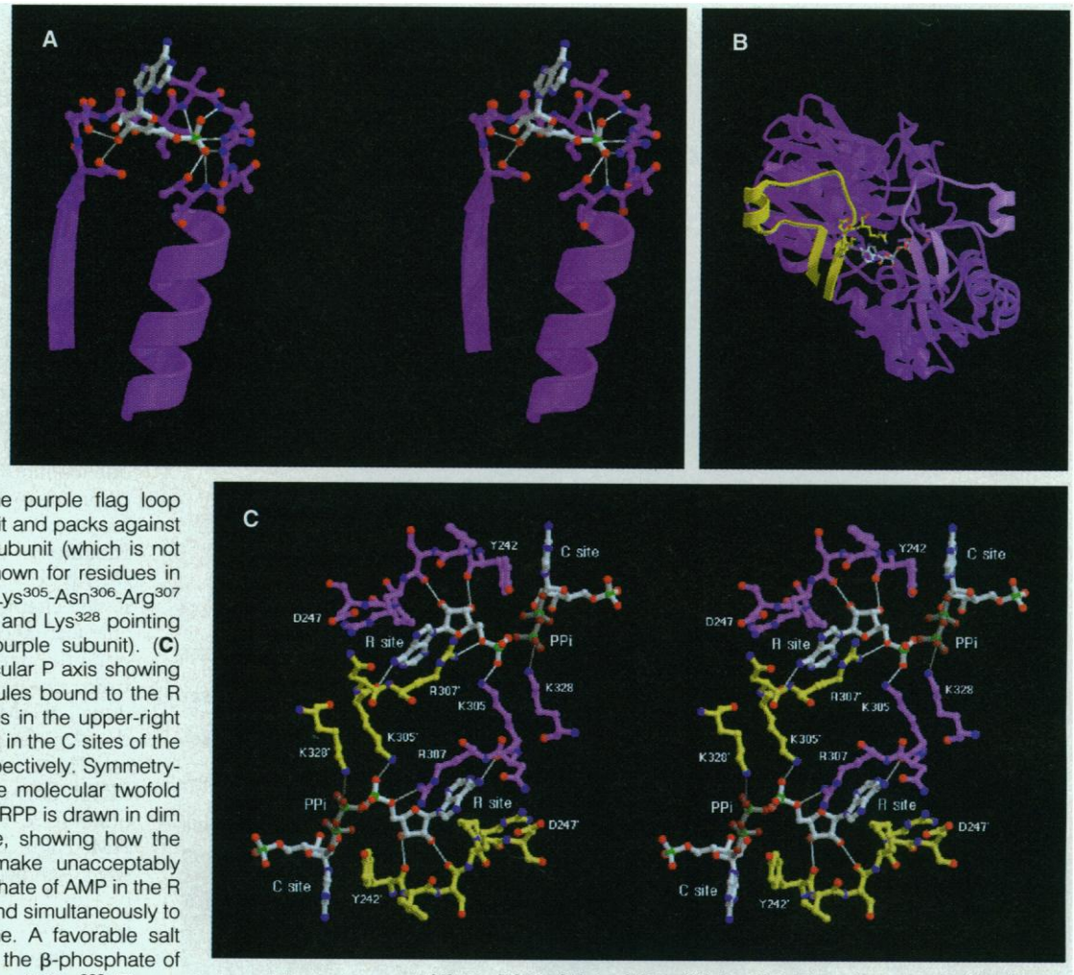


**Fig. 5.** (**A**) Stereo diagram of the catalytic (C) site with inhibitor AMP bound. The peptide shown is residues 340 to 365; atomic positions are shown for $Asp^{345}$-$Asp^{346}$-$Ser^{347}$-$Ile^{348}$-$Val^{349}$-$Arg^{350}$-$Gly^{351}$-$Thr^{352}$-$Thr^{353}$. The PRPP-binding fingerprint is residues 341 to 353. Hydrogen bonds are drawn in thin white lines. (**B**) Nucleotide regulatory (R) site with inhibitor AMP. Flag loops from the purple and yellow subunits, colored with reference to Fig. 2, both contribute residues to a single R site and are related by the molecular P axis. The flag loop from the yellow subunit is packed against the $NH_2$-domain of the purple subunit; the purple flag loop protrudes from the purple subunit and packs against the $NH_2$-domain of the yellow subunit (which is not shown). Atomic positions are shown for residues in the flag loop that contact AMP ($Lys^{305}$-$Asn^{306}$-$Arg^{307}$ in the yellow subunit and $Lys^{305}$ and $Lys^{328}$ pointing away from the viewer in the purple subunit). (**C**) Stereo diagram along the molecular P axis showing the positions of the AMP molecules bound to the R and C sites. The AMP molecules in the upper-right and lower-left corners are bound in the C sites of the purple and yellow subunits, respectively. Symmetry-related R sites are closer to the molecular twofold axis. The α1-pyrophosphate of PRPP is drawn in dim bonds, modeled into the C site, showing how the β-phosphate of PRPP might make unacceptably close contacts with the 5′-phosphate of AMP in the R site, if PRPP and AMP were to bind simultaneously to this conformation of the enzyme. A favorable salt bridge may be formed between the β-phosphate of PRPP and the conserved amino acid $Lys^{328}$. The central peptides in the drawing, $Lys^{305}$-$Asn^{306}$-$Arg^{307}$ and $Lys^{328}$, are from flag loops of the purple and yellow subunits. The outer peptide from each subunit is $Try^{242}$-$Phe^{243}$-$Ser^{244}$-$Arg^{245}$-$Pro^{246}$-$Asp^{247}$. For all parts of this figure, protein carbon atoms are colored according to their subunit of origin; AMP carbon atoms are white; other atoms are colored as in Fig. 3A.

the backbone O of $Lys^{305}$. The contact with the O of $Lys^{305}$ is a nucleotide-specific hydrogen bond to N6 of the adenine base. The peptide $Lys^{305}$-$Asn^{306}$-$Arg^{307}$ in the flag loop contacts AMP molecules in two regulatory sites. The long, positively charged side chains of $Lys^{305}$ and $Arg^{307}$ are extended perpendicular to one another and contact 5′-phosphates of AMP molecules in separate R sites (Fig. 5C).

Binding of the inhibitor AMP to both the active site of the amidotransferase and to a second site explains why the kinetics of AMP inhibition are complex. Specifically, the kinetics are highly cooperative, and they are competitive with respect to PRPP at low AMP concentrations but are not competitive at high AMP concentrations (*4*). When AMP is bound to the C site, the enzyme is obviously inhibited by direct steric blockage. However, the R site for AMP appears also to overlap part of the catalytic site for PRPP because of the proximity and positions of the C and R sites. The α-pyrophosphoryl group of PRPP bound in the same orientation as the ribose 5-phosphate moiety of AMP in the C site would overlap the 5′-phosphate of AMP in the R site (Fig. 5C). This suggests that the R site could function as a site for competitive inhibition by AMP. However, glutamine PRPP amidotransferase is a classic allosteric enzyme by biochemical criteria, and the finding of adjacent regulatory and catalytic sites was unexpected.

The structural hallmark of allosteric regulation in several well studied systems is a quaternary structural change in response to effector binding (*27*). The binding site for an allosteric effector by definition does not overlap the active center, is generally between subunits, and is associated with cooperativity in effector, substrate or ligand binding. Although the R site for AMP is contiguous with the active site in the amidotransferase, it is between subunits of the tetrameric enzyme, and AMP binding is cooperative, as is binding of substrate PRPP in the presence of inhibitory nucleotides (*4*, *5*). These data suggest that AMP binding to the R site may be associated with an allosteric structural transition.

Cooperativity in AMP binding at R sites is readily understood, since these sites are formed at the dimer interface and include residues contributed by both subunits. The positively charged side chains of $Lys^{305}$, $Lys^{328}$, and $Arg^{307}$ would not be brought so close together without interaction with the 5′-phosphate of AMP. Cooperative interactions between nucleotides bound at the R and the C sites are also implied by the large Hill coefficients [3.3 to 3.8 for AMP (*4*)] and by synergistic inhibition exhibited by some pairs of nucleotides (*4*, *5*). The structural basis for such R:C site interactions is not obvious; it may involve the aromatic interaction between the adenine of AMP in the C site and the side chain of $Tyr^{242}$, which is required for catalysis (*28*) and is in the 240 to 246 loop that forms part of the R site (Fig. 5C). Synergistic inhibition presumably also involves positive interactions between different nucleotides bound to the R and C sites.

AMP-inhibited glutamine PRPP amidotransferase seems poised for a structural transition akin to those of allosteric proteins. The catalytic half-sites of the glutamine and transferase domains are too far from one another for effective catalysis (Fig. 3A). The catalytic thiol of $Cys^1$ in the glutamine domain is 16 Å from the ribose C1′ of AMP in the transferase catalytic site. The C1′ edge of the ribose ring is not surrounded by protein groups of potential catalytic import; rather, it is fully solvent exposed. It appears that a large conformational change is required to juxtapose the catalytic half-sites and form an active enzyme. The most obvious structural transition, for which ample precedent exists, is a hinge bending between domains. However, subunit packing at the molecular P axis precludes domain movement in the inhibited structure. Here the AMP molecule in the R site and the flag loop appear to be critical players. The R site for AMP contains four residues from the flag loops of two subunits. The flag loop protrudes from the globular transferase domain of one subunit but is packed tightly against the glutamine domain of the neighboring subunit (Fig. 5B). Through interactions with AMP in the R site, the flag loop of one subunit appears to hold apart the glutamine and transferase domains of the other subunit. Our working hypothesis is that removal of the nucleotides, and perhaps also binding of the substrate PRPP, triggers a conformational change that places the glutamine and transferase halves of the active site close enough for catalysis. The involvement of the flag loop both in subunit interactions and in nucleotide binding suggests that the activating conformational change will also move the subunits relative to one another. Analysis of the structure of the active enzyme should confirm (or deny) this hypothesis and should provide details of the functionally important local structural changes accompanying activation.

## REFERENCES AND NOTES


1. H. Zalkin, *Adv. Enzymol. Relat. Areas Mol. Biol.* **66**, 203 (1993).
2. W. D. L. Musick, *CRC Crit. Rev. Biochem.* **11**, 1 (1981).
3. R. L. Switzer, in *Allosteric Enzymes*, G. Herve, Ed. (CRC Press, Boca Raton, FL, 1989), pp. 129–151.
4. E. Meyer and R. L. Switzer, *J. Biol. Chem.* **254**, 5397 (1979).
5. L. J. Messenger and H. Zalkin, *ibid.*, p. 3382.
6. For *B. subtilis*: C. A. Makaroff, H. Zalkin, R. L. Switzer, S. J. Vollmer, *ibid.* **258**, 10586 (1983). For *E. coli*: J. Y. Tso, H. Zalkin, M. van Cleemput, C. Yanofsky, J. M. Smith, *ibid.* **257**, 3525 (1982); G. Sampei and K. Mizobuchi, *Nucleic Acids Res.* **16**, 8717 (1988). For yeast: P. Mäntsälä and H. Zalkin, *J. Biol. Chem.* **259**, 8478 (1984). For human: K. A. Brayton *et al.*, *ibid.* **269**, 5313 (1994).
7. For chicken, G. Zhou, J. E. Dixon, H. Zalkin, *J. Biol. Chem.* **265**, 21152 (1990). For rat: H. Iwahana *et al.*, *ibid.* **268**, 7225 (1993). For human: H. Iwahana *et al.*, *Biochem. Biophys. Res. Commun.* **190**, 192 (1993).
8. R. L. Switzer, *BioFactors* **2**, 77 (1989).
9. D. A. Bernlohr and R. L. Switzer, *Biochemistry* **20**, 5675 (1981).
10. J. A. Grandoni, R. L. Switzer, C. A. Makaroff, H. Zalkin, *J. Biol. Chem.* **264**, 6058 (1989).
11. M. C. Kennedy and C. D. Stout, *Adv. Inorg. Chem.* **38**, 323 (1992).
12. B. A. Averill *et al.*, *J. Biol. Chem.* **255**, 6007 (1980).
13. R. D. Klausner, T. A. Rouault, J. B. Harford, *Cell* **72**, 19 (1993).
14. R. P. Cunningham *et al.*, *Biochemistry* **28**, 4450 (1989); M. L. Michaels, L. Pham, Y. Nghiem, C. Cruz, J. H. Miller, *Nucleic Acids Res.* **18**, 3841 (1990).
15. C.-F. Kuo *et al.*, *Science* **258**, 434 (1992).
16. To date, structures solved by MAD phasing have been crystallographic problems of 30 kD or less. Most are cited by J. L. Smith [*Curr. Opin. Struct. Biol.* **1**, 1002 (1991)] and by W. A. Hendrickson [*Science* **254**, 51 (1991)]. Others include those of W. I. Weis, R. Kahn, R. Fourme, K. Drickamer, and W. A. Hendrickson (*ibid.*, p. 1608); D. J. Leahy, W. A. Hendrickson, I. Aukhil, and H. P. Erickson [*ibid.* **258**, 987 (1992)]; V. Ramakrishnan, J. T. Finch, V. Graziano, P. L. Lee, and R. M. Sweet [*Nature* **362**, 219 (1993)].
17. W. A. Hendrickson, *Trans. Am. Crystallogr. Assoc.* **21**, 11 (1985).
18. A. Pähler, J. L. Smith, W. A. Hendrickson, *Acta Crystallogr.* **A46**, 537 (1990).
19. The model for one subunit was built into a skeletonized electron density map by means of the program O [T. A. Jones, J.-Y. Zou, S. W. Cowan, M. Kjeldgaard, *ibid.* **A47**, 110 (1991)], and the other three subunits were generated by the noncrystallographic symmetry operators. The initial model fit the 3.0 Å experimental map with a real-space correlation coefficient of 0.84 (calculated by program O with parameters $C = 0.75$ and $A_0 = 0.90$). Model refinement was performed with the program XPLOR [A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987)] with data between 7.0 and 3.0 Å with $|F| > 2\sigma(|F|)$. Exact molecular symmetry was enforced during three rounds of simulated-annealing refinement of one monomer. Rounds two and three were followed by refinement of group B factors for main chain and side chain atoms of each residue. In a final round of conventional crystallographic refinement, molecular symmetry was unrestrained. The final model, including 14,156 protein atoms, four [4Fe-4S] clusters, eight AMP molecules, and no solvent, has an *R* factor of 19.0 percent for all measured data between 7.0 and 3.0 Å and a root-mean-square (rms) deviation from ideality of 0.015 Å in covalent bonds and 3.3° in bond angles. The rms deviation of the model from *D*2 molecular symmetry is 0.39 Å. The real-space correlation coefficient of the refined model to the experimental map is 0.85. Coordinates are available in the Brookhaven Protein Databank with accession code 1GPH.
20. J. Y. Wong, D. A. Bernlohr, C. L. Turnbough Jr., R. L. Switzer, *Biochemistry* **20**, 5669 (1981).
21. D. Suck, C. Oefner, W. Kabsch, *EMBO J.* **3**, 2423 (1984).
22. G. Backes *et al.*, *J. Am. Chem. Soc.* **113**, 2055 (1991).
23. C. A. Makaroff, J. L. Paluh, H. Zalkin, *J. Biol. Chem.* **261**, 11416 (1986).
24. J. Y. Tso, M. A. Hermodson, H. Zalkin, *ibid.* **257**, 3532 (1982); S. J. Vollmer, R. L. Switzer, M. A. Hermodson, S. G. Bower, H. Zalkin, *ibid.* **258**, 10582 (1983); P. Mäntsälä and H. Zalkin, *ibid.* **259**, 14230 (1984); J. L. Souciet, M. A. Hermodson, H. Zalkin, *ibid.* **263**, 3323 (1988).
25. G. E. Schulz, *Curr. Opin. Struct. Biol.* **2**, 61 (1992); J. E. Walker, M. Saraste, M. J. Runswick, N. J. Gay, *EMBO J.* **1**, 945 (1982).
26. B. Hove-Jensen, K. W. Harlow, C. J. King, R. L. Switzer, *J. Biol. Chem.* **261**, 6765 (1986); H. V. Hershey and M. W. Taylor, *Gene* **43**, 287 (1986).
27. M. Perutz, *Mechanisms of Cooperativity and Allosteric Regulation in Proteins* (Cambridge Univ. Press, Cambridge, 1989).
28. G. Zhou, H. Charbonneau, R. F. Colman, H. Zalkin, *J. Biol. Chem.* **268**, 10471 (1993).
29. Purified protein was prepared from *B. subtilis* grown under conditions of purine starvation (*20*). Because of the oxygen sensitivity of the protein, crystals were grown by batch methods in an inert-atmosphere ($N_2$) glove box from solutions containing protein at 7 to 8 mg/ml, 50 mM tris-HCl (pH 7.9), 3 to 4 mM $MgCl_2$, 5 mM dithioerythritol (DTE), 2 mM AMP and 12 to 13 percent (w/v) polyethylene glycol (PEG) 8000 and were stabilized in protein-free solutions of the same buffer with 10 mM $MgCl_2$, 5 mM DTE, 2 mM AMP, and 15 percent PEG 8000. Crystals were polymorphic, predominantly of space group $P2_1$ with $a = 158.8$ Å, $b = 75.7$ Å, $c = 94.1$ Å, and $\beta = 91.4°$ with one tetramer of amidotransferase in the asymmetric unit.
30. J. L. Smith, E. J. Zaluzec, J.-P. Wery, Y. Satow, in *Isomorphous Replacement and Anomalous Scattering*, W. Wolf, P. R. Evans, A. G. W. Leslie, Eds. (Science and Engineering Research Council, Warrington, United Kingdom, 1991), pp. 96–106.
31. W. A. Hendrickson, J. L. Smith, R. P. Phizackerley, E. A. Merritt, *Proteins: Struct. Funct. Genet.* **4**, 77 (1988).
32. Z. Otwinowski, in *Data Collection and Processing*, L. Sawyer, N. Isaacs, S. Bailey, Eds. (Science and Engineering Research Council, Warrington, United Kingdom, 1993), pp. 56–62.
33. W. Yang, W. A. Hendrickson, R. J. Crouch, Y. Satow, *Science* **249**, 1398 (1990).
34. J. T. Bolin, J. L. Smith, S. W. Muchmore, *Abstracts of the American Crystallographic Association*, p. 51 (1993).
35. Figures 2, 3A, 4A, 4C, and 5 were computed with Molscript [P. J. Kraulis, *J. Appl. Cryst.* **24**, 946 (1991)]. Figures 2, 3A, 4C, and 5 were subsequently rendered with Raster3D [E. A. Merritt (merritt@u.washington.edu), personal communication].
36. Single-letter codes are used for amino acids as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.


37. We thank S. J. Vollmer for purification of the protein and J. T. Bolin for use of the inert-atmosphere box and for critical reading of the manuscript. Supported by NIH DK-42303 to JLS, by NIH GM-47112 to RLS, by NIH GM-24658 to HZ, and by a grant from the Lucille P. Markey Foundation.

7 December 1993; accepted 22 April 1994