Single Amino Acid Substitutions Disrupt Tetramer Formation in the Dihydroneopterin Aldolase Enzyme of *Pneumocystis carinii*†

M. Carmen Thomas,‡ Stuart P. Ballantine,‡ Susanne S. Bethell, Satty Bains, Paul Kellam,§ and Chris J. Delves*

Glaxo Wellcome Research and Development, Medicines Research Centre, 6S224, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY, U.K.

Received March 9, 1998; Revised Manuscript Received June 3, 1998

ABSTRACT: In the opportunistic pathogen *Pneumocystis carinii*, dihydroneopterin aldolase function is expressed as the N-terminal portion of the multifunctional folic acid synthesis protein (Fas). This region encompasses two domains, FasA and FasB, which are 27% amino acid identical. FasA and FasB also share significant amino acid sequence similarity with bacterial dihydroneopterin aldolases. In the present study, this enzyme function has been overproduced as an independent monofunctional activity in *Escherichia coli*. Recombinant FasAB-Met23 (amino acids 23−290 of the predicted open reading frame) was purified and shown to contain dihydroneopterin aldolase activity. The native FasAB-Met23 is a tetramer of the 30-kDa subunit, demonstrating characteristics of an associating−dissociating equilibrium system in which only the multimeric form of the enzyme is active. Multiple sequence alignment of FasA and FasB with other dihydroneopterin aldolases highlights only three positions where the amino acid is invariable between all the predicted proteins. The role of these conserved amino acid residues in enzyme function was investigated using site-directed mutagenesis. Mutant FasAB-Met23 species were overproduced and purified to near homogeneity. Three FasA domain mutants and two FasB domain mutants had little or no detectable dihydroneopterin aldolase activity, implicating both FasA and FasB in the catalytic mechanism. We show that each mutant protein containing an inactivating amino acid substitution has lost its ability to form stable tetramers.

Most microbial cells must synthesize folates de novo since they lack the carrier-mediated active transport system of mammalian cells which allows the use of preformed dietary folates. The de novo synthesis of folate involves several enzyme functions unique to microbial cells. With no mammalian counterparts these enzymes are attractive chemotherapeutic targets since the possibility exists of directing highly selective drugs against them for the treatment of microbial disease. Indeed, dihydropteroate synthase (DHPS),[1] which catalyzes the condensation of *p*-aminobenzoic acid (PABA) and 6-hydroxymethyl-7,8-dihydropterin pyrophosphate to form dihydropteroate, is the target of the sulfonamides, substrate analogues of PABA which have found widespread use as antimicrobial agents.

Dihydroneopterin aldolase (DHNA) is a key enzyme in the folate biosynthetic pathway, catalyzing the conversion of 7,8-dihydroneopterin to 6-hydroxymethyl-7,8-dihydropterin. Gene sequences coding for DHNA have been cloned thus far only from gram-positive bacteria and from the fungal pathogen *Pneumocystis carinii*. Amino acid sequence analysis and functional studies reveal marked diversity in DHNA structure and organization in the different species. DHNA homologues appear to be monofunctional proteins in *Staphylococcus aureus* (P.K., S.P.B. & C.J.D., unpublished) and *Bacillus subtilis* (*1*). In contrast, the *sulD* gene of *Streptococcus pneumoniae* expresses a bifunctional protein which contains both an N-terminal DHNA domain and a C-terminal 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (PPPK) domain (*2*). PPPK converts the product of the DHNA reaction to 6-hydroxymethyl-7,8-dihydropterin pyrophosphate. The greatest degree of complexity has been identified in *P. carinii*, with the cloning of a multifunctional folic acid synthesis gene (*fas*) which encodes DHNA, PPPK, and DHPS, three sequential enzymes in the pathway (*3*, *4*).

The 740 amino acid Fas polypeptide maps into four major domains (*4*), two of which, FasA (predicted residues 39−160) and FasB (residues 161−280), show significant sequence similarity to the bacterial DHNA enzymes and 27% amino acid sequence identity to each other (*5*). FasC (residues 291−449) and FasD (residues 463−740) contain the PPPK (*6*) and DHPS activities, respectively.

Characterization of the native form of the *P. carinii* Fas protein has been precluded by the difficulty of producing enough soluble full-length material in heterologous systems (*4*). Our approach to understanding the structure and properties of this multifunctional enzyme has been to express and isolate the individual domain sequences (*5*, *6*). Indi-

†This work was in part supported by a grant to M.C.T. from CSIC, Spain.

* Author to whom correspondence should be addressed. Phone: +44(0) 1438 764073. Fax: +44(0) 1438 764818. E-mail: cjd33356@glaxowellcome.co.uk.

‡These authors contributed equally to this work.

§ Current address: Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, U.K.

[1] Abbreviations: DHNA, dihydroneopterin aldolase; DHPS, dihydropteroate synthase; DMA, dimethyl adipimate; DMS, dimethyl suberimidate; Fas, folic acid synthesis protein; *fas*, gene encoding Fas; PABA, *p*-aminobenzoic acid; PPPK, dihydropterin pyrophosphokinase; PVDF, poly(vinylidene difluoride); SEC, size exclusion chromatography.

S0006-2960(98)00540-6 CCC: $15.00
Published on Web 07/24/1998

vidual domains of a wide range of multifunctional enzymes have been overproduced to yield active recombinant monofunctional proteins (*7*−*11*). Importantly, the kinetic properties of isolated domains appear unaltered from those of the native state (*7*, *10*). Previously, we showed that both FasA and FasB are required for DHNA function, and proposed two possibilities (*5*). First, both domains are subunits of the DHNA enzyme moiety. Alternatively, only one of the domains codes for DHNA with the synthesis of the other domain, of unidentified function, essential for stabilizing the activity, presumably by interdomain interactions. Two enzymes which catalyze consecutive steps in methionine biosynthesis of *E. coli* are 31% amino acid identical (*12*). Here, we describe the overexpression of *fasAB* sequences in *Escherichia coli* and the subsequent characterization of the purified recombinant species. To understand the functional relationship between FasA and FasB and their roles in DHNA function, we have undertaken site-directed mutagenesis studies. A number of FasAB species, each containing a single amino acid substitution in either the FasA domain or the FasB domain, were overproduced in *E. coli* and purified to homogeneity. Functional analysis of these site-specific mutant proteins is reported.

EXPERIMENTAL PROCEDURES

Expression Constructs. Three constructs were made which allowed the production of various FasAB protein species in *E. coli*. The structure of each *fas* construct was confirmed by DNA sequencing to ensure that no changes had occurred during manipulation and that the correct translational reading frame had been reformed. A recombinant plasmid pKS-.*fasAB*, which contains sequence coding for the N-terminal 290 residues of the predicted Fas protein ligated into between the unique *Bam*HI and *Hin*dIII sites of the vector BluescriptKS (Stratgene), has been described previously (*5*). Here, a suitable restriction fragment of pKS.*fasAB*, renamed pKS.*fasAB-Met1*, was utilized in the making of two further constructs. First, a construct was made from which the sequence coding for the N-terminal 22 amino acids of *fasAB*-Met1 was removed, allowing direct expression from a naturally occurring Met residue at position 23. DNA sequence coding for residues 23−57 was amplified from pKS.*fasAB-Met1* using two oligonucleotides. The primers were 5′-ATTAGGATCCATGATATTTAAGAAAAAAA-TCCATTTATCT, which contained nucleotides encoding residues 23−32 and created a *Bam*HI site (underlined) proximal to the initiation codon, and 5′-GCCCACGAAT-TCTTTCCTACAATAGATTTC, which contained bases complementary to nucleotides coding for residues 48−57 and recreated the natural *Eco*RI site (underlined) interrupting codon 55. Taq polymerase was from Perkin-Elmer Cetus and PCR was performed according to the manufacturer's instructions. The amplified fragment was digested with *Bam*HI and *Eco*RI and ligated to *Bam*HI/*Eco*RI-digested pKS.*fasAB-Met1*, to construct pKS.*fasAB-Met23*. A second construct was made which placed a *Bam*HI site proximal to a synthetically engineered Met residue in front of Asp-39. A double-stranded oligonucleotide with complementing *Bam*HI and *Eco*RI ends, made using two complementary sequences, 5′-GATCCATGGATTTAATTCATATTCAT-TCATTAACTTTGAAATCTATTGTAGGAAAG and 5′-AA-TTCTTTCCTACAATAGATTTCAAAGTTAATGAATGA-ATATGAATTAAATCCATG, was ligated to *Bam*HI/*Eco*RI-digested pKS.*fasAB-Met1*. The new plasmid was designated pKS.*fasAB-Asp39*.

Site-Directed Mutagenesis. The entire 817-bp *fasAB-Met23* coding region was excised from BluescriptKS and subcloned into *Bam*HI + *Hin*dIII-digested M13mp18. Single-stranded DNA template was generated, and mutations were introduced by oligonucleotide-directed mutagenesis (Amersham kit). Eight mutants containing one of the following nucleotide changes, corresponding to a single amino acid substitution, were made: D39E, GAT→GAG; G53A, GGA→GCA; Q63N, CAA→AAT; K96R, AAA→AGA; D161E,GAT→GAG;G175A,GGT→GCT;Q185N,CAG→AAT; K218R, AAA→AGA. Each M13 clone containing the desired mutation was identified by DNA sequencing.

Expression of Fas Sequences in E. coli. Each wild-type *fas* construct, *fasAB-Met1*, *fasAB-Met23*, and *fasAB-Asp39*, was excised from BluescriptKS as a product of a *Bam*HI + *Hin*dIII digest, and ligated to a 91-bp *Hin*dIII-*Bam*HI *tac* genblock promoter (Pharmacia). The genblock contains the strong inducible *tac* promoter and a ribosome-binding site 5-bp upstream of the *Bam*HI site. Resulting *Hin*dIII fragments (i.e., *Bam*HI-*Bam*HI ligated ends) were subsequently ligated into the unique *Hin*dIII site of a previously described modified version of pKK223.3 (*6*). Similarly, each of the eight mutant DNA inserts was excised from M13 replicative form (RF) DNA as a product of a *Bam*HI + *Hin*dIII digestion, ligated first to the *tac* genblock promoter, and subsequently into the unique *Hin*dIII site of a modified pKK223.3. The entire *fasAB* coding region of each mutant was then sequenced to ensure that no other changes, apart from the desired substitution, had occurred during the mutagenesis and cloning steps. Procedures for the overproduction of Fas protein species in *E. coli* XL1-blue cells have been reported elsewhere (*6*).

Purification of FasAB-Met23 and FasAB-Asp39. Bacterial cell paste (25−50 g) was resuspended in 50 mM sodium phosphate (pH 7.5), 0.1 M NaCl, 1 mM EDTA at about 30% (w/v) and broken using a pre-cooled French pressure cell as previously described (*6*). All subsequent steps were performed at 4 °C. Following centrifugation at 15 000*g*, supernatants were collected, diluted with further buffer to a protein concentration of 10 mg/mL, and fractionated with ammonium sulfate. FasAB-Met23 precipitated between 30 and 40% saturation. The precipitate was dissolved in 50 mM sodium phosphate (pH 7.5), 0.2 M NaCl (buffer A), dialyzed overnight, and applied to a 1.6 × 8 cm Chelating Sepharose column (Pharmacia) charged with zinc ions, and equilibrated with buffer A. After loading, the column was washed at pH 6.0 (50 mM Mes, 0.2 M NaCl) to remove contaminants, and the recombinant species eluted with 50 mM imidazole in buffer A. Following dialysis into 50 mM sodium phosphate (pH 8.0), 10% (w/v) glycerol, 0.1 mM EDTA, the eluate was applied to a 1.6 × 4 cm Q-Sepharose Fast Flow column (Pharmacia) equilibrated in the same buffer. FasAB-Met23 was collected in the unbound fraction, concentrated by pressure dialysis to 10−15 mL, and applied to a Sephacryl-200 HR column (2.6 × 90 cm), equilibrated with 50 mM sodium phosphate (pH 7.5), 0.2 M NaCl, 10% (w/v) glycerol, 0.1 mM EDTA. From SDS−PAGE analysis the cleanest fractions were pooled, concentrated to at least 4 mg/mL, flash frozen in liquid nitrogen and stored at −70

°C. The eight Fas mutants were overproduced in *E. coli*, and each was purified exactly as described for the wild-type FasAB-Met23 polypeptide.

FasAB-Asp39 differs from FasAB-Met23 in that the protein precipitates at slightly lower $(NH_4)_2SO_4$ saturation, and it does not bind to the zinc chelate column. For the purification of FasAB-Asp39 therefore, cell extracts were adjusted to 25% saturation with $(NH_4)_2SO_4$, clarifed by centrifugation, and the supernatant applied to a column (1.6 × 10 cm) of Butyl Sepharose Fast Flow (Pharmacia), equilibrated in 50 mM sodium phosphate (pH 7.5), 1 M $(NH_4)_2SO_4$, 1 mM EDTA. Proteins were eluted using a 20 column volume gradient from 1 to 0 M $(NH_4)_2SO_4$. Fractions eluting between 0.4 and 0.6 M $(NH_4)_2SO_4$ containing the recombinant species were pooled. All other steps were performed as described for FasAB-Met23.

Analytical size exclusion chromatography (SEC) was performed on a Superdex-200 column (HR10/30) using an FPLC at 4 °C. The mobile phase was buffer A. The calibration standards (Pharmacia) were ribonuclease, chymotrypsin, ovalbumin, albumin, aldolase, catalase, and feritin. Protein concentrations were determined by the method of Bradford (*13*).

Cross-Linking with Bisimidoesters. Reactions were carried out according to the original recommendations of Davies and Stark (*14*). Dimethyl suberimidate (DMS) and dimethyl adipimate (DMA) which cross-link amine groups were obtained from Pierce. FasAB-Met23 was diluted in 0.2 M triethanolamine (pH 8.5) to give a final protein concentration of between 20 and 200 μg/mL. Two-hour reactions were initiated at room temperature by the addition of reagent, made up in triethanolamine buffer to give a final concentration of 1−10 mM. Aliquots were removed at intervals, and reactions were terminated by the addition of glycine to 0.1 M. Following TCA precipitation, samples were analyzed by SDS−PAGE.

Ultracentrifugation. Sedimentation equilibrium experiments were performed on a Beckman Optima XL-A analytical ultracentrifuge equipped with absorbance optics, using an An60Ti rotor and three sample cells, each with a six-sector charcoal-filled epoxy centerpiece. Samples were run at 10K, 12.5K, and 18K rpm and allowed to spin for 24 h at 4 °C prior to data collection. In a separate experiment samples were run at 4 and 30 °C and spun at 11K rpm for 24 h prior to data collection. Ten individual data sets per speed were collected at 2-h intervals and these were the average of 10 scans taken at 280 nm with a step size of 0.003 cm in a step mode. Equilibrium was ascertained by comparing data sets up to 20 h apart. Samples (protein concentration ranging from 2 to 93 μM) were centrifuged against buffer blanks. The buffer used was 0.25 M NaCl, 50 mM sodium phosphate (pH 7.5) and its density (ρ = 1.01869 and 1.0131 g mL^{-1} at 4 °C and 30 °C, respectively) was measured on a Paar DMA58 density meter calibrated against air and water. The partial specific volume of the protein (υ_{25} = 0.754243 mL g^{-1}) was calculated by the method of Cohn and Edsall (*15*) and adjusted to the correct temperature (υ_4 = 0.745743 mL g^{-1} and υ_{30} = 0.756368 mL g^{-1}) as described in Durchschlag (*16*). Data sets were edited with REEDIT (Jeff Lary, National Analytical Ultracentrifugation Center, Storrs, CT) and fit individually and jointly with NONLIN (*17*). NONLIN fits the data to an effective reduced molecular mass, $\sigma = M(1 - \upsilon\rho)\omega^2/RT$ where M is the molecular mass, υ is the partial specific volume of the protein, ρ is the solvent density, $\omega = (2\pi\text{-}(\text{rpm})/60)$, R is the gas constant, and T is the temperature in Kelvin. Data from different speeds and concentrations were combined for a global fit. The smallest observed particle at 4 °C was equivalent to four monomers and thus σ was fixed to a tetramer when fitting to an association scheme. At 30 °C the smallest observed particle was equivalent to a dimer and σ was fixed accordingly. The association constants in NONLIN are fit as ln K to constrain them to positive values. These values were then converted from absorbance to molar units. It is worth noting that in NONLIN K_n represents the association constant for n monomers to an n-mer irrespective of pathway. The calculated extinction coefficient for a DHNA monomer is ϵ_{280} = 19360 M^{-1} cm^{-1}.

Enzyme Assays. DHNA activity was assayed as previously described (*4*, *5*), except that assays were conducted in the presence of 0.05% (w/v) Tween20. The use of Tween20 as a stabilizing agent was found to be superior to 0.1% (w/v) BSA, resulting in around 25% higher activity measurements.

RESULTS AND DISCUSSION

Expression of FasAB-Met1 in E. coli. Previously, recombinant baculoviruses were constructed which directed the production of FasA (residues 1−160), FasB (residues 161−290), and FasAB (residues 1−290) polypeptide species (see Figure 1) in insect cells. While no DHNA activity was associated with either recombinant FasA or FasB, DHNA activity was detected in cell extracts containing the FasAB polypeptide (*5*). Detailed characterization of DHNA function associated with this recombinant species was precluded by low-level expression of FasAB sequences in insect cells. In this study, the FasAB species, here designated FasAB-Met1, was produced to high levels in *E. coli* using an inducible *tac* promoter expression system (Figure 2a). The activity of FasAB-Met1-containing bacterial extracts (0.17−0.28 U/mg) was about 10-fold higher than the endogenous DHNA activity of host *E. coli* extracts (0.015−0.025 U/mg). Analysis of the subcellular distribution revealed however that at least 95% of the recombinant protein was associated with the insoluble fraction (data not shown). Attempts to extract the recombinant species from the pellet fraction with salts or detergent were unsuccessful, as were refolding experiments following extraction of FasAB-Met1 in denaturant. Further studies with this protein were not pursued.

Expression of FasAB-Met23 and FasAB-Asp39 in E. coli. Multiple sequence alignment suggests that the N-terminal 38 residues of FasA have no counterpart in either FasB or DHNA of bacteria (Figure 1). The initial construct, ptac-*.fasAB-Met1*, had included these amino acids, since the exact boundaries of functional domains within a multifunctional enzyme cannot be defined by sequence analysis alone. To determine the importance of the N-terminus of Fas in DHNA function therefore, we generated a construct, ptac.*fasAB-Asp39* (residues 39−290), for expression in *E. coli*. A further construct, ptac.*fasAB-Met23* (residues 23−290), was also produced which utilizes a naturally occurring AUG as the translation initiation codon. We hoped that one of these

```
                                                39
                              23                 E                A          N
FasA 1-MIFKSLKIFP FYQIYGFRFL KGMIFKKKIH LSKLNKNHDL IHIHSLTLKS IVGKNSWAQR LLQPVVLTLS MGINASLSGN MDDLSYSI.D  89
FasB                                        161-DQ FFIKNLSLYT IIGINPEERV NKQNIIIDLI LFKSSINLEC KDDFIINTYN 212
Sp                                            1-MDQ LQIKDLEMFA YHGLFPSEKE LGQKFIVSAI LSYDMTKAAT DLDLTASV.H  52
Sa                                         1-MKMQDT IFLKGMRFYG YHGALSAENE IGQIFKVDVT LKVDLAEAGR TDNVIDTV.H  55
Bs                                            1-MDK VYVEGMEFYG YHGVFTEENK LGQRFKVDLT AELDLSKAGQ TDDLEQTI.N  52

           R
FasA YATVYKEVFK LVENSKFENL LDLSDKISKV VLG..DKCKG NWVKVIAETP KGHLLAETGL QIIRRKDGIR EID-160
FasB IEKLLKEIVK HVEESTFKTI EALALSIARI SCISHNIEKI I.VKVKKSCA ..LAFAESAG VEIVRSRSCF SSNNYIKSEN S-290
Sp   YGELCQQWTT WFQETSEDLI ETVAYKLVER TFESYPLVQE MKLELKKPWA PVHLSLDTCS VTIHRRK-119
Sa   YGEVFEEVKS IMEGKAVNLL EHLAERIANR INSQYNRVME TKVRITKENP PIPGHYDGVG IEIVRENK*-123
Bs   YAELYHVCKD IVEGEPVKLV ETLAERIAGT VLGKFQPVQQ CTVKVIKPDP PIPGHYKSVA IEITRKKS*-120
```

FIGURE 1: Comparison of amino acid sequences of *P. carinii* FasA and FasB with bacterial DHNA. Sp, N-terminus of SulD of *S. pneumoniae* (*2*); Sa, FolQ of *S. aureus* (P.K., S.P.B. & C.J.D., unpublished); Bs, ORF1 of *B. subtilis* (*1*). FasA is 18, 18, and 25% identical with the DHNA of Sp, Sa, and Bs, respectively. FasB shows 23, 23, and 24% identity with Sp, Sa, and Bs DHNA, respectively. The full extent (minus the PCR-engineered start methionine) of Fas sequences included in expression constructs, *fasAB-Met1* (residues 1−290), *fasAB-Met23* (residues 23−290), and *fasAB−Asp39* (residues 39−290), is shown. Filled-in boxes indicate amino acid identity between all five sequences; the open box shows Lys-96 and Lys-218 referred to in the text. The bolded residue above the boxed sequence shows the amino acid change introduced by site-directed mutagenesis into the FasA or FasB sequence. Bolded amino acids in the alignment indicates identity between FasA or FasB, and at least two other sequences. Dots within sequences indicate spaces inserted to give optimal alignment, and stop codons are marked by asterisks.

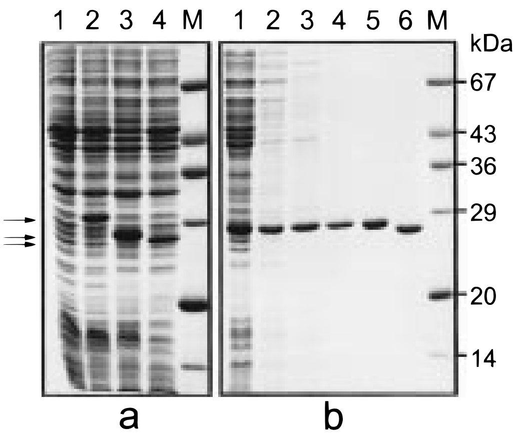

FIGURE 2: Overproduction of FasAB-Met1, FasAB-Met23, and FasAB-Asp39 in *E. coli*, and purification of the recombinant species. Proteins detected by Coomassie-staining after SDS−PAGE. Panel (a) Lane 1, plasmid only transformed *E. coli*; lane 2, ptac.*fasAB-Met1*-transformed *E. coli*, which express FasAB-Met1 (upper arrow); lane 3, ptac.*fasAB-Met23*-transformed *E. coli*, which express FasAB-Met23 (middle arrow); lane 4, ptac.*fasAB-Asp39*-transformed *E. coli*, which express FasAB-Asp39 (lower arrow); lane M, molecular mass markers. Panel (b) Lane 1, ptac.*fasAB-Met23*-transformed *E. coli*, soluble fraction; lane 2, FasAB-Met23 following ammonium sulfate extraction; lane 3, FasAB-Met23 from zinc chelate column; lane 4, FasAB-Met23 from Q-Sepharose column; lane 5, FasAB-Met23 from Sephacryl-200 column; lane 6, purified FasAB-Asp39; lane M, molecular mass markers.

expression constructs would produce a soluble recombinant FasAB species suitable for detailed characterization.

Following a 4-h induction with IPTG an overproduced protein in both, *fasAB-Met23*- and *fasAB-Asp39*-containing bacterial extracts, which migrate on SDS−PAGE as 29-, and 28-kDa bands, respectively, was observed (Figure 2a). The size of the overproduced species is consistent with the synthesis of 30355-Da (FasAB-Met23) and 28375-Da (FasAB-Asp39) polypeptides deduced from the primary structure. Significantly, greater than 90% of both the FasAB-Met23 and FasAB-Asp39 species were found in the soluble fraction of bacterial extracts. Soluble extracts containing FasAB-Met23, and FasAB-Asp39 were analyzed for DHNA activity. The activities of both FasAB domain forms (1.6−2.0 U/mg) were about 100-fold higher than the endogenous *E. coli* DHNA activity.

Purification of FasAB-Met23 and FasAB-Asp39. Recombinant FasAB-Met23 was purified using ammonium sulfate precipitation, metal chelate chromatography, anion exchange, and SEC (Figure 2b). The zinc chelate step binds FasAB-Met23 tightly, with the majority of *E. coli* proteins passing through. The recombinant enzyme is eluted with 50 mM imidazole. Binding to metal chelate columns, requiring this level of imidazole for elution, is usually due to interaction with accessible histidine residues. FasAB-39 did not bind efficiently to zinc chelate presumably reflecting the absence of His-30 and/or His-38 (Figure 1), and so for purification of the smaller species the metal chelate step was replaced by hydrophobic chromatography on Butyl Sepharose. Most residual *E. coli* proteins were removed from FasAB-Met23 or FasAB-Asp39 preparations by passage over an anion-exchange column at pH 8.0, conditions under which the Fas polypeptides do not bind. Although FasAB-Met23 and FasAB-Asp39 are basic with predicted pI's of 9.2 and 8.8, respectively, and both will bind to cation exchangers, neither protein is effectively eluted without employing denaturing conditions.

The two purified recombinant proteins were N-terminal sequenced. FasAB-Asp39 gave the expected sequence including the start methionine. However, FasAB-Met23 yielded the sequence KIHLSKLNKH, indicating the removal of the first five amino acids (MIFKK) of the predicted protein (Figure 1). Mass spectrometry of FasAB-Met23 showed a single species of 29 689 Da, consistent with the sequencing data.

Specific activities for both final enzyme preparations were between 30 and 35 U/mg. The purification of enzymically active FasAB-Asp39 indicates that the hydrophobic N-terminal 38 amino acids are not essential for DHNA function. Furthermore, since the specific activity of both N-terminally truncated FasAB species is comparable, a role, if any, in DHNA function for residues 28−38 must be subtle. Residues 1−23 of the Fas "leader sequence" show 50% sequence identity to the N-terminus of a recently cloned *P. carinii* subtilisin-like serine protease (18). Both predicted proteins start MIFK, a motif which is repeated within the leader sequence of Fas (residues 23−26, see Figure 1). It would be of interest to know how prevalent such hydrophobic leader

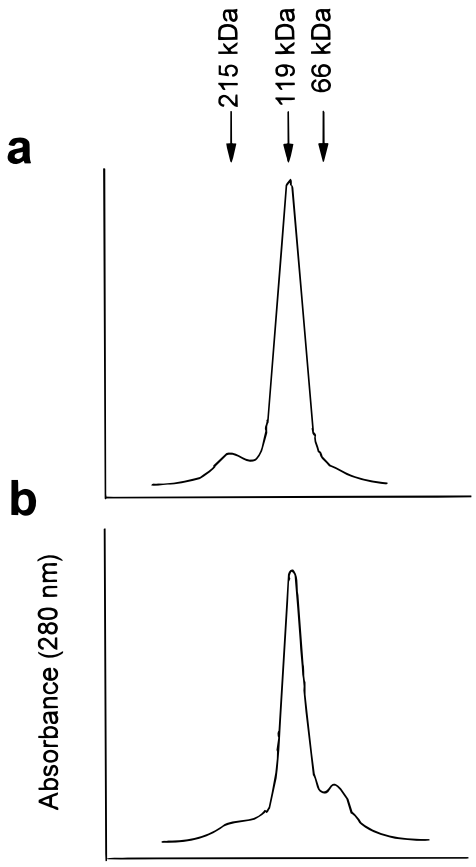

FIGURE 3: Size exclusion chromatography of FasAB-Met23. An amount of 150 μg of purified enzyme (1.5 mg/mL) was chromatographed on a Superdex-200 column (HR10/30) after incubation at (a) 4 °C, or (b) 37 °C, for 20 min. Chromatography was conducted at 4 °C.

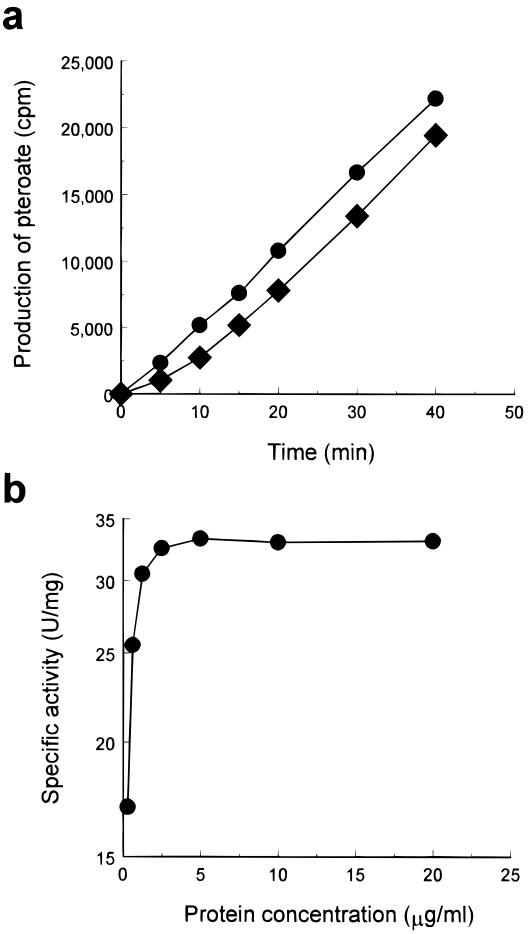

FIGURE 4: Activity of FasAB-Met23. (a) Time course of DHNA activity. Reactions were preincubated at 37 °C for 5 min prior to initiation by the addition of enzyme (●), or substrate (■). Final enzyme concentration was 2.5 μg/mL. (b) Reaction rates (enzyme initiated) were measured with final enzyme concentrations ranging from 0.2 to 20 μg/mL. A plot of specific activity versus enzyme concentration is shown.

sequences are in *P. carinii* proteins. Unfortunately, the lack of a transformation system for *P. carinii* precludes functional analysis studies.

Size Exclusion Chromatography of FasAB-Met23. All subsequent studies described were undertaken using FasAB-Met23. On SEC using a Superdex 200 column, approximately 95% of FasAB-Met23 elutes with an apparent molecular mass of 119 000 at 4 °C, suggesting that the protein is multimeric. A small proportion of a 215 000 species is also observed (Figure 3a). When purified enzyme is transferred from 4 °C to 25 or 37 °C, however, a smaller 66,000 species accumulates (Figure 3b). This effect is reversible, with the smaller species disappearing if the enzyme is returned to 4 °C. Significantly, the 66 kDa species is not produced at 25 or 37 °C if the enzyme is incubated with substrate, under which conditions the 215 000 species is also absent. The data suggests that the protein populates an equilibrium between different oligomers. This equilibrium is slow over the duration of the experiment, otherwise a single broad peak would be observed following SEC analysis.

Properties of the Enzyme. The K_m determined for 7,8-dihydroneopterin was 2.3 ± 0.4 μM, with a V_{max} of 36 ± 1.8 U/mg. Assuming one active site per monomer, this gives a minimum value for the k_{cat} of 1.8 s^{-1}. DHNA function of FasAB-Met23 displays thermolability. Reactions initiated by the addition of ice-cold enzyme to the assay mix show a linear time course of dihydropteroate formation. If the enzyme is held at room temperature prior to initiation of the assay, however, a lag phase in the reaction time course is observed. This lag phase was most evident in substrate-initiated reactions where the enzyme was preincubated in the assay mix at 37 °C prior to addition of dihydroneopterin (Figure 4a). The steady-state rate eventually obtained is comparable to that observed for enzyme held on ice prior to initiation of the assay, indicating that the effect is reversible. The observations suggest that at higher temperature the enzyme either undergoes a conformational change or dissociates to an inactive or less active species. The latter possibility is consistent with SEC data that in the absence of substrate at 25 or 37 °C a smaller species accumulates. Moreover, in a plot of specific activity versus protein concentration, a loss of activity is observed at low protein levels (Figure 4b). Similar findings have been reported for other enzymes and are characteristic of an associating−dissociating system in which only the multimeric form of the enzyme is active (*19*, *20*).

Cross-Linking Experiments with FasAB-Met23. To further investigate the native form of FasAB-Met23 we initially undertook cross-linking experiments with bifunctional reagents, principally the bisimidoesters, DMA and DMS, which have spacer arms of 8.6 and 11 Å, respectively. Figure 5 shows the results of time-course experiments visualized

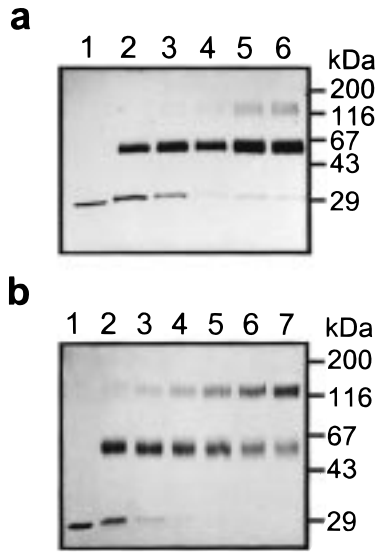

FIGURE 5: Time course of cross-linking FasAB-Met23 with DMA and DMS. Aliquots (5 μg) of protein were removed from reaction mixtures at various time points, and the reactions terminated. Following precipitation samples were subjected to SDS−PAGE and stained with Coomassie Blue. (a) Incubation of FasAB with DMA. Lane 1, 1 μg untreated FasAB-Met23; lanes 2, 3, 4, 5, 6, incubation of FasAB-Met23 (20 μg/mL) with 5 mM DMA for 10, 20, 30, 60, and 120 min, respectively. (b) Incubation of FasAB with DMS. Lane 1, 1 μg untreated FasAB-Met23; lanes 2, 3, 4, 5, 6, 7, incubation of FasAB-Met23 (200 μg/mL) with 5 mM DMS for 10, 20, 30, 60, 90, and 120 min, respectively.

following SDS−PAGE which were generated from the incubation of FasAB-Met23 with the cross-linking agents. In the presence of either DMA or DMS, greater than 90% of the 30-kDa recombinant protein accumulated as a 60-kDa species within 20 min (Figure 5a). Generation of this species was independent of both protein (20−200 μg/mL) and reagent (1−10 mM) concentration. A slower accumulation of a second species of approximately 120 kDa, was also observed. In the presence of DMA, up to 10% of FasAB-Met23 accumulated as the 120-kDa form following a 2-h incubation (Figure 5a). The use of DMS resulted in the accumulation of a much greater proportion of the 120-kDa species, with the formation of this species showing a marked dependence on protein concentration. Following a 2-h incubation, 40−50% of a 20 μg/mL enzyme preparation was observed as the higher MW species, whereas using a 200 μg/mL enzyme preparation up to 70−80% of protein was in this form (Figure 5b).

The results of the cross-linking experiments suggest one of two possibilities. First, that the native form of FasAB-Met23 is a dimer of the 30-kDa polypeptide chain, as suggested by the rapid formation of a 60-kDa species. Although a 120-kDa species with an apparent MW approximating tetramer in size is formed with the different cross-linking agents, the dependence of this event on protein concentration, and the absence of a MW species approximating trimer may indicate an aspecific cross-linking of dimers. The alternative explanation suggested by the results is that the native species is tetrameric. The greater accumulation of the 120-kDa species with DMS may simply reflect a better proximity of amino groups for this event, i.e., the longer chain cross-linking reagent is able to capture the dimer−dimer interaction more efficiently. To distinguish between these two possibilities, we carried out an ultracentrifugation analysis of the FasAB species.

Ultracentrifugation Analysis of FasAB-Met23. Sedimentation equilibrium experiments were conducted at both 4 and 30 °C. Initially, individual data sets were fit to a model allowing for the presence of only a single ideal species, which was calculated to be 4.2 to 5.1 monomers for data at 4 °C, and 5.2 to 6.2 monomers at 30 °C. The molecular mass of the monomer is 29 689 Da as determined by mass spectrometry. The residuals to these fits were poor and the root-mean-square of the fit was unacceptable. This suggested an equilibrium condition where associating species larger than dimers were present in solution. Global fits to data sets spanning the concentrations 2−93 μM at 4 °C best fit to a model in which tetramer was the smallest and most predominant species. A small proportion of octamers was also predicted (Figure 6). Attempts to include dimers into the model prevented the fits converging. The dissociation constant K_{8-4} calculated for the tetramer−octamer equilibrium was 120 μM (95% confidence limits were 80 μM and 200 μM). All attempts to fit the data to trimers, or indeed any other size oligomer, failed. We conclude that the active form of the enzyme is tetrameric, which is consistent with the 119000−120000 species observed following SDS−PAGE of FasAB-Met23 cross-linked with DMS and SEC on Superdex 200. Furthermore, the small proportion of octamer observed following ultracentrifugation correlates with the 215-kDa species observed during SEC.

At 30 °C in addition to tetramer and octamer, dimer was also present. Fitting the data to a dimer−tetramer equilibrium and omitting the octamer failed. The calculated dissociation constant K_{4-2} for the tetramer−dimer equilibrium was 2.6 μM (95% confidence limits were 1.2 μM and 5.6 μM), and for K_{8-4} was 2.8 μM (95% confidence limits were 0.2 μM and 40 μM). Thus at 30 °C the dissociation constant between the dimer and tetramer species is considerably weakened as compared to the data obtained at 4 °C, where no dimer was observed within the detection limits of the instrument. However, the opposite appears true for the equilibrium between tetramer and octomer. In this case, affinity for octomerization increases by 2 orders of magnitude at the higher temperature suggesting that this event is entropy driven. The presence of dimer at elevated temperature following centrifugation is consistent with our SEC data where a 66 kDa species is detected following preincubation of FasAB-Met23 at 25 or 37 °C (Figure 3b) in the absence of substrate.

Analysis of FasAB Mutants. Multiple sequence alignment of FasA and FasB with the DHNA of bacteria highlights only five positions where the amino acid is invariable between all the predicted sequences (Figure 1). We have now cloned a *fas* gene homologue from a second fungal species *Candida albicans* (W. S. Dallas and C.J.D., unpublished). The predicted Fas protein of *C. albicans* also maps into four domains, the N-terminal two domains of which share significant sequence similarity, not only with each other, but also with FasA and FasB of *P. carinii* and the bacterial DHNA enzymes. Importantly, three of the five positions, Asp-39, Gly-53, and Gln-63 in *P. carinii* FasA and the 3 corresponding residues, Asp-161, Gly-175, and Gln-185 in FasB are also found to be invariable in the FasA and FasB of *C. albicans*, and were therefore selected for

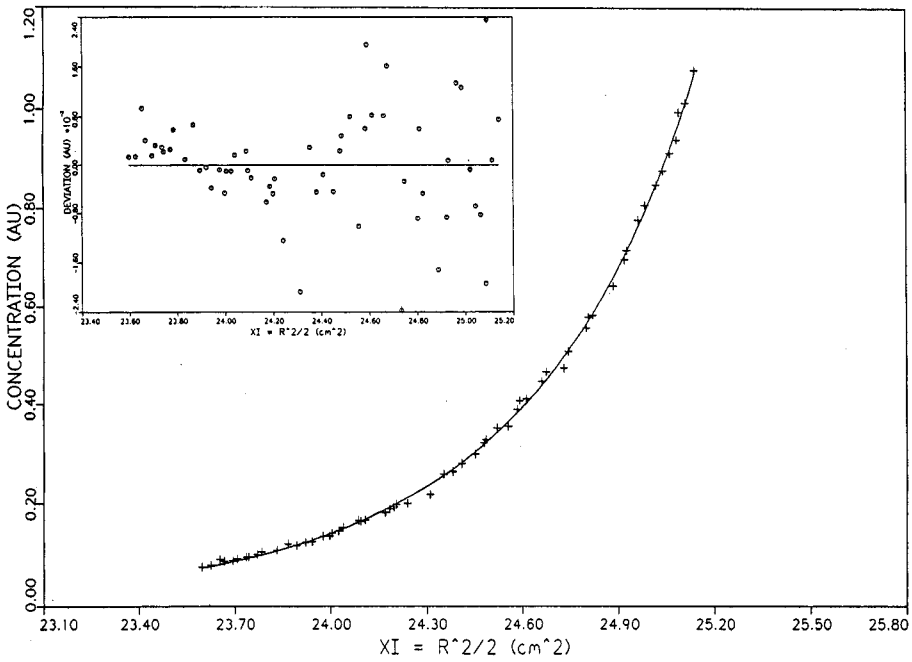

FIGURE 6: Representative sedimentation equilibrium data set of 26 μM initial concentration DHNA at 11K rpm and 4 °C fit to a tetramer−octamer model. The insert presents the residual plot of the data to the model.

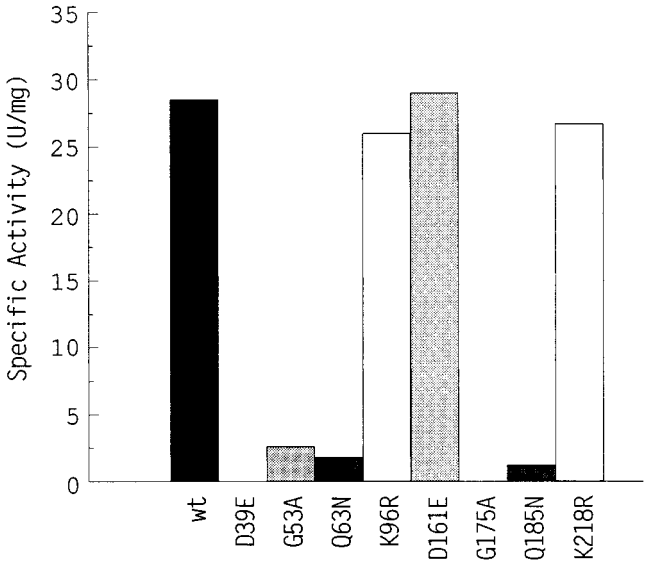

FIGURE 7: Effects of site-specific mutagenesis on DHNA activity of FasAB-Met23 (wt).

site-directed mutagenesis studies. We introduced single conservative amino acid substitutions, Asp→Glu, Gly→Ala, or Gln→Asn, in either the FasA or FasB domain of FasAB-Met23. Two further "control" mutant species, containing single amino acid substitutions, were also produced: Lys-96→Arg in FasA and Lys-218→Arg in FasB. Sequence analysis shows that Lys-96 and Lys-218 align between FasA and FasB, but that the bacterial DHNA sequences show sequence variability at this position in the alignment (Figure 1). We predicted that Lys-96 and Lys-218 would not be essential for DHNA function and could be mutated without significant loss of enzyme activity.

Specific activities of purified wild-type and mutant FasAB proteins are shown in Figure 7. Five of the six amino acid substitutions in positions where the residue is invariable between FasA, FasB, and the bacterial sequences have a dramatic effect on DHNA activity. Indeed, two of these mutant enzyme species, D39E (FasA domain mutant) and G175A (FasB domain mutant), have no detectable DHNA activity in our assay. The two other FasA domain mutants, G53A and Q63N, and the FasB domain mutant, Q185N, show approximately 11-, 16-, and 24-fold decrease, respectively, in specific activity compared to wild-type FasAB-Met23. Only the activity of the FasB domain mutant, D161E, is similar to that of wild-type enzyme. As predicted, the two mutant enzymes, K96R and K218R, have levels of activity comparable to wild-type enzyme. Indeed, K_m values determined for K96R, K218R, and D161E were similar to wild-type FasAB-Met23. It was not possible to determine K_m values for those mutants with slight enzyme activity.

Two of the purified mutant enzyme species, Q63N and Q185N, were sequenced. The N-terminal 10 residues determined for both mutant species, namely KIHLSKLNKH, were consistent with the sequence determined for FasAB-Met23, showing that these two mutants were processed in *E. coli* in a manner similar to the wild-type protein.

Since each mutant species contained only a single conservative amino acid change we anticipated that their chromatographic properties would be similar to the wild-type protein. Indeed, the three mutant species, D161E, K96R, and K218R, whose activities are comparable to wild-type enzyme, were purified in similar yield, and on SEC eluted with an apparent molecular mass of approximately 119 kDa. Surprisingly, however, D39E, G53A, Q63N, G175A, and Q185N, the mutants with low or no detectable DHNA activity, behaved very differently during purification. Although polypeptides of purity similar to that of the wild-type species were obtained in each case, yields were significantly poorer, typically sub-milligram quantities, due to greater losses particularly during the precipitation and ion-exchange steps. Following application to a Superdex 200 column, each eluted with an apparent molecular mass of only 60−70 kDa suggestive of a smaller species than wild-type tetramer. To investigate this change in profile, Q63N and

G175A were subjected to ultracentrifugation analysis. Unfortunately, no interpretable data was generated, the mutant proteins appearing unstable over the time taken for equilibrium to be established in the experiment (data not shown). This finding may also explain the difficulties encountered during the purification of the inactive or less active species. Subsequent experiments involving chemical cross-linking of the mutant species followed by SEC analysis did however prove informative. Following a 1-h incubation with DMA, greater than 90% of D39E, G53A, Q63N, G175A, and Q185N had each accumulated as a 60-kDa species as determined by SDS−PAGE (data not shown), which we interpret as being dimeric. Importantly, each cross-linked mutant species behaved identically to the corresponding untreated enzyme on SEC, i.e., eluting with an apparent molecular mass of approximately 60−70 kDa. We conclude therefore that the five mutants where a conservative substitution results in significant or total loss of function are unable to form stable tetramers, and instead are predominantly dimeric species.

Concluding Remarks. Multifunctional polypeptides have probably evolved by the fusion of ancestral monofunctional genes (*21*). For the *fas* gene of *P. carinii*, this process presumably also involved DHNA gene duplication plus fusion and subsequent divergence of the two DHNA domains, as our work strongly implicates both FasA and FasB in DHNA function. It is common that each subunit of a tetrameric enzyme has only a single recognition site for bonding with a like subunit. The existence of a second set of bonding sites leads to the formation of a tetramer as a dimer of dimers (D_2 symmetry) (*22*). The finding that dimer accumulates above 25 °C in the absence of dihydroneopterin suggests that this is the case for *P. carinii* DHNA. Indeed, the two inactive mutants, D39E and G175A, appear stable as dimers. We suggest that amino acids which are identical between FasA, FasB, and the bacterial DHNA and whose conservative substitution leads to loss of function could be active-site residues, and that it is substrate binding which stabilizes the tetramer form, either because the catalytic sites are located at subunit interfaces or because catalytic activity depends on subunit conformation, which is mediated by dihydroneopterin.

It is relevant to compare the DHNA enzyme of *P. carinii* with its counterpart in *S. pneumoniae*, the only other organism where detailed biochemical data is available. SulD, a bifunctional tetrameric protein, harbors single DHNA and PPPK domains. On purification, the multimer reversibly dissociates into monomeric subunits, which have PPPK activity but no DHNA activity. DHNA function is only expressed by the tetrameric protein (*2*). Thus, although the DHNA of both *P. carinii* and *S. pneumoniae* are tetrameric, the organization of *P. carinii* DHNA contains a further level of complexity since each polypeptide chain already contains two functional domains. Parallels can be drawn with mammalian enzymes such as phosphofructokinase which is twice the size of its bacterial counterpart, being the product of gene duplication plus fusion. Each subunit of the mammalian phosphofructokinase enzyme potentially has therefore two catalytic and two regulatory sites, in contrast to the bacterial enzyme, which has only one catalytic and one regulatory site. The mammalian enzyme is proposed to have further evolved to one catalytic and three regulatory sites per subunit, providing extra sensitivity for regulating this glycolytic enzyme (*23*). Whether *P. carinii* DHNA has evolved, such complex regulation must await further work. We have previously shown, however, that the presence and correct folding of FasAB plays a critical role in stabilizing C-terminal DHPS function (FasD domain) within the Fas multifunctional protein. Indeed, a bifunctional PPPK (FasC)/DHPS species has PPPK activity but no DHPS activity (*4*). If our observations on the isolated FasAB domains are consistent with those in the native state, DHNA activity generated upon assembly of the FasAB tetramer would regulate the function of the larger multienzyme. In the presence of substrate, a stable DHNA tetramer is formed which supports DHPS function, thus DHPS is activated only when 7,8-dihydroneopterin is available.

REFERENCES

1. Slock, J., Stahly, D. P., Han, C.-Y., Six, E. W., and Crawford, I. P. (1990) *J. Bacteriol. 172*, 7211−7226.
2. Lopez, P., and Lacks, S. A. (1993) *J. Bacteriol. 175*, 2214−2220.
3. Volpe, F., Dyer, M., Scaife, J. G., Darby, G., Stammers, D. K., and Delves, C. J. (1992) *Gene 112*, 213−218.
4. Volpe, F., Ballantine, S. P., and Delves, C. J. (1993) *Eur. J. Biochem. 216*, 449−458.
5. Volpe, F., Ballantine, S. P., and Delves, C. J. (1995) *Gene 160*, 41−46.
6. Ballantine, S. P., Volpe, F., and Delves, C. J. (1994) *Protein Expression Purif. 5*, 371−378.
7. Maley, J. A., and Davidson, J. N. (1988) *Mol. Gen. Genet. 213*, 278−284.
8. Pazirandeh, M., Chirala, S. S., Huang, W.-Y., and Wakil, S. J. (1989) *J. Biol. Chem. 264*, 18195−18201.
9. Tauler, A., Lange, A. J., El-Maghrabi, M. R., and Pilkis, S. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7316−7320.
10. Hum, D. W., and MacKenzie, R. E. (1991) *Protein Eng. 4*, 493−500.
11. Hawkins, A. R., and Smith, M. (1991) *Eur. J. Biochem. 196*, 717−724.
12. Belfaiza, J., Parsot, C., Martel, A., Bouthier de la Tour, C., Margarita, D., Cohen, G. N., and Saint-Girons, I. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 867−871.
13. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
14. Davies, G. E., and Stark, G. R. (1970) *Proc. Natl. Acad. Sci. U.S.A. 66*, 651−656.
15. Cohn, E. J., and Edsall, J. T. (1965) in *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions* (Cohn, E. J., and Edsall, J. T., Eds.) p 157, Rheinhold, New York.
16. Durchschlag, H. (1986) in *Thermodynamic Data for Biochemistry and Biotechnology* (Hinz, H. J., Ed.) p 45, Springer-Verlag, New York.
17. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) *Biophys. J. 36*, 575−588.
18. Lugli, E. B., Allen, A. G., and Wakefield, A. E. (1997) *Microbiology 143*, 2223−2236.
19. Anderson, P. M. (1983) *Biochemistry 22*, 3285−3292.
20. Frieden, C. (1981) in *Protein−Protein interactions* (Frieden, C., and Nichol, L. W., Eds.) pp 289−314, Wiley, New York.
21. Hardie, D. G., and Coggins, J. R. (1986) in*Multifunctional proteins and evolution* (Hardie, D. G., and Coggins, J. R., Eds.) pp 229−258, Elsevier, Amstersdam.
22. Traut, T. W. (1994) *Critical Rev. Biochem. Mol. Biol. 29*, 125−163.
23. Poorman, R. A., Randolph, A., Kemp, R. G., and Heinrikson, R. L. (1984) *Nature 309*, 467.

BI980540X