Structure of the Type III Pantothenate Kinase from *Bacillus anthracis* at 2.0 Å Resolution: Implications for Coenzyme A-Dependent Redox Biology[†,‡]

Nathan I. Nicely,[§,||] Derek Parsonage,[§] Carleitta Paige,[§] Gerald L. Newton,[⊥] Robert C. Fahey,[⊥] Roberta Leonardi,[#] Suzanne Jackowski,[#] T. Conn Mallett,[§,Δ] and Al Claiborne*,[§]

Center for Structural Biology, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157, Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0314, and Department of Infectious Diseases, St. Jude Children's Research Hospital, Memphis, Tennessee 38105

Received November 7, 2006; Revised Manuscript Received December 18, 2006

ABSTRACT: Coenzyme A (CoASH) is the major low-molecular weight thiol in *Staphylococcus aureus* and a number of other bacteria; the crystal structure of the *S. aureus* coenzyme A-disulfide reductase (CoADR), which maintains the reduced intracellular state of CoASH, has recently been reported [Mallett, T.C., Wallen, J.R., Karplus, P.A., Sakai, H., Tsukihara, T., and Claiborne, A. (2006) *Biochemistry 45*, 11278−89]. In this report we demonstrate that CoASH is the major thiol in *Bacillus anthracis*; a bioinformatics analysis indicates that three of the four proteins responsible for the conversion of pantothenate (Pan) to CoASH in *Escherichia coli* are conserved in *B. anthracis*. In contrast, a novel type III pantothenate kinase (PanK) catalyzes the first committed step in the biosynthetic pathway in *B. anthracis*; unlike the *E. coli* type I PanK, this enzyme is not subject to feedback inhibition by CoASH. The crystal structure of *B. anthracis* PanK (*Ba*PanK), solved using multiwavelength anomalous dispersion data and refined at a resolution of 2.0 Å, demonstrates that *Ba*PanK is a new member of the Acetate and Sugar Kinase/Hsc70/Actin (ASKHA) superfamily. The Pan and ATP substrates have been modeled into the active-site cleft; in addition to providing a clear rationale for the absence of CoASH inhibition, analysis of the Pan-binding pocket has led to the development of two new structure-based motifs (the PAN and INTERFACE motifs). Our analyses also suggest that the type III PanK in the spore-forming *B. anthracis* plays an essential role in the novel thiol/disulfide redox biology of this category A biodefense pathogen.

CoASH has been classified as a "common" thiol (*1*), as it occurs in all prokaryotes as well as eukaryotes. Although GSH plays the central role in maintaining thiol/disulfide homeostasis and providing an important line of antioxidant defense in, for example, *Escherichia coli* (*2*), GSH is clearly absent in a number of bacteria (*3*, *4*) as well as all archaea (*5*, *6*). CoASH has been shown to be the major low-molecular weight thiol in the human pathogens *Staphylococcus aureus* (*7*) and *Borrelia burgdorferi* (*8*), in the spore-forming *Bacillus megaterium* (*4*) and the radiation-resistant *Deinococcus radiodurans* (*3*), and in the hyperthermophile *Pyrococcus furiosus* (*6*) and other archaea. Coenzyme A-disulfide reductases (CoADR)[1] have been characterized in most of these prokaryotes (*4*, *7−9*), reinforcing the concept that CoASH assumes the intracellular redox function of GSH in these organisms, and the crystal structure of the *S. aureus* CoADR has recently been reported (*10*).

In *E. coli*, pantothenate kinase (CoaA, or type I PanK)[2] is the key regulatory enzyme in CoASH biosynthesis (Scheme 1); kinetic analysis supports an ordered sequential mecha-

Scheme 1: Pathway for CoASH Biosynthesis in *E. coli*[a]
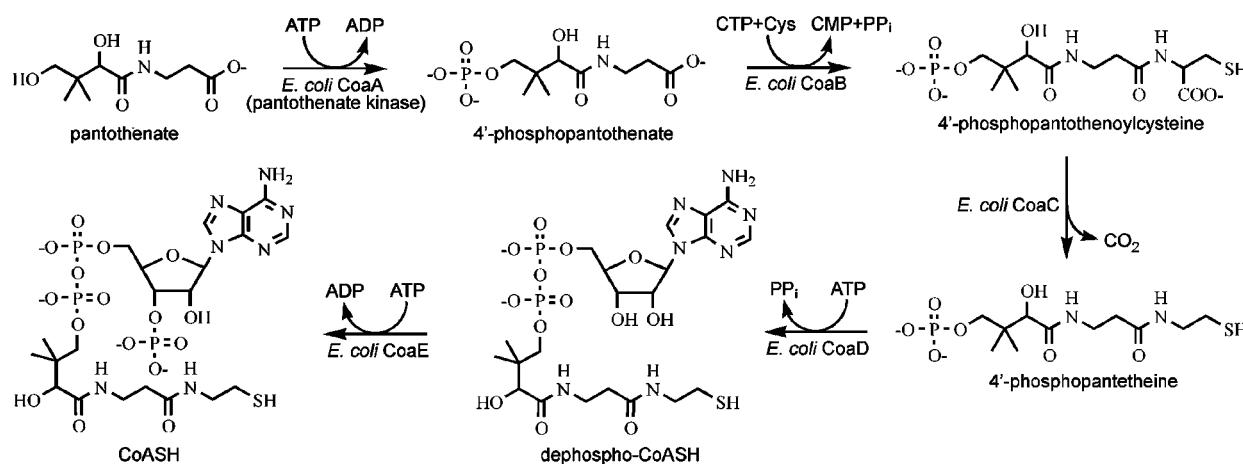
[a] *coaBC* encodes a single bifunctional polypeptide.

nism, with ATP binding first (*11*, *12*). CoASH is a competitive feedback inhibitor ($K_i = 4.3$ μM) with respect to ATP (*13*), and crystal structures are available for the *Ec*PanK complexes with CoASH and with the nonhydrolyzable ATP analog AMPPNP, respectively (*12*), as well as for the ternary complex with ADP and Pan (*14*). There are now three types of bacterial PanK enzymes; the type I PanK found in *E. coli* is a P-loop kinase [kinase family and fold group 2 (*15*)], and Pan binding leads to a significant conformational change, resulting in closure of a "lid" over the bound substrate (*14*). *Ec*PanK effectively phosphorylates the Pan antimetabolites (alternate substrates) *N*-pentyl- and *N*-heptylpantothenamide; K_m values are only 3-fold higher than that of the natural substrate (*14*). The phosphorylated products are carried through the CoASH biosynthetic pathway, resulting in the accumulation of inactive holo-acyl carrier protein (*16*). Modeling studies indicate that the pantothenamides can orient into the Pan binding site, with their extended alkyl chains interacting with a hydrophobic "dome" over the substrate pocket.

In 2005, the bacterial type II and type III PanKs were identified in *S. aureus* (*17*), and in *Helicobacter pylori* and *Bacillus subtilis* (*18*), respectively. Neither isoform is similar in sequence to the type I *Ec*PanK, nor is either subject to feedback inhibition by CoASH. Analysis of the respective kinetic parameters revealed a broad range for K_m(ATP) values, from 34 μM for *Sa*PanK to 3.1 mM for *Bs*PanK. The type III *Bs*PanK also exhibited a 5–6-fold increase in K_m(Pan) relative to *Ec*PanK and *Sa*PanK; this correlates with the observation that N5-Pan is neither an alternate substrate for nor a competitive inhibitor of *Bs*PanK. Limited bioinformatics analyses suggested that both type II and type III PanKs belonged to the ribonuclease H-like family [kinase family and fold group 4 (*15*)] of kinases (*18*). Very recently, Hong et al. (*19*) and Yang et al. (*20*) have confirmed these

predictions with crystal structures for the type II *Sa*PanK-AMPPNP complex and for the type III PanKs from *Pseudomonas aeruginosa* (*Pa*PanK) and *Thermotoga maritima* (*Tm*PanK), as well as the *Pa*PanK-Pan complex.

Global sequence identities between *Pa*PanK and *Tm*PanK, respectively, and the type III PanK from the category A biodefense pathogen *Bacillus anthracis* are 25% and 37%. *P. aeruginosa* is a Gram-negative aerobe that also contains the GSH/GSSG/glutathione reductase thiol/disulfide redox system (*21*–*23*); *T. maritima* is an anaerobic rod-shaped hyperthermophilic bacterium for which the ability to synthesize GSH has not been demonstrated (*23*). The CoADR enzymes, such as that from *S. aureus* (*7*, *24*, *25*), are a key part of the intracellular machinery for those bacteria and archaea in which CoASH serves as the major low-molecular weight thiol. Our structure-based bioinformatics analysis (*10*) suggests that CoADR may be a more common activity than previously recognized and one that is significantly misannotated in databases. In this report we demonstrate that CoASH is the major low-molecular weight thiol in the Gram-positive facultative anaerobe *B. anthracis*. We also describe the crystal structure for the type III PanK from this spore-forming organism, as refined at 2.0 Å resolution. The recent report of Yang et al. (*20*) included a bioinformatics analysis demonstrating that the type III PanK is also widely distributed among bacteria. Here we provide an extended bioinformatics analysis that provides a genome-based link for the absence of GSH biosynthesis, the presence of the type III PanK, and the presence of CoADR-like enzymes in a number of bacteria (Gram-positive and Gram-negative) and spirochetes. Taken together, these observations suggest that the CoASH-refractory type III PanK plays a special role in maintaining higher intracellular CoASH levels to accommodate additional functions for the coenzyme in thiol/disulfide redox homeostasis in a diverse group of bacteria.

EXPERIMENTAL PROCEDURES

Analysis of Thiols from B. anthracis. A 1-L culture of *B. anthracis* Sterne was grown in Brain Heart Infusion medium (Becton, Dickinson, and Co.) and harvested in exponential-phase growth. Four pellets (∼0.5 mL each) were extracted with 2.5 mL of 50% acetonitrile in 20 mM *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid, pH 8, containing

---

[2] The bacterial and eukaryotic pantothenate kinases have traditionally been distinguished as the CoaA and PanK proteins, respectively. With the identification of two distinct new bacterial pantothenate kinase classes in 2005, different investigators have used either type I, -II, and -III CoaAs (e.g., type II *Staphylococcus aureus* CoaA) or PanK-I, -II, and -III nomenclature to identify the three bacterial enzyme classes. In this report we refer to the three bacterial enzyme classes as bacterial type I, type II, and type III PanKs; CoaA is synonymous with the type I PanK.

2 mM mBBr (Calbiochem), and incubated at 60 °C for 15 min (*26*). The extract was cooled on ice and acidified with 5 μL of 5 M methanesulfonic acid. The pellet was removed by centrifugation and dried in a vacuum oven to constant weight. Four more pellets were extracted identically, except that 5 mM NEM was substituted for mBBr. A 0.5-mL aliquot of one NEM-treated pellet was reacted with 2 mM mBBr for 15 min at 60 °C and acidified with methanesulfonic acid, as given above; this sample served as a control for the thiol samples. The remaining $\sim$2 mL of this sample was reacted with 1 equiv of β-mercaptoethanol (based on NEM) and concentrated to 0.6 mL on a Speed Vac. This extract was reduced with 2 mM DTT for 20 min at 23 °C and then reacted with mBBr (8 mM overall). The sample was acidified with methanesulfonic acid prior to HPLC analysis. A DTT sample was prepared as above, without extract, to serve as a control for the disulfide sample.

The extracts were analyzed following HPLC protocols 1 and 2 (*26*) for thiols and disulfides. Single point calibration standards were used for quantitation (50 pmol) of the bimane derivatives of Cys, *N*-acetyl-Cys, GSH, thiosulfate (SSO_3^{-2}), H_2S, mycothiol, CoASH, 3′-dephospho-CoASH, pantetheine, and 4′-phosphopantetheine. The bimane derivative of CoASH hydrolyzes over time in acidic extracts to give the 3′-dephospho-CoASH-adduct and is quantitated as the sum of the two (CoASH + 3′-dephospho-CoASH) from HPLC protocol 2. The bimane derivatives of Cys, U12 (an unidentified *B. anthracis* thiol), pantetheine, 4′-phosphopantetheine, and H_2S were best analyzed with protocol 1.

Expression and Purification of BaPanK and BA2901. Genomic DNA from *B. anthracis* Ames was provided by Dr. Arthur Friedlander, U.S. Army Medical Research Institute of Infectious Diseases. *B. anthracis* PanK (NP_842634; "transcriptional activator, putative, Baf family") was expressed with a C-terminal His-tag provided by the pET28a plasmid vector in *E. coli* C41(DE3) cells. These cells were grown in a 3-L culture of tryptone−yeast extract−phosphate medium (*27*) supplemented with 30 mM glucose and 30 μg/mL kanamycin at 37 °C. Cultures were induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside at an A_{600} of 0.9 and allowed to grow overnight at 37 °C. Following centrifugation at 5000*g* for 15 min, the washed cell pellets were resuspended in $\sim$100 mL of 25 mM sodium phosphate, pH 7.0, containing 2 mM EDTA and 5 mM DTT, and lysed via disruption with a pneumatic cell homogenizer (Avestin EmulsiFlex-C5). After centrifuging at 20000*g*, streptomycin sulfate was added (1% w/v) to the supernatant, with stirring, for 15 min. The supernatant obtained on subsequent centrifugation was filtered and loaded onto a Q-Sepharose HP column equilibrated in 25 mM sodium phosphate, pH 7.0, with 5 mM β-mercaptoethanol. The protein was eluted with a 0 → 1 M NaCl gradient in the loading buffer; fractions were analyzed by SDS−polyacrylamide gel electrophoresis. The *Ba*PanK pool was brought to final concentrations of 0.25 M NaCl and 20 mM imidazole, and the pH was adjusted to 8.0 with 1 M K_2HPO_4. The protein was loaded onto a Ni-NTA Superflow (Qiagen) column equilibrated in 50 mM sodium phosphate, pH 8.0, containing 0.3 M NaCl and 5 mM β-mercaptoethanol. The His-tagged *Ba*PanK was eluted by increasing the imidazole concentration to 100 mM. The pure protein (calculated *m* = 29 915 Da for the His-tagged *Ba*PanK) was buffer-exchanged into 25 mM Tris−HCl, pH

8.0, containing 100 mM KCl, 0.5 mM EDTA, and 0.5 mM DTT and concentrated to $\sim$20 mg/mL, based on a calculated ϵ_{280} = 23 510 $M^{-1}cm^{-1}$ [ProtParam (*28*)]. After adding glycerol to 25% (v/v), 1-mL aliquots were stored at −80 °C. The SeMet derivative of *Ba*PanK was expressed from the pET28a-*Ba*PanK vector in the *E. coli* B834(DE3) strain (*10*); cultures were grown overnight in the auto-inducing medium PASM-5052 (*29*). The purification protocol for SeMet *Ba*PanK was essentially identical to that described for the native protein, except that the SeMet protein was eluted from the Ni-NTA column with 0.25 M imidazole. Yields for the native and SeMet *Ba*PanKs were $\sim$200 and $\sim$30−35 mg/L of culture, respectively.

The *B. anthracis* homolog (NP_845240; "hypothetical protein BA2901") of the type II *Sa*PanK was expressed, with either N-terminal or C-terminal His-tags provided by the pET28a plasmid vector, in *E. coli* B834(DE3) cells. These cells were grown in tryptone−yeast extract−phosphate medium, as described above for the *Ba*PanK expression and purification. N-terminal His-tagged BA2901 was purified in the same way as *Ba*PanK, except that an SP-Sepharose HP column replaced the Q-Sepharose HP column in the first step, given the higher calculated p*I* of 8.7 for BA2901. For the C-terminal His-tagged protein, the pET28a clone was expressed in *E. coli* B834(DE3) at 25 °C. The protein was purified with a combination of Q-Sepharose HP and Ni-NTA column steps and was stored at −80 °C in 50 mM Tris−HCl, pH 7.4 (at 25 °C), containing 1 mM DTT and 20% glycerol (v/v).

PanK Assays. The enzymatic activity of purified *Ba*PanK was measured as recently described for *Pa*PanK (*19*); the standard assay included 60 mM NH_4Cl to allow for activation of the enzyme in the presence of the monovalent cation. The purified BA2901 protein was assayed following the protocol described for *Sa*PanK (*17*). We also tested the ability of BA2901 to complement the *E. coli* mutant strain DV70 [*coaA*(Ts)] (*30*) that is unable to grow at 42 °C due to the inactivation of *Ec*PanK. Control strains were obtained by transforming strain DV70 either with pSC4, a pET28a plasmid harboring the *S. aureus coaA* gene (*17*), or with the empty vector. All strains grew very well at 30 °C, but only pSC4/DV70 grew (36 h) at 42 °C.

Crystallization and Data Collection. Thawed aliquots of the native and SeMet *Ba*PanK proteins were buffer-exchanged into 10 mM sodium *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid to give final protein concentrations of 22−24 mg/mL. The native protein as concentrated (pH 7.0) also included 50 mM KCl; the SeMet protein sample (pH 7.5) contained 10 mM KCl. Crystals were grown in 24-well sitting-drop plates over reservoirs of 0.5 mL of 24−26% ethylene glycol, in drop sizes of either 2 + 2 or 4 + 4 μL. Crystals were flash-frozen in a nitrogen stream at 100 K directly after being removed from the drop. Data sets for MAD phasing were collected at National Synchrotron Light Source beamline X12C.

Phasing and Structure Refinement. Images in each data set were indexed, integrated, and scaled in d*TREK (*31*). Phasing, chain tracing, and density modification were performed with SOLVE/RESOLVE (*32−34*). Model refinement was carried out in CNS (*35*), with manual rebuilding in COOT (*36*). The latter stages of refinement were performed with REFMAC5 (*37*). The final model was
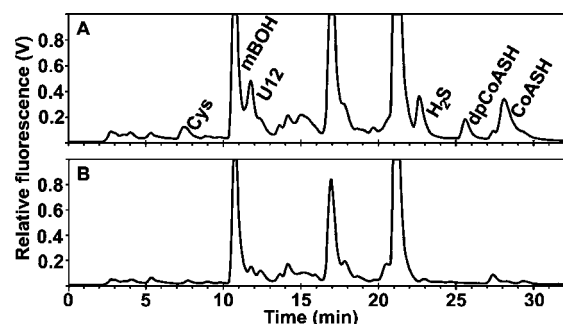
FIGURE 1: Reverse phase HPLC of *B. anthracis* thiols. The two chromatograms correspond to the thiol sample (A), in which cell extracts were reacted with mBBr, and the control sample (B), in which cell extracts were reacted first with NEM and then with mBBr. Peaks that appear in the thiol sample but not the control sample are bimane derivatives of cellular thiols. mBOH, hydrolysis product of mBBr; U12, unidentified *B. anthracis* thiol (bimane derivative) eluting at 12 min in HPLC protocols 1 and 2.

validated with MOLPROBITY (*38*) and with the built-in tools from COOT.

Bioinformatics. Sequence analyses with *Ba*PanK were performed with the NCBI (*39*) and SEED (*40*) databases using *blastp*. Multiple sequence alignments were performed with CLUSTALW (*41*). Secondary structural motifs were identified with DSSP (*42*), and structural homology searches were performed with the Combinatorial Extension [CE (*43*)] and DALI (*44*) web servers. Structural alignments and analyses were carried out in COOT, and structure-based sequence alignments were exported with SSM SUPERPOSE (*45*). The TIGR-Comprehensive Microbial Resource (*46*) Version 2.3 (Genomes Region Comparison tool; Data Releases through 19.0) and the Integrated Microbial Genomes Version 1.6 (phylogenetic distribution tool) were used to identify proteins similar to those in the test set, in other bacteria. The test set included *E. coli* GshA (NP_417173), *E. coli* GshB (NP_417422), *Streptococcus agalactiae* γ-glutamylcysteine synthetase−glutathione synthetase (NP_688811), *E. coli* type I PanK (NP_418405), *S. aureus* type II PanK (NP_646871), *B. subtilis* type III PanK (AAW83041), and *S. aureus* CoADR (YP_493573). Strong evidence for homology is provided by a sequence identity of $\geq 25\%$ (*47*); alignments for determining pairwise identities were performed using LALIGN [default parameters (*48*)].

RESULTS AND DISCUSSION

CoASH in Bacillus anthracis. To analyze thiols from *B. anthracis*, samples were taken from a 1-L culture of *B. anthracis* Sterne harvested in exponential-phase growth; bimane derivatives of low-molecular weight thiols resolved by reverse phase HPLC are given in Figure 1. *B. anthracis* produces CoASH, cysteine, and H_2S; the H_2S is likely derived from iron−sulfur proteins and is not a low-molecular weight thiol. This profile is qualitatively similar to that reported for *B. subtilis* but differs quantitatively in the relative amounts of CoASH and cysteine. GSH, pantetheine, and 4′-phosphopantetheine were not detectable [≤ 0.03 μmol/g (dry residual weight)]. Total CoASH was estimated at 0.87 ± 0.11 μmol/g (dry residual weight), very similar to the value of 1.1 ± 0.1 μmol/g (dry residual weight) reported for *S. aureus* (*7*). A new unknown thiol (U12) was also identified in the *B. anthracis* extract; purification and mass spectro-

metric analysis gave a mass of 398 Da for U12[3] (corrected for the bimane contribution); this thiol could not be identified by comparison with the retention times for known thiols as analyzed in HPLC protocols 1 and 2 (see "Experimental Procedures"). Parallel analyses for the corresponding low-molecular weight disulfides [CoA-disulfide ≤ 0.035 μmol/g (dry residual weight)] gave an intracellular ratio of CoASH/CoA-disulfide > 50. The method used to determine disulfides involves treatment of thiol-blocked extracts with DTT, to reduce the disulfide to its thiol form for mBBr derivatization and analysis. Because DTT can also release CoASH from acyl-CoAs by transacetylation (*49*), the value determined in this way represents an upper limit for the disulfide content; the estimated CoASH/CoA-disulfide ratio therefore represents a lower limit. Using a value of 2 μL/mg dry weight for the cell volume, we estimate the concentration of CoASH in *B. anthracis* cells to be ~ 0.44 mM. The concentration of CoASH in *E. coli* (1.5 μL/mg dry weight) can similarly be estimated at ~ 0.4 mM (*22*).

PanKs in B. anthracis. The conversion of Pan → CoASH requires three gene products, in addition to the type I PanK, in *E. coli* (Scheme 1): the bifunctional phosphopantothenoylcysteine synthetase/decarboxylase (phosphopantothenate + cysteine ⟶ 4′-phosphopantetheine; *coaBC*), phosphopantetheine adenylyltransferase (*coaD*), and dephospho-CoASH kinase (*coaE*). With specific reference to CoASH biosynthesis in *B. anthracis*, Gerdes et al. (*50*) published an analysis of potential broad-spectrum antimicrobial drug targets among cofactor biosynthetic pathways; all four *coa* gene products in *E. coli* were concluded to be individually "essential", as was the *acpS* gene product that synthesizes the holo-acyl carrier protein from CoASH + apo-acyl carrier protein (*51*). Using the ERGO database and the *B. anthracis* Ames genome sequence, the predicted CoASH biosynthetic pathway for the anthrax pathogen was analyzed with the aid of the known *E. coli* pathway (*50*). All components of the de novo Pan biosynthetic pathway, as well as the Na^+/Pan transporter, were identified (presence of the orthologous gene); similar results were found for the pathway from Pan → CoASH, *except* for the type I PanK. The locus identified in this case was the ortholog of the eukaryotic *panK* gene ["hypothetical protein BA2901" (NP_845240)], and the properties of the BA2901 protein are described below. Choudhry et al. (*52*) later demonstrated that the type II *S. aureus* PanK was closely related to BA2901 as well (35% identity), and both Choudhry et al. (*52*) and Genschel et al. (*53*) have suggested that this *B. anthracis* protein, which is also related to the human PanK, has likely arisen as the result of lateral gene transfer. We expressed and purified recombinant BA2901 with N- and C-terminal His-tags, respectively; the mass of the former (33.6 kDa) was verified by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Bruker Autoflex; sinapinic acid matrix). Both purified proteins were inactive in the in vitro PanK assay, under conditions in which the purified *Sa*PanK is fully active. In addition, the BA2901 pET28a expression plasmid

---

[3] A mixed disulfide form of the oxidized OhrR repressor has very recently been identified that is formed in *B. subtilis* cells on treatment with cumene hydroperoxide; the novel thiol component also gives a mass of 398 Da (J. Helmann, personal communication).

Table 1: Data Collection and Phasing Statistics

	edge[a]	peak[a]	remote[a]
wavelength (Å)	0.9801	0.9797	0.9500
resolution range (Å)	44.5−2.0 (2.1−2.0)[b]	44.5−2.0 (2.1−2.0)	44.5−2.1 (2.2−2.1)
no. of reflections	182 091	181 828	160 764
no. of unique reflections	26 148	26 073	22 673
completeness (%)	100 (100)	100 (100)	100 (100)
cell parameters (Å)	67.9, 133.5, 41.2	67.9, 133.4, 41.1	67.9, 133,.6 41.2
cell angles (deg)	90, 90, 90	90, 90, 90	90, 90, 90
redundancy	7.0 (5.8)	7.0 (5.8)	7.1 (7.3)
R_{merge}	0.04 (0.31)	0.06 (0.33)	0.05 (0.29)
I/σ	20.3 (5.7)	16.0 (4.8)	14.9 (4.7)
figure of merit	0.51		

[a] Collected at beamline X12C of the National Synchrotron Light Source (Brookhaven National Laboratory) using an ADSC Q210 CCD detector. [b] Values in parentheses are for the highest resolution shell.

failed to complement the *E. coli coaA*(Ts) strain DV70 under conditions where *S. aureus coaA*-pET28a restored growth at 42 °C. A homology model for BA2901 based on the type II *Sa*PanK structure (*19*) does not provide an explanation for the observed lack of activity. No crystal structure for the *Sa*PanK-Pan complex is available; the model described by Hong et al. (*19*) represents a "loose fit" for bound Pan. The critical residues involved in Pan binding have not been strictly identified or tested. Similarly, a putative mammalian PanK (PanK4) was originally identified on the basis of its very favorable comparison with the mammalian PanK1 sequence, but this protein is also inactive. We conclude that BA2901 is very likely a kinase of the ribonuclease H-like family, but Pan is not the acceptor substrate. Genetic studies already in progress[4] are designed to provide a more definitive answer to this question. If BA2901 does lack compensating PanK activity, deletion of the BA0065 locus should be lethal to *B. anthracis* Sterne (see below).

The absence of a type I PanK (and perhaps a functional type II enzyme as well, see above) in *B. anthracis* is distinct from the situation in *B. subtilis* and *Bacillus halodurans*, where functional type I and III enzymes are found (*18*, *54*). In *B. subtilis*, deletion of the *coaX* (type III) gene gave a normal growing phenotype, but *coaX* could not be deleted from a strain containing a *coaA* deletion; these results suggested that simultaneous deletion of both genes was lethal to *B. subtilis*. The *B. anthracis* Ames genome (*55*) also yields a type III PanK homolog, annotated as "putative Baf family transcriptional activator BA0065" (NP_842634), which is 76% identical to the functional *Bs*PanK. Of those bacterial species referred to earlier, known to lack GSH and containing CoASH as their major low-molecular weight thiol, *B. burgdorferi*, *D. radiodurans*, and *B. megaterium*[5] all have type III PanK homologs (*20*); all lack both type I and II PanK genes. All members of the *Bacillus cereus* group lack *coaA* orthologs, and all completed genomes of this group (*B. anthracis*, *B. cereus*, and *Bacillus thuringiensis*) contain type III PanK homologs 95−100% identical to that of the *B. anthracis* Ames strain. The bacterial type II and III PanK enzymes examined thus far are refractory to inhibition by CoASH (*17−19*), suggesting that bacteria dependent on CoASH as the major intracellular thiol tend to possess CoASH biosynthetic pathways insensitive to feedback inhibition.

---

[4] C. Paige, P. Hanna, and A. Claiborne, unpublished experiments.
[5] J. Ravel, personal communication.

Table 2: Crystallographic Refinement Statistics

resolution range (Å)	35.2−2.0 (2.1−2.0)[a]
no. of reflections total	26 020 (1910)
no. of reflections in work set	24 696 (1822)
no. of reflections in test set	1324 (88)
amplitude cutoff[b]	0
R-factor (%)	21.5 (24.6)
R-free (%)	25.8 (31.2)
no. of amino acid residues	247
no. of water molecules	180
no. of other molecules[c]	1
no. of total non-hydrogen atoms	2121
av *B*-factor (Å^2)	40.9
stereochemical ideality	
bond length rmsd (Å)	0.02
bond angle rmsd (deg)	1.61
ϕ, ψ favored (%)	98.4
ϕ, ψ allowed (%)	100

[a] Values in parentheses are for the highest resolution shell. [b] Refined against all data in the resolution range. [c] Ethylene glycol.

The C-terminal His-tagged BA0065 protein (*Ba*PanK) was expressed and purified from recombinant *E. coli*; in contrast to the BA2901 protein, *Ba*PanK is active in the in vitro enzyme assay, as examined in parallel with the type III *Pa*PanK. *Ba*PanK gives 60% conversion of D[1-^{14}C]pantothenate to 4′-phosphopantothenate under conditions where *Pa*PanK gives 40% conversion.

Structure Solution of BaPanK. The structure of *B. anthracis* PanK was solved by three-wavelength MAD phasing with the SeMet-substituted protein. Statistics for the data sets and final structure model are shown in Tables 1 and 2, respectively. Ten Se sites were identified, and the experimental MAD-phased electron density map calculated at 2.0 Å showed continuous electron density for the protein backbone, with the exception of a stretch of seven residues (Arg164-Arg170), and allowed identification of most side chain positions. The $2F_o − F_c$ map in the region of the PHOSPHATE1 motif (see below) also illustrates the accuracy of the final refined model (Figure 2). The C-terminus, including residues Asn255-Glu262 plus the His$_6$ tag, is not visible in the electron density map; the model is truncated at Ala254. The crystal structure of *Ba*PanK has one monomer in the asymmetric unit. Residues 164−170 are part of a long loop spanning residues Ser163-Asn183; the length of this loop, its location on the surface of the protein, and the absence of electron density for a part of it suggest that the entire loop is flexible. Yang et al. (*20*) have described the equivalent loop in *Tm*PanK as comprising the "Pan cap."
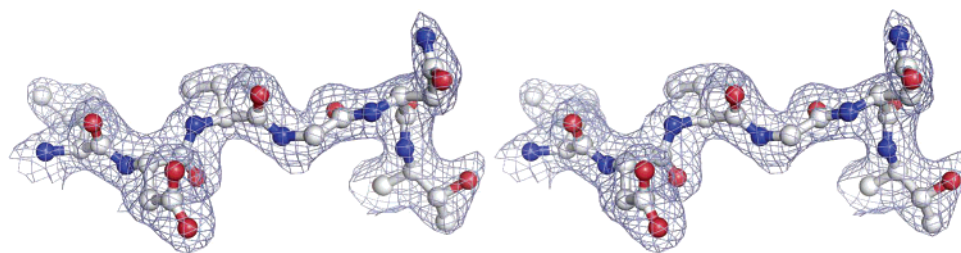
FIGURE 2: Final $2F_o - F_c$ map, in stereo, for residues Leu5-Thr10 of the PHOSPHATE1 motif, together with the refined model. All residues are color-coded by atom type. The depicted contour level is 1σ.
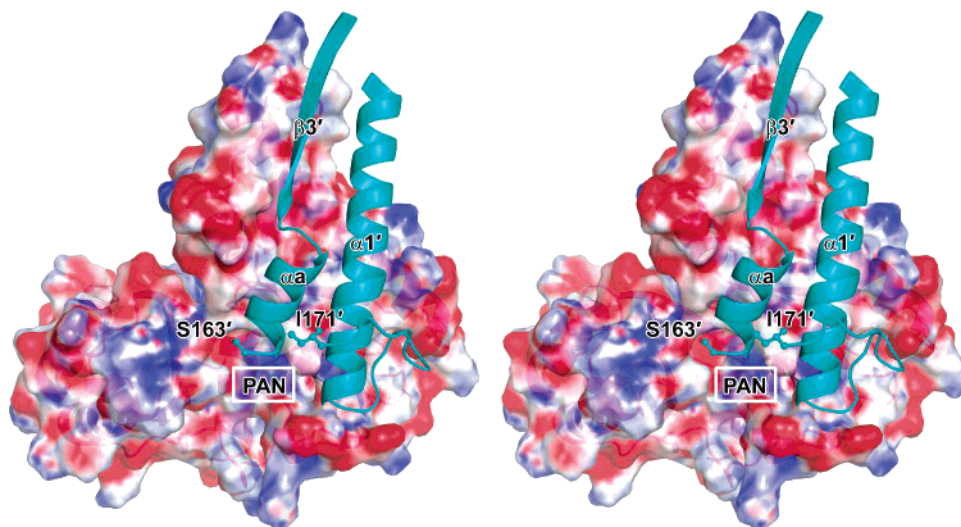

FIGURE 3: Elements involved in *Ba*PanK dimerization, in stereo. Prime symbols designate secondary structural units in domain II [of monomer B (cyan), in this case]; lettered designations (e.g., αa) indicate insertions relative to the shared core fold. These monomer B elements are overlaid with a surface representation (color-coded by electrostatic potential) of monomer A. The segment 164′−170′ missing electron density in monomer B is part of the long loop (163′−183′) connecting αa and α1′ in an area above the monomer A Pan-binding pocket.

The two constituent monomers of the *Ba*PanK dimer are related by crystallographic symmetry (a 2-fold rotation axis). Each monomer interacts with the other through a region consisting of β3′, αa, α1′, and the C-terminal part of the loop between βa and 3₁₀b (Figure 3). As expected, structural superpositions with the *Ba*PanK monomer and either *Pa*PanK (*19*) or *Tm*PanK (*20*) reveal nearly identical folds for the three type III PanKs; C_α rmsd values of 2.1 and 1.6 Å result with the respective A monomers (for 217 and 232 atoms, respectively). The type III PanKs have been identified as members of the ASKHA kinase superfamily [ribonuclease H-like, group 4 (*15*)]. Figure 4 gives a structure-based sequence alignment for the five known functional type III PanKs. As described, all share the three major sequence-structure motifs (*56*) attributed to the ASKHA superfamily (PHOSPHATE1, PHOSPHATE2, and ADENOSINE), as well as the interdomain helix-to-helix CONNECT elements (Ala116 and Gly245 in *Ba*PanK). The three major ASKHA motifs correspond to elements involved in ATP recognition and binding as well as divalent metal coordination for the relevant phosphotransferase (PanK) activity (*15*, *56*).

PAN and INTERFACE Motifs. *Ba*PanK has few insertions to its core ASKHA fold. Other members of the ASKHA superfamily, including the type II *Sa*PanK (*17*) and the eukaryotic PanKs (*15*), have several such insertions that adopt structures and motifs responsible for the formation of unique substrate-binding sites. Based on the structure for the *Pa*PanK-Pan complex [2F9W (*19*)], a model of the *Ba*PanK-Pan complex was developed. Most of the interactions

between protein and Pan occur with residues that are outside the ASKHA motifs; two insertions to the core fold in *Ba*PanK merit special attention. Broadly speaking, Pan binds in a pocket formed between β3′ and the loop connecting 3₁₀b and α3′ (Figure 5); this loop bears the conserved residues Gly107, Asp109, and Arg110 [the *hGh*DR sequence; $h =$ hydrophobic (Figure 4)] and constitutes the new PAN motif. Strand β3′ bears the conserved residues Gly147, Gly148, and Ile150, which play roles in Pan binding, as well as Pro152 and Gly153, which are important in the context of the dimer interface; together, these residues constitute the new INTERFACE motif [the GG*x*I*x*PG sequence (Figure 4)]. These new PAN and INTERFACE motifs are found in all type III PanK proteins (Figure S1) and serve to distinguish this from other ASKHA families; an inspection of the eukaryotic PanK family alignment (*15*) reveals that the PAN and INTERFACE motifs are unique to the type III sequences.

*Ba*PanK Asp109 forms side chain hydrogen bonds to the Pan C2′- and C4′-hydroxyls (Figure 6); this Asp is conserved within the PAN motif in the type III PanK sequences. *Ba*PanK Gly107-N [this Gly is also conserved within the PAN motif (Figure 4)] forms a hydrogen bond with the Pan C1-carboxylate. In addition, the absence of a side chain here avoids a potential steric clash with the bound substrate. *Ba*PanK Arg110 of the PAN motif (conserved in all known functional type III PanKs) and Tyr100 also interact with the Pan carboxylate. The Asp109 and Arg110 side chain rotamers, which form hydrogen bonds with Asp6 and Thr135, respectively, in the apoenzyme, are adjusted in the
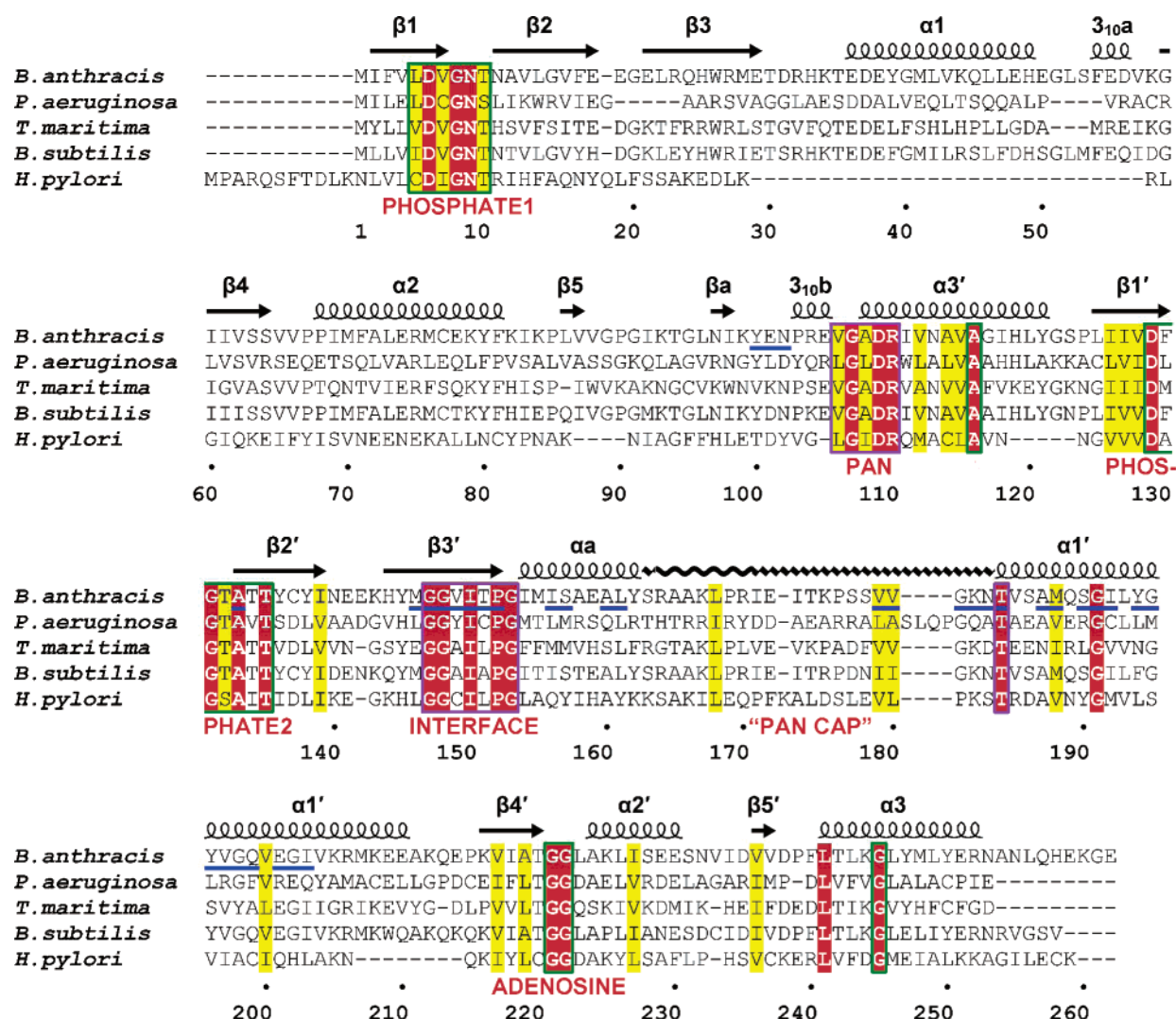
```
                    β1            β2            β3              α1            3₁₀a
                  ━━━━━▶       ━━━━━▶        ━━━━━▶      00000000000000   000    -
B.anthracis    ----------MIFVLDVGNTNAVLGVFE-EGELRQHWRMETDRHKTEDEYGMLVKQLLEHEGLSFEDVKG
P.aeruginosa   ----------MILELDCGNSLIKWRVIEG-----AARSVAGGLAESDDALVEQLTSQQALP----VRACR
T.maritima     ----------MYLIVDVGNTHSVFSITE-DGKTFRRWRLSTGVFQTEDELFSHLHPLLGDA---MREIKG
B.subtilis     ----------MLLVIDVGNTNTVLGVYH-DGKLEYHWRIETSRHKTEDEFGMILRSLFDHSGLMFEQIDG
H.pylori       MPARQSFTDLKNLVICDIGNTRIHFAQNYQLFSSAKEDLK--------------------------RL
                           PHOSPHATE1
                    1        10          20          30        40        50

                 β4            α2            β5           βa       3₁₀b       α3'          β1'
               ━━━▶     00000000000000    ━▶           ━▶      000   00000000000000   ━━━━▶
B.anthracis    IIVSSVVPPIMFALERMCEKYFKIKPLVVGPGIKTGLNIKYENPREVGADRIVNAVAGIHLYGSPLIIVDF
P.aeruginosa   LVSVRSEQETSQLVARLEQLFPVSALVASSGKQLAGVRNGYLDYQRLGLDRWLALVAAHHLAKKACLVIDL
T.maritima     IGVASVVPTQNTVIERFSQKYFHISP-IWVKAKNGCVKWNVKNPSEVGADRVANVVAFVKEYGKNGIIIDM
B.subtilis     IIISSVVPPIMFALERMCTKYFHIEPQIVGPGMKTGLNIKYDNPKEVGADRIVNAVAAIHLYGNPLIVVDF
H.pylori       GIQKEIFYISVNEENEKALLNCYPNAK----NIAGFFHLETDYVG-LGIDRQMACIAVN-----NGVVVDA
                                                                    PAN                  PHOS-
                 60        70        80        90        100       110       120       130

                    β2'           β3'          αa                              α1'
               ━━━▶         ━━━━━▶      00000000∿∿∿∿∿∿∿∿∿∿∿∿∿∿∿∿∿∿∿∿∿00000000000000
B.anthracis    GTATTYCYINEEKHYMGGVITPGIMISAEALYSRAAKLPRIE-ITKPSSVV----GKNTVSAMQSGILYG
P.aeruginosa   GTAVTSDLVAADGVHLGGYICPGMTLMRSQLRTHTRRIRYDD-AEARRALASLQPGQATAEAVERGCLLM
T.maritima     GTATTVDLVVN-GSYEGGAILPGFFMMVHSLFRGTAKLPLVE-VKPADFVV----GKDTEENIRLGVVNG
B.subtilis     GTATTYCYIDENKQYMGGAIAPGITISTEALYSRAAKLPRIE-ITRPDNII----GKNTVSAMQSGILFG
H.pylori       GSAITIDLIKE-GKHLGGCILPGLAQYIHAYKKSAKILEQPFKALDSLEVL----PKSTRDAVNYGMVLS
               PHATE2          INTERFACE               "PAN CAP"
                 140       150       160       170       180       190

                    α1'           β4'        α2'         β5'        α3
               00000000000000   ━━━▶     0000000      ━━▶    00000000000
B.anthracis    YVGQVEGIVKRMKEEAKQEPKVIATGGLAKLISEESNVIDVVDPFLTLKGLYMLYERNANLQHEKGE
P.aeruginosa   LRGFVREQYAMACELLGPDCEIFLTGGDAELVRDELAGARIMP-DLVFVGLALACPIE---------
T.maritima     SVYALEGIIGRIKEVYG-DLPVVLTGGQSKIVKDMIK-HEIFDEDLTIKGVYHFCFGD---------
B.subtilis     YVGQVEGIVKRMKWQAKQKQKVIATGGLAPLIANESDCIDIVDPFLTLKGLELIYERNRVGSV----
H.pylori       VIACIQHLAKN-------QKIYLCGRLVFDGMEIALKKAGILECK---
                                 ADENOSINE
                 200       210       220       230       240       250       260
```

FIGURE 4: Structure-based sequence alignment for *Ba*PanK and four known functional type III PanKs (references given in the text). Structures have now been determined for *Ba*PanK (this work), *Pa*PanK (*19*), and *Tm*PanK (*20*). Residue numbering and assignments correspond to *Ba*PanK and include the 164−170 segment for which electron density is absent. This segment constitutes a part of the loop (extending through Asn183) that exhibits different conformations in the three structures. Red and yellow blocks correspond to conserved residues and conservative substitutions, respectively. ASKHA motifs are highlighted in green boxes (Ala116 and Gly245 represent CONNECT1 and CONNECT2 elements), and the new PAN and INTERFACE motifs, as well as Thr184, are highlighted in magenta boxes. Blue underscores represent residues involved in direct contacts across the dimer interface.

*Ba*PanK-Pan model to optimize the interactions with bound Pan. Inspection of the *Tm*PanK structure (*20*) and the *Tm*PanK-Pan model reveals that a change in the apoenzyme Arg106 rotamer would also be necessary for optimization of that Pan-binding interaction.

Within the INTERFACE motif, Gly147 and Gly148 provide space at the end of the binding pocket for the Pan carboxylate. Any side chain at either of these positions would produce steric clashes with the bound Pan. Ile150 makes van der Waals contacts with bound Pan, but Pro152 and Gly153 are conserved, not for Pan binding, but for their roles in dimerization as described earlier. Conserved regions in the type II *Sa*PanK (*17*) and the eukaryotic PanKs (*15*) do not correlate with either the PAN or INTERFACE motifs; Hong et al. (*19*) have demonstrated that, although the *Sa*PanK and *Pa*PanK monomers have similar structures (SSM SUPERPOSE gives a C_α rmsd = 2.9−3.1 Å for 195 atoms), the respective homodimers assemble with distinctly different architectures. The *Sa*PanK-Pan model suggests that Glu70, which is structurally equivalent to *Pa*PanK Asp101, interacts

with the Pan C4′-OH; this Glu residue does not, however, appear within any PAN-like motif. *Sa*PanK Arg113, implicated in contact with the Pan C1-carboxylate, follows by 43 residues in the sequence (in contrast to the *hGh*DR motif in the type III PanKs). Further inspection of the *Sa*PanK sequence reveals that Arg113 is followed closely by Gly115-Gly116; these Gly residues are structurally equivalent to the conserved Gly147-Gly148 pair described for the INTERFACE motif of the type III PanKs. In the context of the dimer, the importance of another region in the type III PanKs becomes apparent; this region represents a unique insertion element, as compared with the type II *Sa*PanK. The segment corresponding to *Ba*PanK Leu161-Val180 (*Tm*PanK 159−178 and *Pa*PanK 153−176) is observed to exhibit different conformations in the three structures (Figure 7). Helix αa is present in all three structures to some degree, but the subsequent loop leading to helix α1′ shows significant differences. Similarly, residues 164−170 lack electron density in the *Ba*PanK structure. Yang et al. (*20*) describe several residues in this region as part of the "Pan cap" that
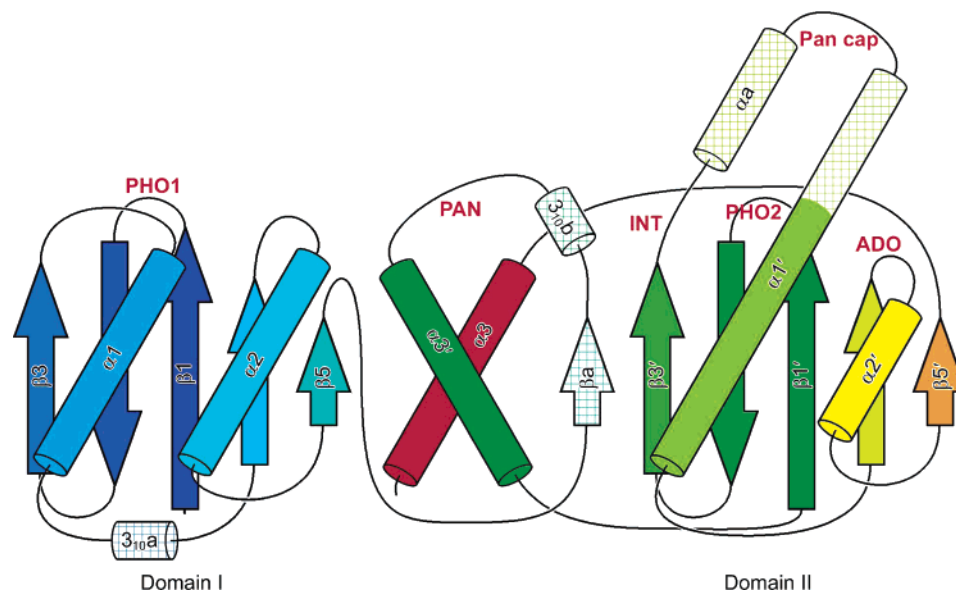
FIGURE 5: Topology diagram for the *Ba*PanK monomer. Domains I and II represent an internal duplication of the RNase H-like fold, and helices α3 and α3′ cross over to form the base of an interdomain cleft. Insertions to the core ASKHA fold are hatched; motifs indicated in Figure 4 are labeled in red (PHO, PHOSPHATE; INT, INTERFACE; ADO, ADENOSINE), as is the Pan cap segment. The diagram is represented as a rainbow from blue (N-terminus) to red (C-terminus).
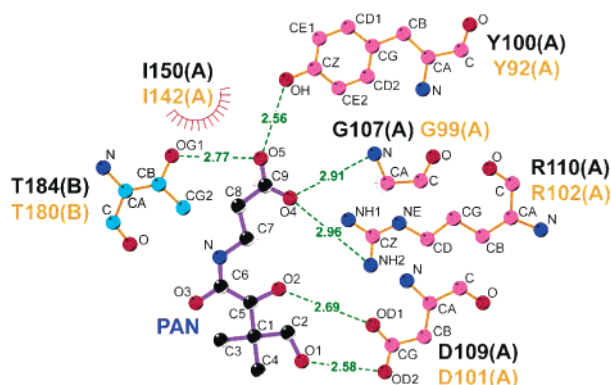

FIGURE 6: LIGPLOT representation of protein interactions with Pan in the *Pa*PanK-Pan complex (amino acid residues labeled in orange; equivalent *Ba*PanK residues in black), using default parameters for geometry and hydrogen-bonding distance. Carbon atoms for monomer A residues are magenta, and those for Thr180′ (monomer B) are cyan. Pan is color-coded by atom type, with carbon atoms in black; note that Pan atom numbering here follows that for the bound Pan ("DPA") in the Protein Data Bank file 2F9W.

contributes to the conserved hydrophobic pocket described earlier (including elements of the INTERFACE motif) that accommodates the two C3′-methyl groups of bound Pan. This cap is also proposed to protect the Pan-binding site—and the bound Pan substrate—from solvent.

The differential ordering/disordering of this region in the three distinct type III PanK structures cannot be explained in the context of the respective crystal lattices. Furthermore, the ordering of the region in monomer B of the apo-*Pa*PanK structure (versus its disordering in monomer A) is nearly identical to that seen in both monomers of the *Pa*PanK-Pan structure, so the conformational differences cannot be explained by the simple absence or presence of bound substrate. These observations have led to a preliminary examination of the conformational dynamics for the Pan cap. Our structural model for the *Ba*PanK-Pan complex is stable

under energy minimization.[6] In the same short molecular dynamics simulations, the Pan substrate diffuses out of its binding site in the *Ba*PanK dimer, suggesting that the unresolved loop nearby plays a role in binding the substrate. In addition to local rearrangements expected in the active site on Pan binding, domain closure may play a role in strengthening the interactions within the enzyme−substrate complex, as well as in excluding solvent from the phospho-transferase reaction center.

A conserved Thr residue near this flexible region shows a specific interaction with the bound Pan. In the model of the *Ba*PanK-Pan complex, though a large part of the long loop is missing electron density, Thr184′ is in nearly the same relative position as *Pa*PanK Thr180′ (Figure 6), suggesting that this residue is conserved to stabilize Pan in its binding site. Thr180′ of *Pa*PanK monomer B forms a hydrogen bond with the Pan carboxylate bound to monomer A.

Interactions with CoASH. CoASH is an important feedback inhibitor of the type I PanK, exhibiting a K_i of 4.3 μM (*13*); binding is competitive with respect to ATP, and crystal structures of the *Ec*PanK-AMPPNP and -CoASH complexes (*12*) demonstrated that the α- and β-phosphates of CoASH occupy the same site as the β- and γ-phosphates of ATP. In both cases, Lys101 is essential in mediating the interactions with the phosphate moieties (*11*). His177, Phe247, and Arg106 are unique and essential to CoASH binding and inhibition; H177Q, F247V, and R106A *Ec*PanK mutants are all refractory to any inhibition at [CoASH] $\leq$ 160 μM, nor do the mutants bind CoASH in equilibrium dialysis experiments (*57*). In the crystal structure of the *Ec*PanK-CoASH complex, the CoASH β-mercaptoethylamido moiety extends into the hydrophobic dome over the Pan-binding pocket; the same hydrophobic residues that interact with the extended alkyl chain of N5-Pan also form the binding groove for the
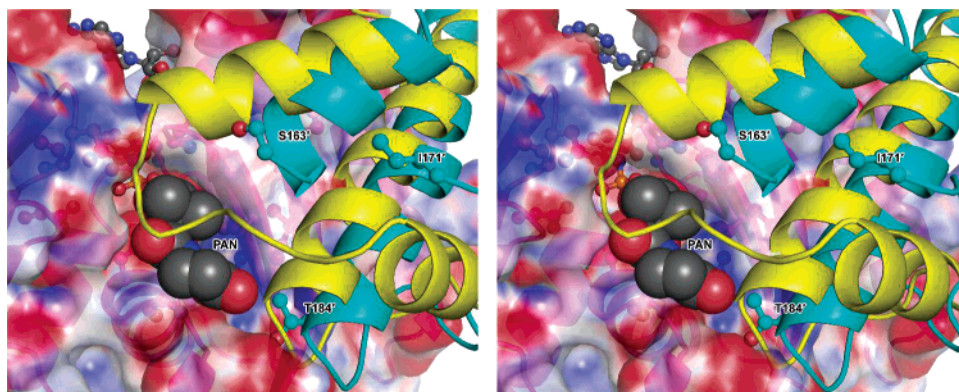

FIGURE 7: Ordered versus disordered Pan caps in *Pa*PanK and *Ba*PanK (stereo). The αa and α1′ helices of the *Ba*PanK-Pan model (cyan) are overlaid with a surface rendering (color-coded by electrostatic potential) of the Pan-binding pocket of monomer A. Bound Pan is represented with a space-filling model. Bound ATP is also shown, as modeled on the basis of the structure of the type II *Sa*PanK-AMPPNP complex. The αa and α1′ helices of *Pa*PanK monomer B (Pan complex, in yellow), with the Leu153′-Pro176′ loop, are shown from a superposition.

β-mercaptoethylamido moiety of CoASH (*12*). The thiol group is tightly sealed from solvent by weakly polar interactions with four aromatic residues from the hydrophobic dome and also participates in a hydrogen bond with the CoASH adenine 6-NH_2 group.

Although many of the protein elements defined in the *Ec*PanK-CoASH interaction are absent in *Ba*PanK, the most direct explanation for the failure of the type III PanKs to exhibit feedback inhibition by CoASH lies in the tight Pan-binding pocket seen in the *Ba*PanK-Pan model (Figure 7). The absence of the hydrophobic dome results in an inability to accommodate the β-mercaptoethylamido moiety of CoASH, in parallel with the failure to accommodate the alkyl substituents of N5-Pan and *N*-heptylpantothenamide. An identical conclusion regarding the absence of CoASH inhibition with the type III *Tm*PanK has recently been given by Yang et al. (*20*). As the type II *Sa*PanK Pan-binding pocket is fully exposed and can bind the pantothenamides (*19*), other factors must explain the resistance of *Sa*PanK to CoASH feedback inhibition. The biological consequences of these analyses are considered in the following section.

Type III PanKs and CoASH-Dependent Redox Biology. In the specific case of *S. aureus*, GSH is absent, as described earlier (*7*). CoASH is the major intracellular thiol, and a unique NADPH-dependent CoADR functions to maintain the appropriate intracellular thiol/disulfide redox balance, analogous to the GSH/GSSG/glutathione reductase system in *E. coli* (*10, 24, 25*); in *S. aureus*, the ratio of CoASH/CoA-disulfide is about 450. The presence of the type II PanK in *S. aureus* (*17, 19*), which is refractory to inhibition by CoASH, allows CoASH levels to rise to an upper limit (calculated to be ∼0.55 mM) likely controlled only by the supply of Pan, thus contributing to this organism's unique CoASH-dependent redox biology. Bacterial type II PanK homologs have been identified only in *Staphylococcus*, in the *B. cereus* group, and in *Oceanobacillus iheyensis* (*20*). In this work, however, we have been unable to demonstrate the catalytic ATP-dependent phosphorylation of Pan in vitro, using two different His-tagged versions of the purified recombinant type II PanK homolog from *B. anthracis* (BA2901), nor does the BA2901 expression plasmid complement the temperature-sensitive *E. coli coaA* mutant strain DV70. Our working conclusion for *B. anthracis* is that the type III PanK may represent the sole functional enzyme for

synthesis of 4′-phosphopantothenate; genetic studies already in progress[4] are designed to provide a conclusive answer to this question. The bioinformatics analysis of Copley and Dhillon (*23*), combined with biochemical analyses reported by the Fahey laboratory over the past 30 years (*1*), demonstrate that GSH biosynthesis is found "only rarely" in Gram-positive bacteria. The demonstration that functional CoADR enzymes are present in *S. aureus* (*7*), in *B. megaterium* (*4*) and *B. anthracis*,[7] and in *B. burgdorferi* (*8*) correlates perfectly with the presence of CoASH as the major low-molecular weight thiol, and with the presence of a type II or type III PanK.

In *B. megaterium* spores (*58*), 43% of the total CoASH was shown to be present as soluble protein-SSCoA disulfides; >75% of this protein-SSCoA disulfide pool was reduced within the first minutes of germination. Two likely functions for these protein-SSCoA disulfides were considered: (1) reversible inactivation of key thiol-dependent spore enzymes, maintaining metabolic dormancy in the spore and (2) reversible protection of labile protein thiols in the spore, thereby contributing to heat and radiation resistance. The protein-SSCoA disulfide pool in vegetative *B. megaterium* accounted for <2% of the total CoASH but increased dramatically to ∼45% of the total CoASH during sporulation. Formation and reduction of these protein-SSCoA disulfides (during spore formation and germination, respectively) has been attributed to a CoASH-mediated redox control process. The effective absence of reduced pyridine nucleotides (NADH and NADPH) in dormant spores of *B. megaterium* (*59*) corresponds to a dramatic decrease in the intracellular [CoASH]/[CoA-disulfide] redox state, which corresponds in turn to maximal protein-SSCoA disulfide formation. The rapid accumulation of NADH during the first minutes of spore germination (*59*) has been proposed to lead to reduction of the protein-SSCoA disulfide pool, either directly by a CoADR-like enzyme or indirectly, via a CoASH-mediated process (*4*).

The potential importance of such a CoASH-mediated germination control mechanism in *B. anthracis* is clear, as spore germination and outgrowth are fundamental to proliferation (*60*). The analysis of the transcriptome activated

---

[7] J.R. Wallen and A. Claiborne, unpublished experiments.

Table 3: Codistribution of CoASH-Refractory PanKs and CoADR-Like Proteins in Bacteria[a]

Bacillus anthracis[b,c]	*Oceanobacillus iheyensis*[b]
Bacillus cereus[b,c]	*Pelobacter carbinolicus*
Bacillus clausii[d]	*Rhodospirillum rubrum*
Bacillus halodurans[d]	*Staphylococcus aureus*[c,g]
Bacillus megaterium[e]	*Staphylococcus epidermidis*[g]
Bacteroides thetaiotamicron[f]	*Staphylococcus haemolyticus*[g]
Borrelia burgdorferi[f]	*Staphylococcus saprophyticus*[g]
Borrelia garinii[f]	*Streptomyces avermitilis*[d,h]
Clostridium tetani	*Streptomyces coelicolor*[d,h]
Deinococcus radiodurans[c]	*Symbiobacterium thermophilum*
Fusobacterium nucleatum	*Thermoanaerobacter tengcongensis*
Geobacillus kaustophilus	*Thermotoga maritima*
Geobacter sulfurreducens	*Thermus thermophilus*
Nocardia farcinica[d]	

[a] Analyzed with TIGR-Comprehensive Microbial Resource as described in "Experimental Procedures". [b] Contains homologs to both type II and III PanKs. [c] Known to lack GSH and to contain CoASH as major low-molecular weight thiol. [d] Contains homologs to both type I and III PanKs. [e] J. Ravel, personal communication. [f] Reference 20. [g] Contains only type II PanK. [h] Known to lack GSH; mycothiol is major low-molecular weight thiol.

during growth and sporulation of *B. anthracis* Sterne has provided additional support for this hypothesis (*61*). The *coaX* gene encoding *Ba*PanK is up-regulated in wave II, as is the *coaD* locus encoding phosphopantetheine adenylyltransferase; the *coaBC* and *coaE* genes corresponding to the bifunctional phosphopantothenoylcysteine synthetase/decarboxylase and dephospho-CoASH kinase, respectively, are up-regulated in waves I and III. Bergman et al. (*61*) have suggested that, because waves I and II are the phases of the *B. anthracis* life cycle that occur within the infected host, genes up-regulated specifically within this temporal frame may be particularly useful as therapeutic targets. The *B. anthracis* CoADR (BA1263) and CoADR-rhodanese homology domain (BA0774) proteins were expressed in waves III and V, respectively, and the presence of CoADR in the fractured spore was confirmed by proteomics analysis (*62*). At least four gene products essential for biosynthesis of the CoASH precursors Pan (BA1562, BA1563, and BA1564) and cysteine (BA1831) were also expressed in wave V. As described by Bergman et al. (*61*), waves IV and V seem to represent the final stages of sporulation in *B. anthracis* and include a large number of sporulation-associated loci as well as loci involved in oxidative stress responses.

In collaboration with Dr. Jacques Ravel of TIGR, we have undertaken a bioinformatics analysis using (1) the *E. coli* *gshA* and *gshB* genes (*23*), along with the bifunctional γ-glutamylcysteine synthetase-glutathione synthetase (SAG1821) gene from *Streptococcus agalactiae* (*63*), (2) the *S. aureus* CoADR gene (*24*), and (3) the *E. coli* type I (*64*), *S. aureus* type II (*17*), and *B. subtilis* type III (*18*) PanK genes, together with the TIGR Comprehensive Microbial Resource (version 2.3; Data Releases through 19.0). Table 3 lists those bacteria (genomes available as of April, 2006) that lack the ability to synthesize GSH, contain a type II or type III PanK, and contain a CoADR-like protein. Of the 27 species represented, for which we conclude that a CoASH-dependent redox biology pertains, only the four staphylococcal species lack the type III PanK. The absence of GSH has been confirmed for seven of the 27 species, and CoASH is known to be the major low-molecular weight thiol in six of these. It is of interest to note that even though mycothiol

is the major low-molecular weight thiol in *Streptomyces coelicolor* and other streptomycetes (*22*), the level of intracellular CoASH is similar to that in *S. aureus*. CoADR-like proteins (E values = $4e^{-41}$–$6e^{-39}$) are present in both *S. coelicolor* and *Streptomyces avermitilis*, suggesting a redundancy in low-molecular weight thiol/disulfide redox systems for this organism. As CoASH has been identified as an important component of the sporulation process in *B. megaterium* (see above), the CoASH/CoA-disulfide/CoADR system may also function in spore formation for streptomycetes.

In a parallel analysis, using the Integrated Microbial Genomes resource (*65*), 14 phylogenetic classes of bacteria (among the 30 classes analyzed) were identified as lacking the three marker proteins for GSH biosynthesis (GshA, GshB, and γ-glutamylcysteine synthetase-glutathione synthetase), and we conclude that these organisms lack the ability to synthesize GSH. All 14 classes also lack the type I PanK, and 13 of these classes have the CoASH-refractory type III PanK; eight classes have both the type III PanK and a CoADR-like protein. It should, of course, be emphasized that other important bacteria such as *Pseudomonas aeruginosa* have been demonstrated to contain millimolar levels of GSH (*22*); still they express only the type III PanK (*19*, *20*). Analysis of the *Helicobacter pylori* genome has been reported to reveal the absence of *gshA*, *gor*, and *grx* genes (*66*), implying the lack of a GSH/GSSG/glutathione reductase thiol/disulfide redox system. Thus, of the five known functional type III PanKs, only one is derived from a GSH-dependent microorganism (*P. aeruginosa*).

ACKNOWLEDGMENT

SUPPORTING INFORMATION AVAILABLE

Figure S1 depicting a structure-based alignment of 13 type III PanK sequences. This material is available free of charge via the Internet at http://pubs.acs.org

REFERENCES

1. Fahey, R. C. (2001) Novel thiols of prokaryotes, *Annu. Rev. Microbiol.* 55, 333−56.
2. Ortenberg, R., and Beckwith, J. (2003) Functions of thiol-disulfide oxidoreductases in *E. coli*: redox myths, realities, and practicalities, *Antioxid. Redox Signal.* 5, 403−11.
3. Fahey, R. C., and Newton, G. L. (1983) Occurrence of low molecular weight thiols in biological systems, in *Functions of Glutathione: Biochemical, Physiological, Toxicological, and Clinical Aspects* (Larsson, A., et al., Ed.) pp 251−60, Raven Press, New York.
4. Swerdlow, R. D., and Setlow, P. (1983) Purification and characterization of a *Bacillus megaterium* disulfide reductase specific for disulfides containing pantethine 4′,4″-diphosphate, *J. Bacteriol.* 153, 475−84.

5. Newton, G. L., and Fahey, R. C. (1990) Glutathione in prokaryotes, in *Glutathione: Metabolism and Physiological Functions* (Vina, J., Ed.) pp 69−77, CRC Press, Boca Raton, FL.

6. Hummel, C. S., Lancaster, K. M., and Crane, E. J., III. (2005) Determination of coenzyme A levels in *Pyrococcus furiosus* and other Archaea: implications for a general role for coenzyme A in thermophiles, *FEMS Microbiol. Lett. 252*, 229−34.

7. delCardayre, S. B., Stock, K. P., Newton, G. L., Fahey, R. C., and Davies, J. E. (1998) Coenzyme A disulfide reductase, the primary low molecular weight disulfide reductase from *Staphylococcus aureus*. Purification and characterization of the native enzyme, *J. Biol. Chem. 273*, 5744−51.

8. Boylan, J. A., Hummel, C. S., Benoit, S., Garcia-Lara, J., Treglown-Downey, J., Crane, E. J., III, and Gherardini, F. C. (2006) *Borrelia burgdorferi* bb0728 encodes a coenzyme A disulfide reductase whose function suggests a role in intracellular redox and the oxidative stress response, *Mol. Microbiol. 59*, 475−86.

9. Harris, D. R., Ward, D. E., Feasel, J. M., Lancaster, K. M., Murphy, R. D., Mallett, T. C., and Crane, E. J., III. (2005) Discovery and characterization of a Coenzyme A disulfide reductase from *Pyrococcus horikoshii*. Implications for the disulfide metabolism of anaerobic hyperthermophiles, *FEBS J. 272*, 1189−200.

10. Mallett, T. C., Wallen, J. R., Karplus, P. A., Sakai, H., Tsukihara, T., and Claiborne, A. (2006) Structure of Coenzyme A-Disulfide Reductase from *Staphylococcus aureus* at 1.54 Å Resolution, *Biochemistry 45*, 11278−89.

11. Song, W. J., and Jackowski, S. (1994) Kinetics and regulation of pantothenate kinase from *Escherichia coli*, *J. Biol. Chem. 269*, 27051−8.

12. Yun, M., Park, C. G., Kim, J. Y., Rock, C. O., Jackowski, S., and Park, H. W. (2000) Structural basis for the feedback regulation of *Escherichia coli* pantothenate kinase by coenzyme A, *J. Biol. Chem. 275*, 28093−9.

13. Vallari, D. S., Jackowski, S., and Rock, C. O. (1987) Regulation of pantothenate kinase by coenzyme A and its thioesters, *J. Biol. Chem. 262*, 2468−71.

14. Ivey, R. A., Zhang, Y. M., Virga, K. G., Hevener, K., Lee, R. E., Rock, C. O., Jackowski, S., and Park, H. W. (2004) The structure of the pantothenate kinase·ADP·pantothenate ternary complex reveals the relationship between the binding sites for substrate, allosteric regulator, and antimetabolites, *J. Biol. Chem. 279*, 35622−9.

15. Cheek, S., Ginalski, K., Zhang, H., and Grishin, N. V. (2005) A comprehensive update of the sequence and structure classification of kinases, *BMC Struct. Biol. 5*, 6.

16. Zhang, Y. M., Frank, M. W., Virga, K. G., Lee, R. E., Rock, C. O., and Jackowski, S. (2004) Acyl carrier protein is a cellular target for the antibacterial action of the pantothenamide class of pantothenate antimetabolites, *J. Biol. Chem. 279*, 50969−75.

17. Leonardi, R., Chohnan, S., Zhang, Y. M., Virga, K. G., Lee, R. E., Rock, C. O., and Jackowski, S. (2005) A pantothenate kinase from *Staphylococcus aureus* refractory to feedback regulation by coenzyme A, *J. Biol. Chem. 280*, 3314−22.

18. Brand, L. A., and Strauss, E. (2005) Characterization of a new pantothenate kinase isoform from *Helicobacter pylori*, *J. Biol. Chem. 280*, 20185−8.

19. Hong, B. S., Yun, M. K., Zhang, Y. M., Chohnan, S., Rock, C. O., White, S. W., Jackowski, S., Park, H. W., and Leonardi, R. (2006) Prokaryotic type II and type III pantothenate kinases: The same monomer fold creates dimers with distinct catalytic properties, *Structure 14*, 1251−61.

20. Yang, K., Eyobo, Y., Brand, L. A., Martynowski, D., Tomchick, D., Strauss, E., and Zhang, H. (2006) Crystal structure of a type III pantothenate kinase: insight into the mechanism of an essential coenzyme A biosynthetic enzyme universally distributed in bacteria, *J. Bacteriol. 188*, 5532−40.

21. Perry, A. C., Ni Bhriain, N., Brown, N. L., and Rouch, D. A. (1991) Molecular characterization of the *gor* gene encoding glutathione reductase from *Pseudomonas aeruginosa*: determinants of substrate specificity among pyridine nucleotide-disulphide oxidoreductases, *Mol. Microbiol. 5*, 163−71.

22. Newton, G. L., Arnold, K., Price, M. S., Sherrill, C., delCardayre, S. B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R. C., and Davis, C. (1996) Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes, *J. Bacteriol. 178*, 1990−5.

23. Copley, S. D., and Dhillon, J. K. (2002) Lateral gene transfer and parallel evolution in the history of glutathione biosynthesis genes, *Genome Biol. 3*, research0025.1−0025.16.

24. delCardayre, S. B., and Davies, J. E. (1998) *Staphylococcus aureus* coenzyme A disulfide reductase, a new subfamily of pyridine nucleotide-disulfide oxidoreductase. Sequence, expression, and analysis of *cdr*, *J. Biol. Chem. 273*, 5752−7.

25. Luba, J., Charrier, V., and Claiborne, A. (1999) Coenzyme A-disulfide reductase from *Staphylococcus aureus*: evidence for asymmetric behavior on interaction with pyridine nucleotides, *Biochemistry 38*, 2725−37.

26. Fahey, R. C., and Newton, G. L. (1987) Determination of low-molecular-weight thiols using monobromobimane fluorescent labeling and high-performance liquid chromatography, *Methods Enzymol. 143*, 85−96.

27. Parsonage, D., Miller, H., Ross, R. P., and Claiborne, A. (1993) Purification and analysis of streptococcal NADH peroxidase expressed in *Escherichia coli*, *J. Biol. Chem. 268*, 3161−7.

28. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server, in *The Proteomics Protocols Handbook* (Walker, J. M., Ed.) pp 571−607, Humana Press Inc., Totowa, NJ.

29. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein Expr. Purif. 41*, 207−34.

30. Calder, R. B., Williams, R. S., Ramaswamy, G., Rock, C. O., Campbell, E., Unkles, S. E., Kinghorn, J. R., and Jackowski, S. (1999) Cloning and characterization of a eukaryotic pantothenate kinase gene (*panK*) from *Aspergillus nidulans*, *J. Biol. Chem. 274*, 2014−20.

31. Pflugrath, J. W. (1999) The finer things in X-ray diffraction data collection, *Acta Crystallogr. D. Biol. Crystallogr. 55*, 1718−25.

32. Terwilliger, T. C. (2000) Maximum-likelihood density modification, *Acta Crystallogr. D. Biol. Crystallogr. 56*, 965−72.

33. Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension, *Acta Crystallogr. D Biol. Crystallogr. 59*, 38−44.

34. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution, *Acta Crystallogr. D Biol. Crystallogr. 55*, 849−61.

35. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D. Biol. Crystallogr. 54*, 905−21.

36. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr. 60*, 2126−32.

37. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D Biol. Crystallogr. 53*, 240−55.

38. Lovell, S. C., Davis, I. W., Arendall, W. B., III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Cα geometry: ϕ,ψ and Cβ deviation, *Proteins 50*, 437−50.

39. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *J. Mol. Biol. 215*, 403−10.

40. Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., de Crecy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E. D., Gerdes, S., Glass, E. M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N., Linke, B., McHardy, A. C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G. D., Rodionov, D. A., Ruckert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., and Vonstein, V. (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes, *Nucleic Acids Res. 33*, 5691−702.

41. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−80.

42. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577−637.

43. Shindyalov, I. N., and Bourne, P. E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path, *Protein Eng. 11*, 739−47.

44. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices, *J. Mol. Biol. 233*, 123−38.

45. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, *Acta Crystallogr. D Biol. Crystallogr. 60*, 2256−68.

46. Peterson, J. D., Umayam, L. A., Dickinson, T., Hickey, E. K., and White, O. (2001) The Comprehensive Microbial Resource, *Nucleic Acids Res. 29*, 123−5.

47. Karplus, P. A. (1999) Structural flavinology on the brink, in *Flavins and Flavoproteins 1999* (Ghisla, S., Kroneck, P., Macheroux, P., and Sund, H., Eds.) pp 233−8, Agency for Scientific Publications, Berlin.

48. Huang, X., and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm, *Adv. Appl. Math. 12*, 337−57.

49. Fenton, S. S., and Fahey, R. C. (1986) Analysis of biological thiols: determination of thiol components of disulfides and thioesters, *Anal. Biochem. 154*, 34−42.

50. Gerdes, S. Y., Scholle, M. D., D'Souza, M., Bernal, A., Baev, M. V., Farrell, M., Kurnasov, O. V., Daugherty, M. D., Mseeh, F., Polanuyer, B. M., Campbell, J. W., Anantha, S., Shatalin, K. Y., Chowdhury, S. A., Fonstein, M. Y., and Osterman, A. L. (2002) From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways, *J. Bacteriol. 184*, 4555−72.

51. Flugel, R. S., Hwangbo, Y., Lambalot, R. H., Cronan, J. E., Jr., and Walsh, C. T. (2000) Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in *Escherichia coli*, *J. Biol. Chem. 275*, 959−68.

52. Choudhry, A. E., Mandichak, T. L., Broskey, J. P., Egolf, R. W., Kinsland, C., Begley, T. P., Seefeld, M. A., Ku, T. W., Brown, J. R., Zalacain, M., and Ratnam, K. (2003) Inhibitors of pantothenate kinase: novel antibiotics for staphylococcal infections, *Antimicrob. Agents Chemother. 47*, 2051−5.

53. Genschel, U. (2004) Coenzyme A biosynthesis: reconstruction of the pathway in archaea and an evolutionary scenario based on comparative genomics, *Mol. Biol. Evol. 21*, 1242−51.

54. Yocum, R. R., and Patterson, T. A. (December 14, 2004) U.S. Patent 6,830,898.

55. Read, T. D., Peterson, S. N., Tourasse, N., Baillie, L. W., Paulsen, I. T., Nelson, K. E., Tettelin, H., Fouts, D. E., Eisen, J. A., Gill, S. R., Holtzapple, E. K., Okstad, O. A., Helgason, E., Rilstone, J., Wu, M., Kolonay, J. F., Beanan, M. J., Dodson, R. J., Brinkac, L. M., Gwinn, M., DeBoy, R. T., Madpu, R., Daugherty, S. C., Durkin, A. S., Haft, D. H., Nelson, W. C., Peterson, J. D., Pop, M., Khouri, H. M., Radune, D., Benton, J. L., Mahamoud, Y., Jiang, L., Hance, I. R., Weidman, J. F., Berry, K. J., Plaut, R. D., Wolf, A. M., Watkins, K. L., Nierman, W. C., Hazen, A., Cline, R., Redmond, C., Thwaite, J. E., White, O., Salzberg, S. L., Thomason, B., Friedlander, A. M., Koehler, T. M., Hanna, P. C., Kolsto, A. B., and Fraser, C. M. (2003) The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria, *Nature 423*, 81−6.

56. Bork, P., Sander, C., and Valencia, A. (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins, *Proc. Natl. Acad. Sci. U.S.A. 89*, 7290−4.

57. Rock, C. O., Park, H. W., and Jackowski, S. (2003) Role of feedback regulation of pantothenate kinase (CoaA) in control of coenzyme A levels in *Escherichia coli*, *J. Bacteriol. 185*, 3410−5.

58. Setlow, B., and Setlow, P. (1977) Levels of acetyl coenzyme A, reduced and oxidized coenzyme A, and coenzyme A in disulfide linkage to protein in dormant and germinated spores and growing and sporulating cells of *Bacillus megaterium*, *J. Bacteriol. 132*, 444−52.

59. Setlow, B., and Setlow, P. (1977) Levels of oxidized and reduced pyridine nucleotides in dormant spores and during growth, sporulation, and spore germination of *Bacillus megaterium*, *J. Bacteriol. 129*, 857−65.

60. Inglesby, T. V., O'Toole, T., Henderson, D. A., Bartlett, J. G., Ascher, M. S., Eitzen, E., Friedlander, A. M., Gerberding, J., Hauer, J., Hughes, J., McDade, J., Osterholm, M. T., Parker, G., Perl, T. M., Russell, P. K., and Tonat, K. (2002) Anthrax as a biological weapon, 2002: updated recommendations for management, *JAMA 287*, 2236−52.

61. Bergman, N. H., Anderson, E. C., Swenson, E. E., Niemeyer, M. M., Miyoshi, A. D., and Hanna, P. C. (2006) Transcriptional profiling of the *Bacillus anthracis* life cycle *in vitro* and an implied model for regulation of spore formation, *J. Bacteriol. 188*, 6092−100.

62. Liu, H., Bergman, N. H., Thomason, B., Shallom, S., Hazen, A., Crossno, J., Rasko, D. A., Ravel, J., Read, T. D., Peterson, S. N., Yates, J., III, and Hanna, P. C. (2004) Formation and composition of the *Bacillus anthracis* endospore, *J. Bacteriol. 186*, 164−78.

63. Janowiak, B. E., and Griffith, O. W. (2005) Glutathione synthesis in *Streptococcus agalactiae*. One protein accounts for γ-glutamylcysteine synthetase and glutathione synthetase activities, *J. Biol. Chem. 280*, 11829−39.

64. Song, W. J., and Jackowski, S. (1992) Cloning, sequencing, and expression of the pantothenate kinase (*coaA*) gene of *Escherichia coli*, *J. Bacteriol. 174*, 6411−7.

65. Markowitz, V. M., Korzeniewski, F., Palaniappan, K., Szeto, E., Werner, G., Padki, A., Zhao, X., Dubchak, I., Hugenholtz, P., Anderson, I., Lykidis, A., Mavromatis, K., Ivanova, N., and Kyrpides, N. C. (2006) The integrated microbial genomes (IMG) system, *Nucleic Acids Res. 34*, D344−8.

66. Baker, L. M., Raudonikiene, A., Hoffman, P. S., and Poole, L. B. (2001) Essential thioredoxin-dependent peroxiredoxin system from *Helicobacter pylori*: genetic and kinetic characterization, *J. Bacteriol. 183*, 1961−73.