Articles

Insights into the Evolution of Allosteric Properties. The NADH Binding Site of Hexameric Type II Citrate Synthases[†,‡]

Robert Maurus,[§] Nham T. Nguyen,[§] David J. Stokell,[||] Ayeda Ayed,[||] Philip G. Hultin,[||] Harry W. Duckworth,[||] and Gary D. Brayer*[,§]

Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada, and Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

Received October 9, 2002

ABSTRACT: Study of the hexameric and allosterically regulated citrate synthases (type II CS) provides a rare opportunity to gain not only an understanding of a novel allosteric mechanism but also insight into how such properties can evolve from an unregulated structural platform (the dimeric type I CS). To address both of these issues, we have determined the structure of the complex of NADH (a negative allosteric effector) with the F383A variant of type II *Escherichia coli* CS. This variant was chosen because its kinetics indicate it is primarily in the T or inactive allosteric conformation, the state that strongly binds to NADH. Our structural analyses show that the six NADH binding sites in the hexameric CS complex are located at the interfaces between dimer units such that most of each site is formed by one subunit, but a number of key residues are drawn from the adjacent dimer. This arrangement of interactions serves to explain why NADH allosteric regulation is a feature only of hexameric type II CS. Surprisingly, in both the wild-type enzyme and the NADH complex, the two subunits of each dimer within the hexameric conformation are similar but not identical in structure, and therefore, while the general characteristics of NADH binding interactions are similar in each subunit, the details of these are somewhat different between subunits. Detailed examination of the observed NADH binding sites indicates that both direct charged interactions and the overall cationic nature of the sites are likely responsible for the ability of these sites to discriminate between NADH and NAD^+. A particularly novel characteristic of the complex is the horseshoe conformation assumed by NADH, which is strikingly different from the extended conformation found in its complexes with most proteins. Sequence homology studies suggest that this approach to binding NADH may arise out of the evolutionary need to add an allosteric regulatory function to the base CS structure. Comparisons of the amino acid sequences of known type II CS enzymes, from different Gram-negative bacteria taxonomic groups, show that the NADH-binding residues identified in our structure are strongly conserved, while hexameric CS molecules that are insensitive to NADH have undergone key changes in the sequence of this part of the protein.

It is axiomatic that the regulatory properties of an enzyme should evolve after the catalytic ones, but there are few cases in which concrete insights into the evolution of regulatory properties can be obtained. One example is provided by the citrate synthases (CS),[1] ancient enzymes that catalyze the entry of acetate carbon into the tricarboxylic acid cycle. In eukaryotes, archaea, and Gram-positive bacteria, CS is a homodimer, with two active sites, each one formed mainly from one subunit but with a few residues contributed from the other. This kind of CS, which does not show regulatory properties, is classified as type I. Its structure has been established, with a number of studies of such enzymes from vertebrates (*3*−*5*) and archaea (*6*−*8*).

In most Gram-negative bacteria, on the other hand, a strikingly different CS is found, designated type II. Type II CS is a hexamer of identical subunits, and is strongly and specifically inhibited by NADH by an allosteric mechanism (*9*−*14*). Amino acid sequence comparisons suggest that type I and type II subunits have similar overall folds (*15*), and this was recently confirmed when we determined the structure of the type II CS from *Escherichia coli* (*2*). In this

[†] This work was supported by a joint operating grant to H.W.D., G.D.B., and P.G.H. from the Canadian Institutes of Health Research (CIHR) and the Manitoba Health Research Council, under the Regional Partnership Program of the CIHR.

[‡] Coordinates for the structures described in this work have been deposited in the Protein Data Bank 1NXE for F383A and 1NXG for F383A−NADH.

* To whom correspondence should be addressed. Telephone: (604) 822-5216. Fax: (604) 822-5227. E-mail: brayer@interchange.ubc.ca.

[§] University of British Columbia.

[||] University of Manitoba.

[1] Abbreviation: CS, citrate synthase(s). Amino acid numbering is according to the sequence of *E. coli* citrate synthase (*2*).

10.1021/bi020622s CCC: $25.00
Published on Web 04/25/2003

enzyme, the hexamer is formed by three dimers, each of which is folded globally in a manner similar to that of a type I enzyme, in a two-domain arrangement containing 18 α-helices. The major structural differences are the presence of a novel N-terminal domain of 52 residues, rich in β-sheet, which is absent in type I CS, and differing lengths for a number of interhelical loops (*2*).

Our structure of *E. coli* CS has suggested a potential pathway by which an allosteric type II enzyme could have evolved from a nonallosteric type I ancestor. In total, three additional structural features would seem to be required: (1) development of contact surfaces so that three dimers could assemble into a hexamer, (2) elaboration of a binding site for NADH, and (3) adjustment of the free energies of possible conformational states so that the binding of NADH inhibits enzyme activity, that is, the creation of an R state−T state equilibrium [using the terminology of Monod et al. (*16*)].

In earlier work, we addressed the issue of how the type II CS have developed hexameric contact surfaces (*2*). The new structures in this paper allow us now to characterize the NADH binding sites in the hexameric form of type II CS and thereby provide some insight into the evolution and nature of NADH regulation over the catalytic activities of these enzymes.

EXPERIMENTAL PROCEDURES

Citrate Synthase Mutagenesis and Variant Protein Purification. Mutagenesis techniques used to construct F383A variant CS have been described previously (*17*). This protein was prepared using the same method that was used for wild-type *E. coli* CS (*18*). The homogeneity and subunit mass were verified by electrospray time-of-flight mass spectrometry (*19*).

Structural Analyses. Recombinant *E. coli* F383A variant CS was crystallized using the hanging drop vapor diffusion method from 2 to 2.2 M ammonium sulfate, 2% (v/v) PEG 400, and 0.1 M Na Hepes at pH 7.5. The starting protein concentration was 20 mg/mL. Crystals of NADH-bound F383A CS were obtained by soaking variant crystals in a solution of 0.7 mM NADH containing 2.8 M ammonium sulfate, 2% (v/v) PEG 400, and 0.1 M Na Hepes at pH 7.5 for 4 h.

Diffraction data for single crystals of the F383A variant and the F383A−NADH complex were collected at 100 K on a Rigaku R-AXIS IIC imaging plate area detector system equipped with Osmic mirrors and using Cu Kα radiation supplied by a Rigaku RU300 rotating anode generator operating at 50 kV and 100 mA. A total of 100 frames (1.0° oscillation, 8 min exposures) of data were collected from each crystal. Intensity data were integrated, scaled, and reduced to structure factor amplitudes with the HKL (*20*) and CCP4 (*21*) program suites.

As shown in Table 1, the diffraction data that were collected indicated that crystals of the F383A variant protein and the F383A−NADH complex were isomorphous with those of wild-type *E. coli* CS. This allowed for the use of the wild-type structure (with residue 383 truncated to an alanine) as the starting refinement model for the F383A variant protein. Similarly, the final structure determined for the F383A variant protein served as the starting refinement model for the F383A−NADH complex. For both structure determinations, it was necessary to apply a detwinning approach (perfect merohedral), in conjunction with positional and thermal factor refinement (*22*, *23*). During this process, the complete polypeptide chain of each refinement model was examined periodically with $F_o - F_c$, $2F_o - F_c$, and fragment-deleted difference electron density maps. Where necessary, manual adjustments were made to the refinement model using O (*24*). The position of bound NADH in the F383A−NADH complex was determined on the basis of a difference electron density map calculated once the protein portion of this complex had been refined. Further refinement of both NADH and protein atomic positions and thermal factors was then carried out to convergence. For both the unliganded and NADH-bound variant structures, solvent peaks were identified by a final series of difference electron density maps, and the validity of these was monitored on the basis of hydrogen bonding potential for bonding to protein or NADH atoms and the refinement of a thermal factor of <75 Å^2.

Table 1: Summary of Structure Determination Analyses

parameter	F383A structure	F383A−NADH structure
data collection		
space group	*R*3	*R*3
unit cell dimensions (Å)		
$a = b$	165.2	164.9
c	156.6	158.9
total no. of measurements	485765	375414
no. of unique reflections	54292	41081
mean $I/\sigma I$[a]	15.1 (8.6)	19.9 (11.6)
multiplicity[a]	3.8 (1.6)	4.1 (2.2)
R_{merge} (%)[a,b]	6.0 (15.4)	8.0 (12.9)
maximum resolution (Å)	2.3	2.5
structure refinement		
no. of reflections	53934	39737
resolution range (Å)	10−2.3	10−2.5
completeness within range (%)[a]	77 (54)	73 (66)
no. of protein atoms	6719	6719
no. of solvent atoms	726	583
no. of ligand atoms	−	88
average thermal factor (Å^2)		
protein atoms	40.0	36.6
solvent atoms	37.8	29.6
ligand atoms[c]	−	23.3/42.1
final crystallographic *R*-factor (%)	17.0	18.7
final structure stereochemistry		
rms deviation		
bonds (Å)	0.007	0.008
angles (deg)	1.6	1.7

[a] Values in parentheses are for the highest-resolution shell of each data set (2.3−2.34 Å for the F383A variant and 2.5−2.54 Å for the F383A−NADH complex). [b] $R_{merge} = \sum_{hkl} \sum_{i=0}^{n} |I_i - \bar{I}_{hkl}| / \sum_{hkl} \sum_{i=0}^{n} I_{ihkl}$. [c] For the F383A−NADH complex structure, the two values correspond to the NADH molecules bound to the A and B subunits, respectively.

Assessment of the final refined structures of the F383A variant protein and the F383A−NADH complex showed both to have excellent polypeptide chain geometry (Table 1). However, in both structures, substantial positional disorder is observed for residues 1−9 and 262−299. The coordinate error was estimated to be 0.23 Å for the F383A variant protein and 0.26 Å for the F383A−NADH complex (*25*).

RESULTS AND DISCUSSION

Initially, we attempted to form the complex of NADH with wild-type *E. coli* CS for structural studies both by cocrys-

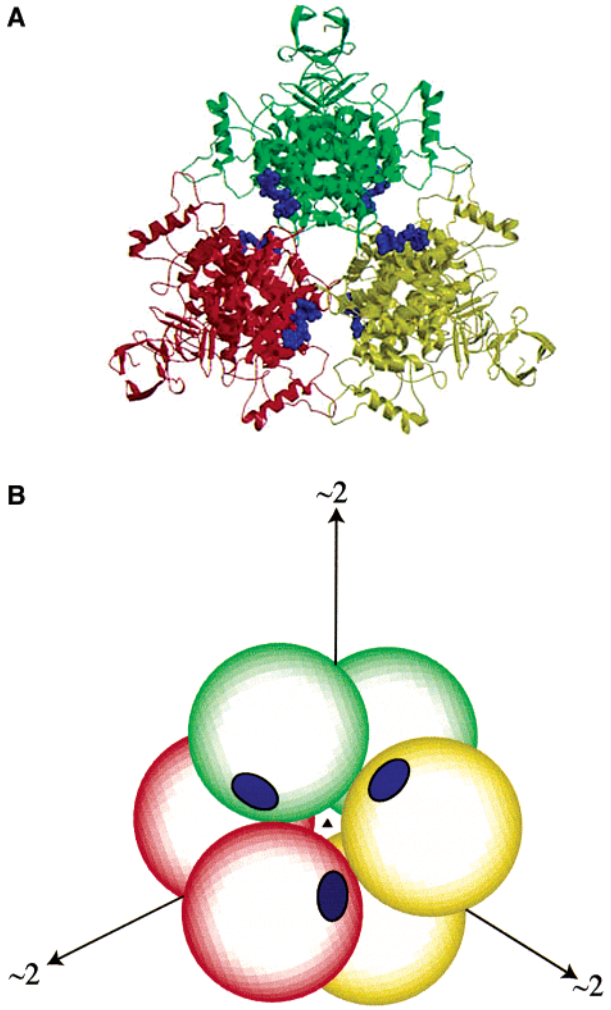


FIGURE 1: (A) Ribbon diagram showing the hexameric structure of F383A type II *E. coli* citrate synthase when NADH is bound. Shown in green, red, and yellow are the three equivalent dimers of this complex which are related by a central 3-fold axis. The locations of the six bound NADH molecules (one per subunit) are shown in a space filling format in blue. This hexameric conformation is also observed for the wild-type enzyme (*2*). This figure was prepared with the assistance of Molscript (*40*) and Raster3D (*41*). (B) Schematic representation of the hexameric complex in the same orientation, with each enzyme subunit represented as a sphere. Notably, this complex nearly exhibits 32 symmetry. The 3-fold axis goes directly down through the complex in the orientation shown and is indicated by the ▲ symbol. Since the subunits of each unique dimer differ slightly in structure, the 2-fold axis between them is only approximate and has been designated by the notation "~2". The locations of bound NADH in this schematic are shown with blue patches.

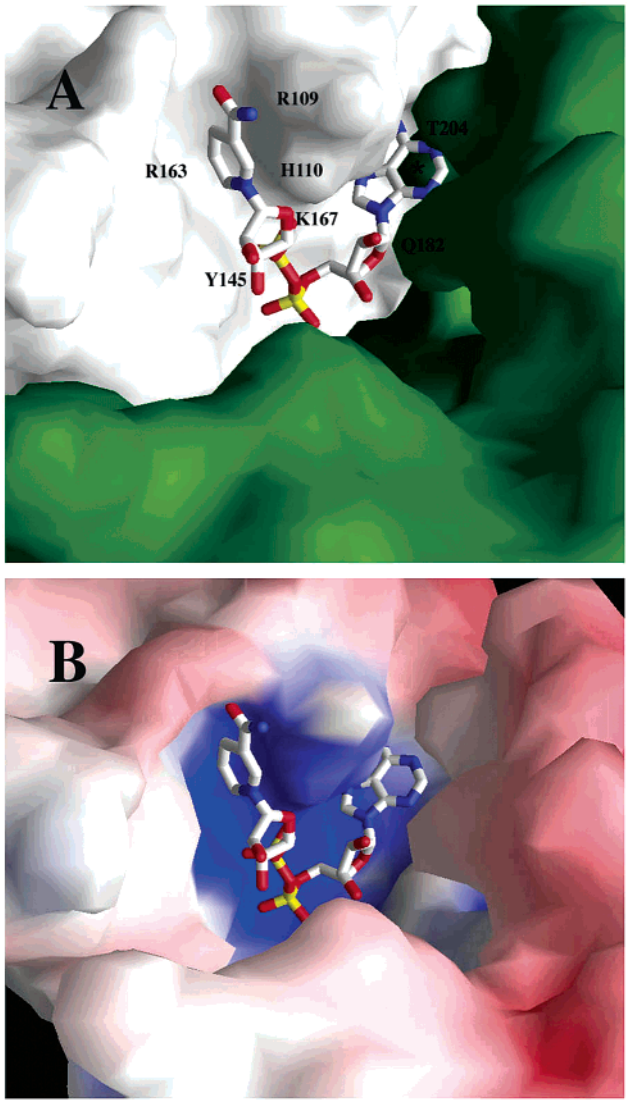


FIGURE 2: (A) Close-up view of the contour of the enzyme surface in the vicinity of one NADH binding site in hexameric *E. coli* CS. Also illustrated are the observed unique horseshoe-like conformation of bound NADH and the positions of some of those residues that form key interactions with this group (also see Table 2). The asterisk indicates the position of Cys206 which is back behind the adenine ring. This binding site is made up of structural elements from two different subunits (colored white and green) which are part of two different enzyme dimers in the overall hexameric conformation of *E. coli* CS (Figure 1). (B) Electrostatic surface diagram of this NADH binding site showing its strong cationic character. This feature, coupled with the presence of the side chain of Arg109 adjacent to the bound dihydronicotinamide group, may serve to explain the ability of this binding site to discriminate between NADH and NAD^+. This figure was prepared with the assistance of GRASP (*42*).

tallization approaches and, alternatively, by soaking the enzyme crystals in various concentrations of NADH solutions. These approaches proved to be unsuccessful. Instead, we used a variant form of CS, in which the active site phenylalanine, Phe383, is replaced with alanine. The kinetic properties of this variant suggest that the equilibrium between the R (active) and T (inactive) conformational states is strongly shifted toward the T state (*17*). Since NADH is an allosteric inhibitor, which should bind selectively to the T state, we reasoned that F383A CS would bind NADH more tightly so that good crystals of the complex would be more readily obtained. To form the groundwork for these studies, we first crystallized the F383A CS variant by itself, and determined its three-dimensional structure (Table 1). Diffraction studies of F383A CS crystals soaked in an NADH solution then allowed for the identification of well-defined electron density corresponding to bound NADH.

NADH Binding Site. The locations of NADH binding sites within the overall hexameric CS structure are shown in Figure 1. NADH occupies six approximately equivalent sites close to the CS dimer−dimer contacts (Figure 2A). There are three identical sites on one face of the hexamer, arranged with 3-fold symmetry, and three more identical, slightly different from the first three, on the other face. The overall placement and symmetry of these sites is schematically represented in Figure 1B. The binding site pockets that are

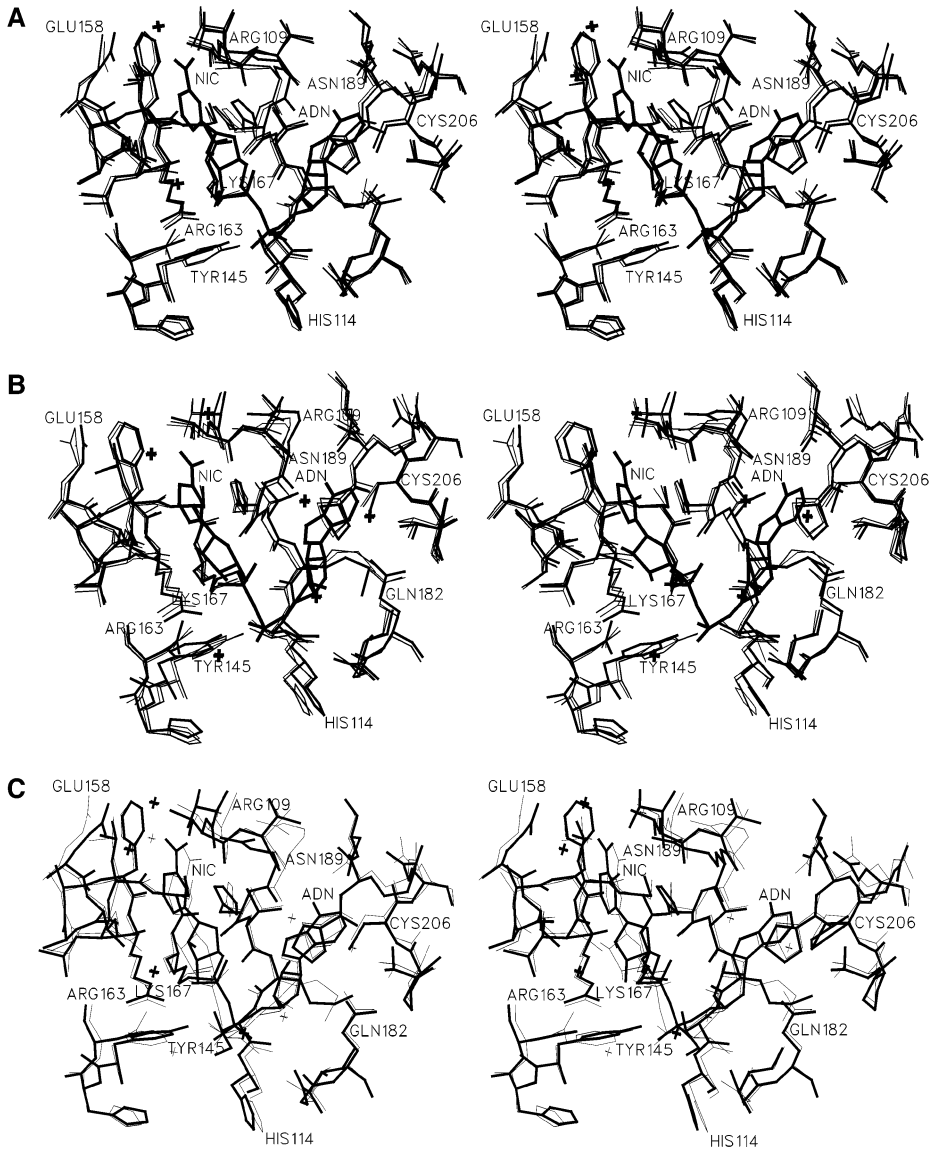


FIGURE 3: Stereodiagrams of the structures of the NADH binding sites on (A) subunit A and (B) subunit B of the type II *E. coli* CS dimer. Each drawing represents the superpositioning of the wild-type (thin lines), F383A variant (medium lines), and NADH-bound F383A variant complex (thick lines) enzymes as observed in the overall hexameric complex. In these diagrams, the NADH bound on the surface of the F383A variant is shown in thick lines and the adenine and nicotinamide ends of this moiety are labeled ADN and NIC, respectively. Also drawn with a corresponding line thickness (as crosses) are the positions of bound water molecules in the vicinity of each NADH binding site. Panel C shows a comparison overlap of the F383A variant CS dimer polypeptide chains for subunits A (thick lines) and B (thin lines) in the vicinity of their respective NADH binding sites when this ligand is bound. The positions of bound NADH and water molecules are also drawn using the same line thickness scheme.

involved are markedly cationic (Figure 2B) and bind NADH in an unusual, roughly horseshoe conformation. This is significantly different from the extended conformation which NADH exhibits when bound to the most common nucleotide-binding protein motif, the Rossman fold (see further below).

The dihydronicotinamide ring of CS-bound NADH sits in that portion of the binding site formed by the roughlyperpendicular junction of helices E (residues 97−111; see Table 3 in ref *2* for the definition of helices in CS) and I (residues 153−181). The pyrophosphate moiety binds in the most cationic part of the site, with positive charge contributed by two residues from helix I (Arg163 and Lys167) and two from helix E (Arg109 and His110). The adenine end of NADH is at the dimer−dimer interface (Figure 2A), with contacts from helix E and the I−J corner (residues 182−191) of one dimer, plus residues from the I−J corner and J−K loop (residues 204−211) of the other. Thus, the NADH sites are fully formed only in the hexamer, by the interaction of subunits from two different dimers. Most of the binding pocket, including the pyrophosphate-binding positive charges, is provided by one subunit, but residues from the adjacent dimer are needed in particular to complete formation of the adenine subsite. This arrangement markedly contrasts with the NADH binding sites in oxidoreductases, which are contained within single subunits.

Between subunits A and B of each unique enzyme dimer, the structures of the NADH−protein complexes are somewhat different. However, as Figure 3 illustrates, most of the same specific ligand−protein interactions are observed in both subunits. A complete list of hydrogen bonds stabilizing the NADH−protein complex is given in Table 2, and a diagram summarizing this information can be seen in Figure 4. As is evident in Figures 3 and 4, some parts of the bound NADH structure, the adenine ring and the pyrophosphates

Table 2: Hydrogen Bond and Charged Interactions between NADH and Neighboring Amino Acids[a,b]

NADH atom	interacting residue	A subunit	B subunit
adenine group			
AN1	Asn189 ND2	3.8	3.8
AN3	Gln182 OE1[c]	(5.3)	3.2
	Wat	–	3.5
AN6	Thr108 O	3.0	3.0
	Arg109 O	(4.2)	3.4
	Thr111 OG1	3.0	3.4
	Asn189 OD1	3.2	3.2
	Cys206 SG[c]	(4.0)	3.7
AN7	Met112 N	2.8	3.0
AN1	*Thr204 CG[c]*	*3.7*	*4.2*
	Glu207 CG[c]	*4.8*	–
	Glu207 OE[c]	*4.9*	*4.8*
AN6	*Pro187 CB*	*4.9*	–
adenine ribose			
AO2′	Wat	–	2.9
AO3′	Gln182 NE2[c]	3.8	3.8
	Wat	3.4	3.3
AO5′	Tyr145 OH	3.8	(5.3)
	Lys167 NZ	3.5	(5.4)
	His114 N	3.3	(4.1)
AC5′	*Ile113 C*	*3.7*	*4.4*
	His114 N	*4.0*	*3.9*
AO5′	*Ile113 C*	*3.4*	*4.9*
AC4′	*Gly181 O[c]*	*4.6*	*4.7*
phosphate groups			
AO1	Ile180 O[c]	3.6	(5.1)
	Wat	2.4	3.1
AO2	Tyr145 OH	3.0	2.9
	His114 N	4.0	3.8
AO2	*Ile113 C*	*4.8*	*4.7*
NO1	*His110 ND1*	*3.3*	*3.5*
	Lys167 NZ	*3.1*	*2.8*
NO2	*Tyr145 OH*	*3.6*	*(4.1)*
	Arg163 NH1	*3.8*	*3.8*
	Arg163 NE	*3.0*	*3.1*
	Lys167 NZ	*3.1*	*3.2*
nicotinamide ribose			
NO2′	Wat	2.9	–
NO3′	Arg163 NH1	3.2	(4.4)
NO5′	Arg163 NH1	3.7	3.7
NC2′	*Ile159 CG2*	*4.4*	*3.8*
nicotinamide group			
NO7	Arg109 NH2	3.0	(5.4)
	Wat	3.8	2.8
	Wat	–	3.5
NN7	Arg109 NH2	3.5	(5.1)
	Wat	2.5	3.3
	Wat	2.5	2.6
NN1	*His110 NE2*	*3.5*	*3.9*
	His110 CE1	*3.8*	*3.5*
NO7	*Thr106 CG2*	*3.4*	*3.6*
	Glu158 CD	*4.3*	–
NC5	*Ile159 CA*	*3.9*	*4.1*
	Phe162 CB	*3.8*	*3.8*
NC6	*Arg163 CG*	*4.8*	*4.4*
	Ile159 CG2	*3.8*	*3.6*

[a] Interaction distances (angstroms) are given for the NADH binding sites on both the A and B subunits of a type II *E. coli* CS dimer as found in an overall hexamer complex of this enzyme. Where an interaction occurs in one NADH binding site but is too distant in the other, the distance for the long interaction is shown in parentheses. A dash is for those cases where no interaction appears to occur. [b] Listed in italics for each NADH grouping are additional distances of <5 Å that indicate other residues that appear to be important in forming this portion of the binding pocket. [c] These interactions are contributed by residues on a neighboring CS dimer.

in particular, form a substantial number of hydrogen bond and salt bridge interactions to the protein. The dihydronicotinamide portion, on the other hand, has a limited number

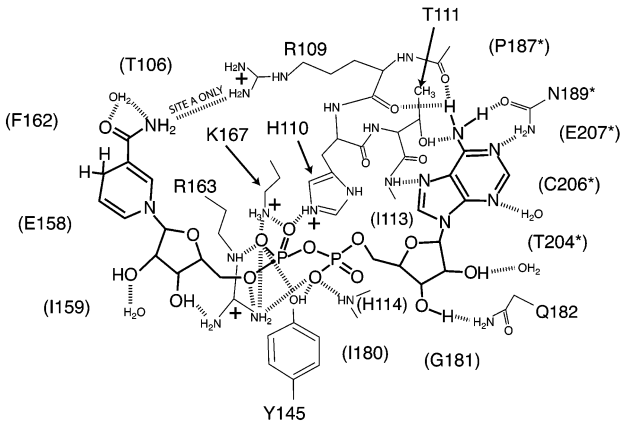


FIGURE 4: Schematic diagram summarizing interactions with NADH when this group is bound on the surface of hexameric *E. coli* CS. NADH is shown with thick lines, and parts of the protein sequence are shown with thin lines. All interactions listed in Table 2 are shown. Probable hydrogen bonds are indicated by thick broken lines. Other residues within van der Waals contact distances are shown in parentheses, in approximately correct locations relative to NADH. Residues contributed by the second dimer are marked with asterisks.

of such bonding interactions. Only one obvious hydrogen bond is formed between this part of the ligand and theprotein, that involving the carboxamido group of NADH and the guanidinium group of Arg109, and this bond is seen only in the sites involving the A subunits.

In the NADH-binding motif most commonly encountered, the Rossman fold, the adenine subsite includes aromatic residues that can contribute π-stacking interactions with the ligand (*26*). In another significant departure, no aromatic residues are found in the adenine-binding part of the NADH site of type II citrate synthase, but instead, we find a "sulfur patch", formed by the sulfur atoms of Met112 and Cys206, which are in van der Waals contact with the adenine ring (Table 2).

Bound Conformations of NADH. The bound conformation of NADH in our structure takes on an unusual horseshoe-like arrangement that is quite different from the extended geometries typically found in NADH−oxidoreductase complexes associated with the Rossman fold (*26*). An example of the normally extended geometry of NADH, as observed in horse liver alcohol dehydrogenase (*27*), is shown in Figure 5, as a comparison with the NADH conformation found with type II CS. In the CS−NADH complex, the adenine and dihydronicotinamide rings are some 11 Å apart, separated by the side chains of Arg109 and His110, which form a "hub" in the binding pocket around which the NADH is wrapped (Figures 2B and 3).

There are only a few other cases where a folded "horseshoe-like" conformation of NADH is observed. One of these is the crystallized free acid form of NAD^+ where the nicotinamide and adenine rings are only 9.6 Å apart (*28*), although notably the Li^+ salt of NAD^+ adopts an extended geometry when crystallized (*29*). In terms of protein complexes, binding of NADH or NADPH in a folded conformation has been described in only three other systems. These include glycogen phosphorylase *b* (*30*) and catalase (*31*); however, in both these cases, the functional significance of the bound nucleotide is unknown. In addition, the folded conformation of $NADP^+$ as it binds to isocitrate dehydrogenase (*32*) is a

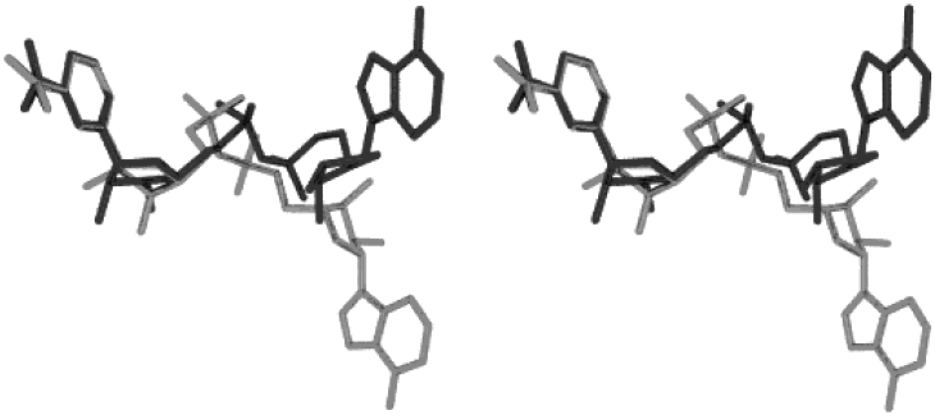
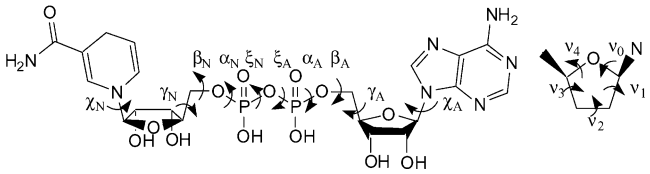


FIGURE 5: Stereoview of NADH ligands from *E. coli* type II CS (black) and horse liver alcohol dehydrogenase [gray (*27*); PDB entry 6ADH]. The ligands are aligned on their dihydronicotinamide rings, clearly showing the unique horseshoe conformation found when NADH is bound to *E. coli* CS and the more extended conformation generally found in the binding sites of other types of enzymes.

Table 3: NADH Conformations in Type II CS



NADH Ribose Pseudorotational Parameters[a]

	ν_0	ν_1	ν_2	ν_3	ν_4	P	ψ_m	conformer
adenosine								
A subunit	−11	23	−26	20	−6	175	26	${}^{2}_{3}T$
B subunit	−21	37	−37	26	−4	167	38	${}^{2}E$
NMN								
A subunit	0	−16	24	−24	15	17	26	${}^{3}E$
B subunit	−45	34	−12	−13	35	106	43	${}^{o}_{1}T$

Glycosidic, Pyrophosphate, and Amide Torsional Angles in NADH[a]

	adenosine					NMN					
	χ_A	α_A	β_A	γ_A	ζ_A	χ_N	α_N	β_N	γ_N	ζ_N	amide
A subunit	−123	60	28	−67	43	−113	121	177	−33	162	180
B subunit	−130	−92	177	−55	97	−132	9	179	109	179	180
NAD^+ [b]	−146			54		−150			48		*trans*
Li-NAD^+ [c]	−165	−125	163	48	133	−121	79	179	47	72	9

[a] Torsional angles and pseudorotational descriptors are defined according to the IUPAC-IUB recommendations (1982) for polynucleotide conformations (http://www.chem.qmw.ac.uk/iupac/misc/pnuc2.html). [b] Data from ref *28*. [c] Data from ref *29*.

unique departure from the extended conformation seen for all other complexes of NAD or NADP with oxidoreductases.

Further analysis shows the ribosyl rings of enzyme-bound NADH are typically found in either the ${}^{2}E$ or ${}^{3}E$ conformation, whereas in Li-NAD^+ by itself, the adenosine ribose is in the ${}^{2}E$ conformation and the NMN ribose is in the ${}^{3}_{2}T$ twist geometry (Table 3). In our type II CS structure, the A chain and B chain NADH ligands adopt slightly differing pseudorotational conformations. In the A chain ligand, both sugar rings are rather flat, as indicated by the relatively small degree of pucker (ψ_m). The adenosine sugar ring is found in the ${}^{2}_{3}T$ conformation, although the resolution of our structure does not permit us to rule out the more common ${}^{2}E$ alternative. The NMN ribofuranose adopts the ${}^{3}E$ geometry, which is uncommon (although not unknown) in oxidoreductase−NADH complexes. In the B chain, the sugar rings are more strongly puckered. The adenosine ribofuranose has the ${}^{2}E$ conformation, while the NMN fragment is found to have the unusual ${}^{o}_{1}T$ geometry.

Other aspects of the observed NADH torsion angles are of interest (Table 3). For example, the N-glycosidic torsion angles χ_N and χ_A of the dihydronicotinamide and adenine moieties in NADH bound to type II CS are negative and *anticlinal*, being somewhat smaller than those observed in isolated NAD^+ structures. Also, the χ_A values are more negative than the values typically seen for NADH bound to oxidoreductases where the adenosine ribose group adopts the ${}^{2}E$ conformation, and are closer to those observed for ${}^{3}E$ geometries. On the other hand, the χ_N angles for CS-bound NADH are within the typical range of values.

Our electron density maps did not permit us to choose between the two possible conformations of the carboxamido group of the dihydronicotinamide moiety, since the carbonyl oxygen and the amide nitrogen cannot be distinguished at

this resolution. Therefore, this amide group was oriented in the ample available electron density so that the most suitable hydrogen bonding interaction could be made with the adjacent side chain of Arg109 (see Table 2). The validity of this placement was supported by the results of a semiempirical (AM1) energy calculation which indicated that this (*trans*) geometry is energetically favored. Interestingly,enzyme-bound NAD^+ or NADH typically has the *trans* amide geometry (*33*), as does the crystalline form of the NAD^+ free acid (*28*).

Effects of NADH Binding on Protein Conformation. Our structure determination for the unliganded F383A variant of type II *E. coli* citrate synthase indicates that overall it is similar to that of the wild-type protein. The average deviation for main chain atoms between these structures is ~0.3 Å. The largest positional deviations in subunit A involve residues 63, 331−340, and 402, whereas in subunit B, those residues most affected are 264−268, 325−335, and 401−403. In both subunits, these residues are localized to exposed surface loops or areas having high thermal factors. No notable conformational changes are found in the vicinity of the F383A substitution site.

The largest displacements induced by bound NADH involve the positions of residues 112−120. Interestingly, there is also a large shift in the position of residue 291, which is far removed from the NADH site, but in a functionally significant location. As our earlier determination of the wild-type structure of type II *E. coli* CS has shown, this residue is within a loop that must shift radically, likely through a refolding mechanism, to form the binding site for the substrate, acetyl-CoA. The ability of bound NADH to induce a shift in the position of residue 291 would seem to provide a preliminary indication as to how NADH exerts its allosteric inhibition on acetyl-CoA binding.

Rationalizing Current Structural Results with Previous Functional Studies. These structural studies can also be used to explain a number of disparate observations made in the past concerning interaction of NADH with type II *E. coli* CS. For example, our data clearly indicate that the NADH site is distinct and remote from the active site. This explains why various treatments of *E. coli* citrate synthase, such as chemical modification (*34*) or adjustment of the pH or salt content of the assay medium (*10*), abolish NADH inhibition while preserving activity. It is also in agreement with observations from a variant form of the enzyme having residues 264−287 deleted. In this case, there is a loss of catalytic activity, while the ability to bind NADH remains normal (*35*). Finally, if *E. coli* citrate synthase is studied at concentrations at which a dimer−hexamer equilibrium is present, NADH binds selectively to the hexamer, and induces a shift in the equilibrium to the hexameric form (*19*). This result is also explained by our structure, which shows that the NADH binding sites are fully formed only in the hexamer.

Cys206 and the "Sulfur Patch". It has long been known that NADH inhibition of *E. coli* CS can be abolished when this enzyme is treated with sulfhydryl group reagents (*34*). Modification of a single residue, Cys206, is involved, and not only NADH itself but also other adenylates can prevent the reaction (*36*, *37*). Our current structural results show that Cys206 is located in the adenine part of the NADH binding site, where, along with the nearby Met112, it creates what may be described as a sulfur patch. Studies of protein and small-molecule crystal structures, as well as theoretical calculations, have yielded considerable evidence for strong sulfur−aromatic interactions (*38* and references therein). Both "edge" and "face" interactions have been identified. These studies have suggested that the optimal aryl−S distance in face interactions is 4.4−4.9 Å, at an angle of roughly 55° from the aromatic ring plane, and that each such interaction could account for 1−3 kcal/mol of stabilization. In the subunit A NADH sites of type II CS, the Cys206 sulfhydryl group is 4.58 Å from the adenine ring centroid, at a 58° angle to the mean ring plane, while the corresponding Met112 sulfur is 4.84 Å away, at a 71° angle. In the subunit B NADH sites, the corresponding values for the Cys206 and Met112 sulfurs are 4.51 Å and 53° and 4.92 Å and 80°, respectively. Thus, the position of the CS sulfur patch within the adenine subsite matches the requirements for a significant interaction. Consistent with this possibility, the C206S variant of CS does not bind NADH as well (*37*), likely because the much "harder" hydroxyl oxygen in serine would not provide a van der Waals attraction as strong as that of the cysteine sulfhydryl group.

Specificity studies have shown that the NADH site can bind a variety of adenylic acid derivatives, but only NADH and very closely related compounds are inhibitors (*14*, *36*). Thus, the dihydronicotinamide portion of NADH must be present for allosteric inhibition to occur. Furthermore, it is also known that NAD^+ binds poorly to the NADH site (*14*). Our structure suggests that it is Arg109 that makes the most important specific interaction with the dihydronicotinamide ring (Figure 3). The presence of this positively charged amino acid, along with the strongly cationic nature of the NADH binding site as a whole (Figure 2B), likely explains the ability of *E. coli* CS to discriminate between NADH and NAD^+. If, as has been proposed (*39*), NADH inhibition of type II CS provides a way for the redox state of cells to regulate the entry of carbon into the citric acid cycle, the ability of this enzyme to exclude NAD^+ from the NADH binding site would be essential.

NADH Site of Type II CS as a Product of Recent Evolution. All CS molecules, whatever organism contains them, appear to be descendents of an ancient enzyme whose overall fold, consisting of 18 α-helices, has been retained throughout evolution. The minimum functional unit is the dimer, in which both subunits contribute to each active site. Each subunit contains a small and a large domain, and motion of the small domain, to open and close the active site, is probably an essential feature of the catalytic cycle. This domain motion has been thoroughly demonstrated in the pig (type I) CS system, using structures of the enzyme with and without active site ligands (*5*). Although we do not yet have a structure for *E. coli* type II CS with active site ligands present, our finding that His264 is ~11 Å from where it should be when acetyl-CoA binds is an indication that domain motion is important in the *E. coli* enzyme as well (*2*). So there is every reason to regard the allosteric properties of type II CS as evolutionary add-ons to the minimum functional unit.

Our earlier determination of the wild-type *E. coli* type II CS structure suggested what evolutionary developments were required to produce new dimer−dimer interactions so that a functional hexamer unit could be obtained. Notably, the

Table 4: Correlation of Gram-Negative Bacterial Citrate Synthase Subunit Size and NADH Sensitivity with the Presence of NADH-Interacting Residues As Identified in the *E. coli* CS−NADH Complex[a]

Taxonomic group	Species (code)	Alignment of sequences (*E. coli* CS numbering)	No. of identities	NADH inhibition	Size
		105-----116 143-147 156-------169 178-------191 202-208			
Alpha subdivision					
Acetobacteraceae	AAC	…NILTNHTLLHEQ……AFYPD……NRDLAAMRLIAKIP……YTQGEAFIYPRNDL……FARMSPP…	12	—	—
	GEU	- sequence not available -	-	NO	]
	GXY	- sequence not available -	-	NO	]
Caulobacter group	CCR	…HNITYHTMLHAQ……AFYSD……ERETSAHRLIAKMP……YTVGQPFVSPRNDL……FAVPAPED…	15	—	—
Rhizobiaceae group	RTR	…YRVVHHTMVHEQ……AFYHD……QRMVASLRMIAKMP……YHIGQPFVYPKNDL……FAVPCEE…	14	—	—
	SME	…YRVTHHTMVHEQ……AFYHD……QRMVASLRMIAKMP……YHIGQPFVYPKNDL……FAVPCED…	15	—	—
	BJA	…DRVIHHTMVHEQ……AFYHD……QRMIASMRMIAKIP……YTIGQPFVYPKNSL……FAVSCEE…	15	—	—
	MLO	…YRVTRHTMVHEQ……AFYHD……QRMVASMRLIAKMP……YHIGQPFIYPKNDL……FAVPCEE…	16	YES	LARGE
	MTR	- sequence not available -	—	NO	
	MEX	…YRVTRHTMVHDQ……AFYHD……QRMIASLRMIAKMP……YTIGQPFVYPKNDL……FAVPCEE…	17	NO	LARGE
	MRH	- sequence not available -	—	YES	
	BHE	…RCIMQHTMVHEQ……AFYHD……QRMIASIRLISKVP……YSIGQAFVYPRNDL……FSVPCEE…	15	—	—
Rhodobacter group	RSP	…YRITRHTMLHEQ……AFYHD……QREVAAIRMIAKLP……YSIGQPFNYPQNHL……FAVPAPP…	16	YES	LARGE
	RCA	…FRITRHTMLHEQ……AFYHD……QREVAAMRMIAKLP……YSIGQPFVYPRNDL……FSVPAPP…	16	YES	—
Rhodospirillaceae	RRU	…KIITYHTMVHEQ……AFYHD……QRMIAQHRLVAKMP……YSQGQPFMYPRNNL……FTTPCEE…	16	YES	—
Rickettsiales	RPR	…KKVAHHSLVNER……AFYPD……DYELTAIRMAIKIP……YSIGQPFIYPDNSL……FATPCTK…	12	NO	—
	SAR	…RIITRHTMLHEQ……AFYHD……HRKISSHRLIAKMP……YSVGQPFVYPDNKL……FGVPAEE…	16	YES	—
Beta subdivision					
Alcaligenaceae	BPE	…SQVTHHTMVNEQ……AFYHD……HRHVSAIRLIAKMP……YSQGQPFIYPQNDL……CATPCEE…	13	—	—
Ammonia-oxidizing	NEU	…DNIKNHTMLHEQ……AFYHD……HREQSAFRLIAKLP……YNIGQPFIYPHNDL……FATPAEQ…	16	—	—
Burkholderia group	BPS	…AIVTKHTMVHEQ……AFYHD……HREVSAIRMIAKLP……YSIGQPFVYPRNDL……FANPCEE…	17	—	—
Neisseriaceae	NME	…NIVRRHTMVHEQ……AFYQD……HRKIAIYRLISKIP……YSNGLPFNYPKNNL……FATPCED…	16	—	—
	NGO	…NIVRRHTMVHEQ……AFYQD……HRKIAIYRLISKIP……YSNGLPFNYPKNNL……FATPCED…	16	—	—
	CVI	- sequence not available -	—	YES	
Ralstonia group	RME	…NSVMNHTMVHEQ……AFYHD……QREISAIRLIAKMP……YTMGQPYIYPQNDL……FATPCAP…	15	YES	LARGE
Gamma subdivision					
Aeromonadaceae	APU	- sequence not available -	—	YES	LARGE
Alteromonadaceae	SPU	…HTVKTHTMVHEQ……AFYQD……HREIAAYRLVSKMP……YSSGQPFVYPRNDL……FAVPCEE…	16	—	—
Enterobacteriaceae	ECO	…TTVTRHTMIHEQ……AFYHD……HREIAAFRLLSKMP……YSIGQPFVYPRNDL……FSTPCEP…	22	YES	
	STY	…TTVTRHTMIHEQ……AFYHD……HREIAAFRLLSKMP……YSIGQPFVYPRNDL……FSTPCET…	22	YES	LARGE
Legionellaceae group	CBU	…RTIKEHTSVYEQ……AFYHD……DRELSAIRLIAKMP……YSIGQPFMHPRRAM……FGTPYEE…	13	NO	—
	LPN	…SLINNHTMVHQQ……AFYHE……DRFISAIRLVAKMP……YSIGMPYMYPKNKM……FGVPSED…	13	—	—
Methylococcaceae	MAL	- sequence not available -	—	NO	—
Moraxellaceae	ACI	…AKVRAHTMVHDQ……AFYHN……HREITAIRLIAKIP……YTVGQPFIYPRNDL……FATPADR…	14	YES	LARGE
Pasteurellaceae	PMU	…QLVRRHTLVHEQ……AFYHD……HREITAYRLLAKMP……YSIGQPFMFPQNHL……FATPCEP…	17	—	—
Pseudomonadaceae	PAE	…GTIKNHTMVHEQ……AFYHD……HREVSAHRLIAKMP……YSKGEPMMYPRNDL……FNTPCET…	15	YES	LARGE
	PPU	…STVKNHTMVHEQ……AFYHD……HREISAVRLVAKMP……YSMGQPMMYPRNDL……FNTPCEI…	17	YES	LARGE
	AVI	…NNIKNHTMVHEQ……AFYHD……HREISAIRLVAKMP……YSLGQPLMYPRNDL……FNTPCEV…	16	YES	LARGE
Vibrionaceae	VCH	…KIVTRHTMVHEQ……AFYHD……HREIAAIRLLSKMP……YSVGQPFIYPRNDL……FATPCEE…	18	—	—
Xanthomonas group	XFA	…DEITHHTMMHES……AFYHD……QRRLAAIRLIAKVP……YSIGWPIRYPRNNL……FEVPSEP…	13	—	—
	XAX	…HEVTHHTMMHES……AFYHD……QRRQAAIRLIAKVP……YSIGWPIRYPRNNL……FEVPSEP…	13	YES	-
Epsilon subdivision					
Campylobacter group	CJE	…RYEMKKRSFIHE……AFYPD……EYMEMAARIVAKIP……YKHGFPMAYPNLDR……RTYPYDH…	5	—	—
Helicobacter group	HPY	…ELELRHRSFVHE……TLYST……DYQTMARRIVAKIP……NEVGAPIIYPDIAR……RGYPYSR…	7	NO	—

[a] Alignment of CS sequences from Gram-negative bacteria (proteobacteria), illustrating conservation of NADH-binding residues. Sequences are arranged according to taxonomic groupings. The residues that are shown are those corresponding to *E. coli* CS residues 105−116, 143−147, 156−169, 178−191, and 202−208, which include all 22 NADH-interacting residues found in our structure and listed in Table 2. Positions that are identical with these 22 *E. coli* CS residues are highlighted in black, and the total of these identities, for each sequence, is given to the right of that sequence. Where data are available, the sensitivity of the CS to NADH inhibition and its size (LARGE, probable hexamer) are also shown. Sequence data and NADH inhibition and size data are often not available for the same species within a taxonomic grouping; where a sequence is not available, this is indicated. Abbreviations for organisms: AAC, *Acetobacter aceti*; ACI, *Acinetobacter anitratus*; APU, *Aeromonas punctata*; AVI, *Azotobacter vinelandii*; BHE, *Bartonella henselae*; BJA, *Bradyrhizobium japonicum*; BPE, *Bordetella pertussis*; BPS, *Burkholderia pseudomallei*; CBU, *C. burnetii*; CCR, *Caulobacter crescentus*; CJE, *Campylobacter jejuni*; CVI, *Chromobacterium violaceum*; ECO, *E. coli*; GEU, *Gluconacetobacter europaeus*; GXY, *Gluconacetobacter xylinus*; HPY, *H. pylori*; LPN, *Legionella pneumoniae*; MAL, *Methylomicrobium album*; MCI, *Mesorhizobium ciceri*; MEX, *M. extorquens*; MLO, *Mesorhizobium loti* (NADH sensitivity and size data for *Me. ciceri*); MRH, *Methylobacterium rhodesianum*; MTR, *Methylosinus trichosporium*; NEU, *Nitrosomonas europaea*; NGO, *Neisseria gonorrheae*; NME, *Neisseria meningitidis*; PAE, *Pseudomonas aeruginosa*; PMU, *Pasteurella multocida*; PPU, *Pseudomonas putida*; RCA, *Rhodobacter capsulatus*; RME, *Ralstonia metallidurans*; RPR, *R. prowazekii*; RRU, *Rhodospirillum rubrum*; RSP, *Rhodobacter sphaeroides*; RTR, *Rhizobium tropici*; SAR, *Sphingomonas aromaticivorans*; SME, *Sinorhizobium meliloti*; SPA, *Sphingomonas paucimobilis*; SPU, *Shewanella putrefaciens*; STY, *Salmonella typhimurium*; VCH, *Vibrio cholerae*; XFA, *Xylella fastidiosa*; XAX, *Xanthomonas axonopodis* (NADH sensitivity data for *Xanthomonas hyacinthi*). Sources of CS sequences: AAC (*43*), ACI (*44*), AVI (unpublished, from unfinished microbial genomes, DOE-JGI, www.jgi.doe.gov/JGI_microbial/html/index.html), BHE (*45*; GenBank entry L38987), BJA (Ledvier and Guerinot, unpublished; GenBank entry U76375), BPE (unpublished, *B. pertussis* sequencing group, Sanger Centre), BPS (unpublished, *B. pseudomallei* sequencing group, Sanger Centre), CBU (*46*), CCR (*47*; GenBank entry AE005864, protein CC1906), CJE (*48*; GenBank entry AL139079, protein Cj1682c), ECO (*15*), HPY (*49*; GenBank entry NC_000915, gene HP0026), LPN (unpublished, from unfinished microbial genomes database at NCBI, www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html), MEX (unpublished, from ERGO site at Integrated Genomics, ergo.integratedgenomics.com/Genomes), MLO (*50*; GenBank entry NC002678, gene mlr0629), NEU (unpublished, from unfinished microbial genomes, DOE-JGI), NGO (unpublished, from unfinished microbial genomes database at NCBI), NME (*51*; GenBank entry AL162755, gene NMA1148), PAE (*52*), PMU (*53*; GenBank entry AE004439), PPU (unpublished, from incomplete genome sequence data at TIGR, www.tigr.org), RCA (unpublished, from ERGO site at Integrated Genomics), RME (unpublished, from unfinished microbial genomes, DOE-JGI), RPR (*54*), RRU (unpublished, from unfinished microbial genomes, DOE-JGI), RSP (unpublished, from unfinished microbial genomes,

Table 4 (Continued)

DOE-JGI), RTR (*55*), SAR (unpublished, from unfinished microbial genomes, DOE-JGI), SME (*56*; GenBank entry U75365), SPU (unpublished, from unfinished microbial genomes database at NCBI), STY (*57*; GenBank entry NC_003197, gene gltA), VCH (*58*; GenBank entry AE004283, gene VC2092), XAX (*59*; GenBank entry AE011985, protein XAC3388), XFA (*60*; GenBank entry AE003983, gene XF1535). Sources of data for NADH sensitivity and size: ACI (*11*), APU (*12*; ref *11*, as for *A. formicans*), AVI (*61*), CBU (*62*), CVI (*11*), ECO (*11*), GEU (ref *63*, as for *Acetobacter europaeus*), GXY (ref *64*, as for *Acetobacter xylinum*), HPY (*65*), MAL (ref *66*, as for *Methylomonas albus* BG8), MLO (ref *67*, data actually for *Me. ciceri*), MEX (*68*), MRH (*69*), MTR (ref *66*, as strain OB3b), PAE (*11*), PPU (*11*), RCA (ref *70*, as *Rhodopseudomonas capsulata*), RME (ref *71*, as *Alcaligenes eutrophus*), RPR (*72*), RRU (*73*), RSP (*74*; ref *11*, as *Rhodopseudomonas sphaeroides*), SAR (ref *11*, data actually for *Flavobacterium devorans*, now classified as *S. paucimobilis*), STY (*11*), XAX (ref *11*, data actually for *X. hyacinthi*). All the sequences used in this table are typical type II CS sequences; that is, they are ~420−440 residues in length and have a sequence corresponding to the N-terminal domain of *E. coli* CS (*2*). No type II CS sequence has yet been reported from bacteria in the delta subdivision. Some Gram-negative bacterial genomes also contain a second CS-like molecule, ~380 amino acids in length, which resembles the type I CS of Gram-positive bacteria and archaea. Where this short CS has been studied carefully, it is found to function as a methylcitrate synthase (*2*).

actual contact surfaces between dimers are small, and the key feature is the J−K loop, which is seven amino acids (in one case, eight) longer in type II CS than in type I molecules (*2*). From our current structural data (Figure 3), it can be seen that the underlying template for the NADH binding site consists of helix E and helix I, with portions of adjacent loops. These form a binding pocket for the dihydronicotinamide moiety, and provide a surface upon which the cationic features for binding of the pyrophosphate moiety have been elaborated. These secondary structural elements superimpose surprisingly closely on their equivalents in known type I CS structures, indicating that this portion of the NADH binding site was developed primarily through the adjustment of the character and placement of side chains. Thus, the structural arrangement here is a very ancient one, predating the divergence of Gram-negative bacteria from other organisms, and strongly conserved ever since.

Formation of the adenine end of the NADH binding site appears to have been more complicated. Two main chain hydrogen bonds to this moiety come from elements in helix E and the E−F corner, and these residues are part of the CS fold regardless of organism. However, the rest of the adenine site is specific to type II CS and required the development of side chain contacts with the I−J corner of the same subunit, and residues from the adjacent dimer (the I−J corner and the J−K loop). The fact that the NADH site is formed by two subunits is the most unusual feature of the site, making it significantly different from the more commonly observed NADH binding sites of oxidoreductases.

In the case where enzymes, like oxidoreductases, bind NADH to function as a coenzyme, nucleotide binding would appear to have evolved at the same time as other parts of the active site, with the Rossman fold usually selected as the most suitable approach, and bound NADH in an extended conformation. For the allosteric NADH site of type II CS, on the other hand, we now can say that the evolutionary starting point was almost certainly a compact, fully functional type I CS dimer. Such a structure would have been the product of much prior evolutionary refinement and therefore unlikely to tolerate tampering of the kind needed to create a new, highly specific nucleotide-binding site. Instead, it seems that the evolutionary process took a different path whereby type I dimers were brought together into a novel hexameric structure, and the NADH binding sites were developed using the new pockets created at the interfaces. Our structural results indicate that to accommodate this approach and satisfy conformational restrictions in the resulting complex, it was necessary to bind NADH in an alternative horseshoe-like conformation.

Relationship between Observed NADH Binding Site Features and NADH Sensitivity in Other Gram-Negative CS Molecules. Alignments of type II CS with the various kinds of type I CS sequences (from Gram-positive bacteria, archaea, and eukaryotes) show that, of the amino acid side chains involved in NADH binding, very few are present in the dimeric type I enzymes. Those few residues conserved in both type I and II CS include a glycine homologous with Gly181 and a proline homologous with Pro187 (both probably important for the main chain conformation), as well as one of the cationic side chains involved in binding pyrophosphate, Lys167 (found as either Arg or Lys in the different organisms).

In contrast, almost all the amino acids involved in the type II *E. coli* NADH binding site (Table 2) are conserved in the sequences of other Gram-negative enzymes. An alignment of the relevant regions, for a variety of type II CS molecules from different subdivisions of Gram-negative bacteria, is given in Table 4, together with available information about NADH sensitivity and the probability of a particular enzyme being hexameric. Of the 22 amino acids whose side chains interact directly with NADH in the *E. coli* enzyme (Table 2), other NADH-sensitive Gram-negative CS molecules are always found to have residues equivalent to His110, Thr111, Met112, His114, Tyr145, Arg163, Lys167, Gly181, Pro187, Asn189, and Glu207. It is interesting that Met112, one component of the sulfur patch that is proposed to assist with binding in the adenine subsite, is always found in NADH-sensitive type II CS. The other component, Cys206, is found in many NADH-sensitive enzymes (Table 4) and, when not present, is replaced with alanine or serine, residues that at least are no larger than cysteine.

Table 4 also includes sequence information for four Gram-negative CS molecules that are not sensitive to NADH, presumably because of further evolutionary pressure, in special metabolic circumstances, on the original, NADH-sensitive type II enzyme. In three of these cases, the available structural information from the *E. coli* CS−NADH complex suggests why NADH sensitivity has been lost. In the CS from *Rickettsia prowazekii* (RPR in Table 4), replacements of Met112 with Leu and His114 with Asn and the deletion of a residue equivalent to *E. coli* residue 154 probably combine to prevent allosteric NADH binding. For the CS from *Coxiella burnetii* (CBU in Table 4), Met112, His114, and Asn189 are all replaced with other residues and Cys206, located near the adenine binding pocket, is replaced with a bulky tyrosine that almost certainly would prevent adenine binding. In the CS from *Helicobacter pylori* (HPY in Table 4), Thr111, Met112, His114, and Asn189 have all been

replaced and tyrosine is again found in place of Cys206. In the case of the fourth of the NADH-insensitive Gram-negative CS enzymes, that from *Methylobacterium extorquens* (MEX in Table 4), the sequence comparison gives no clue as to why NADH sensitivity should have been lost; the sequence is very similar to those from other members of the alpha subdivision of Proteobacteria, which are sensitive to NADH. With this exception, our structural results and these sequence alignment observations suggest the basic requirements needed to express NADH sensitivity in the type II CS, and as such could be valuable in predicting the expected functionalities of other Gram-negative CS as the sequences of these become available.

ACKNOWLEDGMENT

Access to unpublished partial sequences of bacterial genomes, made freely available by the Sanger Centre, the Department of Energy Joint Genome Institute, the National Center for Biotechnology Information, The Institute for Genomic Research, and the ERGO site at Integrated Genomics, as specified in Table 4, is gratefully acknowledged.

REFERENCES

1. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res. 28*, 235−242.
2. Nguyen, N. T., Maurus, R., Stokell, D. J., Ayed, A., Duckworth, H. W., and Brayer, G. D. (2001) *Biochemistry 40*, 13177−13187.
3. Remington, S., Wiegand, G., and Huber, R. (1982) *J. Mol. Biol. 158*, 111−152.
4. Wiegand, G., and Remington, S. J. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 97−117.
5. Remington, S. J. (1992) *Curr. Top. Cell. Regul. 33*, 209−229.
6. Russell, R. J., Hough, D. W., Danson, M. J., and Taylor, G. L. (1994) *Structure 2*, 1157−1167.
7. Russell, R. J., Ferguson, J. M., Hough, D. W., Danson, M. J., and Taylor, G. L. (1997) *Biochemistry 36*, 9983−9994.
8. Russell, R. J., Gerike, U., Danson, M. J., Hough, D. W., and Taylor, G. L. (1998) *Structure 6*, 351−361.
9. Weitzman, P. D. J. (1966) *Biochim. Biophys. Acta 128*, 213−215.
10. Weitzman, P. D. J. (1966) *Biochem. J. 101*, 44c−45c.
11. Weitzman, P. D. J. (1981) *Adv. Microbiol. Physiol. 22*, 185−244.
12. Weitzman, P. D. J., and Jones, D. (1968) *Nature 219*, 270−272.
13. Weitzman, P. D. J., and Dunmore, P. (1969) *Biochim. Biophys. Acta 171*, 198−200.
14. Duckworth, H. W., and Tong, E. K. (1976) *Biochemistry 15*, 108−114.
15. Ner, S. S., Bhayana, V., Bell, A. W., Giles, I. G., Duckworth, H. W., and Bloxham, D. P. (1983) *Biochemistry 22*, 5243−5249.
16. Monod, J., Wyman, J., and Changeux, J.-P. (1965) *J. Mol. Biol. 12*, 88−118.
17. Pereira, D. S., Donald, L. J., Hosfield, D. J., and Duckworth, H. W. (1994) *J. Biol. Chem. 269*, 412−417.
18. Ayed, A., and Duckworth, H. W. (1999) *Protein Sci. 8*, 1116−1126.
19. Ayed, A., Krutchinsky, A., Ens, W., Standing, K. G., and Duckworth, H. W. (1998) *Rapid Commun. Mass Spectrom. 12*, 339−344.
20. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
21. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
22. Yeates, T. O. (1997) *Methods Enzymol. 267*, 344−358.
23. 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) *Acta Crystallogr. D54*, 905−921.
24. Jones, T. A., Zhou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
25. Luzzati, P. V. (1952) *Acta Crystallogr. 5*, 802−810.
26. Eklund, H. (1989) *Biochem. Soc. Trans. 17*, 293−296.
27. Eklund, H., Samama, J.-P., Wallen, L., Brändén, C.-I., Aakeson, A., and Jones, T. A. (1981) *J. Mol. Biol. 146*, 561−587.
28. Parthasarathy, R., and Fridey, S. M. (1984) *Science 226*, 969−971.
29. Reddy, B. S., Saenger, W., Mühlegger, K., and Weimann, G. (1981) *J. Am. Chem. Soc. 103*, 907−914.
30. Stura, E. A., Zanotti, G., Babu, Y. S., Sansom, M. S., Stuart, D. I., Wilson, K. S., Johnson, L. N., and Van de Werve, G. (1983) *J. Mol. Biol. 170*, 529−565.
31. Ko, T. P., Day, J., Malkin, A. J., and McPherson, A. (1999) *Acta Crystallogr. D55*, 1383−1394.
32. Mesecar, A. D., Stoddard, B. L., and Koshland, D. E., Jr. (1997) *Science 277*, 202−2066.
33. Eklund, H., and Brändén, C.-I. (1987) in *Pyridine nucleotide coenzymes: Chemical, biochemical and medical aspects.* (Dolphin, D., Avramovic, O., and Poulson, R., Eds.) pp 51−98, John Wiley and Sons, New York.
34. Danson, M. J., and Weitzman, P. D. J. (1973) *Biochem. J. 135*, 513−524.
35. Anderson, D. H., and Duckworth, H. W. (1988) *J. Biol. Chem. 263*, 2163−2169.
36. Talgoy, M. M., and Duckworth, H. W. (1979) *Can. J. Biochem. 57*, 385−395.
37. Donald, L. J., Crane, B. R., Anderson, D. H., and Duckworth, H. W. (1991) *J. Biol. Chem. 266*, 20709−20713.
38. Zauhar, R. J., Colbert, C. L., Morgan, R. S., and Welsh, W. J. (2000) *Biopolymers 53*, 233−248.
39. Weitzman, P. D. J., and Danson, M. J. (1976) *Curr. Top. Cell. Regul. 10*, 161−204.
40. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
41. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
42. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.
43. Fukaya, M., Takemura, H., Okumura, H., Kawamura, Y., Horinouchi, S., and Beppu, T. (1990) *J. Bacteriol. 172*, 2096−2104.
44. Donald, L. J., and Duckworth, H. W. (1987) *Biochem. Cell Biol. 65*, 930−938.
45. Norman, A. F., Regnery, R., Jameson, P., Greene, C., and Krause, D. C. (1995) *J. Clin. Microbiol. 33*, 1797−1803.
46. Heinzen, R. A., Frazier, M. E., and Mallavia, L. P. (1991) *Gene 109*, 63−69.
47. Nierman, W. C., Feldblyum, T. V., Laub, M. T., Paulsen, I. T., Nelson, K. E., Eisen, J., Heidelberg, J. F., Alley, M. R. K., Ohta, N., Maddock, J. R., Potocka, I., Nelson, W. C., Newton, A., Stephens, C., Phadke, N. D., Ely, B., DeBoy, R. T., Dodson, R. J., Durkin, A. S., Gwinn, M. L., Haft, D. H., Kolonay, J. F., Smit, J., Craven, M., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S. L., Venter, J. C., Shapiro, L., and Fraser, C. M. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 4136−4141.
48. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, S., Jagels, K., Karlyshev, A., Moule, S., Pallen, M. J., Penn, C. W., Quail, M., Rajandream, M. A., Rutherford, K. M., VanVliet, A., Whitehead, S., and Barrell, B. G. (2000) *Nature 403*, 665−668.
49. Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L. M., Lee, N., Adams, M. D., Hickey, E. K., Berg, D. E., Gocayne, J. D., Utterback, T. R., Peterson, J. D., Kelley, J. M., Cotton, M. D., Weidman, J. M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W. S., Borodovsky, M., Karp, P. D., Smith, H. O., Fraser, C. M., and Venter, J. C. (1997) *Nature 388*, 539−547.
50. Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., and Tabata, S. (2000) *DNA Res. 7*, 331−338.
51. Parkhill, J., Achtman, M., James, K. D., Bentley, S. D., Churcher, C., Klee, S. R., Morelli, G., Basham, D., Brown, D., Chillingworth, T., Davies, R. M., Davis, P., Devlin, K., Feltwell, T., Hamlin, N., Holroyd, S., Jagels, K., Leather, S., Moule, S., Mungall, K., Quail, M. A., Rajandream, M. A., Rutherford, K. M., Simmonds,

M., Skelton, J., Whitehead, S., Spratt, B. G., and Barrell, B. G. (2000) *Nature 404*, 502−506.
52. Donald, L. J., Molgat, G. F., and Duckworth, H. W. (1989) *J. Bacteriol. 171*, 5542−5550.
53. May, B. J., Zhang, Q., Li, L., Paustian, M. L., Whittam, T. S., and Kapur, V. S. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 3460−3465.
54. Wood, D. O., Atkinson, W. H., Sikorski, R. S., and Winkler, H. H. (1987) *J. Bacteriol. 155*, 412−416.
55. Pardo, M. A., Lagunez, J., Miranda, J., and Martinez, E. (1994) *Mol. Microbiol. 11*, 315−321.
56. Mortimer, M. W., McDermott, T. R., York, G. M., Walker, G. C., and Kahn, M. L. (1999) *J. Bacteriol. 181*, 7608−7613.
57. McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., and Wilson, R. K. (2001) *Nature 413*, 852−856.
58. Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Umayam, L., Gill, S. R., Nelson, K. E., Read, T. D., Tettelin, H., Richardson, D., Ermolaeva, M. D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R. D., Nierman, W. C., and White, O. (2000) *Nature 406*, 477−483.
59. da Silva, A. C., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., Quaggio, R. B., Monteiro-Vitorello, C. B., Van Sluys, M. A., Almeida, N. F., Alves, L. M., do Amaral, A. M., Bertolini, M. C., Camargo, L. E., Camarotte, G., Cannavan, F., Cardozo, J., Chambergo, F., Ciapina, L. P., Cicarelli, R. M., Coutinho, L. L., Cursino-Santos, J. R., El-Dorry, H., Faria, J. B., Ferreira, A. J., Ferreira, R. C., Ferro, M. I., Formighieri, E. F., Franco, M. C., Greggio, C. C., Gruber, A., Katsuyama, A. M., Kishi, L. T., Leite, R. P., Lemos, E. G., Lemos, M. V., Locali, E. C., Machado, M. A., Madeira, A. M., Martinez-Rossi, N. M., Martins, E. C., Meidanis, J., Menck, C. F., Miyaki, C. Y., Moon, D. H., Moreira, L. M., Novo, M. T., Okura, V. K., Oliveira, M. C., Oliveira, V. R., Pereira, H. A., Rossi, A., Sena, J. A., Silva, C., de Souza, R. F., Spinola, L. A., Takita, M. A., Tamura, R. E., Teixeira, E. C., Tezza, R. I., Trindade dos Santos, M., Truffi, D., Tsai, S. M., White, F. F., Setubal, J. C., and Kitajima, J. P. (2002) *Nature 417*, 459−463.
60. Simpson, A. J., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bove, J. M., Briones, M. R., Bueno, M. R., Camargo, A. A., Camargo, L. E., Carraro, D. M., Carrer, H., Colauto, N. B., Colombo, C., Costa, F. F., Costa, M. C., Costa-Neto, C. M., Coutinho, L. L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A. P., Ferreira, A. J., Ferreira, V. C., Ferro, J. A., Fraga, J. S., Franca, S. C., Franco, M. C., Frohme, M., Furlan, L. R., Garnier, M., Goldman, G. H., Goldman, M. H., Gomes, S. L., Gruber, A., Ho, P. L., Hoheisel, J. D., Junqueira, M. L., Kemper, E. L., Kitajima, J. P., and Marino, C. (2000) *Nature 406*, 151−157.
61. Rault-Leonardon, M., Atkinson, M. A., Slaughter, C. A., Moomaw, C. R., and Srere, P. A. (1995) *Biochemistry 34*, 257−263.
62. Heinzen, R. A., and Mallavia, L. P. (1987) *Infect. Immun. 55*, 848−855.
63. Sievers, M., Stockli, M., and Teuber, M. (1997) *FEMS Microbiol. Lett. 146*, 53−58.
64. Swissa, M., and Benziman, M. (1976) *Biochem. J. 153*, 173−179.
65. Pitson, S. M., Mendz, G. L., Srinivasan, S., and Hazell, S. L. (1999) *Eur. J. Biochem. 260*, 258−267.
66. Colby, J., and Zatman, L. J. (1975) *Biochem. J. 150*, 141−144.
67. Tabrett, C. A., and Copeland, L. (2000) *Arch. Microbiol. 173*, 42−48.
68. Belova, L. L., Sokolov, A. P., Morgunov, I. G., and Trotsenko, Yu. A. (1997) *Biochemistry (Moscow) 62*, 71−76.
69. Mothes, G., Rivera, I. S., and Babel, W. (1996) *Arch. Microbiol. 166*, 405−410.
70. Eidels, L., and Preiss, J. (1970) *J. Biol. Chem. 245*, 2937−2945.
71. Henderson, R. A., and Jones, C. W. (1997) *Arch. Microbiol. 168*, 486−492.
72. Phibbs, P. V., and Winkler, H. H. (1982) *J. Bacteriol. 149*, 718−725.
73. Massarini, E., Higa, A. I., and Cazzulo, J. J. (1976) *Experientia 32*, 426−428.
74. Borriss, R., and Ohmann, E. (1972) *Biochem. Physiol. Pflanz. 163*, 328−333.

BI020622S