



Crystal Structures of an NAD Kinase from *Archaeoglobus fulgidus* in Complex with ATP, NAD, or NADP

Jinyu Liu¹, Yun Lou¹, Hisao Yokota¹, Paul D. Adams¹, Rosalind Kim¹ and Sung-Hou Kim^{1,2*}

¹Berkeley Structural Genomics Center, Lawrence Berkeley National Laboratory, Physical Bioscience Division, 350 Calvin Lab, Mail Stop 5230, Berkeley CA, 94720-5230, USA

²Department of Chemistry University of California Berkeley, CA 94720, USA

NAD kinase is a ubiquitous enzyme that catalyzes the phosphorylation of NAD to NADP using ATP or inorganic polyphosphate (poly(P)) as phosphate donor, and is regarded as the only enzyme responsible for the synthesis of NADP. We present here the crystal structures of an NAD kinase from the archaeal organism *Archaeoglobus fulgidus* in complex with its phosphate donor ATP at 1.7 Å resolution, with its substrate NAD at 3.05 Å resolution, and with the product NADP in two different crystal forms at 2.45 Å and 2.0 Å resolution, respectively. In the ATP bound structure, the AMP portion of the ATP molecule is found to use the same binding site as the nicotinamide ribose portion of NAD/NADP in the NAD/NADP bound structures. A magnesium ion is found to be coordinated to the phosphate tail of ATP as well as to a pyrophosphate group. The conserved GGDG loop forms hydrogen bonds with the pyrophosphate group in the ATP-bound structure and the 2' phosphate group of the NADP in the NADP-bound structures. A possible phosphate transfer mechanism is proposed on the basis of the structures presented.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: crystal structure; NAD kinase; ATP-binding; NADP-binding; phosphate transfer

*Corresponding author

Introduction

Nicotinamide adenine dinucleotide (NAD) and its phosphorylation product nicotinamide adenine dinucleotide phosphate (NADP) are two of the most important coenzymes in the cell. The importance of NAD and NADP as electron carriers in energy metabolism has long been known. Recent studies have shown that NAD and NADP are key molecules involved in a plethora of different biochemical processes, such as signal transduction, transcription, DNA repair, and detoxification reactions.^{1–3} The only enzyme that catalyses the phosphorylation of NAD to NADP in the presence of a phosphoryl donor, generally ATP or inorganic polyphosphate, is the ubiquitous enzyme NAD kinase (EC 2.7.1.23) (Figure 1). Since NAD kinase is the only enzyme leading to the biochemical formation of NADP, it is the key enzyme that

regulates the NADP level and the NADP-dependent anabolic and biosynthetic pathways in the cell. This enzyme has been proposed to be also a potential novel antimicrobial drug target.⁴

To date, this enzyme has been identified, purified and characterized in bacteria, plants, vertebrates and human since its discovery several decades ago.^{5–11} All NAD kinases form a multimeric state in solution and divalent metals (usually magnesium) are essential for carrying out their activities. Within the characterized NAD kinases, enzymes from some bacterial organisms such as *Mycobacterium tuberculosis* H37Rv (Ppkn), *Mycobacterium flavus* and *Bacillus subtilis* have been shown to utilize both inorganic polyphosphate (poly(P)) and nucleoside triphosphates as phosphate donors for catalysis, and were designated as poly(P)ATP-NAD kinase,^{5,7} while NAD kinases from other bacterial organisms such as *Escherichia coli* and higher organisms such as *Saccharomyces cerevisiae* and human utilize only nucleoside triphosphates as phosphate donors for catalysis and are designated as ATP-NAD kinase.^{6,8,11} It is interesting to note that the poly(P) is also proposed as an “ancient” energy carrier preceding ATP.¹²

All NAD kinases contain a conserved catalytic region and the NAD kinases from eukaryotic

Abbreviations used: MAD, multi-wavelength anomalous dispersion; CCD, charge-coupled device; Ppkn, *M. tuberculosis* NAD kinase; Afnk, *A. fulgidis* NAD kinase.

E-mail address of the corresponding author: SHKim@lbl.gov

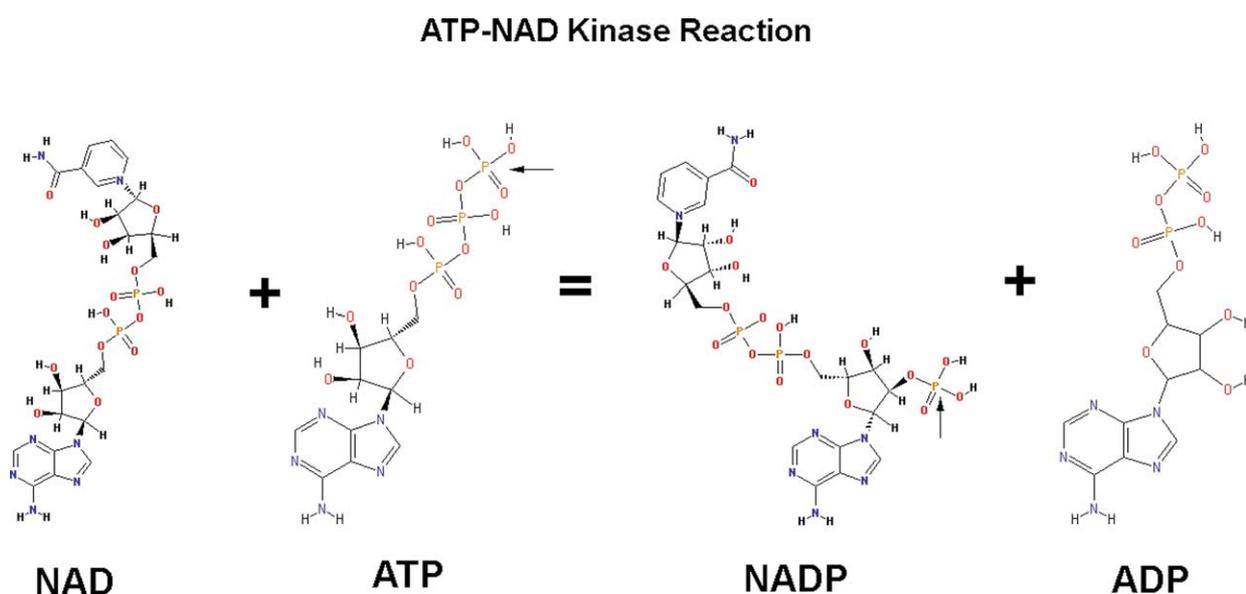


Figure 1. NAD kinase reaction. The γ -phosphate of ATP and the 2' phosphate group of NADP are indicated by arrows.

organisms have an extra N-terminal sequence. Two highly conserved and functionally important motifs have been identified within the NAD kinase family: a GGDG motif and a Gly-rich motif (Figure 2).^{13,14} The role of the Gly-rich motif in NAD binding has been confirmed by mutagenesis and structural studies.^{13,15,16} The GGDG motif has been proposed to be a highly conserved motif in a super-family that includes diacylglyceride kinase, sphingosine kinase, NAD kinase and 6-phosphofructokinase.¹⁴ The 6-phosphofructokinase protein family has been studied extensively by crystallography, where the conserved GGDG motif has been observed to be involved in ADP-binding.¹⁷ This motif (SGDG in the sphingosine kinase family) is proposed to participate in nucleotide-binding in human sphingosine kinase.¹⁸ For NAD kinases, the mutation of D to A in the GGDG motif in Ppnk led to complete loss of NAD kinase activity.¹⁶ However, the mechanistic role of the conserved GGDG motif in NAD kinases has not been established.

Although NAD kinase was discovered several decades ago, its structural basis was not determined until very recently. In 2004, the crystal structures of Ppnk and its complex with NAD were determined.^{15,16} These structures revealed a new structural fold involved in oligomerization of Ppnk and the NAD-binding mode. However, several key questions related to NAD kinases are still not answered. One important question is how NAD kinase binds its phosphate donor ATP or poly(P). There is very little information as to where the phosphate donor binds NAD kinase except that the GGDG motif might be involved in nucleotide substrate binding.^{17,18} Here, we report four structures of an NAD kinase from the archaeal organism *Archaeoglobus fulgidus* (Afnk): (1) in complex with its substrate NAD (Afnk-NAD); (2) in complex with phosphate donor ATP (Afnk-ATP); and (3) in

complex with the product NADP in two crystal forms (Afnk-NADP and Afnk-NADP2). The structures of this NAD kinase in complex with its three different ligands revealed detailed structural information about how NAD kinases bind the phosphate donor ATP, and how the conserved GGDG motif interacts with ATP or NADP. On the basis of the structures presented, a phosphate transfer mechanism for NAD kinases is proposed.

Results

Structure determination

The structure of Afnk was first solved in an NADP-bound form. The crystals of Afnk-NADP were obtained using PEG 6 K as precipitant and they belong to space group $P4_12_12$, with unit cell dimensions of $a=122.1$ Å, $c=198.6$ Å with four molecules in the asymmetric unit. The crystal structure was determined by the multi-wavelength anomalous dispersion (MAD) method using seleno-L-methionine (SeMet) substituted protein. The structure was refined at 2.45 Å resolution with $R_{\text{free}}=26.4\%$ and an R -factor of 22.1%. The crystals contained NADP, although no NADP was added during the purification and crystallization steps. The final model is of good quality: It includes all residues from the four monomers in the asymmetric unit, four well-defined NADP molecules and 389 water molecules (Table 1). PROCHECK shows 90% of the residues are in the most favored regions of the ϕ - ψ plot and there is no outlier.¹⁹

A different crystal form of the Afnk-NADP complex (Afnk-NADP2) was found when using ammonium sulfate as the precipitant. These crystals belong to space group $P2_12_12$, with unit cell dimensions of $a=62.9$ Å, $b=76.4$ Å, $c=120.0$ Å

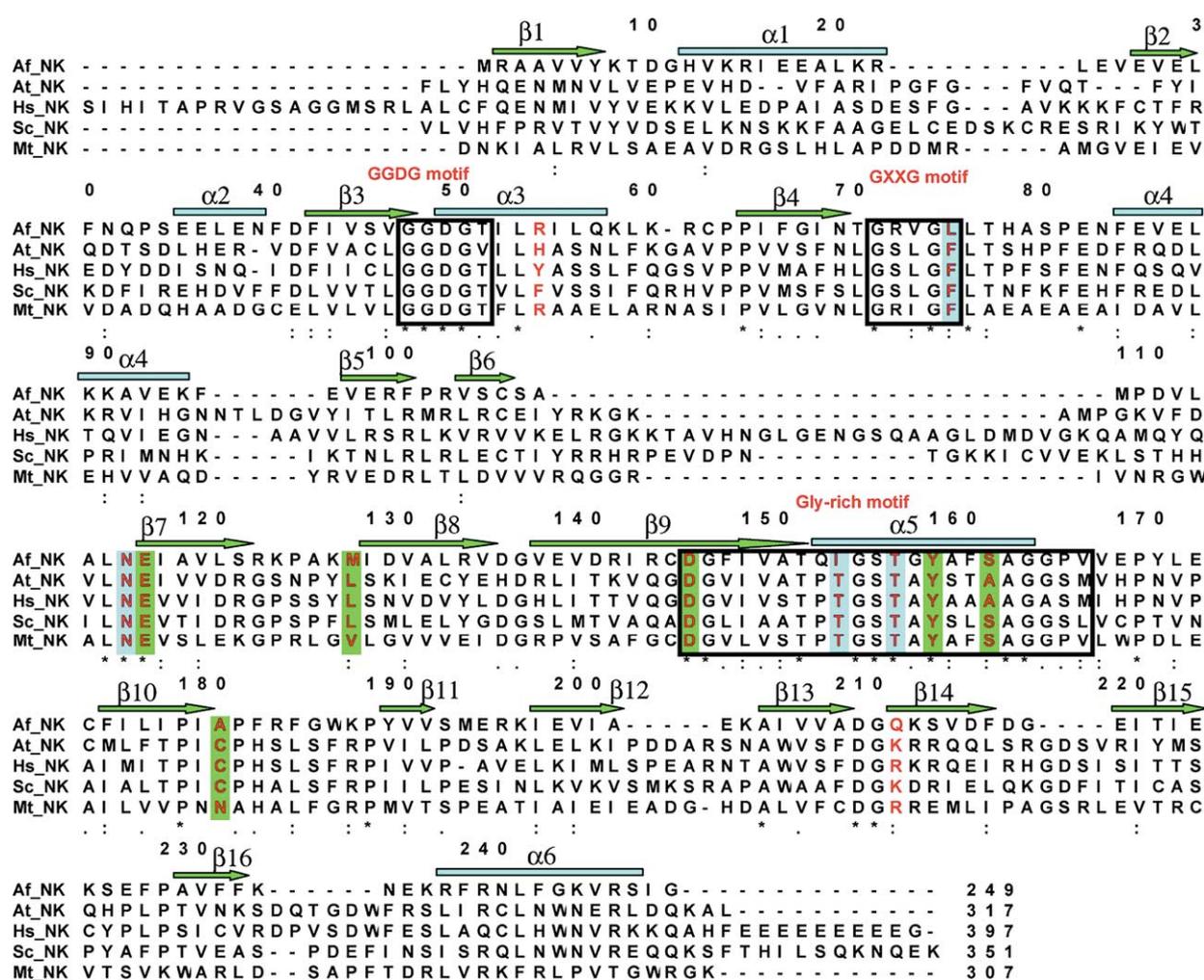


Figure 2. Multiple sequence alignment of NAD kinases from different organisms. Sequences of NAD kinase from *A. fulgidus* (Af_NK), *Arabidopsis thaliana* (At_NK), *Homo sapiens* (Hs_NK), *Saccharomyces cerevisiae* (Sc_NK) and *M. tuberculosis* (Mt_NK) are aligned together using Clustal W.²⁹ The secondary structure of Afnk is labeled on top of the sequence (arrows, β strands; squares, α -helices). Residues that are involved in NAD(P) or ATP binding are shown in red. Residues that interact with the AMP portion of the NADP molecule are labeled in cyan, and residues that interact with the AMP portion of the ATP molecule and nicotinamide ribose portion of the NADP molecule are labeled in green. The GGDG motif, Gly-rich-motif, and GXXG conserved motifs are labeled in boxes.

with two molecules in the asymmetric unit. The structure of Afnk-NADP2 was determined by molecular replacement using a dimer of the Afnk-NADP tetramer model (AA' dimer, see below) as a search model. The final model of Afnk-NADP2 was refined at 2.0 Å resolution with $R_{\text{free}} = 26.4\%$ and an R -factor of 21.2%. In the final model, three sulfate ions were modeled in accordance with the difference electron-density maps. The final model of Afnk-NADP2 includes all residues of the two Afnk molecules in the asymmetric unit, two NADP molecules, three sulfate ions and 299 water molecules. PROCHECK shows 92% of the residues are in the most favored regions of the ϕ - ψ plot and there is no outlier.¹⁹

In order to obtain the Afnk apo protein, the purified Afnk-NADP protein was denatured and refolded (see Materials and Methods). The crystals of Afnk-NAD were obtained by co-crystallizing

the refolded Afnk protein with 10 mM NAD. These crystals are in the same crystal form and unit cell as the Afnk-NADP crystals. The structure of the NAD complex was refined at 3.05 Å using the apo portion of the Afnk-NADP tetramer as a starting model. The final model of Afnk-NAD structure contains a tetramer of Afnk in the asymmetric unit with one NAD molecule bound to each monomer. PROCHECK shows 84% of the residues are in the most favored regions of the ϕ - ψ plot and there is no outlier.¹⁹

The crystals of Afnk-ATP were obtained by co-crystallizing the refolded Afnk protein with 25 mM ATP and 50 mM magnesium acetate. The Afnk-ATP crystals belong to space group $P1$ with unit cell dimensions of $a = 55.3$ Å, $b = 69.7$ Å, $c = 76.3$ Å, $\alpha = 106.4^\circ$, $\beta = 111.3^\circ$, $\gamma = 101.3^\circ$ with four molecules in the asymmetric unit. The structure of Afnk-ATP was determined by molecular replacement using

Table 1. Crystallographic data processing and refinement statistics

	Afnk-NADP				Afnk-NADP2	Afnk-NAD	Afnk-ATP
	Se-Met peak	Se-Met inflection	Se-Met remote (H)	Native	Native	Native	Native
A. Data collection							
Cell dimensions							
<i>a</i> , <i>b</i> , <i>c</i> (Å)		122.1, 122.1, 198.6			62.96, 76.38, 120.00	122.1, 122.1, 198.6	55.25, 69.73, 76.28
α , β , γ (deg.)							106.35, 111.32, 101.30
Space group		<i>P</i> ₄ ₁ ₂ ₁ ₂			<i>P</i> ₂ ₁ ₂ ₁ ₂	<i>P</i> ₄ ₁ ₂ ₁ ₂	<i>P</i> ₁
No. molecules in a.u. ^a		4			2	4	4
Wavelength (nm)	0.9796	0.9798	0.9600	1.0000	1.0000	1.0000	1.0000
Resolution	3.0	3.0	3.0	2.45	2.0	3.05	1.7
No unique reflections	56,829	56,045	55,881	62,727	45,278	29,339	100,976
Completeness (%) (final)	99.7 (98.5)	98.8 (96.8)	98.5 (96.8)	99.3 (99.9)	97.8 (91.9)	98.5 (96.2)	97.0 (95.3)
<i>I</i> / σ (final)	23.8 (2.4)	21.3 (2.1)	20.2 (2.0)	13.2 (1.0)	17.4 (1.36)	13.9 (1.2)	13.4 (2.0)
<i>R</i> _{sym} (%) (final)	0.102 (0.84)	0.100 (0.83)	0.103 (0.84)	0.093 (0.99)	0.065 (0.76)	0.116 (1.0)	0.086 (0.7)
		Afnk-NADP	Afnk-NADP2	Afnk-NAD	Afnk-ATP		
B. Refinement statistics							
Structural determination method		MAD	MR	MR	MR		
No. molecules in asymmetric unit		4	2	4	4		
Resolution range (Å)		20.0–2.45	20.0–2.0	20.0–3.05	20.0–1.7		
No. reflections (working/test)		59,749/2977	43,006/2271	23,248/1208	95,862/5111		
<i>R</i> _{free} / <i>R</i> _{factor}		0.264/0.221	0.264/0.212	0.272/0.206	0.231/0.200		
No. protein atoms (nonhydrogen)		7856	3628	7856	7856		
Ligand type and number		NADP, 4	NADP, 2; SO ₄ , 3	NAD, 4	ATP, 4; Mg ²⁺ , 4 Pyrophosphate, 4		
No. water molecules		389	299	133	404		
Rms deviations from ideality							
Bond lengths (Å)		0.016	0.013	0.011	0.009		
Bond angles (deg.)		1.638	1.523	1.354	1.213		
Ramachandran plot							
Most favored regions (%)		90	92	84	92.4		
Additionally allowed regions (%)		10	8	12	7.6		
Generously allowed regions (%)		0	0	0	0		
Disallowed regions (%)		0	0	0	0		

^a a.u., Asymmetric unit.

the apo portion of the Afnk-NADP tetramer as a search model. The final structure was refined at 1.7 Å with *R*_{free} = 23.1% and an *R*-factor of 20.0%. The final model contains all residues of the Afnk tetramer in the asymmetric unit with each monomer bound to an ATP molecule, a magnesium ion and a pyrophosphate group. PROCHECK shows 92.4% of the residues are in the most favored regions of the ϕ - ψ plot and there is no outlier.¹⁹

Overall structure

The monomer structure of Afnk is composed of six α -helices and 16 β -strands organized into two domains: the N-terminal domain (NTD) and the C-terminal domain (CTD) (Figure 3(a)). The NTD (Met1-Ala100 and Ala229-Gly249) contains an α / β fold where a central parallel four-stranded β -sheet (strand order: β 2, β 1, β 3, β 4) is flanked by four α -helices, two on one side (α 1 and α 4), and two on the other side (α 2, α 3). Another C-terminal α -helix (α 6) is parallel with the central β -sheet. The β 5 and β 16 form a hinge region between the N and C-terminal domains, with β 5 coming from the NTD and β 16 coming from the CTD. The CTD (Met101-Pro228) is

composed of nine β strands (from β 6- β 15) and one α -helix (α 5) located between β 9 and β 10. The nine β -strands form a two layer β -sandwich. NAD, NADP and ATP are all bound in the deep cleft between NTD and CTD (Figure 3(a)–(c)).

The overall fold of Afnk is similar to the Ppnk structure reported previously.^{15,16} The superposition of 215 C α atoms between Ppnk and Afnk-NADP led to an rmsd of 1.6 Å. The major difference between the two structures is that Ppnk has a 30 amino acid residue insertion between β 2 and α 2, which includes an extra β -strand and a disordered region. Ppnk also has an extra loop at the C-terminal end (data not shown).

Superposition of all 249 C α atoms among the four structures of Afnk complexes led to an rmsd of 0.24–0.32 Å, suggesting different ligand-binding has very little effect on the overall protein conformation (Figure 3(d)). The only large conformational change observed was the loop region between β 4 and α 4, which encompasses residues Thr69 to Leu74. This part of the loop includes two highly conserved residues, Gly71 and Gly74, which we define as the GXXG conserved box (Figure 2). The functional role of this loop is not known.

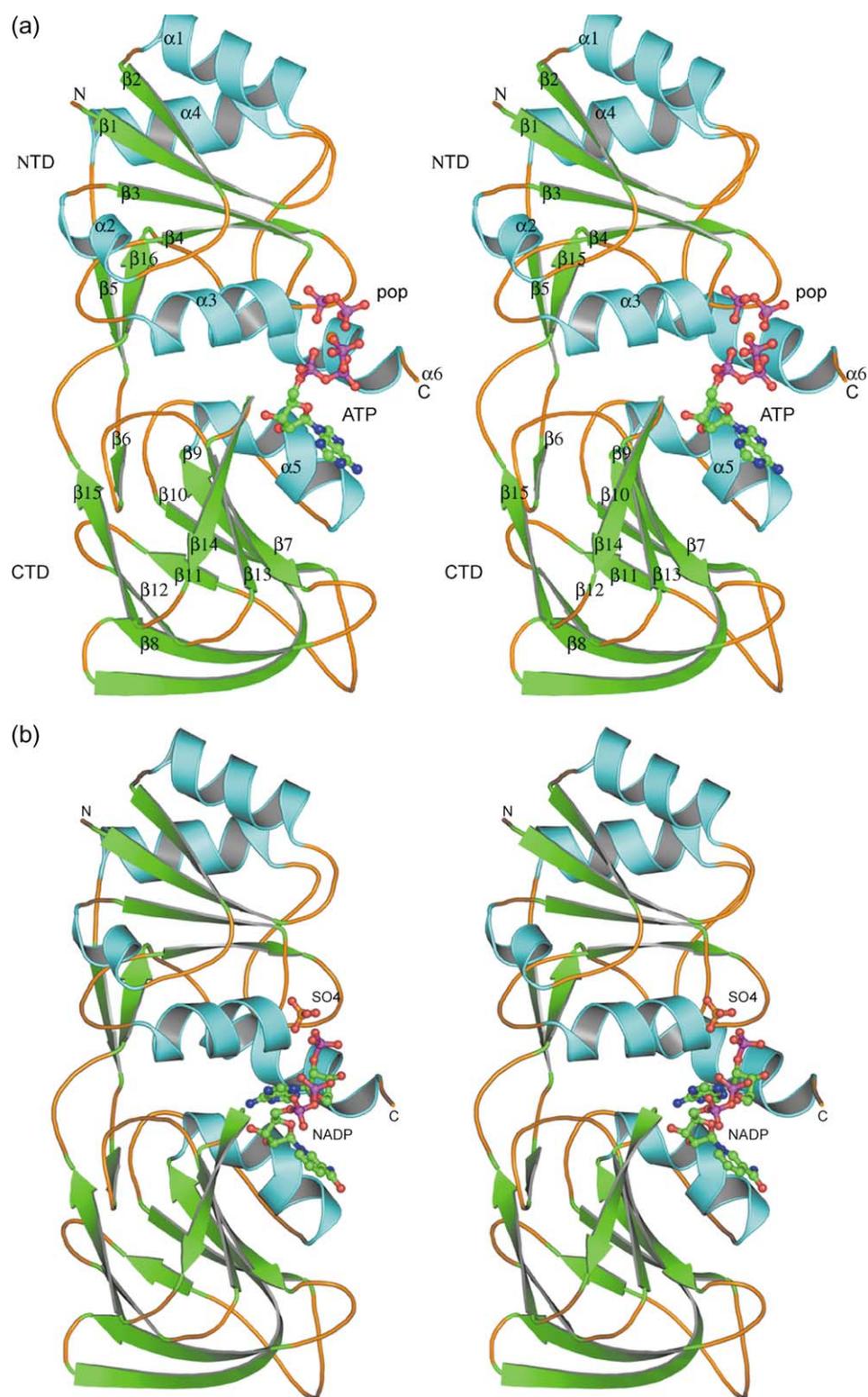


Figure 3 (legend next page)

Oligomeric structure of Afnk

NAD kinases are known to form a multimeric state with a tetrameric state observed for both the Ppkn apo and NAD complex structures.^{15,16} Afnk also exists in a tetrameric state in both the refolded

Afkn protein and its NADP complex according to size-exclusion chromatography and dynamic light-scattering data (data not shown). In the crystal structures, the Afkn-NAD, Afkn-NADP and Afkn-ATP all form a tetramer, while the Afkn-NADP2 forms a tetramer through the crystallographic dyad

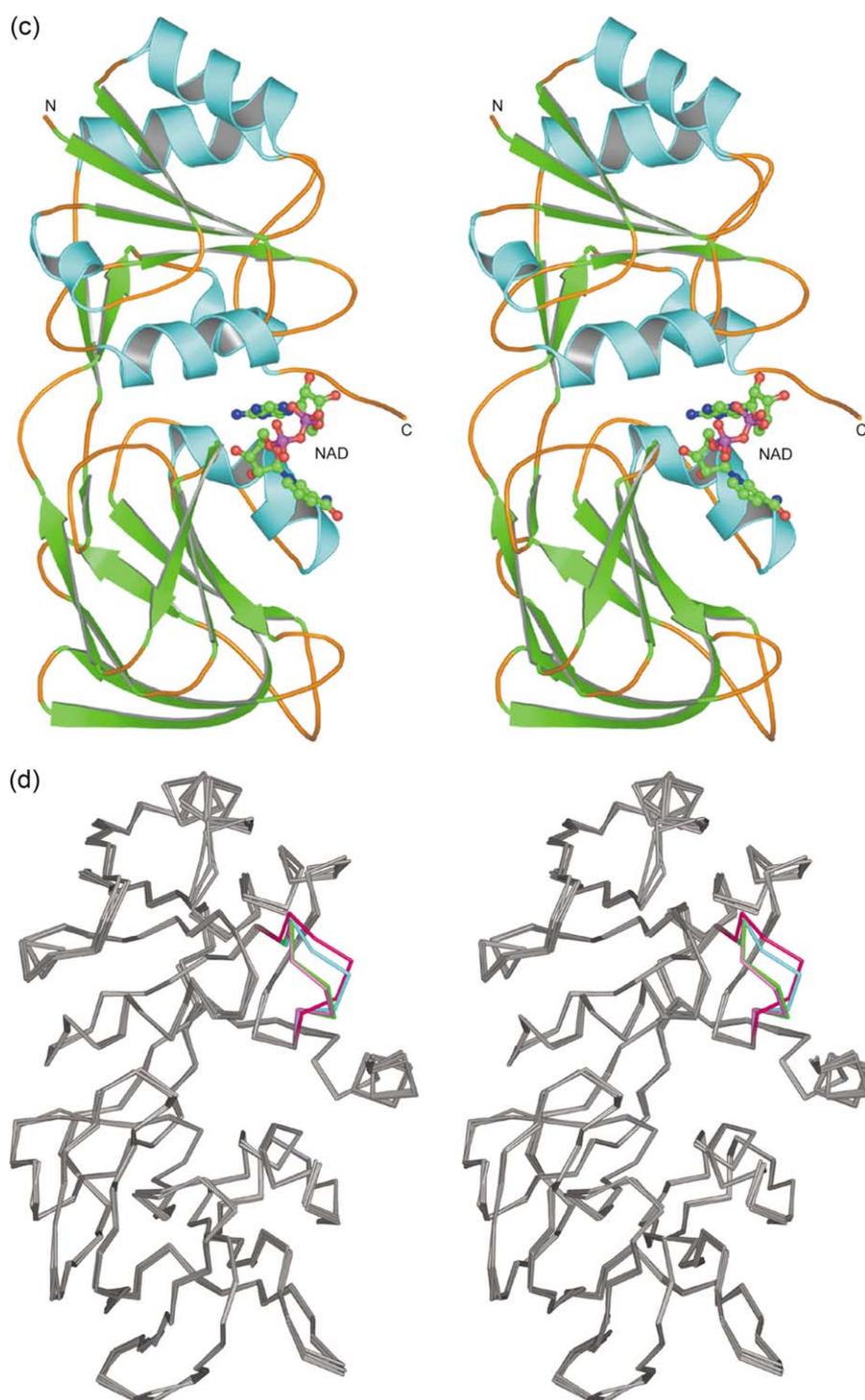


Figure 3. Stereo view of the ribbon diagram of Afnk in complex with ATP, NADP and NAD. The α -helices, β -strands, and loops are shown in cyan, green and orange, respectively. ATP, pyrophosphate (POP), NAD, NADP, and SO_4 are shown as ball-and-stick models. The magnesium ion is shown as a sphere. The atom labeling is as follows: C, green; N, blue; O, red; P, pink; Mg, orange; S, yellow. N and C termini are indicated. (a) Stereo view of the ribbon diagram of Afnk in complex with ATP, magnesium and a pyrophosphate group. (b) Stereo view of the ribbon diagram of Afnk in complex with NADP and SO_4 . (c) Stereo view of the ribbon diagram of Afnk in complex with NAD. (d) Stereo view of the superposition of structures of the Afnk-NAD, Afnk-NADP, Afnk-NADP2, Afnk-ATP. The loop region of Thr69-Gly74 is shown in different colors; Afnk-NAD, red; Afnk-NADP, cyan; Afnk-NADP2, green; Afnk-ATP, salmon.

(Figure 4). Superposition of all four tetramers revealed that the four tetramers are almost identical (data not shown). The Afnk tetramer includes two types of dimers: the AA' dimer and the AB dimer

(Figure 4). The AA' dimer involves extensive interactions between C-terminal domains and the same type of monomer–monomer interactions are observed in all four structures. A similar dimer

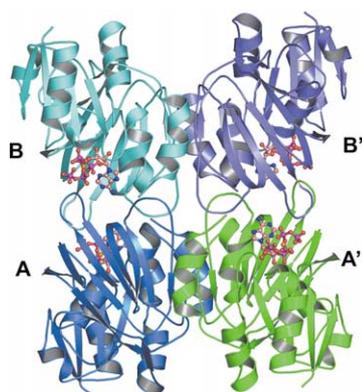


Figure 4. Ribbon diagram of the Afnk tetramer in complex with ATP. The individual monomers are labeled as A, A', B and B'.

interface is found in the Ppnk structures, except that the Afnk does not contain the long C-terminal tail that anchors the dimerization. Taken together, the AA' dimer is probably the most basic oligomeric state for NAD kinases. The AB dimer involves substrate NAD-binding, phosphate donor ATP-binding and final product NADP-binding. Functionally important conserved residues from both monomers of the AB dimer have been found to play an important role in NAD binding in the Ppnk NAD bound structure.¹⁶ In the NAD, NADP or ATP-bound structures presented here, those interactions are found to be conserved (discussed below). All together, the tetrameric state is observed in all stages required for Afnk to carry out its function (including NAD, ATP and NADP binding), and functionally important conserved residues from both monomers in the AB dimer are critical for all ligand-binding. This suggests that the tetrameric state is likely to be the functional oligomeric state for NAD kinases.

NAD and NADP binding

NAD is the substrate and NADP is the final reaction product of NAD kinases. In the Afnk NAD and NADP-bound structures, all NAD and NADP molecules are well-defined (Figure 5(a) and (b)). The NAD and NADP molecules bound in all three structures have almost identical conformations, except NADP molecules have an extra 2' sugar phosphate group. In addition, the NAD and NADP in these structures share an almost identical chemical environment, suggesting that the substrate NAD and the product NADP share the same binding mode. In the following description, the interactions between Afnk and its product NADP are based on the structure of Afnk-NADP2, since it was refined at the highest resolution.

NAD/NADP is bound in the deep cleft between the N and the C-terminal domains, and in the dimer interface between molecule A and molecule B

(Figures 3(b) and (c) and 5(c)). The NAD and NADP-binding mode for the adenine base and the nicotinamide ribose moiety in these three structures are similar to that described in the Ppnk-NAD complex structure. However, a higher resolution and the presence of NADP molecules allow us to look in more detail at the interactions between NAD/NADP and NAD kinases.

In the Afnk-NAD/NADP structures, the adenine ring is bound to a hydrophobic cavity inside the binding pocket and this is observed in all NAD/NADP-bound structures (Figure 5(a)). The AN6 atom of the adenine base forms a hydrogen bond with the OG1 atom of Thr156, the AN7 atom forms a hydrogen bond with the carbonyl atom of Ile153 and the OD1 atom of Asn115, and AN1 forms a hydrogen bond with the amide atom of Asn115 (Figure 5(c)). Similar interactions for these three residues are observed in the Ppnk-NAD complex structure, and the importance of Thr156 and Asn115 has been confirmed by mutagenesis studies.¹¹ Another possibly important residue for adenine binding that was not mentioned previously is Leu75, which forms strong hydrophobic interactions with the adenine ring. In most NAD kinase proteins, this residue is Phe or Tyr. The main-chain N atom of Leu75 forms a hydrogen bond with the OD2 atom of Asp49 in the GG DG motif. This hydrogen bond interaction is observed for all three NAD/NADP bound structures. Asp49 may be involved in stabilizing the Leu75 residue, since it is located right after the Thr69-Gly74 flexible loop.

The nicotinamide ribose portion of NAD/NADP is located at the dimer interface and is coordinated by residues from both molecule A (Mol A) and molecule B (Mol B) (Figure 5(c)). The nicotinamide ring, located in the hydrophobic pocket formed by Met127 from Mol B, and residue Y158, is parallel with the benzene ring of Tyr158 and the side-chain of Met127, and forms possible π - π interactions or hydrophobic interactions between them. The NN7 atom of nicotinamide forms a hydrogen bond with the OG1 atom of Ser161 and the NO7 atom of nicotinamide forms a hydrogen bond with the OD1 atom of Asp145 from Mol B and is close to the carbonyl atom of Ala180 from Mol B. The nicotinamide ribose forms bifurcated hydrogen bonds with the carboxylate group of Glu116. Residues Glu116, Ser161, Asp145 and Y158 are all conserved and their importance has been confirmed by mutagenesis studies.^{13,16} Similar interactions are found in the Ppnk-NAD structure.¹⁶

In the high-resolution Afnk-NADP2 structure, two additional well-defined internal water molecules (W1 and W2) are found to form hydrogen bonds to the OE1 and OE2 atoms of Glu116 (Figure 5(c)). One of the water molecules forms a hydrogen bond with the carbonyl oxygen atom of Gly154 and the main-chain N atom of Ala159. Another water molecule forms hydrogen bonds with main-chain N atoms of Asp209 and Gly210, and the carbonyl oxygen atom of Val207 (not

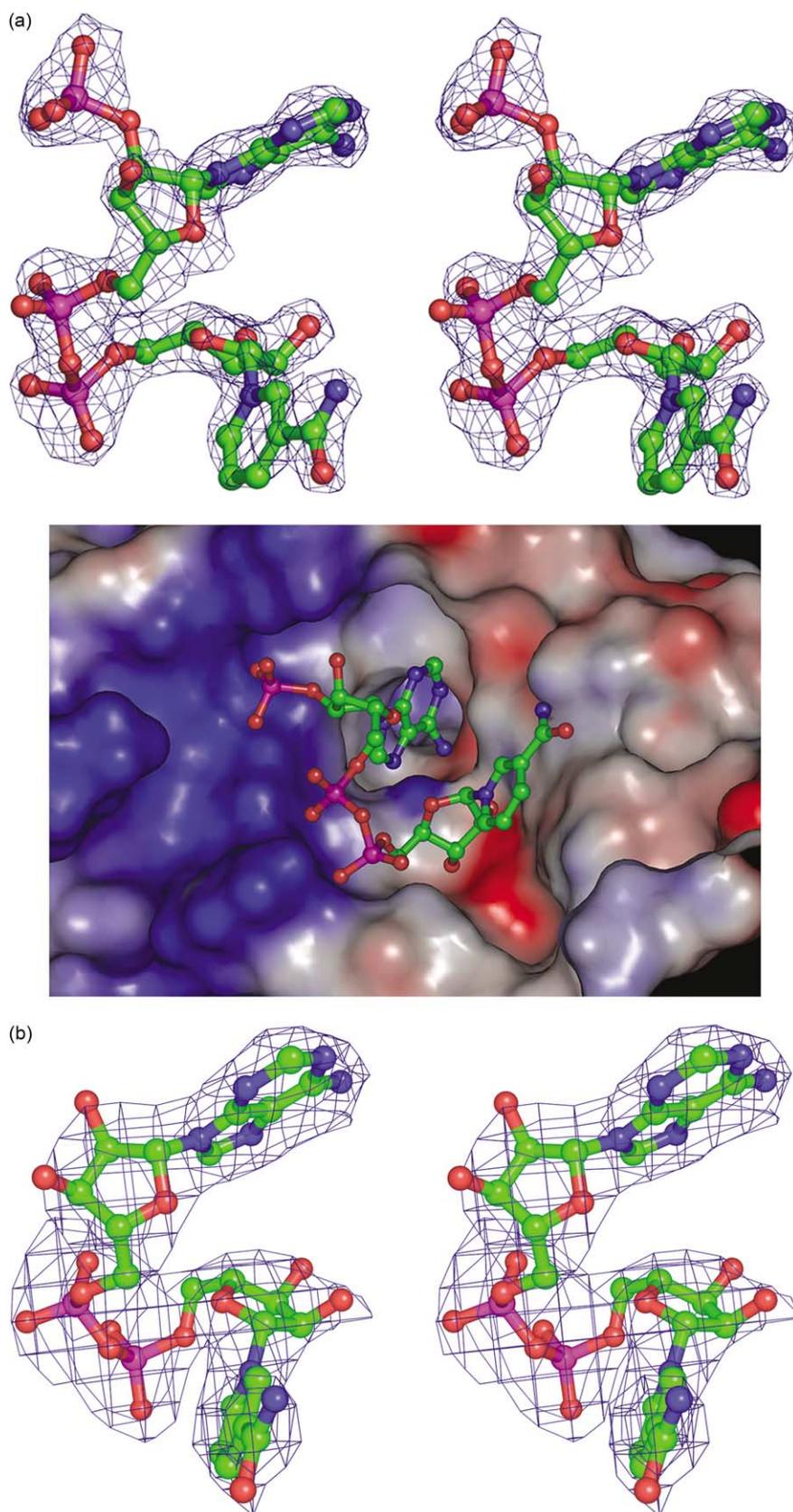


Figure 5 (legend next page)

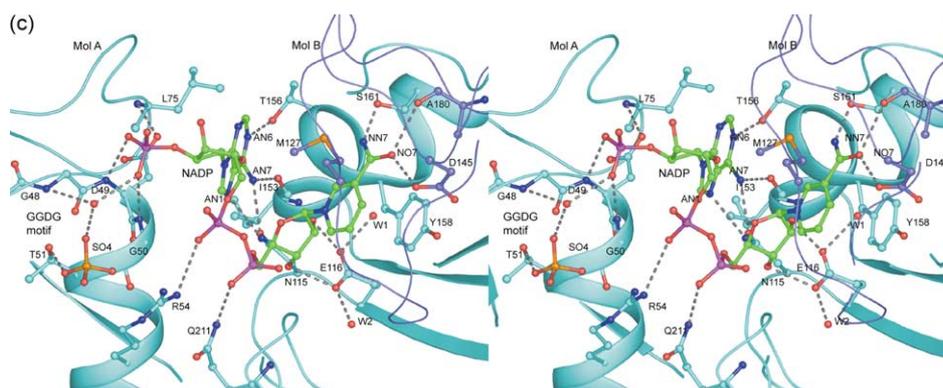


Figure 5. NAD/NADP binding site for Afnk protein. The atom labeling is as follows: C, green; N, blue; O, red; P, pink; S, yellow. (a) Top: stereo view of the $2F_o - F_c$ electron density map of the NADP molecule in the Afnk-NADP2 structure. This map is contoured at 1.2σ . Bottom: electrostatic surface presentation of the NADP binding site in the Af-NADP2 structure. NADP is shown as a ball-and-stick model. (b) Stereo view of the $2F_o - F_c$ electron density map of the NAD molecule in the Afnk-NAD structure. This map is contoured at 1.0σ . (c) Stereo view of the NADP-binding site in the Afnk-NADP2 structure. Molecules A and B are shown in cyan and marine, respectively. NADP, SO_4 , and residues that interact with NADP and SO_4 are shown as a ball-and-stick model. Hydrogen bonds are shown as dashes in black.

shown). Similar water molecules were found in another monomer in the Afnk-NADP2 structure. Interestingly, these two water molecules are observed in all four monomers in the Afnk-ATP structure with the same hydrogen bond coordination. These observations suggest strongly that these two water molecules may be important for correctly positioning the carboxylate group of Glu116, which forms bifurcated hydrogen bonds to the two oxygen atoms of the nicotinamide ribose.

In both the NADP-bound Afnk structures, the 2' sugar phosphate groups form hydrogen bonds with the main-chain N atoms of Asp49 and Gly50 (Figure 5(c)). These two residues belong to the highly conserved GG DG motif, which is also conserved in the super kinase family. The GG DG motif is proposed to be involved in nucleotide binding (see Introduction). In the Afnk-NADP2 structure, a SO_4 ion was built into the structural model according to the difference density. One of the oxygen atoms in the SO_4 ion forms a strong hydrogen bond to the OG1 atom of Thr51, and another oxygen atom of the SO_4 ion forms a hydrogen bond to a well-ordered water molecule. This water molecule is, in turn, hydrogen bonded to the 2' sugar phosphate oxygen atom and the main-chain N atom of Gly48 in the GG DG motif. The same SO_4 ion and the water molecules are found in another molecule in the asymmetric unit in the Afnk-NADP2 structure.

ATP binding

The structure of the Afnk-ATP complex was determined by molecular replacement using the apo structure as the search model and refined at 1.7 \AA resolution. The initial difference electron density map clearly shows the presence of an ATP molecule and a magnesium ion bound to each monomer. An extra difference density was found at

the other side of the magnesium ion. This density was interpreted as a pyrophosphate molecule. In the final refined structure, the ATP molecule, magnesium ion and pyrophosphate molecule were well defined in the 1.7 \AA resolution structure (Figure 6(a)).

The AMP portion of the ATP molecule was found to bind in the same binding site as the nicotinamide ribose portion of the NAD/NADP molecule in the Afnk-NAD/NADP structures (Figures 5(c) and 6(b)). The phosphate tail sticks out of the binding pocket towards the solvent region. The adenine ring of ATP is positioned in the same orientation as the nicotinamide ring in the NAD/NADP-bound structures, forming similar π - π interactions with the conserved residue Tyr158 and strong hydrophobic interaction with the Met127 from molecule B (Figure 6(b)). N1, N6, and N7 of the adenine base are involved in hydrogen bonding to the OD1 atom of Asp145 from Mol B, the carbonyl oxygen atom of Ala180 from Mol B and the OG atom of Ser161, respectively. These three residues are involved also in hydrogen bonding with the nicotinamide moiety of the NAD/NADP-bound structures. The ATP ribose also forms bifurcated hydrogen bonds with the carboxylate moiety of Glu116 and, interestingly, the two water molecules that were observed to form hydrogen bonds to the carboxylate group of Glu116 in the Afnk-NADP2 structure are conserved also and form the same coordinations with the Afnk protein. Superposition of the Afnk-ATP and Afnk-NADP structures indicated that the AMP portion of the ATP molecule in the Afnk-ATP structure shares almost the same binding mode as the nicotinamide ribose portion of NAD/NADP (Figure 6(c)). The phosphate tail of the ATP molecule protrudes out of the binding pocket and is coordinated by a magnesium ion (Figure 6(b)). All three phosphate groups are involved in coordination to the magnesium ion.

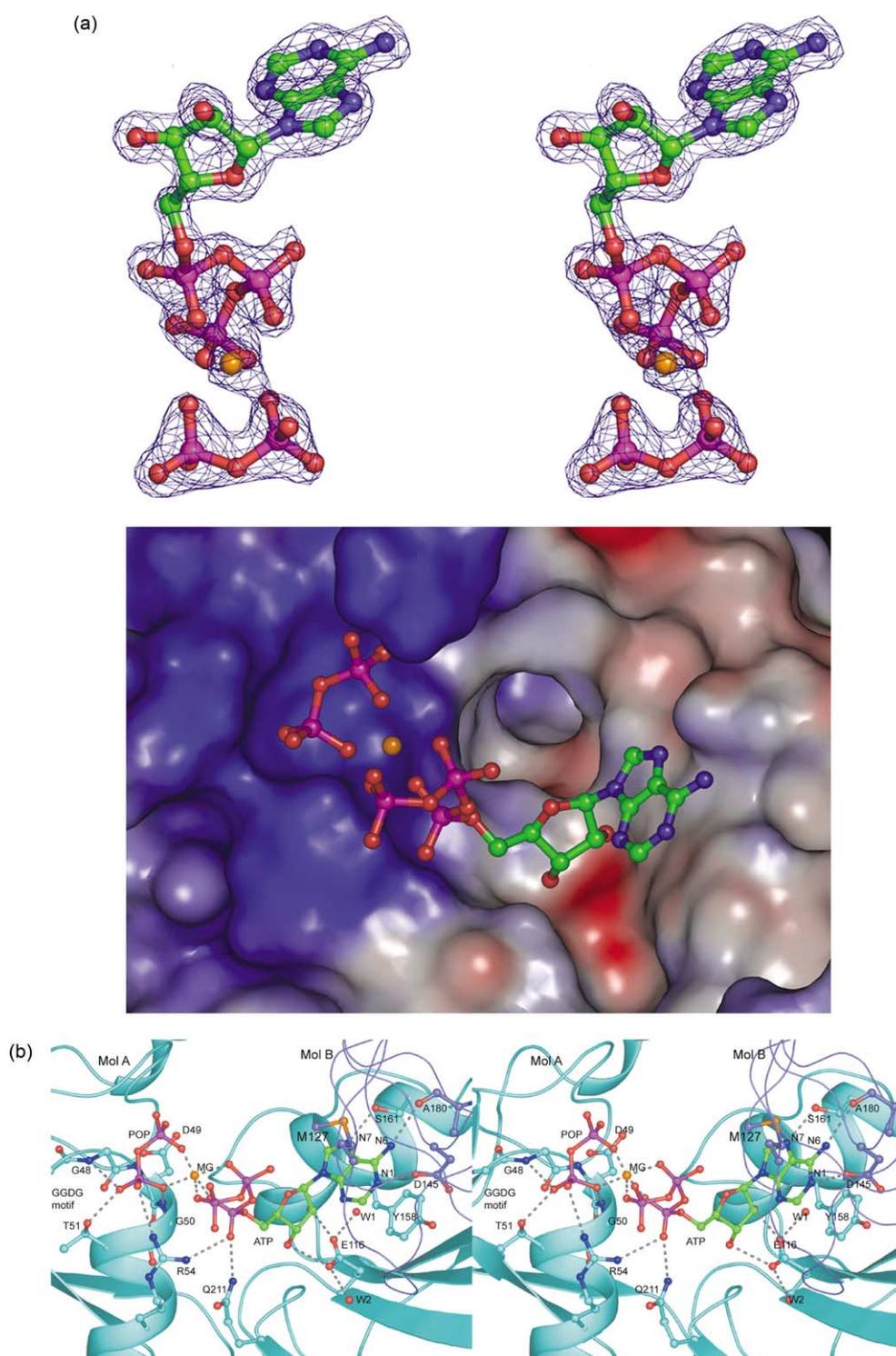


Figure 6 (legend next page)

On the other side of the phosphate tail, the magnesium ion is coordinated by a pyrophosphate moiety (of a second ATP, of which the AMP moiety is partially disordered and occupies a part of the adenosine-binding pocket of the Afnk-NAD/NADP structures, see Discussion). There are a total of five oxygen atoms from five phosphate groups involved in the coordination of the magnesium ion. The average distance between the

magnesium ion and these oxygen atoms is 2.01 Å. The pyrophosphate is involved in interactions with the GGDG motif (Figure 6(b)). One phosphate oxygen atom is involved in hydrogen bonding to the main-chain N atom of Gly48, and another oxygen atom from the same phosphate is involved in hydrogen bonding to the main-chain N atom of Gly50. Besides the interactions with the GGDG motif, the phosphate oxygen that forms hydrogen

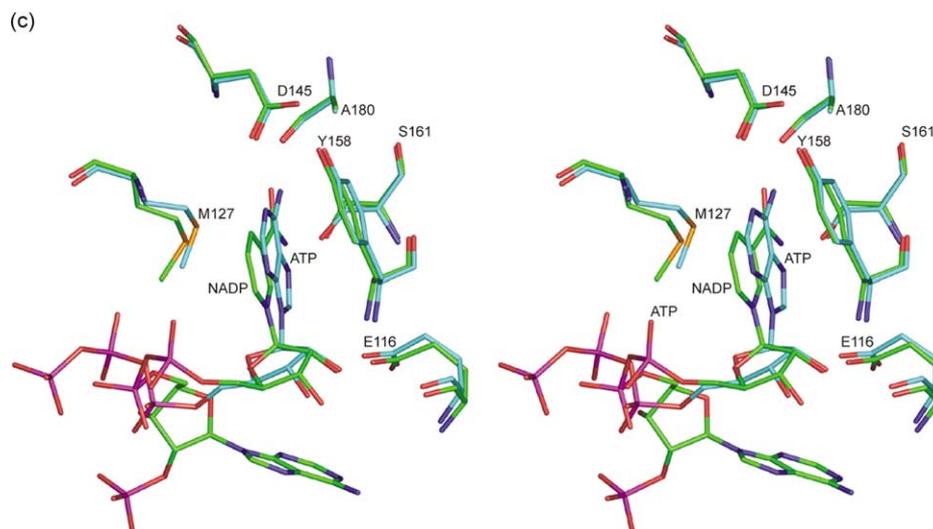


Figure 6. The ATP-binding site for Afnk protein. (a) Top: stereo view of $2F_o - F_c$ electron density map of the ATP molecule in the Afnk-ATP structure. This map is contoured at 1.6σ . Bottom: electrostatic surface presentation of the ATP-binding site. ATP is shown as a ball-and-stick model. (b) Stereo view of the ATP-binding site in the Afnk-ATP structure. Molecules A and B are shown in cyan and marine, respectively. ATP, pyrophosphate (POP) and residues that interact with NADP and SO_4 are shown as a ball-and-stick model. Magnesium and water are shown as spheres. Hydrogen bonds are shown as dashes in black. (c) Stereo view of the superposition of the AMP-binding site of the ATP molecule in the Afnk-ATP structure (in cyan) and the nicotinamide ribose-binding site of the NADP molecule in the Afnk-NADP2 structure (in green). ATP, NADP and residues M127, D145, A180, S161, Y158 and E116 are shown as sticks.

bonding with the N atom of Gly48 is also involved in strong hydrogen bonding to the OG1 atom of Thr51. Note, in the Afnk-NADP2 structure, the OG1 atom of Thr51 is involved in strong hydrogen bonding to a SO_4 ion.

The pyrophosphate molecule, magnesium, and the phosphate tail of the first ATP molecule are bound in a pocket with a highly basic chemical environment suggested by the electrostatic surface (Figure 6(a)). The basic chemical environment is contributed mostly by several basic residues, including Arg54, Asn211, Arg72, Lys8 and Arg143. Similar residues are found in Ppkn, except that the corresponding residue for Arg143 is a Gly. In the Afnk-ATP structure, the side-chains of Arg72 and Lys8 are disordered; Arg54 forms a hydrogen bond with the AMP phosphate group and the pyrophosphate molecule, and Asn211 forms a hydrogen bond to the AMP phosphate group (Figure 6(b)). These two residues are found to interact also with the phosphate groups in the two NADP-bound structures (Figure 5(c)).

Discussion

The modeling of pyrophosphate (of the second ATP molecule) was based on the high-quality 1.7 \AA resolution difference electron-density map, and on its coordination with the magnesium ion and the conserved GGDG motif, which had been proposed to be involved in nucleotide binding. The modeled pyrophosphate moiety was well defined in the final σ_A -weighted $2F_o - F_c$ electron-density map. Since there was no pyrophosphate included in the protein refolding process or in the crystallization condition,

we propose that the pyrophosphate moiety is the β and γ -phosphate groups of the second bound ATP molecule with the AMP portion disordered. This interpretation is based on the following reasons. (1) The pyrophosphate moiety forms hydrogen bonds with the GGDG motif, which has been proposed to be involved in nucleotide binding. (2) In the Afnk-ATP structure, weak difference density is observed extending from the pyrophosphate moiety to the AMP-binding site of the NAD/NADP molecules in the NAD/NADP-bound structures. (3) The pyrophosphate moiety is located in a more favorable position for a phosphate group to be transferred to the phosphorylation site of the substrate NAD. Since the second ATP molecule of the Afnk-ATP structure forms hydrogen bonds with the GGDG motif, which is also hydrogen bonded by the 2' sugar phosphate moiety of the NADP molecule in the Afnk-NADP/NADP2 structures, and the γ -phosphate group of the second ATP molecule is located in a favorable position for phosphate transfer to the phosphorylation site of the NAD molecule, this ATP molecule (or its pyrophosphate moiety representing polyphosphate) is likely to be the actual phosphate donor. In this respect, the first ATP molecule may be a fortuitously bound ATP molecule in the absence of NAD. However, we cannot exclude the possibility that the first ATP molecule is the phosphate donor.

The precise role of the conserved GGDG motif in NAD kinases has not been established. It has been proposed to be involved in nucleotide binding.^{17,18} In both Afnk-NADP structures, the GGDG motif is hydrogen bonded to the 2' sugar phosphate moiety of NADP molecules. In the Afnk ATP-bound structure, the GGDG motif is hydrogen bonded to

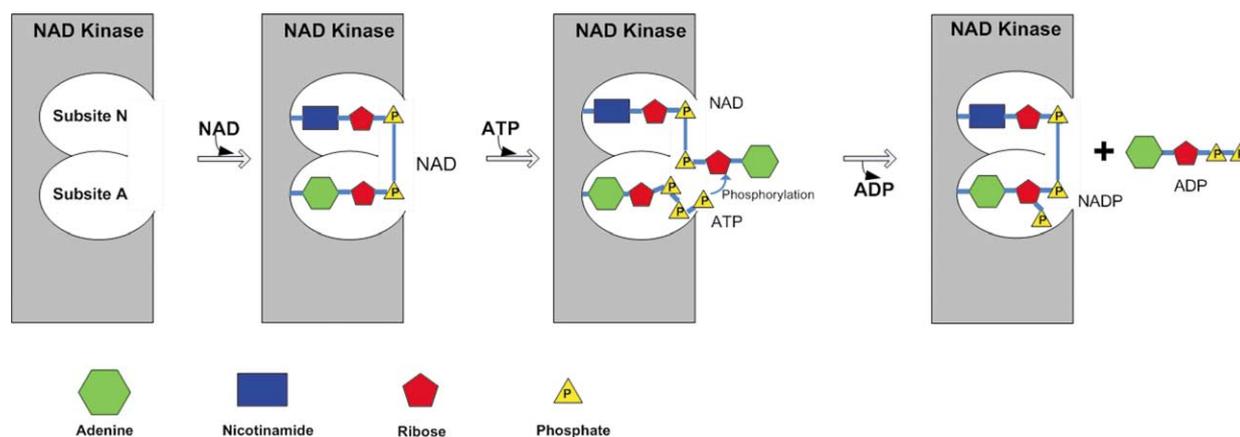


Figure 7. Proposed phosphorylation mechanism for NAD kinase, which kinase is shown as a rectangle. The nicotinamide subsite and the AMP subsite are shown as Subsite N and Subsite A, respectively.

the pyrophosphate. As we stated above, this pyrophosphate moiety could be the β and γ -phosphate portion of the second bound ATP molecule, which may be involved in phosphate transfer. These observations suggest that the GGDG motif may play a role in phosphate transfer.

Residue Thr51 is immediately after the GGDG motif in the Afnk protein sequence. The OG1 atom of this residue was found to form a hydrogen bond with the oxygen atom of the SO_4 ion in the Afnk-NADP2 structure. The SO_4 ion is involved in a hydrogen bond network that includes the SO_4 ion, a well-ordered water molecule, the main-chain N atom of Gly48 of the GGDG motif, and the 2' sugar phosphate moiety of the NADP molecule. The identical coordination was observed for another molecule in the asymmetric unit in the Afnk-NADP2 structure. The SO_4 ion in these structures may be mimicking a phosphoryl group of the phosphoryl donor molecule, either ATP or polyphosphate. In the Afnk-ATP structure, the Thr51 OG atom forms a strong hydrogen bond with the phosphate oxygen atom of the pyrophosphate. Although this residue is not absolutely conserved in the NAD kinase family (35 out of 40 are Thr, two others are Asn in the Pfam listed NAD kinase sequences[†]), we propose that this residue together with the GGDG motif may play a role in phosphate transfer.

On the basis of the structures of Afnk in complex with NAD, NADP or ATP, we propose a possible phosphate transfer mechanism model (Figure 7). In this model, the NAD kinase contains a dinucleotide-binding site composed of two subsites: the AMP-binding site (subsite A) and the nicotinamide binding site (subsite N). Before the phosphorylation occurs, the substrate NAD binds to the subsite A using its AMP portion and the subsite N using its nicotinamide ribose portion. During the phosphorylation, the AMP moiety of the phosphate donor ATP partially occupies subsite A, and

the AMP portion of NAD is partially displaced from subsite A. After the phosphorylation, the AMP portion of NADP binds back to subsite A and ADP is released. Our model of phosphate transfer mechanism of NAD kinase is structure-based and putative: Although our model suggests Afnk binds substrate NAD first, phosphate donor ATP then displaces the AMP portion of NAD, the exact order of binding is not known. Our model does not suggest a unique mechanism for NADP release. Further studies are needed to fully understand the mechanisms of the whole process of NAD kinase.

Materials and Methods

Cloning, expression and purification

The Afnk (gi number 2650718) gene was amplified by PCR using *A. fulgidus* genomic DNA template and primers designed for ligation-independent cloning (LIC).²⁰ The amplified PCR product was prepared for vector insertion by purification, quantification and treatment with phage T4 DNA polymerase (NEB) in the presence of 1 mM dTTP. The prepared insert was annealed to the LIC expression vector pB3, a derivative of pET21a (Novagen, San Diego, CA) that expresses the cloned gene with an N-terminal His₆ tag-TEV cleavage sequence, and transformed into chemically competent DH5 α .

The SeMet isoform of Afnk protein was expressed in *E. coli* B834 (DE3). Cell paste was prepared using Studier's auto-induction method (Dr William Studier, Brookhaven National Laboratory, personal communication).

For protein purification, cell paste was resuspended in buffer A (50 mM Hepes (pH 7.5), 0.1 M NaCl, 10 mM β -mercaptoethanol) supplemented with protease inhibitors (Roche Applied Science, Indianapolis, IN). Cells were opened using a Microfluidizer (Microfluidics, Newton, MA). After centrifugation, the cleared cell lysate was loaded onto a 5 ml HiTrap MC column (GE Healthcare, Piscataway, NJ). The column was washed with ten column volumes of buffer A, ten column volumes of buffer B (50 mM Hepes (pH 7.5), 1 M NaCl, 10 mM β -mercaptoethanol) and ten column volumes of buffer A. The target protein was eluted with an imidazole gradient

[†] <http://www.sanger.ac.uk/Software/Pfam/>

from 40 mM–300 mM in 20 CV. The target eluted at about 300 mM imidazole. For a final purification step, the eluted protein was concentrated to 30 mg/ml and loaded onto a Superdex-75 size-exclusion column (GE Healthcare, Piscataway, NJ) and eluted with buffer C (20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10 mM β -mercaptoethanol, 300 mM imidazole). The peak fractions were then concentrated to 15 mg/ml in buffer C prior to crystallization.

Protein refolding

The Afnk expressed in *E. coli* was found to be in complex with its product NADP. In order to obtain the apo protein, the purified Afnk protein was denatured in buffer D (buffer C + 6 M guanidinium-HCl). The denatured protein was dialysed three times against buffer D and then twice against buffer C. The supernatant of the dialysed protein solution was concentrated to 10 mg/ml and loaded onto a Superdex-75 and the peak fractions were concentrated to 15 mg/ml.

Crystallization and data collection

The Afnk-NADP complexes were crystallized using the vapor-diffusion method in 2 M NaCl, 10% (w/v) PEG 6000 at 4 °C. Large crystals were obtained by mixing 1 μ l of protein solution (15 mg/ml in buffer C) with 1 μ l of reservoir solution and equilibrated over 1 ml of reservoir solution for a week at 4 °C. A second crystal form of the Afnk and NADP complexes (Afnk-NADP2) was found using 0.1 M Hepes (pH 7.5), 2 M ammonium sulfate at room temperature. Large rod-shaped crystals were obtained in two weeks.

Crystals of Afnk-NAD were obtained by mixing 1 μ l of the refolded protein (15 mg/ml in buffer C containing 20 mM NAD) with 1 μ l of reservoir solution containing 0.1 M Tris-HCl (pH 8.5), 0.2 M proline, 10% (w/v) PEG 3350 and equilibrated over 1 ml of reservoir solution for one day at room temperature.

The crystals of Afnk-ATP were obtained by mixing 1 μ l of refolded protein (15 mg/ml in buffer C containing 50 mM ATP and 100 mM magnesium acetate) with 1 μ l of reservoir solution containing 0.2 M disodium hydrogen phosphate, 20% PEG 3350 and equilibrated over 1 ml of reservoir solution for about two weeks at room temperature.

For data collection, the Afnk-NADP crystals were transferred to a cryoprotectant solution (reservoir solution plus buffer C and 20% ethylene glycol) and equilibrated for at least 15 min before they were frozen directly in liquid nitrogen. Native data were collected at the Advanced Light Source (ALS) Beamline 5.0.1 (Lawrence Berkeley National Laboratory, Berkeley, CA) at 100 K with a Quantum 210 charge-coupled device (CCD) detector and processed with HKL 2000.²¹ The crystals belonged to space group $P4_32_12$ or $P4_12_12$ with unit cell dimensions of $a=b=122.1$ Å, $c=198.6$ Å. These crystals contain four monomers per asymmetric unit, corresponding to a solvent content of 56% (v/v). The MAD data for Afnk-NADP SeMet crystals were collected at ALS Beamline 5.0.2 at 100 K with a Quantum 315 CCD detector. A single SeMet crystal was used for fluorescence scans and data collection. Data sets were collected at the peak wavelength, inflection point wavelength and a high-energy reference wavelength. All data were processed with HKL2000.²¹ Data statistics are summarized in Table 1.

Data collections for the Afnk-NAD and the Afnk-NADP2 crystals were both carried out at the ALS Beamline 8.2.2 at 100 K with a Quantum 315 CCD detector and processed with HKL 2000.²¹ The Afnk-NAD crystals were transferred to a cryoprotectant solution (reservoir solution + buffer C + 10 mM NAD plus 20% glycerol) and equilibrated for about 5 min before they were directly frozen in liquid nitrogen. The Afnk-NADP2 crystals were transferred to a cryoprotectant solution (reservoir solution plus buffer C plus 25% glycerol) and equilibrated for about 2 min before they were frozen directly in liquid nitrogen. The Afnk-NAD crystals belonged to the same space group and unit cell as the Afnk-NADP crystals but diffracted to only 3.05 Å resolution. The Afnk-NADP2 crystals belonged to space group $P2_12_12$ with unit cell dimensions of $a=62.9$ Å, $b=76.4$ Å, $c=120.0$ Å with two molecules in the asymmetric unit. These crystals diffracted to 2.0 Å. Data statistics are summarized in Table 1.

Data collection for the Afnk-ATP crystals were carried out at the ALS Beamline 8.2.1 at 100 K with a Quantum 210 CCD detector and processed with HKL 2000.²¹ The plate-shaped crystals were transferred to a cryoprotectant solution (reservoir solution plus buffer C, 25 mM ATP, 50 mM magnesium acetate and 25% glycerol) and equilibrated for 5 min before they were frozen directly in liquid nitrogen. The Afnk-ATP crystals belonged to space group $P1$ with unit cell dimensions of $a=55.3$ Å, $b=69.7$ Å, $c=76.3$ Å, $\alpha=106.4^\circ$, $\beta=111.3^\circ$, $\gamma=101.3^\circ$ with four molecules in the asymmetric unit. These crystals diffracted to 1.7 Å resolution. Data statistics are summarized in Table 1.

Structure determination

The structure of the Afnk-NADP complex was solved by MAD using SeMet-substituted protein. The location of all 16 selenium atoms was determined using the HySS program and confirmed $P4_12_12$ to be the correct space group.^{22,23} Phasing, and phase extension to 2.45 Å resolution with native data using density modification were performed with the CNS program.²⁴ The resultant electron density map was readily interpretable. An initial partial model was built by RESOLVE and completed manually using the program O.^{25,26} Well-defined difference density that chemically resembled NADP was identified in each protein subunit during the initial model building. The coordinates for these NADP molecules were not included in model refinement until the final stages. Model refinement was carried out with the CCP4 program REFMAC5 using the TLS method.^{27,28} The final model was refined to $R_{\text{free}}=26.4\%$ and an R -factor of 22.1% at 2.45 Å resolution, and contains all residues for the four Afnk molecules, four NADP molecules and 389 water molecules. The final refinement statistics are given in Table 1.

The structure of Afnk-NADP2 was solved by molecular replacement by using a dimer of the Afnk-NADP tetramer model (AA' dimer) as a search model (no NADP or water molecules were included). The molecular replacement solution was found using the program AMoRe with a correlation coefficient of 78.4% and an R -factor of 33.6%.²⁸ The final model was refined at 2.0 Å resolution with an R -factor of 21.2% and $R_{\text{free}}=26.4\%$ (Table 1). The final model of the Afnk-NADP2 structure contains two Afnk molecules, two NADP molecules, three sulfate ions and 299 water molecules.

Since the Afnk-NAD crystals were crystallized in the same space group and unit cell as the Afnk-NADP, the tetramer model of Afnk-NADP was used as the initial

model for structural refinement (no water or NADP molecules were included in the initial model). Initial difference density maps clearly indicated the presence of an NAD molecule in each monomer. The final model was refined to an R-factor of 20.6% and $R_{\text{free}} = 27.2\%$, using Refmac5 with the TLS method applied.^{27,28} The final model of the Afnk–NAD complex contains four molecules of Afnk, four NAD molecules and 133 water molecules. The final refinement statistics are given in Table 1.

The structure of the Afnk–ATP complex was solved by molecular replacement by using the Afnk–NADP tetramer model as a search model (no NADP or water molecules were included). The molecular replacement solution was found using the AMoRe program with a correlation coefficient of 74.3% and an R-factor of 31.1%.²⁸ Initial difference density maps clearly indicated the presence of an ATP molecule and a magnesium ion in each monomer. The final model was refined at 1.7 Å resolution to an R-factor of 20.0% and $R_{\text{free}} = 23.1\%$. The final model of Afnk–ATP includes four Afnk molecules, four ATP molecules, four magnesium ions, four pyrophosphate moieties and 404 water molecules.

Protein Data Bank accession codes

The atomic coordinates of *A. fulgidus* NAD kinase in complex with NAD, NADP and ATP have been deposited with the Protein Data Bank[†] with accession codes 1Z0Z, 1SUW, 1Z0U and 1Z0S.

Acknowledgements

We thank Barbara Gold for cloning, Marlene Henriquez and Bruno Martinez for expression studies and cell paste preparation, Ramona Pufan for crystallization, and John-Marc Chandonia for bioinformatics calculations. We are grateful to the staff of the Berkeley Center for Structural Biology at the Advanced Light Source (ALS). The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under contract no. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory. The work described here was supported by the National Institutes of Health grant GM 62412.

References

- Ziegler, M. (2000). New functions of a long-known molecule. Emerging roles of NAD in cellular signalling. *Eur. J. Biochem.* **267**, 1550–1564.
- Rizzi, M. & Schindelin, H. (2002). Structural biology of enzymes involved in NAD and molybdenum cofactor biosynthesis. *Curr. Opin. Struct. Biol.* **12**, 709–720.
- Berger, F., Ramirez-Hernandez, M. H. & Ziegler, M. (2004). The new life of a centenarian: signalling functions of NAD(P). *Trends Biochem. Sci.* **26**, 111–118.

- Gerdes, S. Y., Scholle, M. D., D'Souza, M., Bernal, A., Baev, V. M. & Farrell, M. (2002). From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J. Bacteriol.* **184**, 4555–4572.
- Kawai, S., Mori, S., Mukai, T., Suzuki, S., Yamada, T., Hashimoto, W. & Murata, K. (2000). Inorganic polyphosphate/ATP-NAD kinase of *Micrococcus flavus* and *Mycobacterium tuberculosis* H37Rv. *Biochem. Biophys. Res. Commun.* **276**, 57–63.
- Kawai, S., Mori, S., Mukai, T., Hashimoto, W. & Murata, K. (2001). Molecular characterization of *Escherichia coli* NAD kinase. *Eur. J. Biochem.* **268**, 4359–4365.
- Garavaglia, S., Galizzi, A. & Rizzi, M. (2003). Allosteric regulation of *Bacillus subtilis* NAD kinase by quinolinic acid. *J. Bacteriol.* **185**, 4844–4850.
- Kawai, S., Suzuki, S., Mori, S. & Murata, K. (2001). Molecular cloning and identification of UTR1 of a yeast *Saccharomyces cerevisiae* as a gene encoding an NAD kinase. *FEMS Microbiol. Letters*, **200**, 181–184.
- Turner, W. L., Waller, J. C., Vanderbeld, B. & Snedden, W. A. (2004). Cloning and characterization of two NAD kinases from Arabidopsis: identification of a calmodulin binding isoform. *Plant Physiol.* **135**, 1243–1255.
- Apps, D. K. (1969). The reaction of pigeon liver NAD kinase with bromoacetyl pyridine. *FEBS Letters*, **5**, 96–98.
- Lerner, F., Niere, M., Ludwig, A. & Ziegler, M. (2001). Structural and functional characterization of human NAD kinase. *Biochem. Biophys. Res. Commun.* **288**, 69–74.
- Kornberg, A. (1999). Inorganic phosphate: a molecule of many functions. *Prog. Mol. Subcell. Biol.* **23**, 1–18.
- Raffaelli, N., Finaurini, L., Mazzola, F., Pucci, L., Sorci, L., Amici, A. & Magni, G. (2004). Characterization of *Mycobacterium tuberculosis* NAD kinase: functional analysis of the full-length enzyme by site-directed mutagenesis. *Biochemistry*, **43**, 7610–7617.
- Labesse, G., Douguet, D., Assairi, L. & Gilles, A. M. (2002). Diacylglyceride kinases, sphingosine kinases and NAD kinases: distant relatives of 6-phosphofructokinases. *Trends Biochem. Sci.* **27**, 273–275.
- Garavaglia, S., Raffaelli, N., Finaurini, L., Magni, G. & Rizzi, M. (2004). A novel fold revealed by *Mycobacterium tuberculosis* NAD kinase, a key allosteric enzyme in NADP biosynthesis. *J. Biol. Chem.* **279**, 40980–40986.
- Mori, S., Yamasaki, M., Maruyama, Y., Momma, K., Kawai, S. & Hashimoto, W. (2004). NAD-binding mode and the significance of intersubunit contact revealed by the crystal structure of *Mycobacterium tuberculosis* NAD kinase–NAD complex. *Biochem. Biophys. Res. Commun.* **327**, 500–508.
- Shirakihara, Y. & Evans, P. R. (1988). Crystal structure of the complex of phosphofructokinase from *Escherichia coli* with its reaction products. *J. Mol. Biol.* **204**, 973–994.
- Pitson, S. M., Moretti, P. A., Zebol, J. R., Zareie, R., Derian, C. K., Darrow, A. L. *et al.* (2002). The nucleotide-binding site of human sphingosine kinase 1. *J. Biol. Chem.* **277**, 49545–49553.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
- Aslanidis, C. & de Jong, P. J. (1990). Ligase independent cloning (LIC) system. *Nucl. Acids Res.* **20**, 6069–6074.

[†] www.rcsb.org

21. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collection in oscillation mode. *Methods Enzymol.* **276**, 307–326.
22. Grosse-Kunstleve, R. W. & Adams, P. D. (2003). Substructure search procedures for macromolecular structures. *Acta Crystallog. sect. D*, **59**, 1966–1973.
23. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallog. sect. D*, **58**, 1948–1954.
24. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
25. Terwilliger, T. C. & Berendzen, J. (1999). Automated structure solution for MIR and MAD. *Acta Crystallog. sect. D*, **55**, 849–861.
26. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
27. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacement in macromolecular refinement. *Acta Crystallog. sect. D*, **57**, 122–133.
28. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
29. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673–4680.

Edited by I. Wilson

(Received 20 May 2005; received in revised form 28 August 2005; accepted 2 September 2005)
Available online 30 September 2005