PROTEINS: Structure, Function, and Genetics 28:117–130 (1997)

Structural Basis for Thermostability and Identification of Potential Active Site Residues for Adenylate Kinases From the Archaeal Genus *Methanococcus*

P. Haney,[1]* J. Konisky,[1] K.K. Koretke,[2] Z. Luthey-Schulten,[2] and P.G. Wolynes[2]
[1]*Department of Microbiology, University of Illinois, Urbana, Illinois*
[2]*School of Chemical Sciences, University of Illinois, Urbana, Illinois*

ABSTRACT **Sequence comparisons of highly related archaeal adenylate kinases (AKs) from the mesophilic *Methanococcus voltae*, the moderate thermophile *Methanococcus thermolithotrophicus*, and two extreme thermophiles *Methanococcus igneus* and *Methanococcus jannaschii*, allow identification of interactions responsible for the large variation in temperatures for optimal catalytic activity and thermostabilities observed for these proteins. The tertiary structures of the methanococcal AKs have been predicted by using homology modeling to further investigate the potential role of specific interactions on thermal stability and activity. The alignments for the methanococcal AKs have been generated by using an energy-based sequence–structure threading procedure against high-resolution crystal structures of eukaryotic, eubacterial, and mitochondrial adenylate and uridylate (UK) kinases. From these alignments, full atomic model structures have been produced using the program MODELLER. The final structures allow identification of potential active site interactions and place a polyproline region near the active site, both of which are unique to the archaeal AKs. Based on these model structures, the additional polar residues present in the thermophiles could contribute four additional salt bridges and a higher negative surface charge. Since only one of these possible salt bridges is interior, they do not appear significantly to the thermal stability. Instead, our model structures indicate that a larger and more hydrophobic core, due to a specific increase in aliphatic amino acid content and aliphatic side chain volume, in the thermophilic AKs is responsible for increased thermal stability. Proteins 28:117–130, 1997.**

Key words: structure prediction; threading; energy function alignment; hydrophobicity; salt bridges; ATP binding

INTRODUCTION

There has recently been a drastic increase in the isolation and investigation of microorganisms living under conditions of extreme temperature, pH, or salinity. Investigation of these microorganisms promises to increase our knowledge of protein stability.[1,2] Previous studies have shown that structurally homologous mesophilic and thermophilic proteins have only minor differences in their free energies of stabilization, often due to the cumulative effect of minor changes in structural features throughout the protein. Increased protein hydrophobicity,[3,4] decreased chain flexibility,[5] additional salt bridges,[6–8] increase helix stability,[9,10] or entropic factors,[11] have all been suggested as methods for increasing protein thermostability (reviewed in Refs. 12–14). How thermophilic proteins employ these methods varies considerably, and the ability to predict or engineer substantial changes in protein stability remains a major problem in protein biochemistry.

Sequence and structural comparisons between functionally related mesophilic and thermophilic proteins are often used in attempts to identify or further characterize thermal stabilizing interactions.[5,8,15] Unfortunately, the significant sequence divergence between most mesophilic and thermophilic proteins can often prevent these comparisons from producing reliable or unambiguous results.[12,16] Mesophilic/thermophilic protein sets characterized by a high degree of sequence identity, along with structural information, are essential for reliable investigations of protein thermostability, although such systems have not been readily available for examination.

Adenylate kinases (AKs), with 68–81% sequence identity, have been recently isolated from a mesophilic and three thermophilic members of the archaeal genus *Methanococcus:* mesophile *M. voltae* (MVO), moderate thermophile *M. thermolithotrophicus* (MTH), and extreme thermophiles *M. jannaschii* (MJA) and *M. igneus* (MIG), with optimal activity

*Correspondence to: P. Haney, Department of Microbiology, University of Illinois, Urbana, IL 61801.
Received 7 June 1996; Accepted 8 November 1996

temperatures of 37°C, 68°C, 88°C, and 85°C, respectively.[17,18] AKs are essential for maintaining the energy balance in the cell by interconverting the adenine nucleotides: $Mg^{2+}ATP + AMP \rightleftharpoons Mg^{2+}ADP + ADP$. While very little is known about archaeal AKs, the eukaryotic and eubacterial AKs and structurally homologous uridylate kinases (UKs) have been extensively studied both biochemically and structurally.[19–24] The methanococcal AKs have only low levels of sequence identity (15–20%) to the eukaryotic and eubacterial AKs and UKs and lack several—previously thought essential—active-site residues.[18,25] Nevertheless, eukaryotic and archaeal adenylate kinases display significant similarities in sequence length, effective inhibitors,[25] predicted secondary structure,[17,18] and secondary structure composition based on circular dichorism[18] and Fourier transform infrared spectroscopy.[26] This suggests that archaeal AKs have a similar tertiary fold, and that modeling procedures utilizing threading algorithms should generate a comparative model structure for the methanococcal AKs. Still, obtaining an accurate alignment between sequences that are lower than 30% identical is very difficult.

Koretke and colleagues[27] have shown that alignments based on self-consistently optimized energy function (SCEF) generally produce better alignments than traditional alignments based on evolutionary scoring matrices when the sequence identity between the sequence and scaffold is lower than 21%. For this reason, the optimized energy function[27–29] was used to probe the sequence–structure compatibility of the four methanococcal AKs to the scaffolds of four AKs and one UK: porcine cytosolic AK, bovine mitochondrial AK, *Escherichia coli* AK, *Saccharomyces cerevisiae* AK, and *S. cerevisiae* UK. As we will show, the resulting models of our test cases for *S. cerevisiae* UK and porcine AK and the methanococcal AKs seem energetically consistent with the known kinase structures and contain structural features required for enzymatic activity. Furthermore, novel active-site interactions and interactions consistent with increased thermostability are identified through sequence comparison and our modeling procedure.

MATERIALS AND METHODS

Protein Sequences and Reference Proteins

High-resolution x-ray structures for four AKs and one UK of varying length have been reported in the literature and deposited in the Brookhaven Protein Data Bank.[30,31] They include porcine AK (PDB code 3ADK) with 194 residues, bovine mitochondrial AK (PDB code 2AK3) with 225 residues, *E. coli* AK (PDB codes 1ANK and 1AKE), with 214 residues, *S. cerevisiae* AK (PDB code 3AKY) with 214 residues, and *S. cerevisiae* (PDB code 1UKZ) with 196 residues. The *E. coli* protein has been crystallized with two different inhibitors; however, the two structures deviate by only 0.53 Å (root mean square [RMS] based on Cα atoms). Since the structures are so similar, we will refer to the *E. coli* scaffold as 1ANK. The methanococcal AKs, each containing 192 residues, are similar in size to the 3ADK and 1UKZ. Their peptide sequences are deposited at Genbank under the following accession numbers: *M. voltae* (U39879), *M. thermolithotrophicus* (U39880), *M. igneus* (U39881), and *M. jannaschii* (U39882).

Alignments Based on Self-Consistently Optimized Energy Functions

The optimized energy function[27–29] used in the threading procedure scores sequence–structure compatibility in terms of contributions from profile (E_p), contact-interactions (E_{ct}), hydrogen-bonding (E_{hb}), gap penalty (E_g), and satisfaction of experimental constraints (E_{exp}):

$$E_T = E_p + E_{ct} + E_{hb} + E_g + E_{exp} \tag{1}$$

where E_p is a measure of the propensity of an amino acid to reside in a particular secondary structure element and to be on the surface or buried inside the protein; E_{ct} is a two-body interaction that depends on the C_β–C_β distances between two particular amino acids; E_{hb} describes the backbone–backbone hydrogen bonds in α helices or β sheets; and E_g is the energetic penalty for placing a gap in either the target sequence, the scaffold, or both. These energy terms can be expressed as a linear function of energy parameters, γ, which have been self-consistently optimized over a set of 29 known proteins. E_{exp} allows inclusion of experimental information about the target protein, such as secondary structure assignment or distances between certain residues. It is an empirically chosen value that adds a stabilizing energy to an alignment when the desired experimental constraint is satisfied in a scaffold.

The five kinase sequences with known structures were aligned to each reference protein (scaffold) using the SCEF and a modified GCG Gap[27,32] alignment programs. Single-sequence alignments of the four archaeal sequences were aligned to each of the five reference proteins using the GCG Gap program, and a multiple sequence alignment of the four archaeal sequences was threaded to the five reference proteins using the SCEF. Each alignment using the optimized energy function was produced within a mean-field dynamic programming algorithm allowing 10 iterations of refinement and a maximal length of 30 residues for each gap. Three core β sheets found in both AK and UK structures were included as experimental constraints within the energy function. Thus alignments that placed residues 6–9, 88–90, and 117–121 of the archaeal sequences into β sheet positions were stabilized by the additional energy contribution E_{exp}. These constraints were included in the energy functions based on the high probability of these residues to be in β sheets as calculated from secondary structure predictions[18] and the conserved hydrophobicity patterns observed in corresponding regions in most AK and UK sequences (Fig. 3). The optimal alignment for the archaeal AKs to each scaffold was determined by finding the maximum energy score averaged over all four sequences.

Many of the active site residues that interact with Mg^{2+} and the phosphate groups of the substrate (ATP-AMP) in eukaryotic and eubacterial kinases appear to be absent in the archaeal AKs, two exceptions are the arginines required for phosphate transfer,[22,26,33] Of the five alignments for the archeal sequences based on the SCEF, only the alignment to the 3ADK scaffold pins these two arginines to the corresponding residues in the scaffold. A second set of alignments were generated by using the optimized energy function by adding a distance constraint between the Cβ atoms of archaeal residues R132 and R138. An additional energy contribution to E_{exp} was given to alignments that placed these two residues within a range of 6.4 to 8.4 Å, a range typical of the corresponding arginines in AK/UK x-ray structures. Again, the optimal alignment for the archaeal AKs to each scaffold was determined by finding the maximum energy score averaged over all four sequences.

Full Atomic Structure

Full atomic models were generated using the program MODELLER.[34] Two models were constructed for each of the following sequences, one based on the SCEF alignment and the other based on the alignment produced with GCG Gap: (1) 3ADK based on its alignments to 2AK3(a); (2) 1UKZ based on its alignments to 3AKY; and (3) each of the archaeal AKs based on their alignments to 2AK3(A), 1UKZ, and 3ADK.

Polyproline Region of Archaeal AKs

An alternate model of the polyproline region unique to the methanococcal AKs (residues 95–109 of the MJA sequence), was constructed in MODELLER using the polyproline region (residues 68–82) of *S. cerevisiae* cytochrome c (PDB code 1CTY[35]) as a template. While there are significant differences between these sequences, better templates are not available in the protein database. Residues 96–108 of the MJA model structure were then replaced with the corresponding residues of the modeled fragment. This altered structure will be referred to as MJAf. Residues 96–105, 96–107, and 97–108 were replaced in MIG, MTH, and MVO, respectively. The number of residues included in the inserted fragment for each model was based on finding the most realistic distance for the two new peptide bonds. These refined structures were then minimized using X-PLOR[36] version 3.1 with the CHARMm22 force field.[37]

RESULTS AND DISCUSSION

Identification of Potential Interactions Involved in Thermal Stability

The high degree of protein sequence identity between the methanococcal AKs (68–81%) make it possible to identify interactions potentially involved in the increased stability of the thermophilic AKs by simple sequence analysis. Several trends previously identified as potential thermodynamic factors can be clearly seen. Table I shows the calculated average protein hydrophobicity,[38,39] aliphatic index,[40] charge content, and chain flexibility[41] for the methanococcal AK enzyme set. The thermophilic methanococcal AKs have a substantially higher average hydrophobicity compared to the mesophilic MVO AK. The average hydrophobicity for each AK was calculated by using several different hydrophobicity scales, based on the energy of transfer of amino acids into different solvents; ethanol, octanol, and cyclohexane.[38,42] Similar results were seen for all scales (data not shown). The hydrophobicity difference between the MJA and MVO AKs is substantial; (202 cal residue^{-1}), approximately equivalent to the substitution of 13 glycines for tryptophans in a hypothetical protein of 192 amino acids. This gain in hydrophobicity is specifically correlated with an increase in the aliphatic side chain volume and content (see aliphatic index Table I). Ikai[40] has previously shown that thermophilic proteins generally have a higher average aliphatic index than mesophilic proteins, and that a 10-unit shift in aliphatic index corresponds to a difference of 5 to 7 kcal mol^{-1} of protein in terms of hydrophobic free energy. The aliphatic indices for the methanococcal AKs are higher then the average aliphatic index for the mesophilic and thermophilic proteins studied by Ikai, 78.8 and 92.6, respectively, but are consistent with values seen between other related extremely thermophilic and mesophilic proteins.

As revealed in Table II there is a strong preference for moderately hydrophobic residues in the mesophilic MVO AK to be replaced by highly hydrophobic aliphatic residues in the thermophilic AKs, with Met to Leu and Val to Ile being the most common. Of the 44 hydrophobic to hydrophobic amino acid differences between the MVO AK and thermophilic AKs, 40 of them result in the more hydrophobic residue being present in the thermophilic protein. This enhancement of overall hydrophobicity through hydrophobic/hydrophobic residue differences is statistically significant even when codon bias is taken into account. Furthermore, conversion of hydrophilic amino acids, mainly Ser and Thr, in the MVO AK to hydrophobic amino acids in the thermophilic AKs is also seen, while the reverse substitution is rare (Table II).

A general trend to increase side chain volume and reduce the serine content in the thermophilic AKs

TABLE I. Properties of the Methanococcal Adenylate Kinases

Organism	MVO	MTH	MIG	MJA
Optimal temp. AK activity (C)	30–40	65–75	80–85	80–90
H0* (cal res^{-1})	1070	1173	1197	1272
Aliphatic index†	85	100	101	104
Chain flexibility‡	1.004	1.000	1.001	0.997
Positively charged residues	24	24	28	29
Negatively charged residues	25	28	31	32

*Average protein hydrophobicity calculated based on the free energy of transfer for amino acids to ethanol from water at 25°C.[38]

†The aliphatic index was calculated according to the method of Ikai and represents the relative volume of a protein occupied by aliphatic side chains.[40]

‡Chain flexibility was calculated by the method of Karplus and Schulz.[41]

can be seen when all amino acid exchanges are examined. This has the cumulative effect of reducing peptide chain flexibility across much of the thermophilic AKs structure (Table I). A reduction in Asn and Gln content is also found in the extreme thermophiles MJA and MIG and may be important in protecting against irreversible denaturation damage resulting from deaminations that can occur at high temperatures.[43]

An exchange of hydrophilic residues between the MVO and the thermophilic AKs is common. These differences are responsible for the increased content of charged amino acids in the thermophilic AKs. The MJA enzyme contains an additional 5 positively and 7 negatively charged residues that in the MVO AK (Table I). The charge increase is due almost exclusively to gaining Lys and Glu residues. This is counter to the preference for Arg over Lys seen in many other thermophilic proteins.[5,44] The potential involvement of hydrophobic and charged residues specific to the thermophile AKs are addressed below.

Alignment and Model Generation of Test Cases

While the high degree of evolutionary conservation between the methanococcal AKs simplified identification of potential thermal stabilizing interactions, the evolutionary distance between the archaeal AKs and other AKs make identification of the active-site residues difficult. Furthermore, the structural role of the potential thermal stabilizing trends identified above are unknown. Generation of three-dimensional models of the methanococcal AKs were attempted in order to identify potential active site residues, and to help substantiate the role of thermostabilizing interactions. Due to evidence suggesting a similar tertiary structure between archaeal and eukaryotic AKs[26] and UKs (George Phillips, personal communication), we utilized an energy-based sequence–structure threading procedure[27] to align a multiple sequence alignment of the four methanococcal AKs to four AK and one UK structures.

To access the accuracy of our threading procedure for AK structures, the five reference protein sequences were aligned to each x-ray structure via the SCEF and GCG Gap. The sequence–structure alignments are nontrivial due to the structural variation in the x-ray structures. The five known structures represent both small and large variant kinases, the main structural difference between these variants being the length of the loop that closes over the active site. Small variant kinases, such as the 3ADK and 1UKZ, have a small 11-residue loop, whereas the large variants, such as 2AK3, 1ANK, and 3AKY, contain a 38-residue insert domain consisting of four β strands. Furthermore, the x-ray structures represent various conformational states of AKs and UKs: a open state or partially closed, which occurs before substrate binding (3ADK and 2AK3), and a closed state, occurring upon substrate binding (1ANK, 3AKY, 1UKZ). The RMS deviations (based on C_α atoms) between the kinase x-ray structures and the alignments generated by the optimized energy function and the GCG Gap programs are shown in Table III, with resulting values ranging from 1.7 to 8.11 Å. The lowest RMS deviations are seen when the scaffold structure is of the same size variant and in the same binding state (e.g., the prediction of the 3AKY structure based on the 1ANK(A) scaffold). It should be noted that in 10 of the 14 test cases, the optimized energy function produces better alignments then GCG GAP, especially for sequences with low percentage identity.

The quality of the alignments for the AK sequences to any scaffold is also measured by a discrimination score D (see Table III) and the position of possible active-site residues (see Table IV). The discrimination score is a dimensionless quantity that measures the ratio of the stability gap (δE) of the predicted structure to the standard deviation in the distribution of energies of misfolded structures (ΔE).

$$D = \frac{\delta E}{\Delta E} = \frac{E_\mathrm{f} - \langle E \rangle_\mathrm{misf}}{\sqrt{\langle E^2 \rangle_\mathrm{misf} - \langle E \rangle_\mathrm{misf} \langle E \rangle_\mathrm{misf}}} \tag{2}$$

where E_f is the energy of the predicted structure based on the optimized energy function and $\langle E \rangle_\mathrm{misf}$ is the average energy of the protein's sequence translated along random structures (i.e., misfolded structures). Table III lists the discrimination scores for alignments generated for each of the known structures. Self-recognition, the alignment of a sequence to its own structure, has values ranging from 10.10 to 12.28 while alignments developed from structural homologues have lower scores, 6.65 to 11.13.

TABLE II. Examination of Amino Acid Difference Between Adenylate Kinases of the Mesophilic *M. voltae* and the Thermophilic *M. thermolithotrophicus, M. igneus,* and *M. jannaschii* Reveals Trends Consistent With Increased Thermal Stability*

Hydrophobic to hydrophobic					Hydrophilic to hydrophobic					Hydrophobic to hydrophilic					Hydrophilic to hydrophilic				
	MTH	MIG	MJA	*Total*		MTH	MIG	MJA	*Total*		MTH	MIG	MJA	*Total*		MTH	MIG	MJA	*Total*
V-I	3	3	6	12	S-G	1	2		3	A-T	1	2	1	4	S-T	3	2	3	8
M-L	3	3	4	10	S-A	1	1	1	3	L-K	1	1	1	3	T-S	1	1	1	3
M-I	1	1	2	4	S-I	1	1		2	P-N		1	1	2	D-E	2	2	2	6
L-I	1	2	1	4	S-V	1		1	2	V-N	1	1		2	E-D	1	2		3
A-V	2		1	3	S-L		1		1	V-Q			1	1	N-E		2	4	6
V-L	1		1	2	S-F			1	1	G-S		1	1	2	R-K	1	2	2	5
I-V†		1	1	2	T-I		1	1	2	G-N	1		1	2	S-N	2	1	1	4
A-I		1		1	T-L		1		1						S-E	1	1	2	4
F-I	1			1	T-V			1	1						S-K	1	1	2	4
L-F			1	1	T-A			1	1						Q-K		2	2	4
G-F			1	1	T-F	1			1						Q-E	1	1	1	3
G-A			1	1	T-G	1			1						Q-N			1	1
C-A†	1			1	N-G		1	1	2						N-K	1	1		2
P-A†		1		1	E-G			1	1						N-D		1	1	2
					D-P			1	1						N-Q	1			1
					E-A		1		1						N-S		1		1
					E-G			1	1						T-E			2	2
					Q-Y	1			1						K-E	1	1		2
					Q-I			1	1						D-H			1	1
															T-D		1	1	2
															R-S	1			1
															K-Q	1			1
															E-Q			1	1
Total	13	12	19	44		7	9	11	27		4	6	6	16		18	22	27	67

*The amino acid differences in unconserved areas of the methanococcal AK sequences were tabulated to identify trends consistent with increased thermostability. The residue on the left is present in the mesophilic *M. voltae,* while the residue on the right is the corresponding residue in a thermophilic methanococci. Amino acid differences are grouped according to the polarity of the two amino acids. For the purpose of this table glycine was considered hydrophobic. A strong preference to increase the size of hydrophobic residues in the thermophilic AKs is evident.

†These hydrophobic-to-hydrophobic residue differences decrease hydrophobicity in the thermophilic AK.

Using the program MODELLER, complete atomic model structures were generated for two test cases, 1UKZ and 3ADK, based on alignments to distant homologues in similar conformational states (3AKY and 2AK3(A), respectively). The overall model structure of 1UKZ (Fig. 1) has an RMS deviation of 4.80 Å (C_α atoms) and 5.23 Å (all equivalent atoms) compared to the x-ray structure, with the greatest deviations located at the N terminus and within peripheral turns (Fig. 1). In fact, the first 10 residues of 1UKZ are not aligned to the 3AKY scaffold, and if removed from the model structure, the RMS deviations are significantly lower (2.76 Å C_α atoms and 3.46 Å all equivalent atoms). The positioning of core and active-site residues within the UK model are nearly identical to those in the x-ray structure (1.51 Å (C_α atoms) (see discussion of core residues below). The structural energetics of this model were analyzed and compared with the x-ray structure of 1UKZ. To do this, the SCEF was used to examine the sequence to structure compatibility by averaging over a window size of 20 residues. As can be seen in Fig. 2A, the model of 1UKZ based on the 3AKY scaffold is energetically comparable to its native state. The RMS deviations and energetics for 3ADK model structure are very similar to the results for the more distantly related 1UKZ homologue (data not shown). Based on these test cases, we expect an accuracy 1.5–2.0 Å for the core residues and 5–6 Å overall between our models and the final x-ray structures of the archaeal AKs.

The core structure of eukaryotic and eubacterial AKs and UKs has been substantially conserved throughout evolution.[19,21–23] The interactions of specific active-site residues with the inhibitor AP_5A, which mimics the ATP/AMP substrate, have been well studied and classified by their potential function in the *E. coli* AK structure. The corresponding active-site residues in the other AK and UK structures, based on the optimized energy function alignments, are reported in Table IV. Most interactions involve hydrogen bonds between the peptide backbone and the substrate, so the specificity of the residue is relatively unimportant except in terms of packing. Key side chains essential to the phosphate transfer are those of K21, T23, R97, R132, R138, and R149 (3ADK numbering). These core residues play an important role in in measuring the reliability of alignments and models generated for the methanococcal AKs (see later discussion).

Alignment and Model Generation of Archaeal AKs

The alignments of methanococcal AK sequences to 3ADK shown in Figure 3A depict the significant differences between the methanococcal AKs and other classes of AKs. The 27 residues underlined in Figure 3A are nearly invariant in all eubacteria and

TABLE III. Comparison of Alignments Produced by Using Self-Consistent Energy Function and Modified GAP Program for Each N-Kinase Sequence to Other Homologous Scaffolds. The Alignments Generated With SCEF Include Discrimination Score of Each Alignment

		SC-EF with constraints				Gap		
PDB code		D score	q score	RMS*	% ident	q score	RMS*	% ident
Target:	3ADK							
Scaffold	3ADK	11.51	1.00	0.0	100.0	1.00	0.0	100.0
	1UKZ	10.58	0.68	3.23	42.7	0.65	3.18	45.45
	3AKY	9.24	0.59	4.10	29.84	0.55	5.03	32.80
	2AK3	9.23	0.57	3.95	27.75	0.44	8.11	30.89
	1ANK	8.3	0.52	3.57	36.61	0.53	3.84	38.12
Target:	1UKZ							
Scaffold	1UKZ	12.28	1.00	0.0	100.0	1.00	0.0	100.0
	3ADK	11.13	0.66	2.74	43.37	0.63	3.18	44.97
	3AKY	10.26	0.67	3.28	24.34	0.66	3.40	26.32
	2AK3	9.88	0.58	3.69	25.91	0.55	4.49	28.50
	1ANK	9.24	0.65	2.09	29.19	0.62	2.74	31.32
Target:	1ANK							
Scaffold	1ANK	11.38	1.00	0.0	100.0	1.00	0.0	100.0
	2AK3	9.82	0.45	7.24	35.05	0.49	7.55	40.19
	3AKY	9.73	0.75	2.05	42.52	0.81	1.70	47.2
Target:	2AK3							
Scaffold	2AK3	10.10	1.00	0.0	100.0	1.00	0.0	100.0
	3AKY	7.35	0.47	6.77	41.23	0.45	7.17	42.92
	1ANK	6.21	0.45	7.88	39.23	0.44	7.55	40.95
Target:	3AKY							
Scaffold	3AKY	10.48	1.00	0.0	100.0	1.00	0.0	100.0
	2AK3	8.55	0.50	6.94	36.70	0.48	7.17	42.13
	1ANK	8.08	0.74	2.19	43.93	0.78	1.70	47.20

RMS deviation is based on Cα atoms only.

TABLE IV. Active-Site Residues for the AKs, UK, and Modeled Methanococcal Structures

Substance	Interaction type	1AKA	3AKY	2AK3	1UKZ	3ADK	Potential MJA
Adenylate of ATP	Backbone	K200	Q204	N203	R187	G177	R176
	Side chain	R119	R128	R116	R138	R128	R131
	Backbone	P201, V202	P205, P206	K204, I205	S188, V189	S178, V179	D177, F178
Adenylate of AMP	Side chain	T31	T19	S36	A47	T39	T91*
	Backbone	V59	V63	I64	V76	V67	L62
	Backbone	G85	G90	G89	G104	G94	T94
	Side chain	Q92	Q103	Q96	Q111	Q101	T97*
Sugar ATP	Backbone	Y131	Y142	Y136	NP	NP	NP
Sugar AMP	Backbone	K57	G61	K62	Q74	Q65	R60
PO_4 of ATP	Side chain	T15	T19	T20	T31	T23	T18
	Side chain	K13	K23	K18	K29	K21	H92, S93*
	Backbone	A8	G18	A13	G24	G16	V11
	Backbone	G14	G24	G19	G30	G22	T17†
	Side chain	R123	R132	R126	R142	R132	R132
	Side chain	R156	R165	R159	R148	R138	R138
	Side chain	R167	R176	R170	R159	R149	R140*
PO_4 of AMP	Side chain	R88	R93	R92	R107	R197	S93, T94, R156*
	Side chain	R36	R40	R41	R52	R44	R56*

*Assignment of these residues based on their approximate spatial positions.
†Potentially interacts with D90 and could interfere with binding.

eukaryotic AKs (over 25 sequences), and have been shown to be structurally or chemically essential by several crystal structures and mutational experiments.[45–48] Of these 27 residues, only 10 are found in the methanococcal AKs and 5 of these are from a conserved phosphate binding loop (P-loop) motif (residues 10–17 in the 3ADK). Several other residues or trends conserved in over 70% of AKs are also

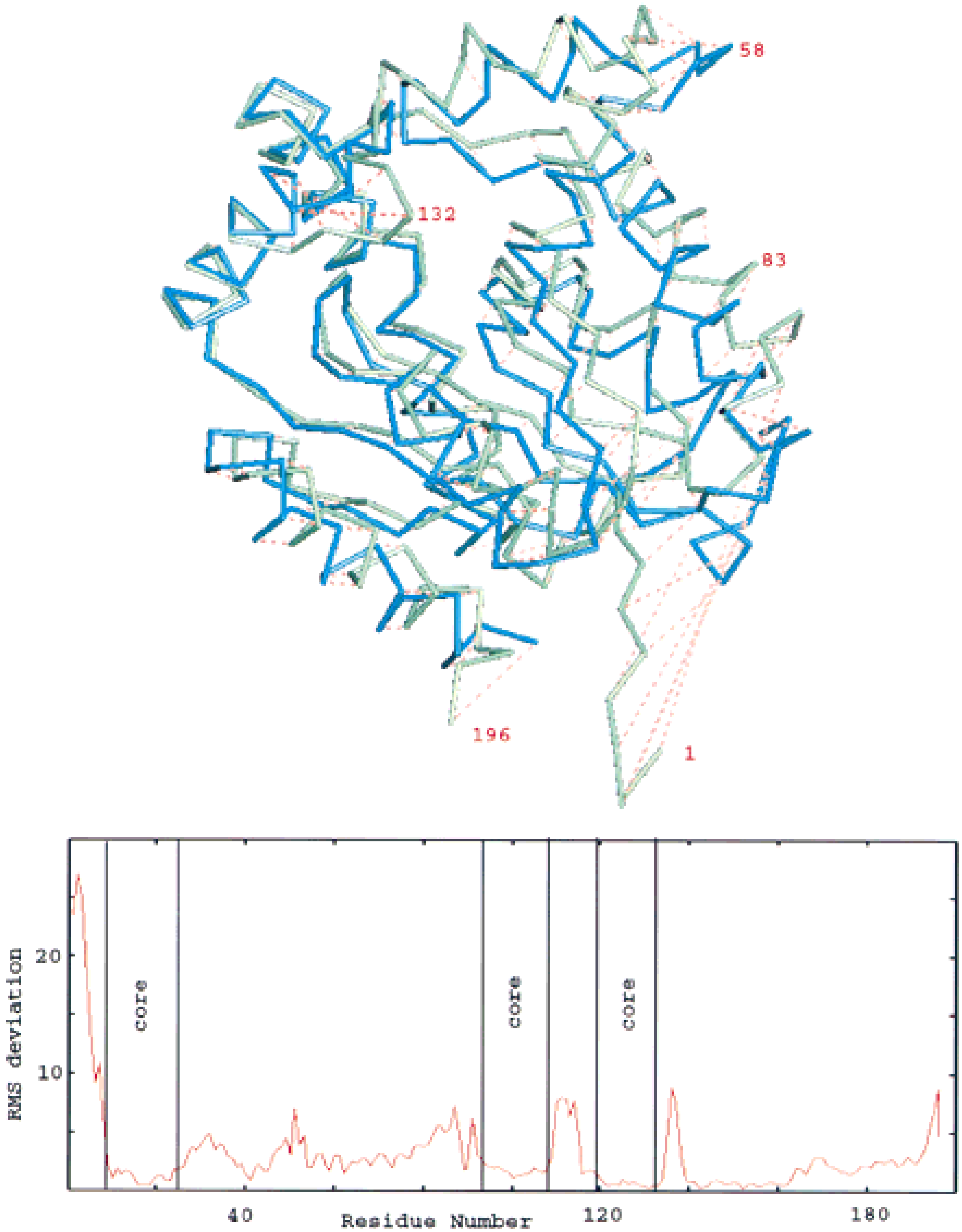


Fig. 1. Comparison of the model and x-ray structures (blue) of 1UKZ. **A:** The model structure (green) was generated by the MODELLER program by using the optimized energy function alignment of 1UKZ to 3AKY as a template. Deviations are indicated with the dashed red line. Without the N terminus segment which MODELLER built in, the overall RMS deviation (for $C\alpha$ atoms) is 2.76 Å. **B:** The RMS deviations per residue (based on $C\alpha$ atoms) for the two structures is approximately 1.5 Å for the core residues.

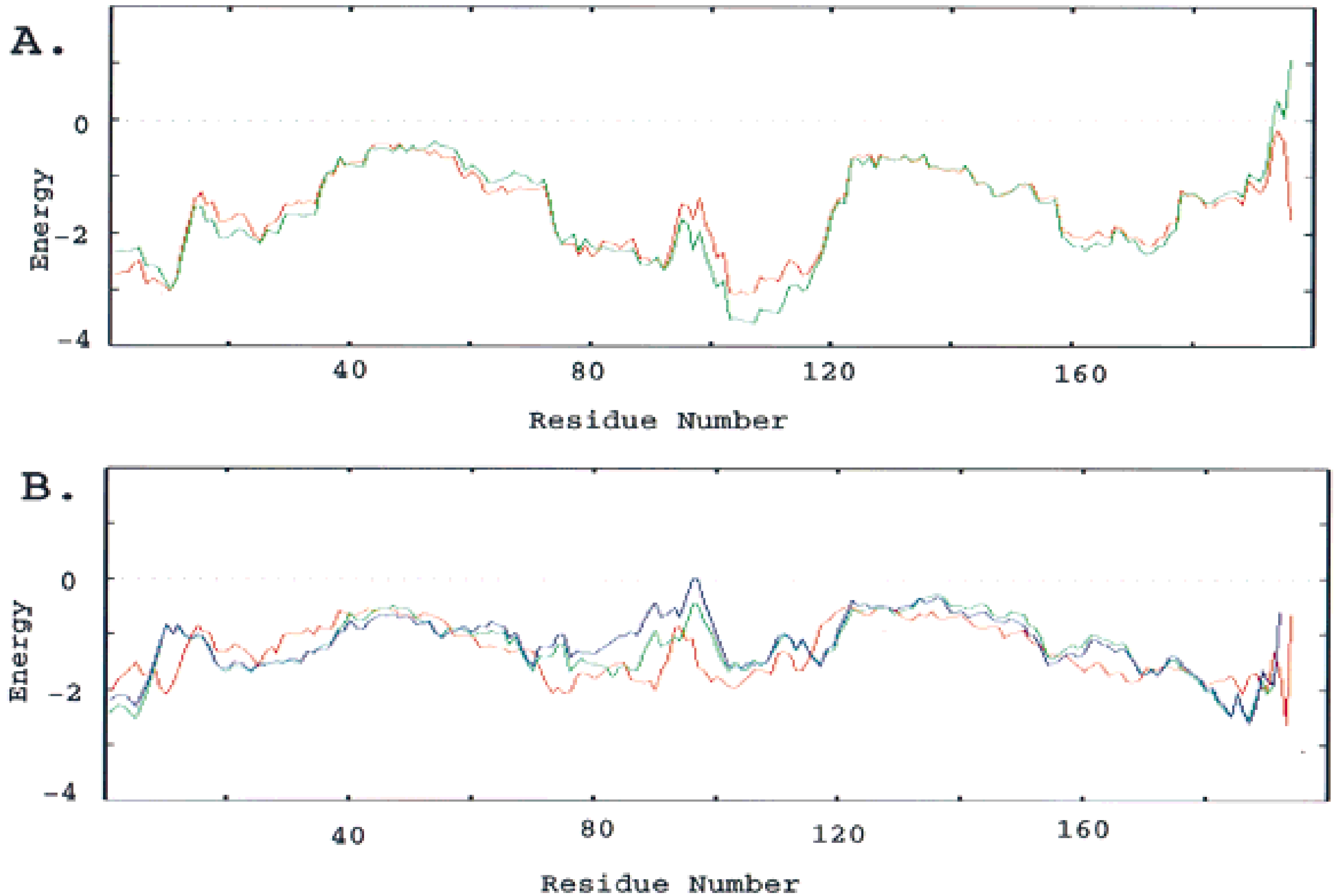
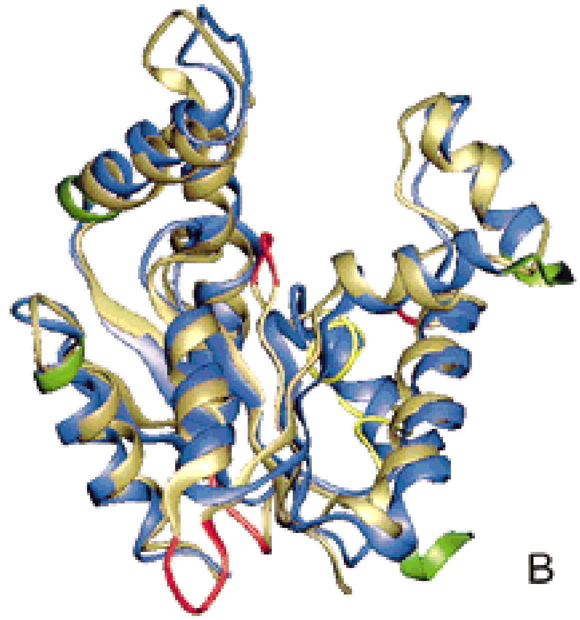


Fig. 2. A detailed examination of the sequence–structure compatibility based on the optimized energy function. Each point represents the energy per residue averaged over a window size of 20 residues. **A:** Energy traces of the 1UKZ x-ray (red) and modeled (green) structures. The modeled structure is based on the alignment of 1UKZ to 3AKY produced from the optimized energy function program. **B:** Energy traces of the 3ADK x-ray (red), modeled MJA (green), and modeled MJAf (blue) structures. Both modeled structures were based on the alignment of MJA to 3ADK scaffold.

```
                1/6                     21/26                       41/41
                 |          β1|           |          |          |
MVO         MKNKV VVVTGVPGVG STTSSQLAMD NLRKEGVNYK MVSFGSVMFE
MTH         MKNKL VVVTGVPGVG GTTITQKAME KLSEEGINYK MVNFGTVMFE
MIG         MKNKV VVVTGVPGVG GTTLTQKTIE KLKEEGIEYK MVNFGTVMFE
MJA         MKNKV VVIVGVPGVG STTVTNKAIE ELKKEGIEYK IVNFGTVMFE
Pig    MEEKLKKSKI IFVVGGPGSG KGTQCEKIVQ K-----YGYT HLSTGDLLRA
Eco               IILLGAPGAG KGT                       GDMLRA

                             61/66                 81/84
                  |          |          |          |      β3
MVO    VAK-----EE NLVSDRDQMR KMDPETQKRI QKMAGRKIAE MAKESPVAVD
MTH    VAQ-----EE NLVEDRDQMR KLDPDTQKRI QKLAGRKIAE MVKESPVVVD
MIG    VAK-----EE GLVEDRDQMR KLDPDTQKRI QKLAGRKIAE MAKESNVIVD
MJA    IAK-----EE GLVEHRDQLR KLPPEEQKRI QKLAGKKIAE MAKEFNIVVD
Pig    EVSSGSARGK MLSEIMEKGQ LVP--LETVL DMLRDAMVAK VDTSKGFLID
Eco          G      MD G  LV                             FLLD

                 101/101             121/118
       |          |          |      β4 |          |
MVO    THSTVSTPKG YLPGLPSWVL NELNPDLIIV VETT---GDE ILMRRMSDET
MTH    THSTIKTPKG YLPGLPVWVL NELNPDIIIV VETS---GDE ILIRRLNDET
MIG    THSTVKTPKG YLAGLPIWVL EELNPDIIVI VETS---SDE ILMRRLGDAT
MJA    THSTIKTPKG YLPGLPAWVL EELNPDIIVL VEAE---NDE ILMRRLKDET
Pig    ---GYPREVK QGEEFERKIG QPT---LLLY VDAGPETMTK RLLKRGETSG
Eco       GFPR    Q                YVLE FD         R   R

       141/141              161/161                     181/183
          |          |          |          |          |
MVO    RVRDLDTAST IEQHQFMNRC AAMSYGVLTG ATVKIVQNRN G--LLDQAVE
MTH    RNRDLETTAG IEEHQIMNRA AAMTYGVLTG ATVKIIQNKN N--LLDYAVE
MIG    RNRDIELTSD IDEHQFMNRC AAMAYGVLTG ATVKIIKNRD G--LLDKAVE
MJA    RQRDFESTED IGEHIFMNRC AAMTYAVLTG ATVKIIKNRD F--LLDKAVQ
Pig    RVDDNEETIK KRLETYYKAT EPVIAFYEKR GIVRKVNAEG SVDDVFSQVC
Eco    R DD       KRL       T  PLI  Y

          |
MVO    ELTNVLR
MTH    ELFQVLR
MIG    ELISVLK
MJA    ELIEVLK
Pig    THLDTLK
```

A



B

Fig. 3. **A:** The block alignment of the methanococcal AKs to the 3ADK scaffold generated by using the energy functions with experimental constraints. The residues constrained to be in β sheets are indicated in bold print. The residues that are conserved in eukaryotic and eubacterial AKs are indicated in the partial sequence of the *E. coli* AK (Eco), where only residues occurring in over 75% of known AKs are included. Underlined residues are invariant in eukaryotic and eubacterial AKs. **B:** Comparison of the Cα trace of the MVOf model (gold) and the 3ADK scaffold (blue), insertions (red), deletions (green), and the polyproline region have been highlighted (yellow).

absent in the methanococcal AKs (Fig. 3A). How the function of these previously conserved residues are replaced in the methanococcal AKs has been partially examined by 3-D protein comparative modeling.

The quality of the alignments based on discrimination scores for the archaeal AK sequences to the various scaffolds is given in Table V. The discrimination scores are lower than the eubacterial and eukaryotic homologues (see Table III) but consistent for distant homologues with less than 21% sequence identity.[27] Using the alignments to 2AK3(A), 1UKZ, and 3ADK, complete models were constructed for the four archaeal AKs, using the program MODELLER. The resulting structures are similar in the core region, and we will only discuss the model based on 3ADK, since it has the highest percentage identity. We should point out that energetically the sequences are more compatible to the 1UKZ structure. A more detailed examination of the sequence–structure compatability is obtained by analyzing the average energy per residue. The energy patterns of interaction are very similar between the modeled methanococcal and 3ADK structures with an exception in the region between methanococcal residues 95–110. Figure 2B shows the energy traces for MJA and 3ADK. The decrease in energy stability in this region could be caused by the alignment of three prolines (residues 93, 103, and 106 of MJA sequence) to a right-handed α helix. These three prolines should be relatively unstable, since polyprolines tend to form left-handed helices or turns.

In an attempt to more accurately model this polyproline region, the Protein Data Bank was searched for structures containing sequence similarity to this region. Residues 68–82 of the cytochrome c structure from *Saccaromyces cerevisiae* had the highest percentage identity with residues 95–105 in the methanococcal AKs. A full atomic model of residues 95–109 from the methanococcal MJA sequence was built by MODELLER using the corresponding structural fragment of cytochrome c as a template. Part of this new fragment was then grafted back into each of the model structures, and the altered models were then minimized (see methods). While energetic stability of the model segment improves, its interaction with the whole model is decreased due to poor packing, resulting in a higher total energy (Fig. 2B). Several other methods to correct this region were tried, such as insertion of a random coil structure or realignment of the sequences in this region, but these methods also failed to increase stability (data

TABLE V. Discrimination Score Evaluated With Self-Consistent Energy Functions and Percentage Identity for Alignments of the Methanococcal AKs to Five Scaffold Proteins*

	Scaffold									
	1ANK		2AK3		3AKY		1UKZ		3ADK	
Target	D score	% ident	D score	% ident	D score	% ident	D score	% ident	D score	% ident
3ADKZ	8.30	36.61	9.23	27.75	9.24	29.84	10.58	42.27	11.51	100.00
MJA	4.77	11.64	6.06	11.46	4.55	7.33	5.82	14.06	5.45	15.10
MIG	4.76	13.76	5.26	10.42	4.96	7.85	5.44	13.54	4.70	14.58
MTH	4.67	13.23	5.02	9.90	4.72	8.90	4.99	11.98	4.63	14.06
MVO	4.35	11.11	4.33	9.90	4.50	8.38	4.28	11.98	3.88	13.54

*A protein's discrimination score is a measure of energy differences between the predicted structure and the average in an ensemble of collapsed, misfolded states over the standard deviation of these misfolded states (see Eq. 2).

not shown). Better methods for predicting the structure or stability of multiple proline structures need to be developed, and the lack of information on these types of structures make it difficult to assess the reliability of current predictive methods. Due to uncertainty in the structure of this region, models with and without the modeled proline loop were used to investigate potential active-site residues and thermostability. Figure 3B shows the final structure of the MVOf model compared to the known structure of 3ADK (Cα trace only), the positions of insertions, deletions, and the polyproline region have been highlighted.

Besides demonstrating energetic characteristic comparable to known AK structures, further inspection of the model structures also supports their reliability. Structurally, the model AKs have few hydrophobic residues exposed to the solvent or hydrophilic residues buried in the core. Additional structural information from circular dichroism (CD) and tryptophan fluorescence spectrophotometric studies also support the model structures (unpublished data). CD spectrum of the methanococcal AKs have negative minima at approximately 222 and 208 nm, which is indicative of a high content of α-helical structure, and are very similar to spectrums of other AKs.[49,47] Fluorescence data shows Trp108 has a high quantum yield and a red-shifted λ_{max} (338 nm), indicating that Trp108 is solvent-exposed, and little quenching by the protein occurs (data not shown). Positioning of Trp108 in the model structures is consistent with these observations. Furthermore, interactions similar to those seen in known AK structures and consistent with the proper functioning of an AK are present in the archaeal AK models (see below).

Identification of Potential Active Site Interactions

An essential feature of the AK active site is the ability to bind and orient the multiple phosphates found on the substrates. This is accomplished through multiple interactions, including several positively charged residues, such as 3ADK residues K21, R97 and R132, R138, and R149, which make up a giant "anion hole."[45,48,50,51] The conservation and importence of these interactions are further discussed in Table IV. The surface charge potential of the 3ADK and MVO AKs (is shown in Fig. 4). Both the 3ADK and the archaeal AK models maintain a positive surface potential throughout much of the central active site cleft; however, the methanococcal AKs derive this charge potential from different sources. While residues corresponding to active-site residue R132 and R138 (3ADK numbering) are present in the methanococcal sequences (R132 and R138, respectively), residues K21, R97, and R149 are not found. However, examination of the models and alignments identifies residues potentially involved in substrate binding (Fig. 5 and last column of Table IV). In the model AKs, H92 is in the vicinity of spaces left unoccupied by the missing 3ADK residue K21. Protonation of the histidine residue would place a positively charged residue in a location able to participate in phosphate binding and transfer. The presence of several negative charges on the substrate should substantially shift the pK_a of H92 to allow protonation at neutral pHs. Even in the deprotonated state, H92 could stabilize substrate binding through hydrogen bonding with a phosphate group or interaction with an essential Mg^{2+} ion.

The 3ADK residue R97 is part of a second strongly conserved phosphate binding loop and is responsible for binding the AMP phosphate.[45,48] In the methanococcal AKs, this region is significantly altered, with the modeled loop (residue 91–94) substantially extended and contains several residues able to hydrogen-bond the substrate phosphates, such as H92, S93, and T94 (see Table IV). No reasonable replacement for porcine residue R149 is evident in the sequence alignment. However, the model structures show methanococcal residue R140 occupying the same approximate location in space and should be able to function similarly. Several other functional important active-site residues are strictly conserved between the methanococcal AKs and 3ADK, such

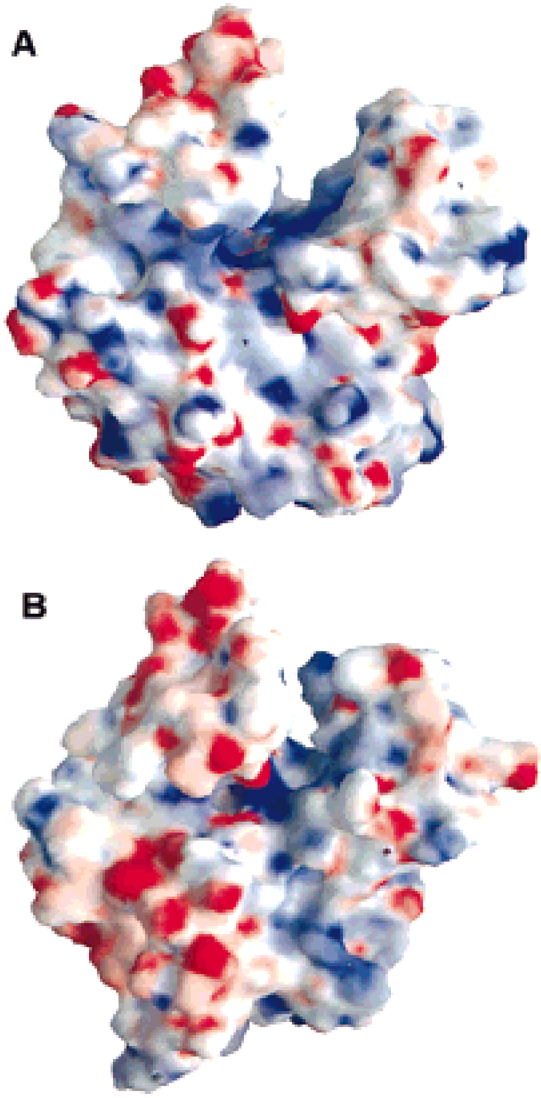


Fig. 4. Comparison of the surface electrostatic potentials. **Top:** 3ADK. **Bottom:** The final structure of MVOf. The active site pockets of the two structures both contain a high positive charge (blue) despite lack of sequence identity in this region. This positive charge is essential in phosphate binding.

as 3ADK residues T23, D93, and multiple glycines. The function of these residues is also believed to be conserved in the methanococcal AKs. While many hypotheses about other substrate binding residues can be made, the open conformation and lack of substrate in the protein models make further identifications unreliable. Despite the significant divergence of the methanococcal AK sequences a plausible active site cleft is seen in our models. Figure 5 displays the inhibitor AP5A bound to the MTH active site cleft, and a comparison of the important residues in this region with the 3ADK AK structure. Alignments of the methanococcal AKs to the five reference scaffolds using other standard methods (e.g., GCG Gap and multiple sequence alignments) were also generated and modeled. However, these structures gave unsatisfactory tertiary models when trying to locate possible active-site residues (data not shown).

While elements consistent with enzymatic function are present in the methanococcal AK models, a few potentially damaging interactions can be seen in the model structures. Residues 91–93 are not aligned and are built in as a loop. Even though these residues are identical in each sequence, MODELLER places the backbone and side chains in different conformations in each protein, which is a common weakness of the method.[34] In the final MJA and MIG structures these residues are placed in the middle of the active site cleft and would potentially block substrate binding. Repositioning of this loop below or to the side of the active-site cleft, similar to the MTH or MVO models, may prevent interference with substrate binding. Other potential problems include an interaction between Thr17 and Asp90 in the methanococcal models. In all eubacterial and eukaryotic AKs an invariant Gly residue is at the position corresponding to methanococcal residue Thr17. Any residue other than Gly at this position would potentially collide with substrate or interact with the Asp residue and prevent Mg^{2+} and substrate binding.[52] In the model AKs, Thr17 can be clearly seen hydrogen bonding to Asp90 (Fig. 5). This Thr is conserved in the seven archaeal AKs for which nucleotide or peptide sequence data is available, (data not shown), and may be essential for maintaining the position of the highly flexible active-site P loop. How this interaction with Asp 90 effects Mg^{2+} or substrate binding is unclear.

Potential Structural Features Related to Thermal Stabilizing

Model proteins also make it possible to investigate the residues believed important for increased thermal stability and support a structural role. Previous studies have shown that salt bridges linking protein subunits or regions distantly separate in the amino acid sequence, may have a significant effect on protein thermal stability.[8,6,53] However, salt bridges between neighboring residues or those that are highly solvent exposed usually contribute very little to protein stability.[54] Potential ionic interactions in the methanococcal AKs were identified by finding the hydrogen-bonding donor and acceptor atoms of charged functional groups seperated by less than 4.0 Å and having proper orientation. A potential internal salt bridge exists between K96 and E112, and while it could have a significant effect on the stability of the thermophilic AKs, this salt bridge is located in the problematic polyproline region of the protein models, and its prediction is very tentative.

A preference for the placement of particular types of amino acid differences involving hydrophobic resi-

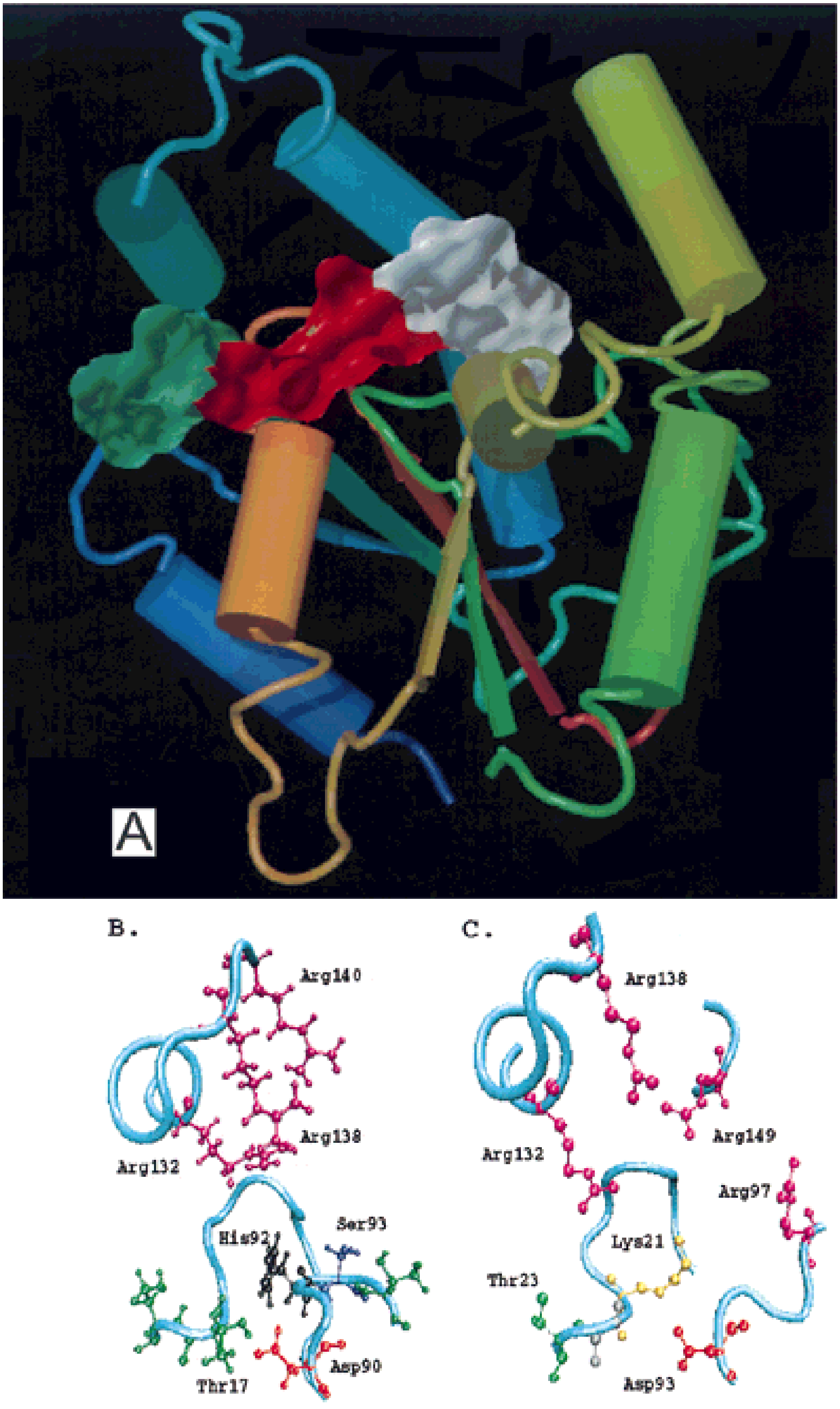


Fig. 5. **A:** The bisubstrate analogue AP5A bound to the active site cleft of the MTH AK model. *Note:* model is in a partially closed state. Residues predicted to be important in substrate binding in modeled structure MTHf (**B**) compared to the known porcine 3ADK (**C**) active site. Small variations in side-chain positioning occur between each of the methanococcal models.

dues can be seen between the mesophilic and thermophilic AK models. A majority of the amino acids exchanges involving only aliphatic amino acids, such as Val-Ile, Leu-Ile, and Ala-Val, occur in close proximity to each other in the protein core β sheet region. Hydrophobic exchanges involving a Met, however, occur nearer the protein's periphery. The large increase in the average hydrophobicity, through aliphatic residue, of the thermophilic methanococcal AKs, together with their location within our models, strongly support a suggestion that hydrophobic branched chain amino acids play a major structural role in conferring thermal stability to the methanococcal AKs. This could be accomplished by altering water accessibility to the protein core, while limiting internal cavity formation and increasing van der Waal interactions. Mutational and crystallization studies on the AK from *M. voltae* and *M. jannaschii* are now underway in order to address the many hypotheses generated for this predictive study.

ACKNOWLEDGMENTS

We thank Dorina Kosztin, Tom Bishop, Bill Humphrey, and Andrew Dalke for their helpful contributions with the molecular dynamics simulations and the VMD visualization tools,[55] and Michael Glaser and George Phillips for comments about the manuscript and our model structures. This work has been supported by grants from the National Institute of Health grant (2 R01 GM44557) and the Roy J. Carver Charitable Trust to Peter Wolynes and Zan Luthey-Schulten, and the Department of Energy grant (DE-FG02-84GR1324) to Jordan Konisky. Computations were carried out in part at the Resources for Concurrent Biological Computing the University of Illinois, Urbana/Champaign, funded by the National Institutes of Health (grant P41RR05969).

REFERENCES

1. Adams, M. Enzymes from microorganisms in extreme environments. Chemical & Engineering News, 18:32–42, 1995.
2. Adams, M., Perler, F., Kelly, R. Extremozymes: Expanding the limits of biocatalysis. Biotechnology 13:662–668, 1995.
3. Britton, K., Baker, P., Borges, K., Engel, P., Pasquo, A., Rice, D., Robb, F., Scandurra, R., Stillman, T., Yip, K. Insights into thermal stability from a comparison of glutamate dehydrogenases from *Pyrococcus furiosus* and *Thermococcus litoralis.* FEBS 229:688–693, 1995.
4. Zuber, H. Temperature adaptation of lactate dehydrogenase: Structural, functional and genetic aspects. Biophys. Chem. 29:171–179, 1988.
5. Argos, P. Engineering protein thermal stability: Sequence statistics point to residue substitutions in alpha helices. J. Mol. Biol. 206:397–406, 1989.
6. Kelly, C., Nishiyama, M., Ohnishi, Y., Beppu, T., Birktoft, J. Determinants of protein thermostability observed in the 1.9a crystal structure of malate dehydrogenase from the thermophilic bacterium *Thermus flavus.* Biochemistry 32: 3913–3922, 1993.
7. Walker, J., Wonacott, A., Harris, J.I. Heat stability of a tetrameric enzyme, *d*-glyceraldehyde-3-phosphate dehydrogenase. Eur. J. Biochem. 108:581–586, 1980.
8. Perutz, M., Raidt, H. Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin 42. Nature 255:256–259, 1975.
9. Serrano, L., Sancho, J., Hirshberg, M., Fersht, A.R. α-Helix stability in proteins. J. Mol. Biol. 227:544–559, 1992.
10. Nicholson, H., Becktel, W.J., Matthews, B.W. Enhanced protein thermostability from designed mutations that interact with α-helix dipoles. Nature 336:651–655, 1988.
11. Matthews, B., Nicholson, H., Becktel, W. Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. Proc. Natl. Acad. Sci. USA 84: 6663–6667, 1987.
12. Jaenicke, R. Protein stability and molecular adaptation to extreme conditions. Eur. J. Biochem. 202:715–728, 1991.
13. Jaenicke, R., Zavodszky, P. Proteins under extreme physical conditions. FEBS 268:344–349, 1990.
14. Matthews, B. Structural and genetic analysis of protein stability. Annu. Rev. Biochem. 62:139–160, 1993.
15. Schlapfer, B., Zuber, H. Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase (gap operon) from mesophilic *Bacillus megaterium:* Comparison with corresponding sequences from thermophilic *Bacillus stearothermophilus.* Gene 122:53–62, 1992.
16. Bohm, G., Jaenicke, R. Relevance of sequence statistics for the properties of extremophilic proteins. Int. J. Peptide Prot. Res. 43:97–106, 1994.
17. Rusnak, P., Haney, P., Konisky, J. The adenylate kinases from a mesophilic and three thermophilic methanogenic members of the archaea. J. Bacteriol. 177:2977–2981, 1994.
18. Ferber, D.M., Haney, P.J., Berk, H., Lynn, D., Konisky, J. The adenylate kinase genes of *Methanococcus jannaschii* and *Mehtanococcus igneus:* Structural and phylogenetic analysis defines a new family of adenylate kinases. Genes (in press).
19. Schulz, G.E. Structural and functional relationships in the adenylate kinase family. Cold Spring Harbor Symp. Quant. Biol. LII:429–439, 1987.
20. Gerstein, M., Schulz, G.E., Chothia, C. Domain closure in adenylate kinase. Joints on either side of two helices close like neighboring fingers. J. Mol. Biol. 229:494–501, 1993.
21. Berry, M., Meador, B., Bilderback, T., Liang, P., Glaser, M., Phillips, G. The closed conformation of a highly flexible protein: The structure of *E. coli* adenylate kinase with bound amp and amppnp. Proteins 19:183–198, 1994.
22. Mullerdieckmann, H.J., Schulz, G.E. The structure of uridylate kinase with its substrates, showing the transition state geometry. J. Mol. Biol. 236:361–367, 1994.
23. Mullerdieckmann, H.J., Schulz, G.E. Substrate specificity and assembly of time catalytic center derived from two structures of ligated uridylate kinase. J. Mol. Biol. 246:522–530, 1995.
24. Scheffzek, K., Kliche, W., Wiesmuller, L., Reinstein, J. Crystal structure of the complex of ump/cmp kinase from *Dictyostelium discoideum* and the bisubstrate inhibitor p1-(5′-adenosyl) p5-(5′-uridyl) pentaphosphate (up5a) and mg2+ at 2.2 angstroms: Implications for water-mediated specificity. Biochemistry 35:9716–9727, 1996.
25. Kath, T., Schmid, R., Schafer, G. Identification, cloning, and expression of the gene for the adenylate kinase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius.* Arch. Biochem. Biophys. 307:405–410, 1993.
26. Bonisch, H., Backmann, J., Kath, T., Naumann, D., Schafer, G. Adenylate kinase from *Sulfolobus acidocaldarius:* Expression in *Escherchia coli* and characterization by Fourier transform infrared spectroscopy. Arch. Biochem. Biophys. 333:75–84, 1996.
27. Koretke, K.K., Luthey-Schulten, Z., Wolynes, P.G. Self-consistently optimized statistical mechanical energy functions for sequence structure alignment. Prot. Sci. 6:1043–1059, 1996.
28. Goldstein, R., Luthey-Schulten, Z., Wolynes, P. Protein tertiary structure recognition using optimized hamiltonians with local interactions. Proc. Natl. Acad. Sci. USA, 89:9029–9033, 1992.
29. Goldstein, R., Luthey-Schulten, Z., Wolynes, P. A bayesian approach to sequence structure alignment algorithms for protein structure recognition. In "Proceedings of the 27th

Hawaii International Conference on System Sciences." Los Alamitoos, CA: IEEE Computer Society Press, 1994:306–315.
30. Abola, E., Bernstein, F., Bryant, S., Koetzle, T., Weng, J. Protein data bank. In "Crystallographic Databases: Information Content, Software Systems, Scientific Applications." Allen, F.H., Bergerhoff, G., Sievers, R. (eds.). Bonn: Data Commissions of the International Union of Crystallography, 1987:107–132.
31. Bernstein, F., Koetzle, T., Williams, G., Meyer Jr., E., Brice, M., Rodgers, J., Kennard, O., Shimanouchi, T., Tasumi, M. The protein data bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
32. Genetics Computer Group, Madison, Wisconsin. "Program Manual for the Wisconsin Package," 8th edit. 1994.
33. Berry, M.B., Meador, B., Bilderback, T., Liang, P., Glaser, M., Phillips, G.N. The closed conformation of a highly flexible protein: The structure of *E. coli* adenylate kinase with bound amp and amppnp. Proteins 19:183–198, 1994.
34. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., Karplus, M. Evaluation of comparative protein modeling by modeller. Protein 23:318–326, 1995.
35. Berghuis, A.M., Guillemette, J.G., Smith, M., Brayer, G.D. Mutation of tyrosine-67 in cytochrome c significantly alters the local heme environment. J. Mol. Biol. 235:1326–1341, 1994.
36. Bunger, A.T. X-PLOR. New Haven, CT: The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, 1988.
37. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. Charmm: A program for macromolecular energy, minimization and dynamics calculations. J. Comput. Chem. 4:187, 1983.
38. Tanford, C. Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. J. Am. Chem. Soc. 84:4240–4247, 1964.
39. Bigelow, C. On the average hydrophobicity of proteins and the relation between it and protein structure. J. Theoret. Biol. 16:187–211, 1967.
40. Ikai, A. Thermostability and aliphatic index of globular proteins. J. Biochem. 88:1895–1898, 1980.
41. Karplus, P., Schulz, G. Prediction of chain flexibility in proteins: A tool for the selection of peptide antigens. Naturwissenschaften 72:212–213, 1985.
42. Sharp, K., Nicholls, A., Friedman, R., Honig, B. Extracting hydrophobic free energies from experimental data: Relationship to protein folding and theoretical models. Biochemistry 30:9686–9697, 1991.
43. Ahern, T., Klibanov, A. The mechanism of irreversible enzyme inactivation at 100°C. Science 228:1280–1284, 1985.
44. Mrabet, N.T., Den Broeck, A.V., Den Brande, I.V., Stanssens, P., Laroche, Y., Van Tilbeurgh, H., Rey, F., Quax, J.J., Laster, I., Maeyer, M.D., Wodak, S.J. Arginine residues as stabilizing elements in proteins. Biochemistry 31:2239–2253, 1992.
45. Reinstein, J., Schlichting, I., Wittinghofer, A. Structurally and catalytically important residues in the phosphate binding loop of adenylate kinase of *Escherichia coli*. Biochemistry 29:7451–7459, 1990.
46. Tain, G., Yan, H., Jiang, R., Kishi, F., Nakazama, A., Tsai, M.-D. Mechanism of adenylate kinase: Are the essential lysines essential? Biochemistry 29:4296–4304, 1990.
47. Rose, T., Glaser, P., Surewicz, W., Mantsch, H., Reinstein, J., Blay, K.L., Gilles, A., Barzu, O. Structural and functional consequences of amino acid substitutions in the second conserved loop of *Escherichia coli* adenylate kinase. J. Biol. Chem. 266:23654–23659, 1991.
48. Dahnke, T., Shi, Z., Yan, H., Jiang, R., Tsai, M.-D. Mechanism of adenylate kinase. structural and functional roles of the conserved arginine-97 and arginine-132. Biochemistry 31:6318–6328, 1992.
49. Monnot, M., Gillies, A., Girons, I., Michelson, S., Barzu, O., Fermandjian, S. Circular dichroism investigation of *Escherichia coli* adenylate kinase. J. Biol. Chem. 262:2502–2506, 1987.
50. Dreusicke, D., Schulz, G. The glycine-rich loop of adenylate kinase forms a giant anion hole. FEBS 208:301–304, 1986.
51. Muller, C.W., Schulz, G.E. Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor ap5a refined at 1.9Å resolution. J. Mol. Biol. 224:159–177, 1992.
52. Dreusicke, D., Karplus, P.A., Schulz, G.E. Refined structure of porcine cytosolic adenylate kinase at 2.1 Å resolution. J. Mol. Biol. 199:359–371, 1988.
53. Anderson, D., Becktel, W., Dahlquist, F. pH-induced denaturation of proteins: A single salt bridge contributes 3–5 kcal/mol to the free energy of folding of t4 lysozyme. Biochemistry 29:2403–2408, 1990.
54. Dao-pin, S., Sauer, U., Nicholson, H., Matthews, B. Contributions of engineered surface salt bridges to the stability of t4 lysozyme determined by directed mutagenesis. Biochemistry 30:7142–7153, 1991.
55. Humphrey, W.F., Dalke, A., Schulten, K. VMD: Visual molecular dynamics. J. Mol. Graphics 14:33–38, 1996.