The Refined Structure of the Complex between Adenylate Kinase from Beef Heart Mitochondrial Matrix and Its Substrate AMP at 1·85 Å Resolution

Kay Diederichs and Georg E. Schulz

Institut für Organische Chemie und Biochemie
Albertstrasse 21, D-7800 Freiburg i. Br., F.R.G.

The crystal structure of the complex between adenylate kinase from bovine mitochondrial matrix and its substrate AMP has been refined at 1·85 Å resolution (1 Å = 0·1 nm). Based on 42,519 independent reflections of better than 10 Å resolution, a final R-factor of 18·9% was obtained with a model obeying standard geometry within 0·016 Å in bond lengths and 3·2° in bond angles. There are two enzyme:substrate complexes in the asymmetric unit, each consisting of 226 amino acid residues, one AMP and one sulfate ion. A superposition of the two full-length polypeptides revealed deviations that can be described as small relative movements of three domains. Best superpositions of individual domains yielded a residual overall root-mean-square deviation of 0·3 Å for the backbone atoms and 0·5 Å for the side-chains. The final model contains 381 solvent molecules in the asymmetric unit, $2 \times 72 = 144$ of which occupy corresponding positions in both complexes.

1. Introduction

Adenylate kinase from beef heart mitochondrial matrix (AK3†, GTP:AMP-phosphotransferase, EC 2.7.4.10) is a monomeric enzyme that catalyzes the reaction $Mg^{2+}GTP + AMP \rightleftharpoons Mg^{2+}GDP + ADP$ (Tomasselli & Noda, 1979; Tomasselli *et al.*, 1979). The enzyme is essential for the biosynthesis of ADP from AMP. The peptide sequence has been reported by Tomasselli *et al.* (1986). The N terminus is heterogeneous, 33% and 18% of all molecules are shortened by one and two amino acid residues, respectively (Wieland *et al.*, 1984). A cDNA-derived sequence has also been published (Yamada *et al.*, 1989). The sequence homology to other members of the adenylate kinase family ranges between 24% and 37% identical residues (Schulz *et al.*, 1986).

Two medium-resolution crystal structures of adenylate kinases (AKyst, AKeco) with bound two-substrate-mimicking inhibitor P^1,P^5-bis(adenosine-5'-)pentaphosphate (Ap₅A) and one high-resolution crystal structure of an adenylate kinase (AK1)

without substrate have been reported (Egner *et al.*, 1987; Mueller & Schulz, 1988; Dreusicke *et al.*, 1988). Since AK3 was crystallized with only one of the substrates (AMP) bound, its crystal structure allows us to distinguish between the tri- and mono-phosphate position of bound Ap₅A. After reporting the initial X-ray analysis (Diederichs & Schulz, 1990), we now describe the final refined structure at a resolution of 1·85 Å (1 Å = 0·1 nm).

2. Methods

The crystal space group is $P2_12_12_1$ with 2 enzyme:substrate complexes per asymmetric unit. The unit cell dimensions are 49·83 Å × 67·99 Å × 155·41 Å as derived from the analysis of rotation photographs. A native data set at 1·85 Å resolution (41,795 reflections) had been collected using the synchrotron source at the EMBL outstation in Hamburg (Diederichs & Schulz, 1990). After data reduction and merging with diffractometer data measured out to 3 Å resolution, a 95% complete data set at 1·85 Å resolution containing 42,869 reflections was produced (Diederichs, 1990) and used in all refinements. The refinement calculations were done with the molecular dynamics program X-PLOR (Bruenger *et al.*, 1987), 1 cycle of which took up to 4 days of CPU time on our μVAX-II. The first 6 rounds of the refinement resulted in a generally correct model with an R-factor of 28·4%, that has been reported previously (Diederichs & Schulz, 1990).

The results of the subsequent refinement are given in Table 1. In round 7 an additional alanine, Ala8, was

---

Table 1

Structural refinement by simulated annealing in the resolution range 10 to 1·85 Å

Round[a]	R-factors (%)			
	min-1[b]	md-300[c]	min-2[d]	min-B[e]
7	28·5	28·8	26·9	25·8
8	25·9	27·3	25·1	24·3
9	—	—	24·1	23·4
10	23·6	25·3	23·2	22·6
11	—	—	22·2	21·7
12	21·7	23·3	21·1	20·7
13	21·9	23·5	21·0	20·7
14	—	—	20·8	20·6
15	21·2	22·5	20·2	20·0
16	20·3	21·9	20·0	20·0
17	—	—	20·0	18·8
18	19·2	19·9	19·0	18·9[f]

[a] The model was rebuilt after each round. At the end of each round the model geometry was always near ideality with r.m.s. deviations of 0·016 Å in bond lengths and 3·2° in bond angles.

[b] First energy minimization (60 cycles), no NCS-restraint.

[c] Molecular dynamics at 300 K for 0·25 ps, always NCS-restrained using a weight of 6 kcal/(mol·Å^2).

[d] Second energy minimization (60 cycles), no NCS-restraint.

[e] Temperature factor refinement (10 cycles), NCS-restrained for rounds 7 through 16.

[f] In round 18 the R-factor increased slightly due to deletion of 51 solvent molecules with low occupancies.

inserted before Ala9. It improved the fit of the preceding residues (Leu-Leu-Arg) to the electron density considerably (see Fig. 2). This correction was confirmed by repeating the N-terminal sequence determination of the protein (E. Schiltz, personal communication). Moreover, the additional alanine was also found in the cDNA sequence published recently (Yamada et al., 1989). To stay consistent with previous publications, we decided to start sequence numbering with 0 instead of 1. This accounts somewhat for the inhomogeneous N terminus and it affects only the 8 N-terminal residues.

Most of the side-chains were positioned correctly after round 7. Minor errors in side-chain orientations, especially for Thr, Val, Leu, Asn, Gln, and His were corrected at later stages. For Asn, Gln and His and Thr we tried to establish the orientation that maximizes hydrogen bonding and that minimizes temperature factor discrepancies between the respective pairs of C, N and O atoms. The orientation of Leu and Val side-chains became unambiguous for temperature factors below about 30 Å^2.

Solvent molecules were usually identified in difference Fourier maps. They were added to the model if they were located at hydrogen bonding distance to the protein. If in later rounds the electron density of a solvent molecule dropped below a given threshold of about 80% of the standard deviation of the $(2F_{obs} - F_{calc})$ electron density map, it was removed from the model. When the R-factor crossed the 20% level at round 17 (Table 1), the temperature factors of NCS-related atoms were no longer restrained, causing another drop of 1·2%. The largest temperature factor differences of NCS-related atoms were observed in domain AMPbd (see Fig. 7).

As can be seen in Table 1, the refinement with X-PLOR proceeded smoothly. We did not apply molecular dynamics runs at 2000 K as proposed by Bruenger (1988) because the initial model was already generally correct. Instead, we kept the dynamics temperature at 300 K.

Even at this temperature, the dynamics increased the radius of convergence of the refinement appreciably, because in all rounds the energy minimization following dynamics (min-2) yielded lower R-factors than the one preceding dynamics (min-1). The final R-factor was 18·9% for 42,519 reflections in the resolution range 10 to 1·85 Å. Here, one should keep in mind that the refinement included *all* measured reflections as observables. The R-factor for the subset of 29,074 reflections with intensities larger than 3 standard deviations is as low as 16·1%.

An error estimate for the atomic co-ordinates can be obtained from a Luzzati plot (Luzzati, 1952), which assumes that the disagreement between observed and calculated structure factor amplitudes is merely caused by atomic displacements. From Fig. 1, an upper error limit of 0·25 Å can be derived for the co-ordinate error. In this Figure we also included the R-factor for the subset of higher intensities, which lowers the error estimate.

Another error estimate is available from equivalent solvent positions in the 2 complexes. Two solvent molecules were considered as equivalent or NCS-related, if the distance between the 1st solvent molecule and the NCS-transformed position of the 2nd solvent molecule is less than 1 Å. Using this criterion and the NCS-relations separately for the domains MAIN, INSERT and AMPbd (Diederichs & Schulz, 1990), we observed 72 pairs of equivalent solvent molecules. The resulting distances were generally below 0·5 Å and the average distance was 0·3 Å although these pairs were not NCS-restrained during the refinement. This value agrees with the r.m.s. main-chain atom distance between complexes no. 1 and no. 2 after
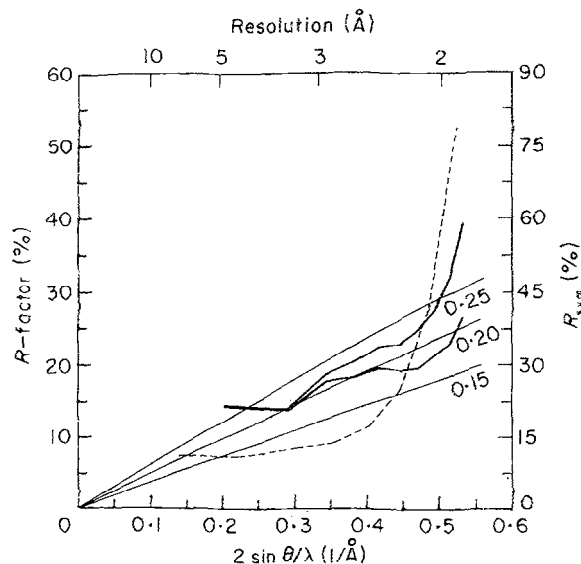

Figure 1. Luzzati plot for the final refined crystal structure of the AK3:AMP complex. The lines of constant co-ordinate error are labeled with the respective value. The upper continuous line (thick) represents R-factors of the final refinement for all 42,519 reflections in the range 10 to 1·85 Å resolution yielding an overall R-factor of 18·9%. The lower continuous line (thick) refers to R-factors calculated without the weak reflections (i.e. intensities smaller than 3 standard deviations) yielding an overall R-factor of 16·1%. The broken line gives the R_{sym} values derived from film data processing (scale at right-hand side). In contrast to the R-factors, the R_{sym} values are calculated on intensities and can therefore not be compared directly. The R_{sym} values show, however, that the data quality decreases appreciably beyond 2 Å resolution.
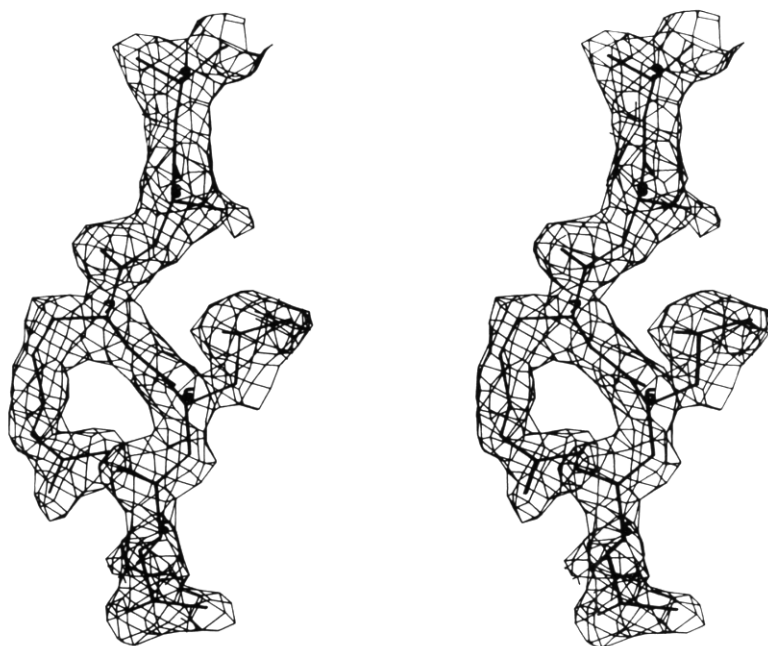
Figure 2. Electron density contour of a $(2F_{obs} - F_{calc})$ map together with the final model. Shown is segment Leu5-Leu6-Arg7-Ala8-Ala9 containing Ala8, which was initially derived from the electron density and is now confirmed by peptide and cDNA sequence analyses. The cut level is at 18% of the density maximum corresponding to 1·3 standard deviations of the map. The CA-atoms are labeled. For clarity, all density at distances of more than 2 Å from any atom of the model has been removed.

separate domain superposition. Considering these residual distances of main-chain atoms and solvent molecules (which are certainly augmented by local deviations from the NCS-relations) together with the Luzzati upper limit we estimate the actual co-ordinate error as about 0·2 Å.

3. Results and Discussion

(a) Main and side-chain conformations

As exemplified in Figure 2, the final electron density is of good quality. However, the five N-terminal (0 to 4) and the six C-terminal residues (220 to 225) are less well defined. The chain conformations at the N termini of complexes no. 1 and no. 2 are at variance, which is presumably caused by contacts to neighboring molecules. Residues 220 to 225 could only be positioned for complex no. 1 (into weak density); whereas no appropriate density was found in the case of complex no. 2.

The final model agrees well with the standard peptide geometry with r.m.s. deviations from ideality of 0·016 Å in bond lengths and 3·2° in bond angles. As shown in Figure 3, the main-chain dihedral angles fall mostly within the energetically favored regions (Ramachandran & Sasisekharan, 1968). Exceptions from this rule are found for three residues of complex no. 1: Ser2 at (ϕ, ψ) $= (-56°, -79°)$, Gln219 at $(46°, 26°)$ and Val223 at $(170°, 95°)$, all of which are in regions of low density. In complex no. 2, Ser2 and Gln219 adopt conformations in allowed regions and Val223 is not included in the model. A general comparison of the (ϕ, ψ) angles of complexes no. 1 and no. 2 shows that for

the structurally related residues 6 through 217 the r.m.s. deviation is 6° in ϕ as well as in ψ. This demonstrates the accuracy of these angles.

The geometry of the side-chains is given in Table 2. The expected trimodal distribution of the first side-chain dihedral angle χ_1 with a strong preference for staggered conformations g^-, t and g^+ is observed. There is almost no eclipsed conformation e. In a general way, the distribution agrees with those of other proteins as for example reported by Karplus & Schulz (1987) and McGregor *et al.* (1987). The statistics for complex no. 2 are similar. The r.m.s. χ_1 difference between the two complexes is 20°. This is three times higher than the r.m.s. difference of the main-chain angles and emphasizes the fact that the main-chain is better defined than the side-chains. For instance, the main-chain is less affected by crystal packing contacts.
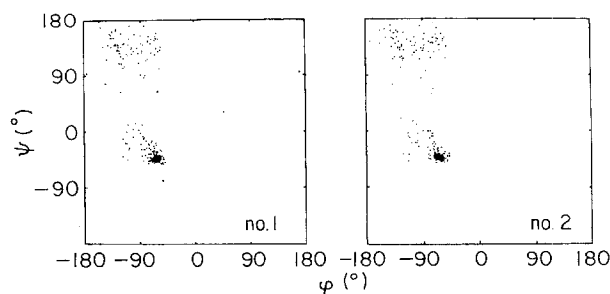

Figure 3. Scatter plot of all main-chain dihedral angles ϕ and ψ (Ramachandran & Sasisekharan, 1968) of the polypeptides of complexes no. 1 (left) and no. 2 (right). All glycine residues are omitted.

Table 2

Statistics of side-chain torsion angles χ_1 of complex no. 1[a]

Residue type	Number of residues					Average values		
	N[b]	g^-	t	g^+	e	g^-	t	g^+
Ser	10	3	1	3	—	57 (6)	162	289 (8)
Thr	19	8	1	8	—	63 (5)	184	302 (5)
Val	15	—	12	2	—	—	173 (8)	301 (5)
Ile	12	—	1	11	1	—	186	287 (12)
Asp	12	1	4	6	—	60	185 (5)	294 (5)
Asn	8	1	3	3	—	62	174 (4)	286 (13)
Glu	15	—	3	7	—	—	201 (15)	298 (7)
Gln	10	1	1	5	1	60	185	291 (7)
Met	4	—	1	2	—	—	191	287 (1)
Lys	14	—	3	8	1	—	177 (4)	300 (17)
Arg	16	—	4	10	—	—	190 (13)	287 (10)
Leu	26	2	9	15	—	94 (1)	185 (5)	294 (7)
His	5	—	3	2	—	—	178 (3)	297 (1)
Phe	7	—	3	3	—	—	179 (4)	291 (4)
Tyr	7	—	3	4	—	—	181 (8)	293 (4)
Trp	3	1	1	1	—	30	173	290
Total	183	17	53	90	3	63 (14)	181 (11)	293 (11)
Total-α[c]	77	3	30	39	2	58 (6)	181 (12)	292 (8)
Total-β[d]	25	1	12	12	1	30	180 (11)	293 (10)
Total-non-$\alpha\beta$[e]	81	13	11	39	—	67 (14)	181 (11)	293 (12)

[a] χ_1 is the torsion angle at the CA–CB bond. For exact definition of χ_1 as well as g^-, t, g^+ and e see Karplus & Schulz (1987). Standard deviations are given in parentheses.

[b] Number of residues present in the sequence. For 20 residues at least 1 of the temperature factors of the CA, CB or CG atoms exceeded 60 Å² so that no conformation was assigned.

[c] Residues in α-helices.

[d] Residues in β-sheets.

[e] Residues neither in α-helices nor β-sheets.

(b) Secondary structures and hydrogen bonds

The secondary structure assignments were made using the simple cut-off criterion (distance between H- and acceptor atom smaller than 2·5 Å, regardless of the angle of approach) suggested by Baker & Hubbard (1984). The assignments are given in
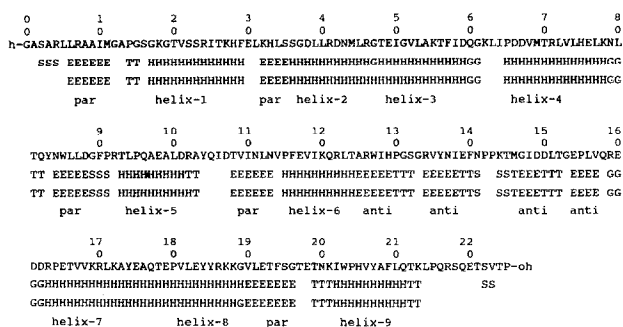
```
        0       1        2         3          4         5        6        7          8
        0       0        0         0          0         0        0        0          0
h-GASARLLRAAIMGAPGSGKGTVSSRITKHFELKHLSSGDLLRDNMLRGTEIGVLAKTFIDQGKLIPDDVMTRLVLHELKNL
     SSS EEEEEE  TT HHHHHHHHHHHH    EEEEHHHHHHHHHHHGHHHHHHHHHHHHGG    HHHHHHHHHHHHHHHGG
         EEEEEE  TT HHHHHHHHHHHH    EEEEHHHHHHHHHHHHHHHHHHHHHHHHGG    HHHHHHHHHHHHHHHGG
         par     helix-1       par  helix-2     helix-3             helix-4

            9       10       11       12        13        14        15          16
            0       0        0        0         0         0         0           0
TQYNWLLDGFPRTLPQAEALDRAYQIDTVINLNVPFEVIKQRLTARWIHPGSGRVYNIEFNPPKTMGIDDLTGEPLVQRE
TT EEEEESSS HHHHHHHHTT     EEEEEE HHHHHHHHHHEEEEETTT EEEEETTS  SSTEEETTT EEEE GG
TT EEEEESSS HHHHHHHHHHT    EEEEEE HHHHHHHHHHEEEEETTT EEEEETTS  SSTEEETTT EEEE GG
   par     helix-5        par    helix-6  anti    anti         anti   anti

        17       18       19       20        21        22
        0        0        0        0         0         0
DDRPETVVKRLKAYEAQTEPVLEYYRKKGVLETFSGTETNKIWPHVYAFLQTKLPQRSQETSVTP-oh
GGHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHEEEEEEE   TTTHHHHHHHHHHHTT         SS
GGHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHNGEEEEEEE   TTTHHHHHHHHHHHTT
   helix-7        helix-8       par   helix-9
```

Figure 4. Amino acid sequence and secondary structure assignment of AK3. Since the model is well refined, we used the rather restrictive cut-off criterion (distance between H and acceptor atom smaller than 2·5 Å) and followed Karplus & Schulz (1987). From top to bottom the lines contain: sequence, secondary structure of complex no. 1, secondary structure of complex no. 2, labels at helices and at the *par*allel and *anti*parallel β-sheets of domains MAIN and INSERT, respectively. The assignment is: H, α-helix; E, β-pleated sheet; G, 3_{10}-helix; T, loop with an $(i+n,i)$ hydrogen bond $(n = 3,4,5)$; S, loop without hydrogen bond.

Figure 4. They differ only slightly for the two poly-peptides. The basic structural motif of all adenylate kinases consists of a five-stranded, parallel β-sheet (Fig. 5) surrounded by nine α-helices, where helices 7 and 8 are separated by a proline, here Pro180. This proline is conserved in all adenylate kinases. It disturbs the hydrogen bonding pattern so that helix 7 changes its direction by about 30° before continuing as helix 8.

While this part of the structure resembles the chain fold of the small variants of adenylate kinases (Schulz et al., 1986), a stretch of 39 residues forms the separate domain INSERT consisting of two β-meanders. These β-meanders are loosely connected to form a six-stranded curved antiparallel β-pleated sheet specified in Figure 5 and illustrated in Figure 6. This domain is further stabilized by side-chain hydrogen bond interactions as listed in Table 3.

The statistics of main-chain dihedral angles and hydrogen bond parameters in α-helices and β-sheets are given in Table 4. They are in general agreement with those of other proteins (Karplus & Schulz, 1987).

(c) Non-crystallographic symmetry

The two complexes in the asymmetric unit are related by NCS (residues 0 to 4 and 220 to 225 are here neglected), which was exploited for density averaging during the initial phases of structure
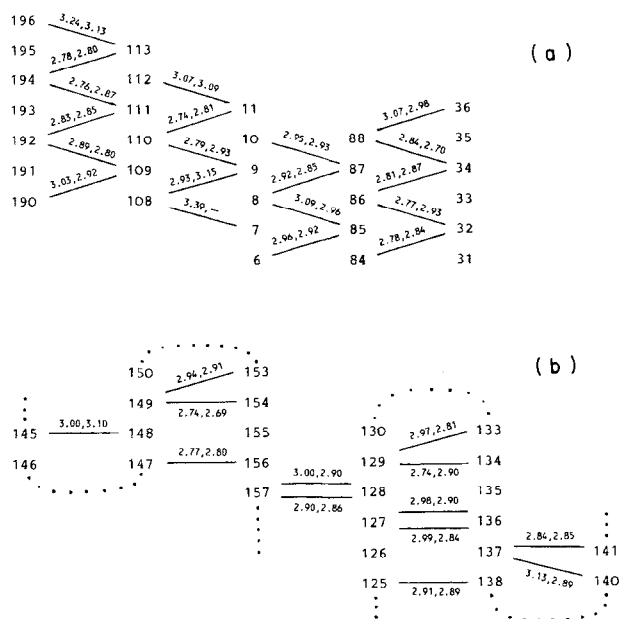
Figure 5. The two β-pleated sheets of AK3. All hydrogen bond lengths of complexes no. 1 and no. 2 are given as donor ... acceptor distances. (a) Central parallel sheet of domain MAIN observed in all adenylate kinases. (b) Antiparallel sheet of the domain INSERT, which is only present in the large variants among the adenylate kinases (Schulz *et al.*, 1986).

elucidation. Soon after beginning the refinement, significant deviations from an overall NCS relation were recognized (Diederichs & Schulz, 1990). They amounted to up to 3 Å for backbone atoms and gave rise to an r.m.s. deviation of 1·1 Å (all atoms) between the two complexes. This number drops appreciably if the domains MAIN, INSERT and AMPbd are separately overlaid. In this case, the all-atom r.m.s. deviation is 0·4 Å. If subdivided into main- and side- (given in parentheses) chain atoms

Table 3
Hydrogen bonds involving side-chains of the domain INSERT

| | | Donor–acceptor distance in complex | |
Atom 1	Atom 2	no. 1 (Å)	no. 2 (Å)
Arg134–NE[a]	Asp149–OD1	2·88	2·89
Arg134–NH1	Asp149–OD2	2·82	2·97
Val135–O	Asn141–ND2	3·30	3·14
Tyr136–OH	Asp149–OD1	2·75	2·76
Tyr136–OH	Asp150–N	2·98	3·20
Asn137–N	Asn141–OD1	2·69	3·07
Lys144–N	Asp150–OD2	2·97	2·89
Asp149–OD2	Thr152–N	3·18	3·07

[a] This is a salt bridge at the surface.

the r.m.s. deviations are: 0·32 Å (0·52 Å) for MAIN (residues 5 to 32, 59 to 123, 163 to 215), 0·38 Å (0·52 Å) for AMPbd (residues 33 to 68) and 0·27 Å (0·50 Å) for INSERT (residues 124 to 162).

The deviations are best characterized by superimposing domain MAIN of complex no. 1 on its counterpart in complex no. 2 and comparing the resulting positions of the other domains. This superposition shows that in complex no. 1 domain INSERT is rotated by 9° toward domain AMPbd and that domain AMPbd is rotated by 7° toward INSERT. Both rotations close the large substrate binding cleft (Diederichs & Schulz, 1990) to some extent, rendering complex no. 1 more "closed" than no. 2. These rotations are almost certainly due to the different contact environments in the crystal. They also demonstrate that these domains are not rigidly glued together. Apart from the domain rotations there are a number of side-chains with differing conformations, most of which are at the surfaces.
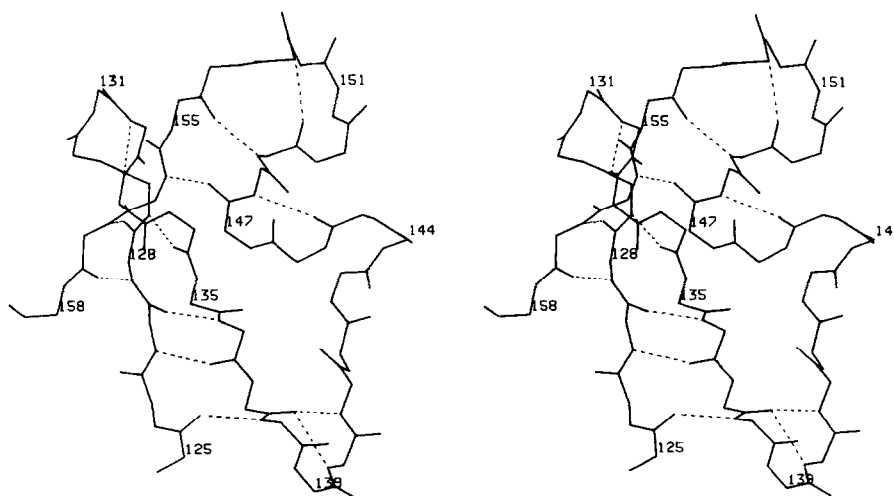

Figure 6. Main-chain conformation of domain INSERT of AK3:AMP complex no. 1 (complex no. 2 is closely similar). The chain contains 2 consecutive β-meanders that form a curved 6-stranded antiparallel sheet. If residues 141 and 145 are considered as being in the same β-strand, the curved 6-stranded sheet can be described as a 5-stranded barrel.

Table 4

Average main-chain dihedral angles and hydrogen bond parameters in α-helices and β-sheets[a]

		β-Sheet		
	α-Helix	Parallel	Antiparallel	Total
$\langle\phi\rangle$[b]	−64 (8)	−113	−110	−112 (18)
$\langle\psi\rangle$[c]	−40 (8)	130	140	133 (15)
$\langle N..O\rangle$[d]	3·00 (0·15)	2·92	2·90	2·91 (0·13)
$\langle H..O\rangle$[e]	2·06 (0·16)	1·98	1·97	1·97 (0·15)
$\langle N-\hat{H}..O\rangle$[f]	158 (10)	159	156	158 (11)
$\langle H..\hat{O}=C\rangle$[g]	151 (8)	156	148	153 (13)

[a] The secondary structures are as defined in Fig. 4. Standard deviations are given in parentheses.
[b] Average angle ϕ (°).
[c] Average angle ψ (°).
[d] Average distance N . . O (Å).
[e] Average distance H . . O (Å).
[f] Average angle N − $\hat{H}$ . . O (°).
[g] Average angle H . . $\hat{O}$ = C (°).

(d) Temperature factors

The mean temperature factor for all atoms of the refined model is 33·0 Å². Subset averages are: 26·5 Å² for the main-chain atoms, 38·3 Å² for side-chain atoms and 49·2 Å² for solvent atoms. The variation of the temperature factors along the main-chain is plotted in Figure 7. A striking feature in this Figure is the high temperature factor of domain AMPbd of complex no. 1 in comparison to complex no. 2. This difference can be explained by the crystal lattice contacts (see below). Towards the N and C termini, the temperature factor rises sharply in both complexes. It seems therefore likely that neither N nor C termini have defined structures in solution.

The occupancies of all solvent molecules have been fixed at 100%. The mean temperature factor of the NCS-related solvent molecules modeled as water is merely 40·0 Å² while that of the other water molecules is 55·0 Å², which indicates a tighter
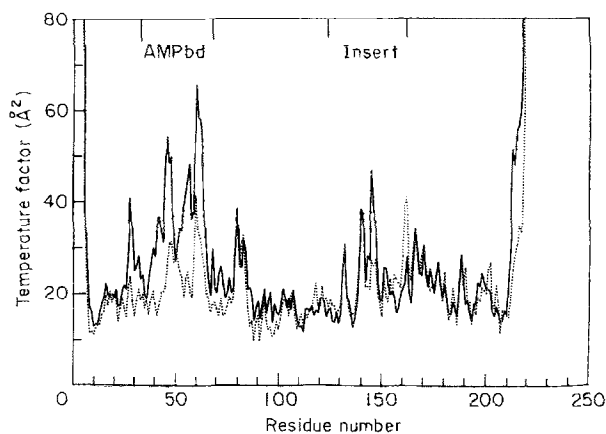

Figure 7. Temperature factors of the main-chain atoms (averaged per residue) of AK3:AMP complex no. 1 (continuous line) and complex no. 2 (dotted line). Domains AMPbd and INSERT are indicated.

Table 5

Contacts between AK3:AMP complexes in the crystal[a]

Packing interaction[b]	Buried area (Å²)[c]	Hydrogen bonds[d]		
		Ref. set[b]	2nd set[b]	D..A distance (Å)[d]
I	550	201–O	62–NZ	3·44
		205–NE2	62–NZ	2·85
II	383	59–OD1	201–NZ	2·95
		158–N	211–OE1	3·00
		158–O	211–NE2	2·84
		160–OE1	203–NE1	2·93
		160–OE2	203–NE1	3·37
		163–NE	207–OH	3·04
		163–NH1	219–NE2	3·18
III	399	49–N	30–OE2	3·06
		49–OE2	217–NH2	2·87
		75–NE2	219–O	2·79
IV	514	7–NE	132–O	3·02
		7–NH1	132–O	3·19
		101–OD1	126–NH1	3·07
		101–O	126–NH1	2·77
		103–O	135–N	2·96
		105–N	133–O	2·84
V	504	41–NE	187–O	3·46
		126–NH2	105–OE1	3·49
		126–NH2	105–NE2	2·74
		132–O	7–NE	3·39
		133–O	105–N	3·03
		134–NH2	104–OH	3·44
		135–N	103–O	3·04
VI	216	116–N	151–O	3·42
		117–N	151–O	2·90
		175–OE2	144–NZ	2·77
VII	181	None		
VIII	57	None		
IX	100	None		
X	172	0–O	192–OE2	3·06
		0–N	189–O	2·86
		1–N	192–OE2	2·65
		2–N	192–OE1	3·08

[a] The given contacts contain at least one interatomic distance below 4·5 Å.
[b] The packing contacts I through X are formed by a residue set of complex no. 1 (res-1) or complex no. 2 (res-2) taken as a reference (operator = $[X, Y, Z]$) and a second residue set of another complex (res-1 or res-2) related by the operator given as OP[. . .]. I: res-1 (200–205, 207, 208, 211, 216, 221, 225) with OP[$X-1, Y, Z$] on res-2 (44, 45, 47, 49, 52, 53, 56, 57, 60, 62, 65). II: res-1 (59, 155, 157–160, 163–165) with OP[$\frac{1}{2}-X, -Y+1, \frac{1}{2}+Z$] on res-2 (28, 201, 203, 207, 208, 211, 216, 219). III: res-1 (45–49, 52, 75) with OP[$\frac{3}{2}-X, -Y+1, \frac{1}{2}+Z$] on res-2 (0, 2, 3, 5, 30–32, 82–84, 215, 217–219). IV: res-1 (2, 3, 5, 7, 81, 101–105, 183, 186, 187) with OP[$X+1, \frac{1}{2}+Y, \frac{1}{2}-Z$] on res-2 (38, 41, 45, 126, 128, 131–135, 141, 159). V: res-1 (38, 41, 126, 131–135, 141) with OP[$-X+1, \frac{3}{2}+Y, \frac{1}{2}-Z$] on res-2 (5, 7, 81, 101–105, 186, 187). VI: res-1 (115–117, 175) with OP[$-X, \frac{1}{2}+Y, \frac{1}{2}-Z$] on res-2 (144, 150–152). VII: res-1 (144, 150–153) with OP[$-X, -\frac{1}{2}+Y, \frac{1}{2}-Z$] on res-2 (78, 115–118, 197). VIII: res-1 (78, 81) with OP[$X-1, Y, Z$] on res-1 (117, 120). IX: res-2 (46, 78) with OP[$X+1, Y, Z$] on res-2 (117, 121, 145, 146). X: res-2 (0–2) with OP[$\frac{1}{2}+X, -\frac{1}{2}-Y, -Z$] on res-2 (189–192, 209).
[c] As determined using program DSSP of Kabsch & Sander (1983).
[d] All hydrogen bonds within the contacts with donor . . . acceptor distances less than 3·5 Å are listed.
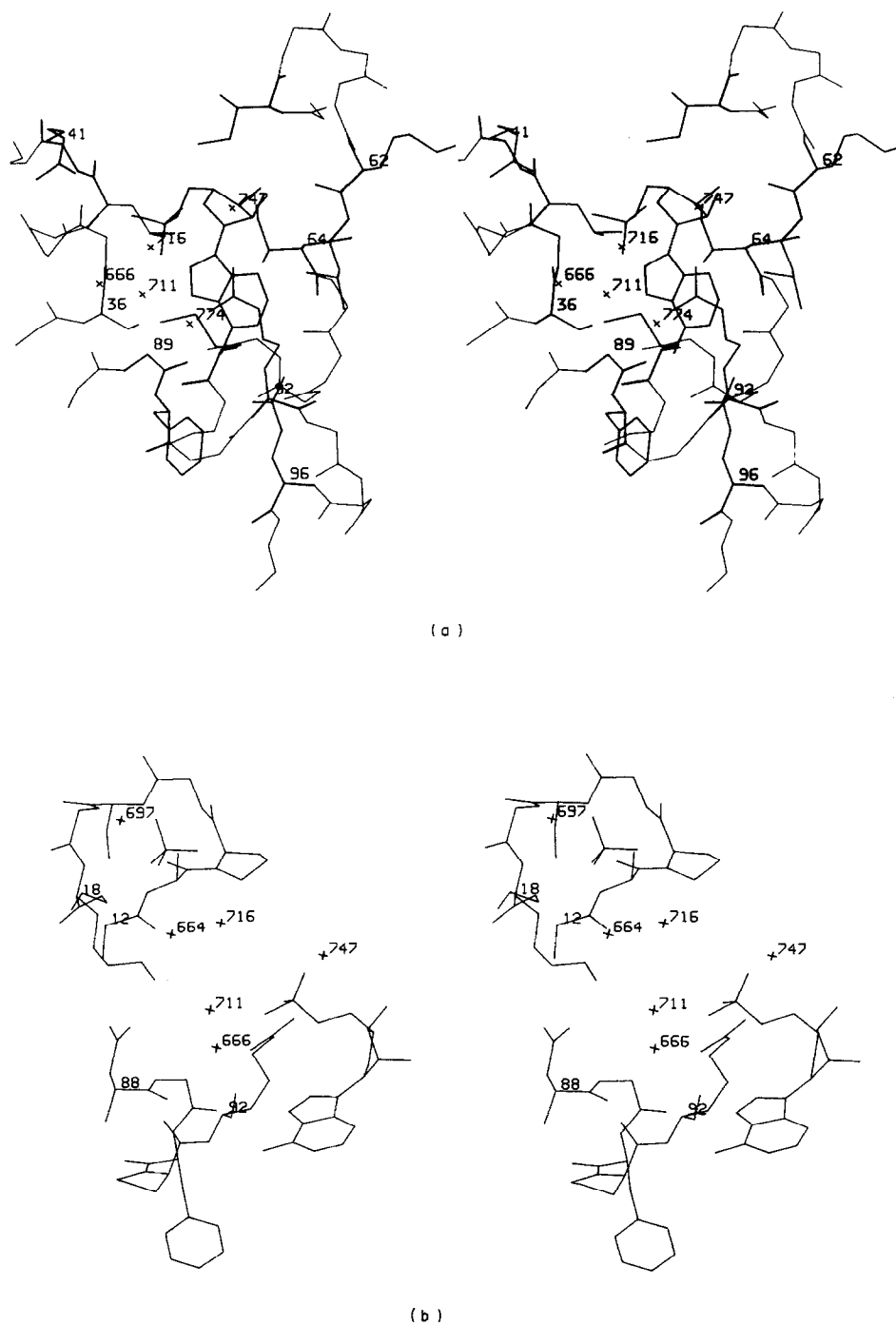
(a)

(b)

Figure 8. The environment of bound AMP in AK3:AMP complex no. 1 (complex no. 2 is very similar). For clarity, all side-chains that do not contact AMP are deleted. Several CA-atoms and water molecules are labeled. (a) The immediate environment of AMP. (b) The geometric relation between the phosphate of AMP and the sulfate molecule (upper left-hand side) bound at the giant anion hole formed by the glycine-rich loop in a position similar to AK1 (Dreusicke & Schulz, 1986).

binding of the NCS-related water molecules. They can be considered as an integral part of the protein. In both complexes, the temperature factor of the bound sulfate is around 80 Å². These sulfates occupy the "giant anion hole" formed by the glycine-rich loop as observed in AK1 (Dreusicke & Schulz, 1986), but they are much less rigorously bound than in AK1.

(e) *Crystal packing*

As demonstrated in Table 5, the strongest packing interactions in the crystal occur between the NCS-related complexes. The contacts between complexes related by crystallographic symmetry are much weaker. Ala103–Gln105 and Val135–Gly113 form a small distorted intermolecular antiparallel

Table 6
Contacts between atoms of the substrate AMP and the enzyme[a]

AMP atom	Protein atom	Distance (Å)	H . . A (Å)	D − Ĥ . . A (°)
N1B	Gln96–NE2	2·91 (2·99)	1·95 (2·03)	166 (163)
N1B	Met69–CG	3·66 (3·66)		
N3B	Ile64–N	3·25 (3·31)	2·35 (2·43)	153 (150)
N3B	Leu63–CD2	3·58 (3·67)		
N7B	Ser36–OG	2·92 (2·92)	1·97 (1·97)	165 (163)
N7B	Arg92–NH2	3·65 (3·08)	3·34 (2·63)	110 (107)
N7B	Gly37–N	4·29 (4·42)	4·09 (4·26)	95 (93)
N9B	Leu40–CD1	3·52 (3·58)		
N10B	Gln96–OE2	2·86 (3·02)	1·89 (2·07)	160 (155)
N10B	Gly89–O	2·87 (2·85)	1·89 (1·87)	162 (160)
N10B	Phe90–CA	3·89 (3·73)		
O1RB	Leu40–CD1	3·66 (3·28)		
O2RB	Lys62–O	2·54 (3·06)	1·78 (2·91)	134 (89)
O3RB	Wat782	2·93	2·55	102
O3RB	Wat958	(4·12)	(3·78)	(103)
C4RB	Ile58–CG1	4·12 (4·16)		
O5RB	Gly37–CA	4·23 (3·63)		
OP5R[b]	Wat747	2·50	1·85	120
OP5R	Wat901	(2·41)	(1·86)	(111)
OP5R	Arg92–NH1	3·85 (4·14)	3·00 (3·25)	143 (150)
OP5R	Wat716	4·34	3·74	121
OP5S[b]	Wat711	2·50	1·87	118
OP5S	Wat903	(3·09)	(2·16)	(154)
OP5S	Wat666[c]	3·33 (2·55)	3·05 (1·83)	97 (126)
OP5S	Arg92–NH2	3·53 (4·37)	2·64 (3·44)	142 (155)
OP5S	Arg92–NH1	3·50 (4·76)	2·60 (3·92)	149 (143)
OP5S	Wat774	4·14	4·07	87
O45[b]	Arg41–NH1	3·53 (2·51)	3·27 (1·97)	83 (111)
O45	Wat666[c]	3·63 (4·28)	2·68 (3·30)	159 (167)
O45	Wat918	(4·41)	(3·68)	(132)
O45	Wat928	(4·22)	(3·39)	(142)
O45	Gly37–CA	3·91 (4·03)		

[a] All contacts with distances smaller than 4·5 Å in any one of the complexes are listed. The values for complex no. 2 are given in parentheses following those of complex no. 1.

[b] The oxygen atoms OP5R, OP5S, O45 are named as in Ap₅A. The numbers denote the phosphate number (1 through 5) in Ap₅A.

[c] The values given for complex no. 2 are those of the NCS-related solvent molecule Wat866.

β-sheet. This contact occurs twice (packing interactions IV and V of Table 5) such that it does not give rise to dimers, but connects subsequent molecules forming a linear string along the b axis of the crystal. These favorable contacts along the b axis are likely to enhance crystal growth along this very direction. This corresponds to the observation that crystals are usually about ten times larger along the b axis than along the a and c axes.

The lattice contacts of domain AMPbd differ appreciably in the two complexes. In complex no. 1, AMPbd contacts only a few residues of an NCS-related molecule, whereas AMPbd of complex no. 2 is stabilized by a contact with the C-terminal helix of complex no. 1. As a consequence, the average temperature factor for AMPbd is lower in complex no. 2 (28 Å²) than in complex no. 1 (40 Å²). The same applies for the mean temperature factor of the bound AMP (29 Å² in complex no. 1, 40 Å² in complex no. 2).

(f) The AMP binding site

The AMP binding site is located in the pocket between the helices 2, 3 and 4. The contacts of AMP are given in Table 6. The binding geometry is illus-

trated in Figure 8. In crystals of AKyst:Ap₅A and AKeco:Ap₅A, one of the adenines binds at the equivalent location. The known high specificity of the adenylate kinases towards AMP, which contrasts their lower specificity for ATP, together with the highly conserved sequence in this region allowed Egner et al. (1987) to assign this site as the AMP site. This hypothesis is confirmed by the crystal structure presented here.

Hydrogen bond interactions between enzyme and substrate AMP involve mostly the nucleotide base of AMP; all nitrogen atoms take part. Adenine is hydrogen-bonded to the side-chains of Ser36 and Gln96 as well as to the amide of Ile64 and the carbonyl of Gly89. Gly89 is located in a loop which interacts with the Gly-rich loop via salt bridge Asp88–OD2:Lys18–NZ. The loop conformation is stabilized by cis-Pro91, which is conserved in all known adenylate kinases. While Gln96 is also a conserved residue, Ser36 is exchanged for Thr in some species (Schulz et al., 1986). But Thr has a hydroxyl in the γ-position like Ser, so that this exchange has probably only minor effects. One may therefore assume that these four residues warrant the high specificity of the adenylate kinases for the base adenine.

The ribose and the phosphate of AMP are much less rigidly bound than the base. The average temperature factors for (adenine, ribose, phosphate) are $(25 \text{ Å}^2, 45 \text{ Å}^2, 63 \text{ Å}^2)$ in complex no. 1 and $(17 \text{ Å}^2, 31 \text{ Å}^2, 48 \text{ Å}^2)$ in complex no. 2. Only the 2′-hydroxyl of the ribose forms a strong hydrogen bond directly to the polypeptide; all other ribose atoms are either hydrogen-bonded to water molecules or they contact non-polar atoms (Table 6). The phosphate contacts water as well as the guanidinium groups of Arg41 and Arg92.

Taken together, the base adenine seems to sit snugly at its final position for catalysis while further movements have to be expected for ribose and phosphate. As shown in Figure 8(b), the sulfate bound in the glycine-rich loop is rather far away from the phosphate of AMP. Since this sulfate corresponds to the β-phosphate of ATP as derived from Ap₅A bound to the adenylate kinases AKyst as well as AKeco (U. Abele & C. W. Müller, personal communications), it is clear that some movement of AMP relative to ATP has to occur to allow transfer of the γ-phosphoryl group of ATP. Such a movement has been derived from a comparison between several different adenylate kinase structures (Schulz *et al.*, 1990).

References

Baker, E. N. & Hubbard, R. E. (1984). Hydrogen Bonding in Globular Proteins. *Progr. Biophys. Mol. Biol.* **44**, 97–179.

Bruenger, A. T. (1988). Crystallographic Refinement by Simulated Annealing. Application to a 2·8 Å Resolution Structure of Aspartate Aminotransferase. *J. Mol. Biol.* **203**, 803–816.

Bruenger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic *R*-factor Refinement by Molecular Dynamics. *Science*, **235**, 458–460.

Diederichs, K. (1990). Die Struktur von zwei Kristallformen der Adenylatkinase aus Rinderherz-mitochondrien bei 1·85 Å bzw. 3·0 Å Auflösung. Dissertation (PhD), Universität Freiburg i.Br.

Diederichs, K. & Schulz, G. E. (1990). Three-dimensional Structure of the Complex Between Mitochondrial Matrix Adenylate Kinase and Its Substrate AMP. *Biochemistry*, **29**, 8138–8144.

Dreusicke, D. & Schulz, G. E. (1986). The Glycine Loop of Adenylate Kinase Forms a Giant Anion Hole. *FEBS Letters*, **208**, 301–304.

Dreusicke, D., Karplus, P. A. & Schulz, G. E. (1988).

Refined Structure of Porcine Cytosolic Adenylate Kinase at 2·1 Å Resolution. *J. Mol. Biol.* **199**, 359–371.

Egner, U., Tomasselli, A. G. & Schulz, G. E. (1987). Structure of the Complex of Yeast Adenylate Kinase with the Inhibitor P¹,P⁵-di(adenosine-5′-) pentaphosphate at 2·6 Å Resolution. *J. Mol. Biol.* **195**, 649–658.

Kabsch, W. & Sander, C. (1983). Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-bonded and Geometrical Features. *Biopolymers*, **22**, 2577–2637.

Karplus, P. A. & Schulz, G. E. (1987). The Refined Structure of Glutathione Reductase at 1·54 Å Resolution. *J. Mol. Biol.* **195**, 701–729.

Luzzati, V. (1952). Traitement Statistique des Erreurs dans la Détermination des Structures Crystallines. *Acta Crystallogr.* **5**, 802–810.

McGregor, M. J., Islam, S. A. & Sternberg, M. J. E. (1987). Analysis of the Relationship Between Side-chain Conformation and Secondary Structure in Globular Proteins. *J. Mol. Biol.* **198**, 295–310.

Mueller, C. W. & Schulz, G. E. (1988). Structure of the Complex of Adenylate Kinase from *Escherichia coli* with the Inhibitor P¹,P⁵-di(adenosine-5′-) pentaphosphate. *J. Mol. Biol.* **202**, 909–912.

Ramachandran, G. N. & Sasisekharan, V. (1968). Conformations of Polypeptides and Proteins. *Advan. Protein Chem.* **23**, 283–437.

Schulz, G. E., Schiltz, E., Tomasselli, A. G., Frank, R., Brune, M., Wittinghofer, A. & Schirmer, R. H. (1986). Structural Relationships in the Adenylate Kinase Family. *Eur. J. Biochem.* **161**, 127–132.

Schulz, G. E., Mueller, C. W. & Diederichs, K. (1990). The Induced-fit Movements in Adenylate Kinase. *J. Mol. Biol.* **213**, 627–630.

Tomasselli, A. G. & Noda, L. H. (1979). Mitochondrial GTP-AMP Phosphotransferase. 2. Kinetic and Equilibrium Dialysis Studies. *Eur. J. Biochem.* **93**, 263–270.

Tomasselli, A. G., Schirmer, R. H. & Noda, L. H. (1979). Mitochondrial GTP–AMP Phosphotransferase. 1. Purification and Properties. *Eur. J. Biochem.* **93**, 257–262.

Tomasselli, A. G., Frank, R. & Schiltz, E. (1986). The Complete Primary Structure of GTP:AMP Phosphotransferase from Beef-heart Mitochondria. *FEBS Letters*, **202**, 303–308.

Wieland, B., Tomasselli, A. G., Noda, L. H., Frank, R. & Schulz, G. E. (1984). The Amino Acid Sequence of GTP:AMP Phosphotransferase from Beef-heart Mitochondria. Extensive Homology with Cytosolic Adenylate Kinase. *Eur. J. Biochem.* **143**, 331–339.

Yamada, M., Shahjahan, M., Tanabe, T., Kishi, R. & Nakazawa, A. (1989). Cloning and Characterization of cDNA for Mitochondrial GTP:AMP Phosphotransferase of Bovine Liver. *J. Biol. Chem.* **264**, 19192–19199.

Edited by R. Huber