ACADEMIC PRESS

Biochemical and Biophysical Research Communications 294 (2002) 173–179

BBRC

www.academicpress.com

Comparative modelling and immunochemical reactivity of *Escherichia coli* UMP kinase☆

Gilles Labesse,[a,*] Nadia Bucurenci,[b] Dominique Douguet,[a] Hiroshi Sakamoto,[c] Stéphanie Landais,[c] Cristina Gagyi,[c] Anne-Marie Gilles,[c] and Octavian Bârzu[c]

[a] *Centre de Biochimie Structurale, Faculté de Pharmacie, Université de Montpellier I, 34000 Montpellier, France*
[b] *Institutul Cantacuzino, 70100 Bucharest, Romania*
[c] *Laboratoire de Chimie Structurale des Macromolécules, Institut Pasteur, 75724 Paris Cedex 15, France*

Received 29 April 2002

Abstract

Bacterial UMP kinases do not exhibit any sequence homology with other nucleoside monophosphate kinases described so far, and appear under oligomeric forms, submitted to complex regulation by nucleotides. We propose here a structural model of UMP kinase from *Escherichia coli* based on the conservation of the fold of carbamate kinase whose crystal structure was recently solved. Despite sequence identity of only 18% over 203 amino acids, alignment of UMP kinase from *E. coli* with carbamate kinase from *Enterococcus faecalis* by hydrophobic cluster analysis and threading suggested the conservation of the overall structure, except for a small subdomain (absent in UMP kinase). The modelled dimer suggested conservation of the dimer interface observed in carbamate kinase while interaction of UMP kinase with a monoclonal antibody (Mab 44-2) suggests a three in-plane dimer subunit arrangement. The model was analyzed in light of various modified forms of UMP kinase obtained by site-directed mutagenesis. © 2002 Elsevier Science (USA). All rights reserved.

Keywords: UMP kinase; Molecular modelling; Structure–function relationship; Site-directed mutagenesis; Remote similarity

UMP kinase is an ubiquitous enzyme which catalyzes the phosphorylation of UMP to UDP, with ATP as preferred donor. The enzyme from eukaryotic organisms is monomeric, uses both UMP and CMP as substrates with similar efficiency, and shares significant sequence homology and similar folding with other nucleoside monophosphate (NMP) kinases [1,2]. UMP kinase from bacteria is specific for UMP, does not share any homology with other NMP kinases, and is composed as a hexamer subjected to regulation by UTP (inhibitor) and GTP (activator) [3]. CMP phosphorylation is accomplished in bacteria by a distinct enzyme which conserves the global folding found in adenylate kinase or eukaryotic UMP/CMP kinase [4,5]. Deciphering the structure and the catalytic mechanism of UMP kinase from bacteria is of interest not only from the fundamental point of view, but also knowledge of its enzymatic properties in different species is useful in designing inhibitors acting on bacteria with little or no effect on host organisms.

UMP kinase from *Escherichia coli* has been characterized by various biochemical, immunochemical, and physico-chemical methods [3,6,7]. Mutants defective in the gene encoding UMP kinase (*pyrH*), or site-directed mutagenesis experiments allowed the identification of residues important for catalysis, allosteric regulation, and thermodynamic stability [3,6,8]. In the absence of the three-dimensional structure of the protein, data on various mutants were previously rationalized on a rather speculative basis. To overcome this difficulty we developed a structural model, based on sequence comparisons with proteins of known 3D structures, despite the rather low sequence similarities detected in the present case. Analysis of a protein structure only distantly related to a given structural family requires additional methods for correct gathering and alignment of all

☆ *Abbreviations:* CKef, Carbamate Kinase from *Enterococcus faecalis*; CPSpf, Carbamoyl Phosphate Synthase from *Pyrococcus furiosus*. The one-letter code is used for amino acids.

* Corresponding author. Fax: +33-4-67-52-96-23.
E-mail address: labesse@cbs.cnrs.fr (G. Labesse).

0006-291X/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S0006-291X(02)00450-3

sequences [4,9]. However, careful analysis of the deduced structural model may suggest specific functional features [10]. This strategy was applied to bacterial UMP kinases, using the enzyme from *E. coli* as a representative model. This led us to propose that the hexameric form of UMP kinase is constructed as a trimer of dimers, prediction based on the binding stoichiometry of a monoclonal antibody to *E. coli* UMP kinase. The overall model is discussed in detail in light of the wealth of experimental data and functional information about wild type and mutant bacterial UMP kinases.

Materials and methods

Chemicals. Nucleotides, restriction enzymes, T4DNA ligase, and coupling enzymes were from Roche-Diagnostics, Mannheim, Germany. T7DNA polymerase and the four deoxynucleoside triphosphates used in sequencing reactions were from Pharmacia. Oligonucleotides were synthesized by the phosphoamidinate method. NDP kinase from *Dictyostelium discoideum* (2000 U/mg prot) was provided by M. Véron. Mab 44-2 was obtained by immunizing mice with a C-terminal truncated form of *E. coli* UMP kinase (amino acids 1–180) [8]. Overproduced as inclusion bodies and requiring urea for solubilization, this fragment induced a significant polyclonal response and allowed a set of Mabs to be obtained. Mab 44-2 is an IgG2a recognized by the native protein as well as by the immunogen with a K_a in the 10^{-9} M range.

Bacterial strains, plasmids, and growth conditions. General DNA manipulations were performed as described by Sambrook et al. [11]. Open reading frames from the complete *pyrH* gene were generated by PCR and inserted into the expression vectors pET22b and pET24a (Novagen). Cloning experiments were carried out with strain NM554/pDIA17 [12,13]. The resulting plasmids were introduced into strain BL21 (DE3)/pDIA17 to overproduce the corresponding proteins. Recombinant strains were grown in 2YT medium supplemented with antibiotics to an optical density of 1 at 600 nm and then overproduction was triggered by isopropyl-β-D-thiogalactoside induction (1 mM final concentration) for 3 h. Bacteria were harvested by centrifugation.

Purification and catalytic assay of UMP kinase from E. coli. UMP kinase was purified as previously described [3]. The wild-type enzyme is almost insoluble at neutral pH. At pH 8.6, the solubility was situated between 2 and 3 mg of protein/ml. The various modified forms of UMP kinase (R62H, D77N, D148N, D154N, D168N, D174N, and D201N) were described previously [8]. The R11H, P141Q, P141L, and L226Q variants will be described elsewhere. The UMP kinase activity was determined at 30 °C and 0.5 ml final volume, using a coupled spectrophotometric assay with an Eppendorf ECOM 6122 photometer at 340 nm. The reaction medium buffered with 50 mM Tris–HCl (pH 7.4) contained 50 mM KCl, 2 mM $MgCl_2$, 1 mM phosphoenolpyruvate, 0.2 mM NADH, 1 mM ATP, 0.3 mM UMP, and 2 U each of lactate dehydrogenase, pyruvate kinase, and NDP kinase. The reaction was started with UMP kinase. One unit of the enzyme corresponds to 1 μmol of product formed per minute.

Other analytical procedures. Protein concentration, amino acid analysis after 6 N HCl hydrolysis for 22 h at 110 °C, and SDS–PAGE were performed as described in [3,6].

Sequence comparison and molecular modelling. Protein sequence database searches were performed with the PSI-BLAST version 2.0.5 program [14] with default parameters. Pairwise, multiple alignments and secondary structure prediction were performed by Hydrophobic Cluster Analysis as previously described [15]. Alignment refinement was subsequently performed using the program TITO [16] using either template: CKef or CPSpf (accession codes in the RCSB Protein Data Bank: 1b7b and 1e19, respectively). Validity of the refined alignment was assessed through pseudo-energy and visual inspection. The secondary structures (α-helix, β-strand) of UMPKs were predicted using Hydrophobic Cluster Analysis (HCA) and JPRED[2] [17] and also were deduced by similarity during TITO processing. These secondary structure predictions were merged by consensus and used as additional restraints in the following modelling steps. Three-dimensional models were built using the CKef and CPSpf as a combined template in MODELLER 4.0 [18] and assessed using PROSA [19], ERRAT [20], and Verify3D [21]. These 3D structures were visualized on a UNIX workstation using XmMol [22].

Results and discussion

Sequence analysis of bacterial UMP kinases

Database screening using the program PSI-BLAST with UMP kinase from *E. coli* as a query first showed sequence similarities with UMP kinases from other bacteria (probability scores: e-66 to e-53). Various other bacterial kinases followed immediately after, in the PSI-BLAST output (pyrroline-5 carboxylate synthase with probability scores: e-49 to e-44; glutamate-5-kinase, e-43 to e-41; aspartokinases, e-41 to e-31 and carbamate kinases, e-30 to e-23). No similarity with eukaryotic UMP kinases was detected. Among the consensus motifs highlighted, UMP kinase from *E. coli* possesses a motif (similar to PROSITE PS00324; in UMP kinase from *E. coli* residues 9–20; underlined in Fig. 1) confirming its classification in the aspartokinase superfamily (Pfam HMM probability score: 1.4 e-52, overlap aa 9–192) [3]. The motif PROSITE: PS00902 corresponds to a loop specific for the pyrroline-5-carboxylate synthase and glutamate-5-kinase subfamily, which is absent in UMP kinases. A second motif was derived to specifically gather the whole aspartokinase superfamily. Two specific motifs with probability scores of roughly e-6 and e-8 were thus defined using PATTINPROT [23], which comprised two predicted β-strands and their following loop. Secondary structure prediction using HCA [15] or JPRED [17] indicated the presence of alternating, mostly hydrophobic β-strands and amphipathic α-helices. Most conserved residues lie at the C-termini of the predicted β-strands. These predictions suggested that bacterial UMP kinases adopt a α/β-fold in agreement with infrared spectroscopy [6] showing that UMPK contains up to 27% of α-helices (in presence of UTP) and both parallel and anti-parallel β-strands (21% and 16%, respectively). One of the motifs identified in UMP kinase superfamily sequences ressemble the signature (/d///; where '/' and lower case stand for any hydrophobic residues {A, C, V, I, M, L, F, Y, W} and for prefered amino acid, respectively) of the central strand β-4 in a standard Rossmann fold which is found in most related, doubly wounded (α/β)-folds including those of two-domain dehydrogenases [24], the RED family [25] or the

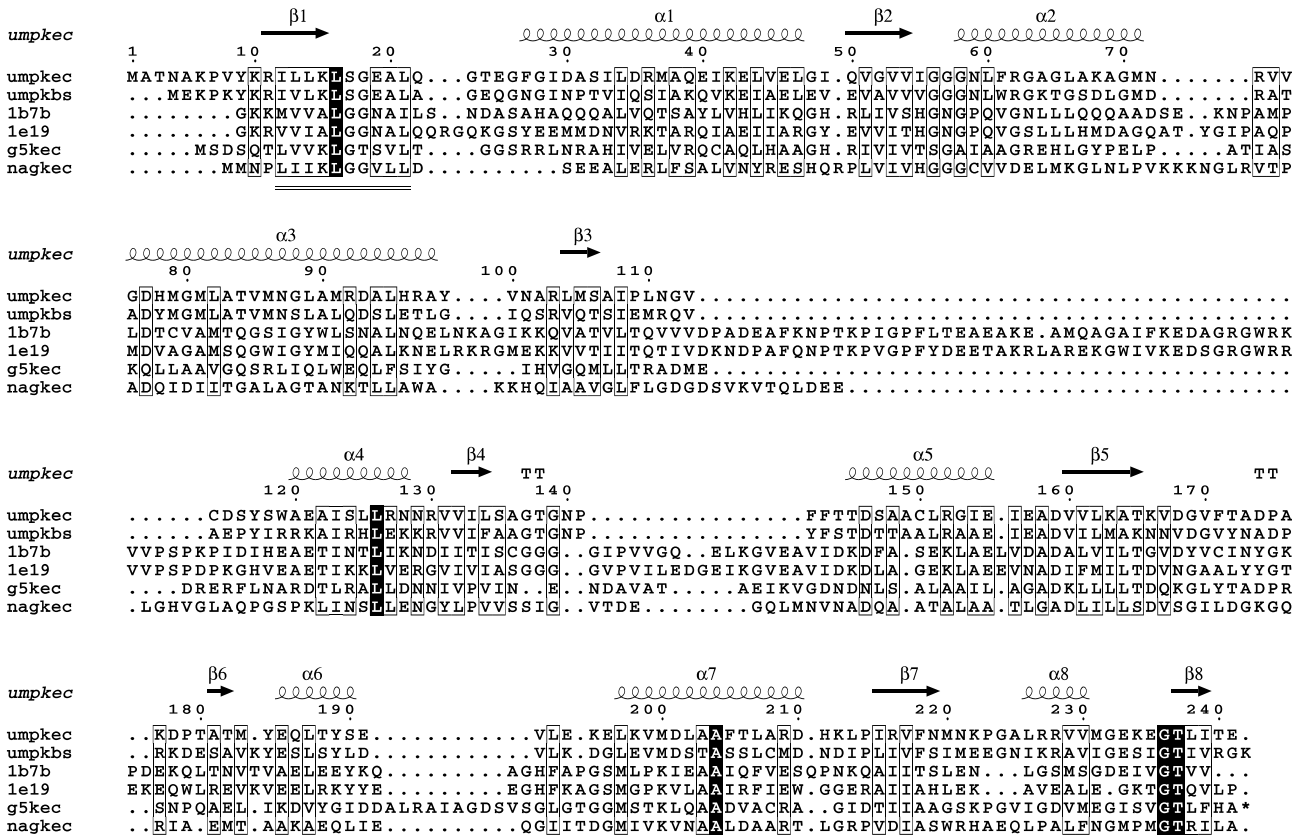

Fig. 1. Alignment of the amino acid sequences of UMP kinase from *E. coli* (umpkec; in SWISSPROT: P29464), UMP kinase from *Bacillus subtilis* (umpkbs; O31749), carbamate kinases (1b7b and 1e19 as in RCSB Protein Data Bank), glutamate-5 kinase from *E. coli* (g5kec; P07005), and *N*-acetyl-glutamate kinase from *E. coli* (nagkec; P11445). The secondary structure elements assigned in the model of UMP kinase from *E. coli* are shown over the alignment. The figure was drawn using ESPRIPT [33].

AdoMet-dependent methyltransferases [26]. In this fold superfamily, the motif of the most N-terminal β-strand is preferentially preceded by a positively charged residue and is followed, after a stretch of mostly hydrophobic residues, by a glycine-rich loop implicated in the anion-binding active site (e.g., phosphate). The N-terminus of UMP kinases and related enzymes appeared to possess a compatible sequence motif (r///x/gGxA/; upper case stands for strictly conserved residue and 'x' for any amino acid). These comparative predictions suggested that UMP kinases might adopt a Rossmann fold. The size of the bacterial UMP kinase and the predicted secondary structures, as well as the conserved sequence motifs, showed no significant similarities to those of other known bacterial or eukaryotic NMP kinases.

Structural comparison of bacterial UMP kinases with carbamate kinases

PSI-BLAST convergence revealed 18% sequence identity over 203 amino acids of UMP kinase from *E. coli* with carbamate kinase from *Enterococcus faecalis* (CKef) and with carbamoyl phosphate synthase from *Pyrococcus furiosus* (CPSpf) whose crystal structures were recently solved [28,29]. Sequence alignment of UMP kinase from *E. coli* with CKef and CPSpf and HCA comparison [15] showed conservation of most secondary structures and of the overall folding, except for a small subdomain which is absent in UMP kinases (Fig. 1). Secondary structures derived by homology were in good agreement with the predicted secondary structure using the program JPRED [17]. Only two errors (one β-strand for α-helix and an α-helix for a β-strand) were detected and corresponded to the two weaker predictions where most methods disagreed (data not shown). The conservation of a carbamate kinase-like fold in UMP kinase from *E. coli* was further confirmed by threading using the recently developed meta-server [9]. The consensus results of threading showed significant structural compatibility between UMP kinases and carbamate kinases. During the manual threading optimization [16], most insertion/deletions were gathered in loops and most hydrophobic stretches were buried. The CKef and CPSpf are over 60 amino acids longer than bacterial UMP kinases. The length difference is due to two deletions of 50 and 15 amino acids in UMP kinases (Fig. 1) and corresponds to the loss of a small subdomain. A similar structural change is also predicted in glutamate-5-kinases (Fig. 1). The pseudo-energy de-

rived from TITO [16] is satisfactory (−58.7 and −65.1, respectively) taking into account the relatively low sequence identity and the quaternary structure of *E. coli* UMP kinase. Most residues crucial for substrate binding and catalysis are strictly conserved (Fig. 1).

Molecular modelling was performed using MODELLER [18], based on the structure of CKef and CPSpf. A dimer was modelled. The final complete model had an energy according to PROSA [19] of −0.8 kcal, compared to −1.7 and −1.4 kcal for the crystal structures of CKef and CPSpf, respectively. The model was also evaluated at the atomic level using the program ERRAT [27] showing that up to 75% (versus 99% for CPSpf) of the residues are in a proper environment. Similarly, the model was validated using the program Verify3D [21]: each monomer scored around 0.35 (versus 0.46 and 0.41 for CPSpf and CKef, respectively) while the dimer got a higher score (0.40) suggesting that the proposed dimerization interface is correct. In the model, the helices α2, α3, and α4 and the strand β3 involved in this interface possess mainly hydrophobic side-chains (α2: L65, A67, M71; α3: M81, L82, V85, M86, L89; β3: L104, M105, A107, P109; α4 Y117, L125) that would be surface-exposed in a monomeric form.

The dimer dimensions are $40 \times 40 \times 80\,\text{Å}^3$ and the surface buried by dimerization is roughly equivalent to that observed in carbamate kinase dimer ($\sim 3700\,\text{Å}^2$). This important interface suggests it is stable in solution. This dimeric form is predicted to be conserved in the various kinases of the superfamily, while the hexameric form has been described only for UMP kinases and glutamate-5-kinases. Three dimers would associate to form the observed hexameric structure. In fact, the transient dimer has been observed for wild-type UMP kinase from *E. coli* at pH 6 [6], while no trimers could be observed. Under other conditions, only monomers, hexamers, and a few higher oligomers are observed in equilibrium, suggesting a mechanism for hexamerization: $6M \leftrightarrow (3D) \leftrightarrow H$. Noteworthy, the dimer of UMP kinases (in the absence of the subdomain observed in carbamate kinases) appeared to form a polyhedral structure with two edges starting from the dimer interface and running to the ATP bound molecules in our model and forming a 120° angle (Fig. 2)

The structure of active site

The present model gathers the sequence stretch best conserved among UMP kinases on one face of each monomer. These faces are mostly composed of β-strand C-termini and α-helix N-termini and the connecting loops. The first motif contributing residues to the active site corresponds to the β1–α1 loop with the consensus r/// k/sGea/ in UMP kinases. In UMP kinase from *E. coli*, the side chains of K15, S17, and E19 would point into the active site. The second motif (β2–α2: ///Gggnx/r) provides

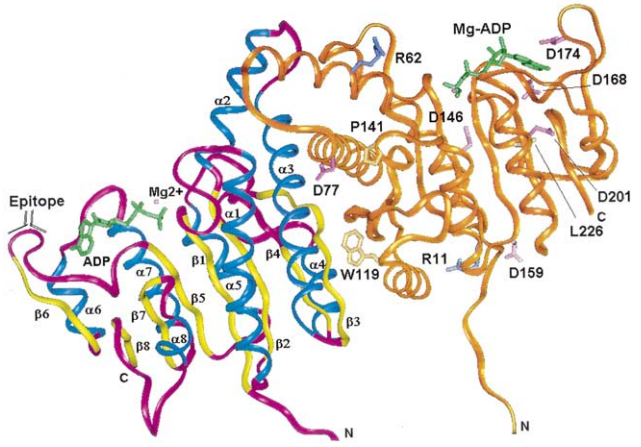

Fig. 2. Ribbon representation of the modelled dimeric structure of UMP kinase from *E. coli*. One monomer is in orange while in the second monomer α-helices are in blue, β-strands in yellow, and loops in purple. The ADP molecules are in green in both monomers. The side chains of the mutated residues discussed in the text are shown (apolar in gray; basic in blue and acidic in purple).

additional glycines to the active site. These residues and the glycine G18 likely interact with a phosphate group, since they correspond to the glycines (G10, G50, and G52) in contact with a sulfate ion in the crystal structure of CKef [27]. A third stretch (r////saGxgP) is present at the end of the fourth strand. It comprises the P141, whose mutation almost abolishes the enzyme activity (H. Sakamoto, unpublished data). It is immediately followed by the motif fftTD including the catalytic residue D146. The next motif (txvdGvftadP) followed the predicted β5-strand (ead///) and provides a threonine and a valine to the catalytic center in UMP kinases, as well as in aspartokinases and glutamate kinases. This motif comprises D168 and D174 (UMP kinase from *E. coli* numbering), whose role in enzyme activity was also probed by site-directed mutagenesis [8]. In CPSpf, this region participates in binding of the phosphate donor [28]. The observed conservation of these motifs suggested that a similar mode of ATP binding is used by this enzyme superfamily. A motif (/k//D//a) conserved among UMP kinases (including D201 in UMP kinase from *E. coli*) corresponds to the N-terminus of the helix α7, which is predicted to contribute residues to the catalytic center. The last conserved motif (gtx//) seems to stabilize the specific structure of the ATP-binding site rather than directly participating in the binding itself in the two solved crystal structures [27,28].

The binding stoichiometry of a monoclonal antibody to UMP kinase from E. coli suggests an in-plane arrangement of the six subunits

We have previously shown that a monoclonal antibody (Mab 44-2) interacts with a 10 amino acid linear

epitope situated between F171 and T180 of *E. coli* UMP kinase [7]. Mab 44-2 inhibited the enzyme activity in a dose-dependent manner, never exceeding 60%. This result suggested the proximity of the F^{171}–T^{180} sequence to the active site, and a particular arrangement of UMP kinase subunits in the hexamer, such that only half of the epitopes are accessible to the monoclonal antibody. When mixtures of UMP kinase and Mab 44-2 in 50 mM Tris–HCl (pH 8.6) were analyzed by gel permeation chromatography, two peaks were separated (Fig. 3). The first, preceded by a shoulder, exhibits a molecular mass of 600 kDa and corresponds to UMP kinase hexamers in complex with three moles of Mab 44-2. The second peak, of 150 kDa, corresponds to free Mab 44-2. SDS–PAGE analysis of fractions resulting from gel permeation chromatography (Fig. 3, inset) confirmed the proposed stoichiometry. Thus, the ratio of enzyme monomer/Mab 44-2 after the scan of Commassie blue stained gels was 2.1/1 for the peak fraction of the 600 kDa band. Fractions in the shoulder indicated a ratio of 1.3/1. The model of the subunit arrangement in the UMP kinase hexamer of three in-plane dimers is shown in Fig. 4. The epitopes in each dimer which react with Mab 44-2 are surface-exposed and on the opposite side of the subunit interface (Fig. 2). Epitopes of the vicinal dimers are in close proximity to each other,

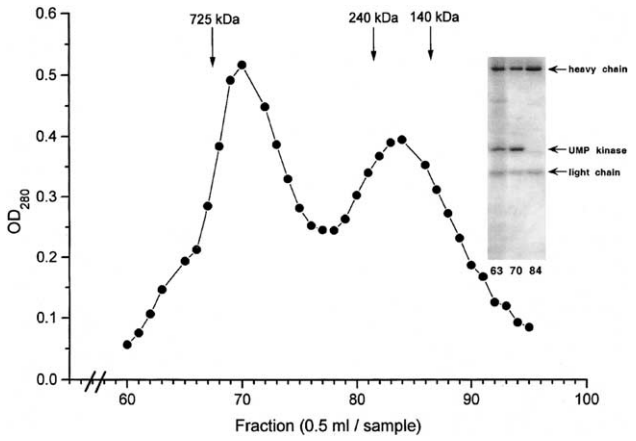

Fig. 3. Identification of molecular species resulting from interaction of *E. coli* UMP kinase with the monoclonal antibody Mab 44-2. One mg of UMP kinase (38 nmol in terms of monomer) and 6 mg of Mab 44-2 (40 nmol) were mixed in 1 ml of 50 mM Tris–HCl, pH 8.6, and loaded onto a $0.9 \times 120\,cm^2$ column of Sephacryl S-300. The column previously equilibrated with 50 mM Tris–HCl, pH 8.6, was calibrated with four molecular mass markers: α_2-macroglobulin (725 kDa), muscle pyruvate kinase (240 kDa), muscle lactate dehydrogenase (140 kDa), and horseradish peroxidase (40 kDa). Proteins were eluted from the column at flow rates of 5 ml/h, collecting samples of 0.5 ml. Samples 63, 70, and 84 (inset) were stained with Commassie blue and scanned for determination of protein concentration. Separate Commassie blue stained gels with increasing concentrations of Mab 44-2 (6, 10, 20, and 40 μg) and UMP kinase (2, 5, 7, and 10 μg) were scanned for calibration curve. In both cases the curves were linear in the mentioned concentration range.

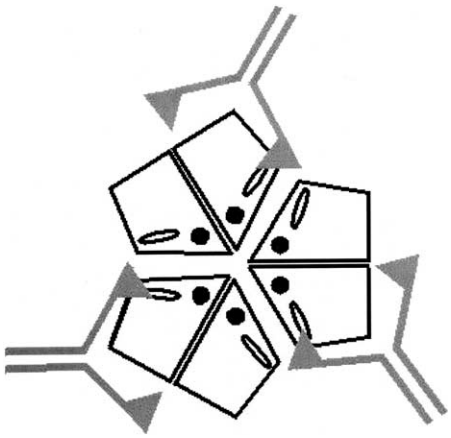

Fig. 4. Schematic representation of the interaction of UMP kinase hexamer with the antibody Mab 44-2. Three identical dimers of UMPK are shown positioned according to our molecular model (see text). The antibody is drawn in gray lines. The binding sites of the substrate and allosteric effectors (filled circle; named C1 and C2 in the text) and that of ATP (open oval, C3) are predicted by similarity and from sequence-structure analysis (see text and Fig. 2).

explaining why only one Mab 44-2 molecule can interact with a dimeric subunit. On the other hand, the active site not far from the F^{171}–T^{180} sequence is no longer accessible to substrates upon binding of Mab 44-2. Two independent observations support this interpretation: (i) electron microscopy of negatively stained UMP kinase confirmed an in-plane arrangement of the hexamer (Larquet et al., unpublished); (ii) substitution of D174 with asparagine abolishes the interaction of UMP kinase with Mab 44-2, affecting at the same time the K_m for UMP [8].

Structure/function relationship and analysis of pyrH gene mutants

Analysis of the UMP kinase model led us to identify three cavities surrounding the catalytic site centered around D146. The first (named herein C1) would be formed by the long β1–α1 loop and one face of the helix α2, but this region has not as yet an assigned functional role. However, the equivalent region in yeast aspartokinase was assigned the substrate-binding site according to site-directed mutagenesis studies [29]. The second one (C2) is formed by the other face of the helix α2, the loop connecting helices α3–α4 (71–77) and the loop β4–α6 (amino acids 140–144) preceding the catalytic residue D146 as well as in the vicinity of the unique tryptophan in *E. coli* UMP kinase (W119). Three mutations in this cavity (R62H, D77N, or P141L) affect the enzyme activity and especially the allosteric regulation. Intrinsic fluorescence of W119 showed that this residue is buried in an hydrophobic environment and is sensitive to the presence of the allosteric effector UTP [3] in agreement with its position in our current model. Modelling of the

E. coli glutamate-5-kinase suggested that it adopts a similar molecular organization. The equivalent region in glutamate-5-kinase (amino acids 100–150) would also comprise an allosteric site, since three mutants (D107N, A117V, and E143A) are impaired in the proline-dependent allosteric regulation [30–32]. Noteworthy, C1 and C2 are located near the dimer interface, C1 from one monomer facing C2 from the second one (Fig. 4). The third cavity, C3, is mainly composed of the C-terminus which in carbamoyl-phosphate synthase crystal structure is involved in ADP-binding [28]. Mutants in this region impaired the function of UMP kinase (D174N and D201N) or its stability (L226Q). While, C3 could be involved in ATP binding, C2 is likely involved in UTP recognition and possibly UMP at high concentrations. The proposed structure of the distinct binding sites of the various ligands (ATP versus allosteric effectors) is in agreement with the observed stabilization effects of these ligands [6] as C1 and C2 appeared deeper and more hydrophobic in nature than C3.

Deletion analysis performed on *E. coli* UMP kinase also agrees with the proposed model since this enzyme is composed of only one domain and no deletion would separate the allosteric function from the active site [7]. Similarly, the stability against thermal or chemical (guanidinium–HCl) denaturation of the various mutants tested [8] is in agreement with the proposed position in the structure of the substituted residues. R62, P141, D174, and D168 are external in our model. R62H, D174N, and D168N mutants are as stable as the wild-type UMP kinase from *E. coli*, while the mutant P141Q is more soluble. On the contrary, L226 is buried and the mutant L226Q is insoluble (H. Sakamoto, unpublished results).

Comparative analysis of the sequence of the UMP kinases from other bacteria suggested that our model is also valid for these enzymes.

Acknowledgments

This work was supported by grants from the Institut Pasteur, the Institut National de la Santé et de la Recherche Médicale (U554, the Centre National de la Recherche Scientifique (URA 2185 and UMR 5048), the Genopole Languedoc-Roussillon, and AstraZeneca R&D Boston Inc. We thank Pedro Alzari and Vicente Rubio for many fruitful discussion, Susan Michelson for critical comments, and Régine Lambrecht for excellent secretarial help.

References

[1] D. Dreusicke, P.A. Karplus, G.E. Schulz, Refined structure of porcine cytosolic adenylate kinase at 2.1 Å resolution, J. Mol. Biol. 199 (1988) 359–371.

[2] H.-J. Müller-Dieckmann, G.E. Schulz, The structure of uridylate kinase with its substrates, showing the transition state geometry, J. Mol. Biol. 236 (1994) 361–367.

[3] L. Serina, C. Blondin, E. Krin, O. Sismeiro, A. Danchin, H. Sakamoto, A.-M. Gilles, O. Bârzu, *Escherichia coli* UMP-kinase, a member of the aspartokinase family, is a hexamer regulated by guanine nucleotides and UTP, Biochemistry 34 (1995) 5066–5074.

[4] N. Bucurenci, H. Sakamoto, P. Briozzo, N. Palibroda, L. Serina, R.S. Sarfati, G. Labesse, G. Briand, A. Danchin, O. Bârzu, CMP kinase from *Escherichia coli* is structurally related to other nucleoside monophosphate kinases, J. Biol. Chem. 271 (1996) 2856–2862.

[5] P. Briozzo, B. Golinelli-Pimpaneau, A.-M. Gilles, J.-F. Gaucher, S. Burlacu-Miron, H. Sakamoto, J. Janin, O. Bârzu, Structures of *Escherichia coli* CMP kinase alone and in complex with CDP: a new fold of the nucleoside monophosphate binding domain and insights into cytosine nucleotide specificity, Structure 6 (1998) 1517–1527.

[6] L. Serina, N. Bucurenci, A.-M. Gilles, W.K. Surewicz, H. Fabian, H.H. Mantsch, M. Takahashi, I. Petrescu, G. Batelier, O. Bârzu, Structural properties of UMP-kinase from *Escherichia coli*: modulation of protein solubility by pH and UTP, Biochemistry 35 (1996) 7003–7011.

[7] S. Landais, P. Gounon, C. Laurent-Winter, J.-C. Mazié, A. Danchin, O. Bârzu, H. Sakamoto, Immunochemical analysis of UMP kinase from *Escherichia coli*, J. Bacteriol. 181 (1999) 833–840.

[8] N. Bucurenci, L. Serina, C. Zaharia, S. Landais, A. Danchin, O. Bârzu, Mutational analysis of UMP kinase from *Escherichia coli*, J. Bacteriol. 180 (1998) 473–477.

[9] D. Douguet, G. Labesse, Enhanced threading efficiency through Web-based comparisons and cross-validations, Bioinformatics 17 (2001) 752–753.

[10] H. Munier-Lehmann, A. Chaffotte, S. Pochet, G. Labesse, Thymidylate kinase of Mycobacterium tuberculosis: a chimera sharing properties common to eukaryotic and bacterial enzymes, Protein Sci. 10 (2001) 1195–1205.

[11] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, second ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.

[12] E.A. Raleigh, N.E. Murray, H. Revel, R.M. Blumenthal, D. Westaway, A.D. Reith, P.W.J. Rigby, J. Elhai, D. Hanahan, McrA and McrB restriction phenotypes of some *Escherichia coli* strains and implications for gene cloning, Nucleic Acids Res. 16 (1988) 1563–1575.

[13] H. Munier, A.-M. Gilles, P. Glaser, E. Krin, A. Danchin, R.S. Sarfati, O. Bârzu, Isolation and characterization of catalytic and calmodulin-binding domains of *Bordetella pertussis* adenylate cyclase, Eur. J. Biochem. 196 (1991) 469–474.

[14] S.F. Altschul, T.L. Madden, A.A. Schaeffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.

[15] I. Callebault, G. Labesse, P. Durand, A. Poupon, L. Canard, J. Chomilier, B. Henrissat, J.-P. Mornon, Deciphering protein sequence information through Hydrophobic Cluster Analysis (HCA): current status and perspectives, Cell. Mol. Life Sci. 53 (1997) 621–645.

[16] G. Labesse, J.-P. Mornon, A Tool for Incremental Threading Optimization (TITO) to help alignment and modelling of remote homologs, Bioinformatics 14 (1998) 206–211.

[17] J.A. Cuff, M.E. Clamp, A.S. Siddiqui, M. Finlay, G.J. Barton, JPred: a consensus secondary structure prediction server, Bioinformatics 14 (1998) 892–893.

[18] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.

[19] M.J. Sippl, Recognition of errors in three-dimensional structures of proteins, Proteins 17 (1993) 355–362.

[20] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (1993) 1511–1519.

[21] D. Eisenberg, R. Luthy, J.U. Bowie, VERIFY3D: assessment of protein models with three-dimensional profiles, Methods Enzymol. 277 (1997) 396–404.
[22] P. Tuffery, XmMol: an X11 and motif program for macromolecular visualization and modeling, J. Mol. Graph. 13 (1995) 67–72.
[23] C. Combet, C. Blancher, C. Geourjon, G. Deleage, NPS: network protein sequence analysis, TIBS 25 (2000) 147–150.
[24] M.G. Rossmann, D. Moras, K.W. Olsen, Chemical and biological evolution of nucleotide-binding protein, Nature 250 (1974) 194–199.
[25] G. Labesse, A. Vidal-Cros, J. Chomilier, M. Gaudry, J.-P. Mornon, Structural comparisons lead to the definition of a new superfamily of NAD(P)(H)-oxydoreductases: the single domain Reductases/Epimerases/Dehydrogenases, the "RED" family, Biochem. J. 304 (1994) 95–99.
[26] G. Schluckebier, M. O'Gara, W. Saenger, X. Cheng, Universal catalytic domain structure of AdoMet-dependent methyltransferase, J. Mol. Biol. 247 (1995) 120–166.
[27] A. Marina, P.M. Alzari, J. Bravo, M. Uriarte, B. Barcelona, I. Fita, V. Rubio, Carbamate kinase: new structural machinery for making carbamoyl phosphate, the common precursor of pyrimidines and arginine, Protein Sci. 8 (1999) 934–940.
[28] S. Ramon-Maiques, A. Marina, M. Uriarte, I. Fita, V. Rubio, The 1.5 Å resolution crystal structure of the carbamate kinase-like carbamoyl phosphate synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus*, bound to ADP, confirms that this thermostable enzyme is a carbamate kinase, and provides insight into substrate binding and stability in carbamate kinases, J. Mol. Biol. 299 (2000) 463–476.
[29] M. Arevalo-Rodriguez, I.L. Calderon, S. Holmberg, Mutations that cause threonine sensitivity identify catalytic and regulatory regions of the aspartate kinase of *Saccharomyces cerevisiae*, Yeast 15 (1999) 1331–1345.
[30] L.N. Csonka, S.B. Gelvin, B.W. Goodner, C.S. Orser, D. Siemieniak, J.L. Slightom, Nucleotide sequence of a mutation in the *proB* gene of *Escherichia coli* that confers proline overproduction and enhanced tolerance to osmotic stress, Gene 64 (1988) 199–205.
[31] A.M. Dandekar, S.L. Uratsu, A single base pair change in proline biosynthesis genes causes osmotic stress tolerance, J. Bacteriol. 170 (1988) 5943–5945.
[32] K. Omori, S. Suzuki, Y. Imai, S. Komatsubara, Analysis of the mutant *proBA* operon from a proline-producing strain of *Serratia marcescens*, J. Gen. Microbiol. 138 (1992) 693–699.
[33] P. Gouet, E. Courcelle, D.I. Stuart, F. Metoz, ESPript: analysis of multiple sequence alignments in PostScript, Bioinformatics 15 (1999) 305–308.