Reprinted from Trends in Cell Biology – *February 1993* *Vol. 3, No. 2*

A new ATP-binding fold in actin, hexokinase and Hsc70

Ken C. Holmes, Chris Sander and Alfonso Valencia

One of a cell biologist's favourite occupations is to discover the proteins that perform newly described functions in the cell. Very often lately, this has resulted in the identification of protein families whose related amino acid sequences reflect similar functions, but can proteins with totally unrelated sequences have similar structures and functions? In this review, Ken Holmes, Chris Sander and Alfonso Valencia describe the structural similarities between three well-known proteins that have no readily detectable primary sequence similarities but for which X-ray crystallography has revealed very similar structures. A comparison of their structures provides insights into their common mechanisms of action and into protein evolution, and has been used to detect related proteins in sequence data bases.

Although a fundamental tenet of molecular biology is that amino acid sequence determines protein structure, we are still unable to predict a protein's structure from its sequence unless the sequence is clearly related to that of a protein of known structure. Therefore, it is still a surprise when proteins with apparently very different sequences turn out to have similar three-dimensional structures, as has recently been found for actin, the N-terminal (Nt) domain of the heat shock cognate protein Hsc70, and hexokinase.

Let us take a look at these three protein structures one by one. Actin is a well-known muscle and cell skeleton protein. The all-atom structure of monomeric actin in a complex with DNase I was solved by Kabsch *et al.*[1] in 1990 by X-ray crystallography (Fig. 1). By combining this with X-ray diffraction data from actin fibres, Holmes *et al.*[2] were able to reconstruct the likely actin filament as it occurs *in vivo*, complete with details of actin–actin contacts (Fig. 2). The crystal structure of the monomer was solved with either ATP or ADP bound in the active site in a complex with Ca^{2+}. The ADP-bound form was produced from the ATP-bound form by slow hydrolysis of ATP in the crystal. Surprisingly, the actin–ADP and actin–ATP crystal structures are very similar. Does this mean that there are no large conformational changes that accompany ATP hydrolysis *in vivo*? Not necessarily. The close packing of the molecules in the crystal and the binding of DNase I may inhibit such changes. Also, there is indirect evidence for a conformational change from the differences in actin–actin association accompanying the hydrolysis of the γ-phosphate of ATP[3].

Hsc70 is a heat shock protein 'cognate', a chaperone that helps keep other proteins from folding prematurely or inappropriately. The crystal structure of the Hsc70 Nt domain, including the nucleotide-binding site, was solved by McKay *et al.*[4] in the presence of Mg^{2+} and ATP or ADP (Fig. 1). The part missing from the crystal, the C-terminal (Ct) domain, may contain the main peptide-binding site, but this has yet to be proven[5]. A full comparison of the ADP- and ATP-binding forms of Hsc70 has not yet been published, so the structural comparisons reviewed here are based only on the ADP-bound form[6,7].

Hexokinase is a major metabolic enzyme. The crystal structures of two isoforms of hexokinase were solved by T. Steitz and colleagues more than 15 years ago[8,9]. At that time the amino acid sequences were not known and therefore the most likely residues were fitted into the density maps. Now that the sequences are available, these two isoforms are known to be identical apart from the first 15 residues. Our knowledge of their ATP- and sugar-binding sites is incomplete. The structure of isoform A was solved with glucose, but not ATP, present and the structure of isoform B was solved in the presence of a sugar analogue lacking biological activity (Fig. 1). The structure of isoform B was also solved in the presence of an ATP analogue (8-bromo-AMP) without sugar[10]. However, the reported orientation of the base appears to be the wrong way around compared to actin, and attempts to bind ATP in hexokinase crystals caused these to shatter. Because of these technical difficulties, it has been impossible so far to identify clearly all residues in the active site(s), although some inferences can be drawn from a comparison of the three hexokinase structures[8–12].

Structural comparison

To the crystallographers' great surprise, the domain organization of all three proteins turned out to be remarkably similar. The similarity between actin and Hsc70 became apparent as their structures were determined and a detailed comparison has been published[6]. Actin and Hsc70 are, in fact, much more similar to one another than either is to hexokinase[7], but the domain organization of all three proteins is remarkably alike.

In the following discussion we have used the nomenclature given for Hsc70 (Ref. 6). The structure is organized into two domains, I and II, with the ATP situated in the interdomain cleft. Each of the domains is divided into two subdomains: a and b. The two largest subdomains, Ia and IIa, each have five-stranded β-sheets joined by two α-helices.

Ken Holmes is at the Max Planck Institut für Medizinische Forschung, Jahnstrasse 29, D-6900 Heidelberg, Germany; and Chris Sander and Alfonso Valencia are at the European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900 Heidelberg, Germany.

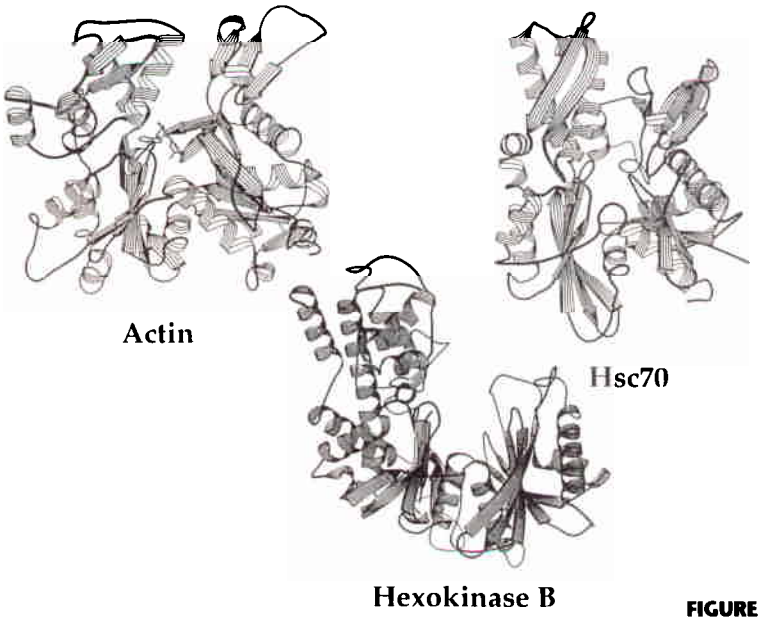


FIGURE 1

Ribbon drawings of actin, Hsc70 and hexokinase isoform B. The protein folds in the lower right and lower left subdomains are topologically identical. ATP is bound in a deep cleft at the centre of the molecule that is lined by residues conserved in all three structures. The external loops and upper subdomains of the three structures are less similar.

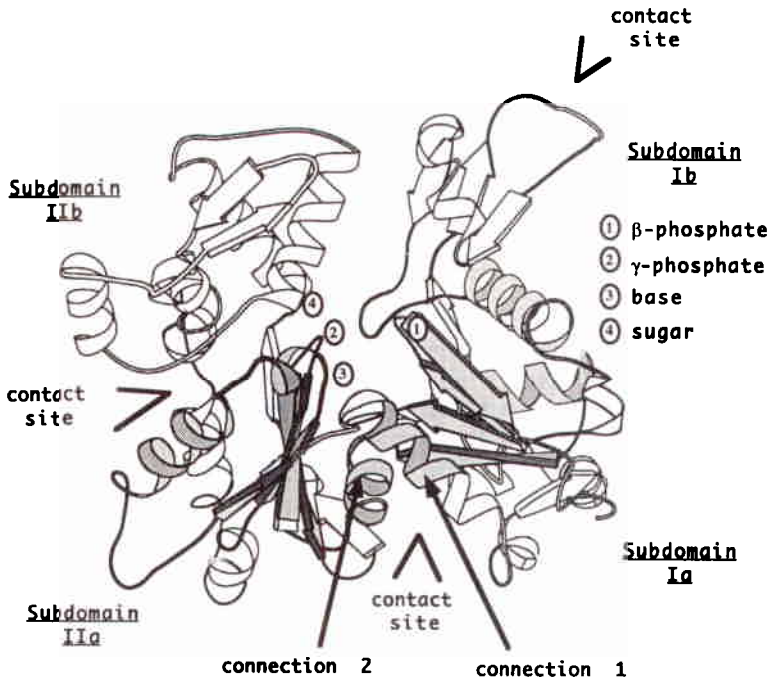


FIGURE 2

Ribbon drawing of an actin monomer. The core of symmetrical subdomains of the structure (shaded) are structurally similar in Hsc70 and hexokinase (the structure is shown in the same orientation as in Fig. 1). The regions that bind different parts of ATP (labelled as β-phosphate, γ-phosphate, base and sugar) and those that make the interdomain connection (labelled as connection 1 and 2) are structurally conserved between the three families. Subdomains are labelled from Ia to IIb in Hsc70 notation[4]. 'Contact sites' are the main sites of interaction of actin with other proteins or other actin monomers[2]. By analogy, these regions in Hsc70 may also contact protein substrates.

They have the same topography and are related by pseudo twofold symmetry[6,7] (Figs 1 and 2). Remarkably, this symmetry also includes the phosphate-binding loops in each domain.

Subdomain Ib (top right in Fig. 1) is an excursion from Ia and is virtually absent from hexokinase. Hsc70 has an extra segment inserted in Ib compared with actin. Subdomain IIb (top left in Fig. 1) is an excursion from IIa and is very similar in actin and Hsc70, except that the β-strands are longer in Hsc70 and Hsc70 lacks one α-helix found in actin. Subdomain IIb in hexokinase has a similar topography but the detailed structure is quite different.

The most obvious difference in the crystal structures of the three proteins is the angle between domains I and II. In actin and Hsc70, the two domains are closed, whereas in both isoforms of hexokinase there is a large opening between the domains (Fig. 1). This difference points to a very interesting possibility: hinged rotation of the two domains, apparently in response to the presence of bound nucleotide (see below).

Given the unexpected similarity between these proteins, what are the key common elements and what are the key differences? The distribution of the structurally equivalent residues of actin, Hsc70 and hexokinase, after rotating the domains into optimal superimposition, is shaded in Fig. 2. The 'common core' includes two cross-over segments that form the connection between subdomains Ia and IIa, two phosphate-binding loops, a base-binding pocket and some residues involved in ribose binding. Outside this core the three structures are quite different with respect to size and disposition of the external regions. There are three loops of particular functional importance. First, a loop inserted in the middle of a long α-helix that connects subdomains IIa and IIb in actin is involved in monomer–monomer binding[2]; second, a loop following the third β-strand in subdomain Ia of hexokinase forms part of the sugar-binding site[13]; and third, a loop in the middle of the outermost (upper in Fig. 1) β-strand of domain Ia in Hsc70 of as yet unknown function[4,6].

The discovery of this structural similarity gives us substantial predictive power. For example, because the very similar active sites of actin and Hsc70 are known to atomic detail, details of the less well resolved active site of hexokinase can be predicted by analogy. The residues in the ATPase active site (see Box 1) and details of the mechanism of ATP hydrolysis (see Box 2) can also be deduced. Predictions of the protein–protein interaction sites on the surface of Hsc70 based on actin monomer–monomer interactions, and of mechanical movements of the proteins upon ATP binding and hydrolysis, are discussed below.

Protein–protein interaction sites

The chaperone function of Hsc70 (Ref. 14) involves both ATPase activity and peptide binding. We know that the ATPase activity is located in the Nt domain, while the Ct domain is thought to be responsible for the main polypeptide-binding ac-

BOX 1 – THE ATPase ACTIVE SITE

The position of the bound nucleotide in actin and Hsc70 is visible in the crystal structures and the arrangements of ligand-binding residues are very similar. For hexokinase, the nucleotide-binding site is less clear, because neither crystal form is in a 'closed' conformation and because of ambiguities in the sequence assignment in the crystal structures. Here, we therefore extrapolate from the known binding sites in actin and Hsc70 to infer the nucleotide-binding site in hexokinase. This requires the domains of hexokinase to be brought into concordance with the domains in actin or Hsc70 to generate a putative closed form of hexokinase (see main text).

Residues involved in phosphate binding

The figure shows schematically the residues involved in binding the nucleotide in actin and Hsc70 (and, by inference, also in hexokinase). The β- and γ-phosphates are bound to symmetrically disposed β-hairpin loops from domains Ia and IIa. The divalent ion forms a bidentate complex with the β- and γ-phosphate groups, which are related by a pseudo twofold axis. Each of the hairpin loops has three amide groups in the turn oriented to coordinate a phosphate oxygen. In hexokinase, the structure of the hairpin loop in domain IIa seems to be slightly different as the loop is one residue shorter. This may reflect a slightly different binding geometry for the nucleotide in hexokinase. The structural difference is reflected in a sequence difference in the γ-phosphate-binding loop (D-X-G-X-G-X-X-P in actin and Hsc70, versus G-T-G in hexokinase, where X is any residue and P is a polar residue)[7].

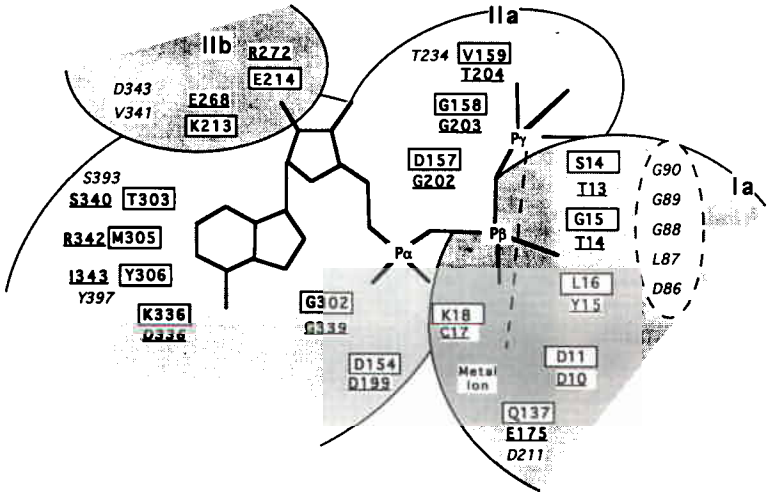
On hydrolysis of the γ-phosphate there are no marked movements in the crystal (although there may be in solution). The metal ion remains complexed to the β-phosphate. The cation-binding site forms a hydrophilic hole in the protein interior between subdomains Ia and IIa. Other groups that coordinate the cation appear to be Gln137 in actin and the corresponding residues Glu175 in Hsc70 and Asp211 in hexokinase.

The adenosine-binding pocket

The adenosine group is bound in a hydrophobic pocket between subdomains IIa and IIb. The relevant part of subdomain IIb consists of a helix-turn-helix motif ascending from the outermost strand of the five-stranded β-sheet in subdomain IIa where the residues at the end of the descending helix form the top of the base-binding pocket. This pocket is formed by the aliphatic chain of Glu214 in actin, and by the aliphatic chain of Arg272 in Hsc70; both of these residues are held in place by salt bridges to residues one turn back along the descending α-helix. The back of the pocket is formed by hydrophobic residues (Tyr306 in actin, Ile343 in Hsc70), which are stabilized by the ascending α-helix. The bottom of the pocket is formed by one turn of a tightly wound (3_{10}) helix, joining the second strand of the β-strand of subdomain IIa with the ensuing α-helix. Contacts are made with the main-chain atoms and for the formation of this structure at least one correctly positioned Gly residue is essential. The front of the pocket is open to the solvent so that, for example, the Asn6 position on the base in actin is available for chemical modification.

Hexokinase differs in detail

While the nucleotide-binding sites in actin and Hsc70 are very similar, there appear to be differences in hexokinase. When subdomains IIa of actin and hexokinase are superimposed, the 3_{10} helix of hexokinase protrudes into the actin base-binding pocket. However, it is possible to accommodate the ATP coordinates of actin in the hexokinase ATP-binding pocket if the ATP molecule is rotated through ~10°, raising the base by about 1.5 Å and lowering the phosphates by the corresponding amount. Alternatively, binding of the nucleotide may be accompanied by a small structural rearrangement of the 3_{10} helix and a 2 Å movement of the following α-helix to make it like that in actin. A refined crystal structure of the 'closed' form of hexokinase with ATP or ADP bound would be required to clear up such ambiguities.



Residues that are in contact with ATP and the metal ion in Hsc70 and actin[1,4,6], and a tentative assignment of the corresponding residues in hexokinase[10,11]. Actin residues are shown in boxes, Hsc70 residues are underlined and hexokinase residues are in italics. Note that the γ-phosphate also contacts the metal ion (dashed line).

tivity[5]. So the two domains have to be coupled in some way. Can we determine how the Nt and Ct domains interact when only the structure of one of the two is known? Fortunately, we have extensive information about regions of protein–protein interactions between actin monomers[2] (Fig. 2). Some of these regions have structural equivalents in Hsc70 (Refs 6,7) and, remarkably, contain clusters of solvent-exposed hydrophobic residues that are conserved in the many known sequences of the Hsc70 family.

By analogy, these residues are excellent candidates for functional interaction of Hsc70 with other proteins. They are located mainly in the

BOX 2 – THE ATPase CATALYTIC MECHANISM

Hexokinase

The structure of hexokinase B (the open form) has been determined with a bound glucose analogue (N-orto-toluoylglucosamine)[8]. If we model the putative closed form of hexokinase on that of actin and Hsc70 by rotating the two domains by 37° about a point between subdomains Ia and IIa and then combine this 'closed' structure with the ATP coordinates from actin, we arrive at a model of the active site of hexokinase. The γ-phosphate is about 3.5 Å from the O6 glucose oxygen in a geometry that would allow in-line attack of the γ-phosphate on the glucose oxygen. In addition, as has been pointed out by Steitz *et al.*[11], D211 is well positioned to act as a general base for the catalysis. This model requires experimental verification.

Hsc70 and actin

A possible catalytic mechanism for Hsc70 can be obtained by imposing the coordinates of ATP from actin onto Hsc70 (because only the ADP form of Hsc70 is available). This shows the γ-phosphate to be 3.5 Å from a water molecule that can be polarized by Asp206 (which could act as a proton acceptor or general base) and held in place by a hydrogen bond from Thr204. In fact, the threonine hydroxyl group comes very close to the γ-phosphate. Moreover, the corresponding threonine in DnaK is a target for autophosphorylation[29], which would be a natural consequence of its proximity to the γ-phosphate and thus lend experimental support to the extrapolation of the actin ATP coordinates to Hsc70. In the case of actin, a catalytic mechanism has yet to be described. No water molecule can be seen in the active site, nor is there a clear candidate for a general base.

lower cleft formed at the junction between subdomains Ia and IIa, and, as they are not far from the ATP-binding site (Fig. 3), they are poised to couple their function to ATP use. These Hsc70 residues are Tyr149, Asn151, Asp152, Glu218 and Val219. Their structural equivalents in actin are involved in the actin–actin contact in the filament[2]. This structural analogy would be confirmed by altered chaperone activity as a result of point mutations in these residues in Hsc70 (or in its *Escherichia coli* homologue DnaK).

Interdomain hinge motion – a mechanism for sensing the ATP state?

An important hypothesis about the mechanism of action of hexokinase is that binding of a sugar molecule leads to the formation of an ATP-binding site[12,15], the so-called 'induced-fit' hypothesis. The structural analysis suggests that the ATP-binding site is created by the closing of domains I and II about a hinge. Indeed, the crystal structure of hexokinase in the presence of an unproductive sugar analogue is different from the structure in the presence of the glucose substrate[8,9,15]. The binding of ATP, in addition to the already bound sugar, is thought to produce a second movement of the domains, which closes the ATP-binding site[10] (Fig. 5).

This second movement of the domains, after the correct positioning of ATP, remains to be demonstrated experimentally. But there is already some supporting evidence: (1) crystals of hexokinase with sugar crack when ATP is added, presumably because major structural changes cannot be accommodated[11,15]; (2) the crystal structures of actin and Hsc70 are more closed than the 'closed' form of hexokinase (A form); and (3) when the position of ATP in the 'open' form of hexokinase (B form)[12] is imposed on the 'closed' form, the β-phosphate is still far from the binding residues in subdomain Ia (Ref. 12), so further closing is necessary for proper ATP binding. What emerges is that ATP plays the role of a structural bridge between domains. Binding ATP tips the energy balance in favour of a closed structure that is stabilized by numerous contacts[10–12].

By analogy to the 'induced-fit' movement in hexokinase, it is plausible that actin and Hsc70 also undergo a structural transition between two states, namely a fully closed structure with ATP bound and an open structure that allows the exchange of ADP for new ATP. The hinge points for the open–closed transition are located in the two

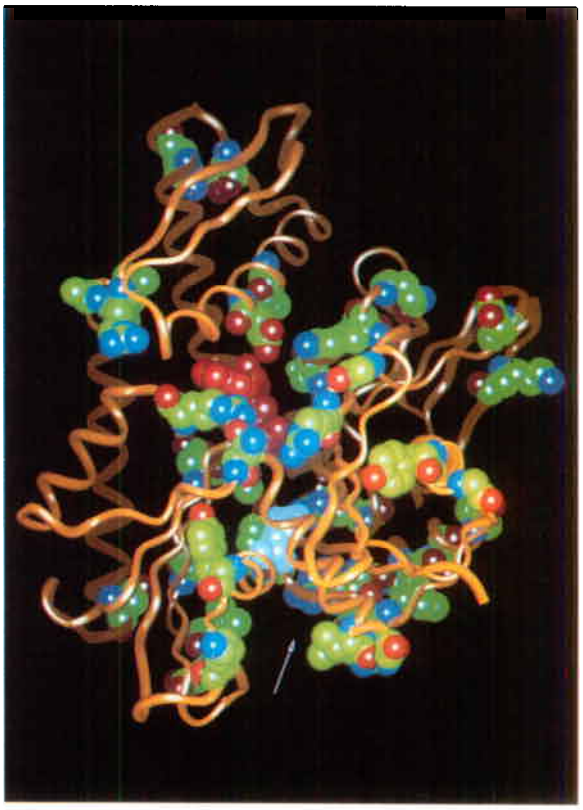


FIGURE 3

Predicting external binding sites on Hsc70. 'Conserved and exposed' residues are candidates for protein–protein and protein–substrate interaction. Here, residues conserved in the Hsc70 family, including *E. coli* DnaK, with a solvent-accessible surface area greater than 50 Å^2 are coloured according to atom type (dark green, carbon; red, oxygen; blue, nitrogen; light green, sulphur) on a ribbon tracing of the protein backbone; ADP is shown in red. Remarkably, many of these residues are located in the lower cleft (arrow) between subdomains Ia and IIa. This cleft extends from near the key residue E175 (light blue, lower middle) to the 'back' of the molecule. These residues are likely candidates for contacts with the Ct domain of Hsc70, the peptide substrate or with other proteins.

'connecting' regions (Fig. 2). A number of small residues in these regions that are conserved in all three families appear to form the key contact points between the two α-helices at the domain interface[7]. The hinged movement can be simulated on the computer by rotating domain I relative to domain II by 37° about an axis defined by two conserved key hinge residues. The predicted hinge residues are Gly372 and Ala182 in Hsc70 and Gly342 and Ala144 in actin.

The proposed movement of the domains has a striking consequence. The 'winged' architecture of the structure (Fig. 2) provides an ATP-dependent mechanical coupling between the upper and lower parts. As the upper cleft opens, separating subdomains IIa and IIb, the lower cleft between subdomains Ia and Ib closes (Figs 4 and 5). Conversely, when the upper cleft is closed as a result of binding ATP, the lower cleft is opened, making its residues accessible from the outside. Could another protein or another domain bind in the lower cleft in this state? Indeed, it is known[16,17] that actin and Hsc70 bind their target proteins more strongly in the ATP-bound form. How would a bound protein dissociate? It is plausible that after ATP hydrolysis the energy balance would favour opening the upper cleft to release ADP, and closing the lower cleft to dissociate any molecule bound in the lower cleft. Can the hinge mechanism help explain the chaperone function of Hsc70? One possibility is that the polypeptide substrate binds to the lower cleft in the ATP state and that a conformational change is induced in the substrate, keeping it from participating in folding. Alternatively, the cleft may be the binding site for the Ct domain and the effect on the bound polypeptide substrate may be indirect.

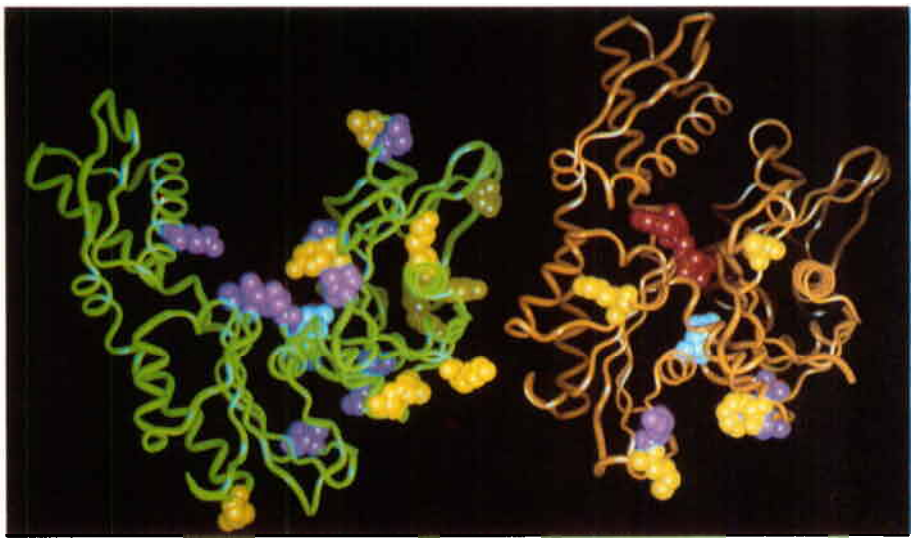


FIGURE 4

Hinged motion of domains in Hsc70. (right) Observed 'closed' Hsc70 structure with bound nucleotide. (left) Proposed 'open' form, without nucleotide. The 'open' form of Hsc70 was obtained by superposition of subdomains Ia and IIa of Hsc70 onto the corresponding domains in the 'open' form of hexokinase. By comparing the open and the closed forms, a set of residues that change their exposure on the surface by more than 30 Å^2 can be identified. Some of these (highlighted on the left) are more exposed in the nucleotide-free, open form. Others (highlighted on the right) become more exposed in the closed form. Of these, the conserved residues (violet) are likely to have functional importance (nonconserved residues are coloured yellow). The conserved residues in the upper cleft become accessible in the open form, whereas the conserved residues in the lower cleft become accessible in the closed form. Residue Glu175 is light blue and ADP is red.

The proposed hinge mechanism coupled to ATP hydrolysis is simple and fascinating, but as yet unproven. However, experimental tests are in progress. For example, mutations are being made in the critical residue Gln137 in actin (Glu175 in Hsc70), located in the first connecting region and involved in binding Ca^{2+}. Loss or gain of an interaction in this position can provide a direct link between the ion–ATP-binding region in the upper cleft and the conserved surface residues in the lower cleft.

Pattern search – are FtsA, MreB and StbA bacterial ATPases?

There is less than 20% sequence identity between actin, Hsc70 and hexokinase and therefore these similarities cannot be detected by standard sequence search methods[6,7]. Yet we know that the sequences contain common information specific for this fold. The challenge is to capture and make use of this common information. A new generation of structure-derived sequence pattern methods (see below) promise progress towards solving this important problem (Refs 18,19; M. Scharf, Diploma thesis, University of Heidelberg, 1989). In one case[18], the structure-derived sequence pattern of actin picks up the sequence of Hsc70 and, in another[19], the sequence of Hsc70 picks up the structure of hexokinase as the most likely fold, but much more work has to be done to improve these methods before all remote homologues can be detected.

Common three-dimensional structures can be used to align what cannot be aligned by sequence methods alone. After bringing the three sequence families into proper register and generalizing sequence information at each position, a common sequence pattern becomes apparent that is characteristic for this fold[7]. The pattern corresponds to five structurally conserved regions: the two phosphate-binding loops, the base-binding region and the two 'connecting' regions. When searching the entire sequence data base with this pattern, not only actin, Hsc70 and hexokinase are detected, but also other candidates, which probably have a similar ATP-binding domain[7]. These include a large family of sugar kinases related to hexokinase and

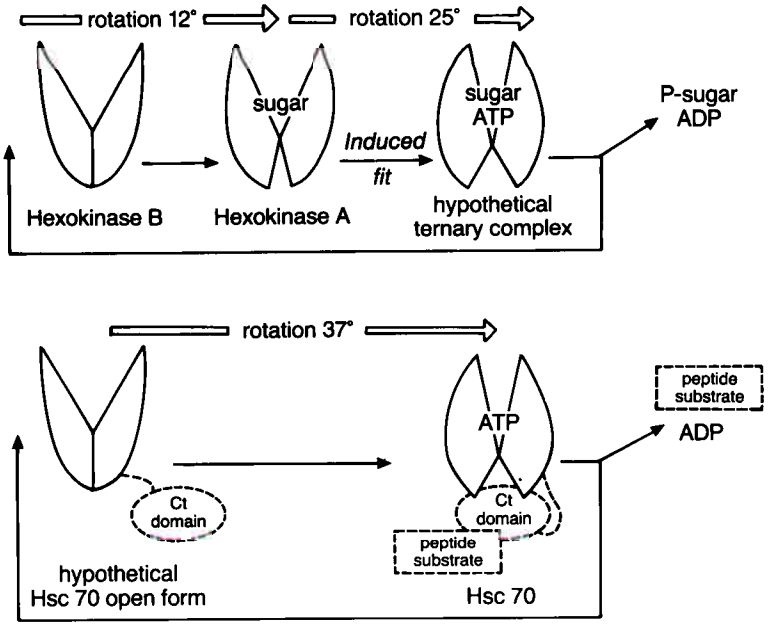


FIGURE 5

Coupling of domain movement and ATP hydrolysis. Since structural information is available for the closed form of actin and Hsc70 and two open forms of hexokinase, the interdomain motion can be modelled for all three families. (a) The structure of hexokinase changes from the open form (hexokinase B) to an intermediate closed form (hexokinase A) when sugar binds, and to the closed form (hypothetical ternary complex) when nucleotide is bound. (b) The structure of Hsc70 also changes from an open nucleotide-free form (left) to the closed form with a bound metal-ion–ATP complex (right) to which the Ct domain or a peptide substrate could bind. Actin probably undergoes a similar interdomain motion. The closing of the upper cleft is likely to be coupled to the opening of the lower cleft in a quasi-mechanical manner because of the strong cohesion of the two halves. The model predicts the following angles for the movement of the domains: 12° between the open and intermediate closed form of hexokinase, 25° between the intermediate closed form and the closed form of hexokinase, and thus 37° between the open form and the closed form of Hsc70 or actin. The actual geometry is slightly more complicated as the rotation is accompanied by a small translation (2 Å).

clearly different from other families of sugar kinases[20] (prokaryotic gluco-, ribulo-, glucono-, xylulo-, glycero-, fructo-, rhamno- and fucokinases) and a group of prokaryotic cell cycle proteins (FtsA[21], MreB[22] and StbA[23]). The best characterized of these is FtsA protein, which is important for the later stages of cell division[21] and probably has a direct role in the formation of the septum complex[24]. Work is in progress to verify that FtsA is indeed an ATPase (M. Vincente, pers. commun.).

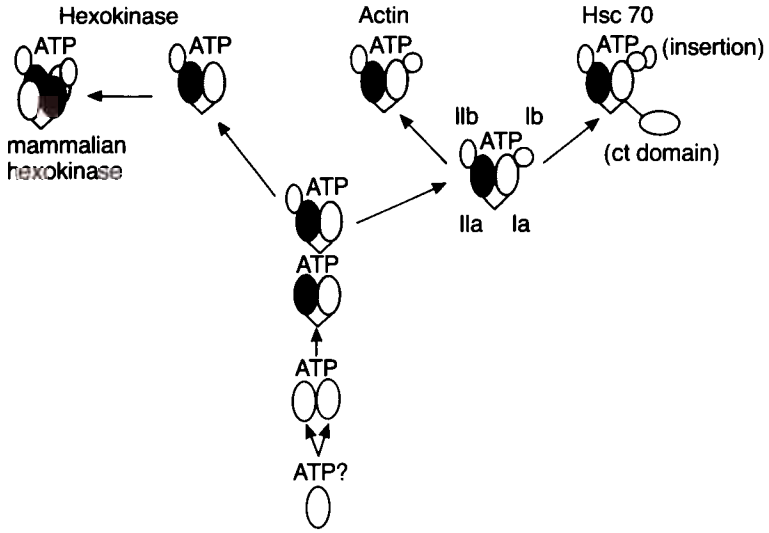


FIGURE 6

Divergent evolution of the ATPase domains. In this proposed evolutionary scenario, the gene for an ancestral ATP-binding monomer (bottom) duplicated and fused, and the sequences encoding domain Ia (right oval) and domain IIa (left oval) then diverged. The structural difference between domains Ia and IIa is more prominent in the hexokinase family than in the actin and Hsc70 family, as indicated by darker shading (left). The small subdomain IIb (top left of the molecule) evolved earlier than subdomain Ib, as it is common to all three families. Subdomain Ib was a later acquisition of the actin/Hsc70 branch. A further gene duplication occurred in the mammalian hexokinase family[26,27]. This scheme finds support in the structural similarity of the three proteins and in the perfect symmetry of the two larger subdomains (Ia and IIa).

Convergent versus divergent evolution

Is there a common evolutionary ancestor of actin, Hsc70 and hexokinase, or has convergent evolution resulted in the same fold and similar ATP-binding sites?

The arguments in favour of divergent evolution are strong. If the striking similarity of subdomains Ia and IIa in each of the three structures is a result of gene duplication, then duplication of a common ancestor is much more likely than independent duplication in each of the three families. Moreover, the intron–exon structures of actin and Hsc70 are similar, and in both cases an intron is present between domains I and II (M. M. Ladjimi, pers. commun.). Remarkably, the *fes* oncogene product contains a sequence fragment equivalent to a single domain I of actin[25], and this is probably the result of 'genetic hijacking'. Finally, the structural similarity covers not only the ATPase site but also the interdomain hinge region and these similarities are reflected, albeit weakly, at the sequence level. These similarities are much greater than that observed in the classic case of convergent evolution: trypsin and subtilisin. These two classes of proteases merely share a few active site residues while their overall folds are different.

We cannot reliably reconstruct evolutionary history, but the following scenario appears likely (Fig. 6). At the root of the tree may have been a single-domain protein. It is plausible that this ancestor may already have had some ATP-binding ability, since hexokinase in the 'open' conformation binds ATP in only one of the domains[10]. A two-domain protein might then have been created by gene duplication, and its ATP-binding ability improved.

Subsequently, actin and Hsc70 may have diverged one way, after a further small adaptation in domain II, while hexokinase may have diverged another way, with somewhat larger changes in domain II, particularly in the β-phosphate-binding loop[7]. A further gene duplication and specialization of function may then have given rise to the hexokinase of higher eukaryotes, with one regulatory and one catalytic subunit[26,27].

More refined adaptation is apparent in the subdomains Ib and IIb, which are more varied. Probably subdomain IIb, present on all branches of the evolutionary tree and with residues that participate in ATP binding, is the more ancient of the two. Subdomain Ib is probably a more recent insertion because it is not present in hexokinase and contains an additional inserted fragment in Hsc70 compared with actin.

The divergence of actin, Hsc70 and hexokinase is most probably a very ancient event[28]. Independently, many other structural types of mononucleotide-binding sites have evolved, such as those of Ras, adenylate kinase, cAMP-dependent protein kinase, recA and others.

Within our limited knowledge of protein structures, the hexokinase-like ATPase domain is one of the most beautiful examples of adaptation of a basic protein fold to different functional tasks, using duplication and variation of an ancient folding theme. It will not be a surprise if genome sequencing uncovers many more protein families that exploit this basic unit to functional advantage.

References

1 KABSCH, W., MANNHERZ, H. G., SUCK, D., PAI, E. F. and HOLMES, K. C. (1990) *Nature* 347, 37–44
2 HOLMES, K. C., POPP, D., GEBHARD, W. and KABSCH, W. (1990) *Nature* 347, 44–49
3 CARLIER, M. F. (1989) *Int. Rev. Cytol.* 115, 139–170
4 FLAHERTY, K. M., DELUCA-FLAHERTY, C. and MCKAY, D. B. (1990) *Nature* 346, 623–628
5 CHAPPELL, T. G., KONFORTI, B. B., SCHMID, S. L. and ROTHMAN, J. E. (1987) *J. Biol. Chem.* 262, 746–751
6 FLAHERTY, K. M., MCKAY, D. B., KABSCH, W. and HOLMES, K. C. (1991) *Proc. Natl Acad. Sci. USA* 88, 5041–5045
7 BORK, P., SANDER, C. and VALENCIA, A. (1992) *Proc. Natl Acad. Sci. USA* 89, 7290–7294
8 ANDERSON, C. M., STENKAMP, R. E. and STEITZ, T. A. (1978) *J. Mol. Biol.* 123, 15–33
9 ANDERSON, C. M., STENKAMP, R. E., MCDONALD, R. C. and STEITZ, T. A. (1978) *J. Mol. Biol.* 123, 207–219
10 SHOHAM, M. and STEITZ, T. A. (1980) *J. Mol. Biol.* 140, 1–14
11 STEITZ, T. A., SHOHAM, M. and BENNETT, W. S. (1981) *Philos. Trans. Roy. Soc. London Ser. B* 293, 43–52
12 BENNETT, W. S. and STEITZ, T. A. (1978) *Proc. Natl Acad. Sci. USA* 75, 4848–4852
13 ARORA, K. K., FILBURN, C. R. and PEDERSEN, P. C. (1991) *J. Biol. Chem.* 266, 5359–5362
14 LEWIS, M. J. and PELHAM, H. R. (1985) *EMBO J.* 4, 3137–3143
15 BENNETT, W. S. and STEITZ, T. A. (1980) *J. Mol. Biol.* 140, 211–230
16 GETHING, M-J. and SAMBROOK, J. (1992) *Nature* 355, 33–45
17 SELVE, N. and WEGNER, A. (1986) *J. Mol. Biol.* 187, 627–631
18 BOWIE, J. U., LUETHY, R. and EISENBERG, D. (1991) *Science* 253, 164–170
19 OUZOUNIS, C., SANDER, C., SCHARF, M. and SCHNEIDER, R. *J. Mol. Biol.* (in press)
20 BORK, P., SANDER, C. and VALENCIA, A. *Protein Sci.* (in press)
21 DONACHIE, W. D. *et al.* (1979) *J. Bacteriol.* 140, 388–394
22 DOI, M. *et al.* (1988) *J. Bacteriol.* 170, 4619–4624
23 TABUCHI, A. *et al.* (1988) *J. Mol. Biol.* 202, 511–525
24 PLA, J., DOPAZO, A. and VICENTE, M. (1990) *J. Bacteriol.* 172, 5097–5102
25 NAHARRO, G., ROBBINS, K. C. and REDY, E. P. (1984) *Science* 223, 63–67
26 EASTERBY, J. S. (1971) *FEBS Lett.* 18, 23–26
27 SCHWAB, D. and WILSON, J. E. (1989) *Proc. Natl Acad. Sci. USA* 86, 2563–2567
28 DOOLITTLE, R. F. (1992) *Protein Sci.* 1, 191–200
29 ZYCLICZ, M., LEBOWITZ, J. H., MCMACKEN, R. and GEORGOPOULUS, C. (1983) *Proc. Natl Acad. Sci. USA* 80, 6431–6435

Acknowledgements

We thank Dave McKay for the coordinates of Hsc70, and the EC Bridge program for financial support.