Identification of sequence similarity between 60 kDa and 70 kDa molecular chaperones: evidence for a common evolutionary background?

Ana I. FLORES and José M. CUEZVA*

Departamento de Biología Molecular, Centro de Biología Molecular 'Severo Ochoa', Consejo Superior de Investigaciones Científicas, Universidad Autónoma de Madrid, 28049 Madrid, Spain

Recent findings support the premise that chaperonins (60 kDa stress-proteins) and α -subunits of F-type ATPases (α -ATPase) are evolutionary related protein families. Two-dimensional gel patterns of synthesized proteins in unstressed and heat-shocked embryonic *Drosophila melanogaster* SL2 cells revealed that antibodies raised against the α -subunit of the F1-ATPase complex from rat liver recognize an inducible p71 member of the 70 kDa stress-responsive protein family. Molecular recognition of this stress-responsive 70 kDa protein by antibodies raised against the

F1-ATPase α -subunit suggests the possibility of partial sequence similarity within these ATP-binding protein families. A multiple sequence alignment between α -ATPases and 60 kDa and 70 kDa molecular chaperones is presented. Statistical evaluation of sequence similarity reveals a significant degree of sequence conservation within the three protein families. The finding suggests a common evolutionary origin for the ATPases and molecular chaperone protein families of 60 kDa and 70 kDa, despite the lack of obvious structural resemblance between them.

INTRODUCTION

The identification of reliable sequence identity between apparently unrelated protein families is used as a tool for the prediction of protein evolution, structure and function [1–3]. However, a problem that arises in the search for homologous proteins is the detection of similar but distantly related sequences [4]. Recently, evidence has been provided that sequences of the α -subunit of F-type ATPases (α -ATPase) are evolutionarily related to the chaperonins [5,6]. Furthermore, it has been shown that antibodies raised against synthetic peptides derived from the rat liver α -ATPase sequence, highly similar to peptides within the chaperonin sequences, recognized stress responsive proteins of both the 60 kDa and 70 kDa protein families [6,7]. The α -subunit of F-type ATPases has been shown to play some other functional roles in the eukaryotic cell, in addition to its involvement in the synthesis of ATP [8–10]. In fact, the α -subunit has been shown both to promote the correct folding/assembly of the β -subunit of the ATPase of *Rhodospirillum rubrum* chromatophores [8] and to be involved in the import and processing of mitochondrial precursor proteins [9], two functional activities that characterize the molecular chaperones.

Molecular chaperones of both the 70 kDa and 60 kDa family of stress proteins are involved in similar functional roles during protein biogenesis [11–13]. In addition, both protein families bind and hydrolyse ATP. The crystal structure of the N-terminal ATPase fragment of the 70 kDa heat-shock cognate protein (hsc70) [14] as well as the secondary structure topology of the C-terminal peptide-binding domain of hsc70 [15] have been resolved. Similarly, the three dimensional structure of the bacterial GroEL chaperonin has been reported [16]. In line with the low overall sequence similarity [11] detected by standard methods of sequence analysis, the two protein families reveal no obvious structural resemblance [14–16]. However, it is also possible, considering the analogous functional roles played by these two protein families, that they might be related in an evolutionary sense. A hypothesis in this regard has already been put forward [17]. In such a situation, the identification of significant sequence

similarity between both protein families could provide an indication of evolutionary relatedness. Herein, we provide evidence to support this contention.

EXPERIMENTAL

Cell culture, heat-shock treatment and protein synthesis rates in embryonic *Drosophila melanogaster* (SL2) cells

SL2 cells were plated in M3 medium supplemented with 10% (v/v) fetal-calf serum and grown to semiconfluence on 60-mm plates at 24 °C. For the heat-shock treatment, plates were placed in a culture chamber at 37 °C for 60 min [6]. To measure rates of protein synthesis, 100 μ Ci of [³⁵S]methionine (1200 Ci/mmol) were added to the incubation mixture and allowed to incorporate for 60 min at either 24 °C or 37 °C. Cells were pelleted, washed twice with PBS, and resuspended in 600 μ l of a 20 mM phosphate buffer/150 mM NaCl/1% (v/v) Triton X-100, pH 7.0, and the following protease inhibitors; 2 μ g/ml leupeptine, 1% (w/v) aprotinin, and 1 mM phenylmethylsulphonyl fluoride. Resuspended cells were freeze-thawed three times and centrifuged for 10 min at 100 000 g. Protein concentrations were determined with the Bio-Rad protein assay.

Two-dimensional (2-D) gel electrophoresis

Analytical 2-D gel electrophoresis [18] to fractionate SL2 cellular proteins was performed using the immobiline DryStrip Kit developed by Görg et al. [19] as indicated in the supplier's instructions (Pharmacia-LKB). Briefly, gels were rehydrated in a solution containing 8 M urea, 0.5% (v/v) Triton X-100, 1.3 mM dithiothreitol and 0.005% (v/v) Pharmalyte 3-10. The first dimension separation (IEF) was performed in 110 $\times$ 3 mm polyacrylamide gels containing an immobilized pH 4–7 gradient, at 300 V for 3 h and 2200 V for 15 h. When a different gradient (pH 3–10.5) was used, the run was carried out at 300 V for 1 h and 1400 V for 17 h. The second-dimension separation (SDS) was in 12.5% polyacrylamide gels (14 $\times$ 13 cm), and at room

Figure 1 Induction of the heat-shock response in cultured embryonic *D. melanogaster* cells: two-dimensional profiles of synthesized proteins

Embryonic *D. melanogaster* SL2 cells were grown at 24 °C (panel **A**). Other cells were subjected to heat-shock treatment at 37 °C for 1 h (panels **B** and **C**). After incorporation of [³⁵S]-methionine for 1 hour, the cells were processed as described in the Experimental section. Equal amounts of radioactive cellular proteins (7×10^5 c.p.m.) were separated by two-dimensional PAGE (pH 4–7 gradient) and processed for fluorography (panels **A** and **B**) or Western blotting with a 1:200 dilution of a rabbit polyclonal serum against DR-1 peptide (panel **C**). The relative synthesis rate of p56 was essentially the same at 24 °C and 37 °C (654 ± 145 and 687 ± 32 arbitrary units respectively). The relative synthesis rate of p71 was significantly increased at 37 °C (18 ± 2 and 912 ± 138 arbitrary units for 24 °C and 37 °C respectively). For each experiment, the synthesis rate of p56 and p71 were normalized for the actin densitometric value (552 ± 151 and 441 ± 75 arbitrary units, for 24 °C and 37 °C respectively). The results are means $\pm$ S.E.M. of four independent experiments. The migration of molecular-mass markers is shown on the left (from top to bottom 97.4 kDa, 69 kDa, 46 kDa and 30 kDa).

temperature overnight. Equal amounts of protein (20 μ g or 40 μ g) or radioactivity incorporated into the protein (approx. 7×10^5 c.p.m.) were applied to the gels. After electrophoresis, the gels were either silver stained [20], or processed for fluorography [21] or Western blotting [6,7].

Synthetic peptide, production of antibodies and Western blotting

Synthesis of peptide DR-1 (NH₂-VGLKAPGIIPRI-COOH), corresponding to the sequence region of the F1-ATPase α -subunit from rat liver mitochondria, comprising residues 162 to 173 of the preprotein and highly similar to a sequence region of the chaperonins [5,6], and production of antibodies against DR-1 peptide have been described in detail previously [6]. For Western blotting, polyacrylamide gels containing 2-D-fractionated SL2 proteins were transferred to poly(vinylidene difluoride) membranes (Millipore) and probed with anti-DR-1 serum as described [6]. Visualization of the immunoreactive p71 and p56 proteins was carried out using a goat anti-rabbit IgG peroxidase conjugate [GAR/IgG (H + L)/PO] (Nordic Immunology) [6].

Protein Sequences

All protein sequences used in this study were obtained from the SwissProt data base. The access code and original reference for these sequences are as follows. Chaperonins: *Saccharomyces cerevisiae* heat shock protein (hsp) 60 (p19882 [22]); *Escherichia coli* GroEL (p06139 [23]); *Triticum aestivum* ribulosebisphosphate carboxylase/oxygenase (Rubisco) subunit-binding protein (p08823 [23]); *Mycobacterium bovis* 65-kDa antigen (p06806 [24]); human mitochondrial P1 protein (p10809 [25]). F-type ATPase α -subunits: *S. cerevisiae* (p07251 [26]); *Zea mays* mitochondria (p05022 [27]); bovine heart mitochondria (p19483 [28]); rat liver mitochondria (p15999 [29]); *E. coli* (p00822 [30]). Hsp70: human (p08107 [31]), chicken (p08106 [32]), *E. coli* (p04475 [33]); *S. pombe* (p22774 [34]) and *S. cerevisiae* (p12398 [35]).

RESULTS AND DISCUSSION

We have recently reported that antibodies raised against a synthetic peptide (DR-1) derived from the mitochondrial F1-ATPase α -subunit sequence from rat liver, a sequence region that is highly conserved in chaperonin sequences [5,6], recognized two stress-responsive proteins in embryonic *D. melanogaster* cells (SL2) [6,7]. The antibodies recognized a constitutively expressed p56 and an inducible p71. Both proteins increase their relative cellular amount during hyperthermic exposure of the cells [6,7]. Figure 1 shows the 2-D profile of synthesized proteins in unstressed (Figure 1A) and heat-shocked (Figure 1B) SL2 cells. The position of p56 and p71 in the 2-D protein maps as recognized by antibodies against the DR-1 peptide in Western blotted 2-D gels of heat-shocked SL2 cells is indicated (Figure 1C). Data from four independent experiments (see legend to Figure 1) indicated that the accumulation of p56 in stressed SL2 cells [6,7] is not the result of an increase in the relative rate of synthesis of the protein (Figure 1, compare panels A and B), which suggests that p56 accumulation results from a stress-related change in its degradation rate. By contrast, the accumulation of p71 [6,7] is brought about by a profound increase in the relative rate of synthesis of the protein (Figure 1, compare panels A and B), indicating that p71 is one of the inducible members of the 70 kDa heat-shock protein family [36–38].

An immunological relationship between F1-ATPase α -subunits and the chaperonins [6] is not surprising, considering the significant sequence identity existing between these two protein families [5,6]. However, what is really striking is the immunological recognition of a member of the 70 kDa stress-responsive family. A possible explanation of this striking result is the existence of a certain structural and/or sequence similarity between the three protein families considered herein. Since current automated methods cannot accurately align sequences that have less than 35 % sequence similarity without the use of reliable anchor points [2], we used the eleven most conserved regions between chaperonins and α -ATPases characterized previously [6] as query sequences for examination of the PILEUP alignment of

Figure 2 Multiple sequence alignment of hsp70s, chaperonins and the α-subunit of F-type ATPases

The alignment includes two sequences of 70 kDa hsp [human (HS70-H; P08107) and yeast (HS77-Y; P12398) mitochondria], two sequences of chaperonins [yeast mitochondria hsp60 (HS60-Y; P19882) and prokaryotic *E. coli* GroEL (CH60-E; P06139)], and two sequences of mitochondrial F1-ATPase α-subunits (ATP0-B; P19483 and ATPA-Y; P07251 from ox and yeast respectively). Previously identified regions of conserved amino acids between chaperonins and ATPases [6] are shown in roman numerals (I to XI) above the alignment. III*, VIII* and X* define additional regions of sequence similarity. The numbers on the left indicate the amino acid position in each sequence. Amino acid identities in the alignment between two or three of the protein families are shown by white lettering in black boxes. Other relevant amino acid similarities are enclosed in boxes. Dots indicate gaps introduced for optimizing the alignment. The hatched squares and triangle within the sequences indicate the position of amino acid insertions. A closed circle above hsp70 sequences identifies a relevant glutamate (E) residue (see text). Conservative amino acids (in one-letter notation) were considered as follows: G=A=S; A=V; V=I=L=M; L=R=H; D=E=N=Q; S=T and P=G. **Insert**, a schematic representation of the three protein families showing the location within their sequence of the conserved protein regions. Numbers above the sequences indicate the length (in amino acids) of the different insertions.

Figure 3 For legend see opposite

Table 1 A common sequence profile between chaperonins, 70 kDa hsp and α -ATPases provides evidence of evolutionary relatedness

Selected sequence regions from five chaperonins (hsp60), hsp70s and α -ATPases were aligned using PILEUP. When required, the alignment obtained was manually adjusted according to that shown in Figure 2. Consensus position-specific sequence profiles were obtained using PROFILEMAKE. Profiles were generated from continuous regions of the alignment avoiding long insertions (see Figure 2). The small insertion of chaperonin (14 residues) and α -subunit (19 residues) sequences were not considered as interruptions of the alignment. The approximate number of amino acid residues of the profile is shown. Comparison of the profiles obtained for each sequence in the Swiss Protein database was done using PROFILESEARCH. The number of chaperonin, hsp70 and α -ATPase sequences obtained above a significant Z score ($Z > 5$) is indicated. The number of SwissProt sequences searched was 36115.

	PROFILEMAKE (as defined in Figure 2)	Number of residues	PROFILESEARCH				
			hsp60	α -ATPase	hsp70	Others	Z score
A	I–VI	175	53	43	69	none	16.84 < Z > 5.16
B	VII–VIII	80	44	46	64	1 (GP84)	10.69 < Z > 5.00
C	VIII*–XI	176	2	46	54	1 (hsp26)	15.40 < Z > 5.01
D	VII–XI	268	53	46	–	none	24.66 < Z > 5.04
E	VIII*–XI	171	50	–	92	1 (hsp26)	17.95 < Z > 5.98

hsp70. The rationale behind this approach is that clusters of similar amino acid residues could define, at least in part, the same functions in the three protein families, α -ATPases, 60 kDa chaperones and 70 kDa chaperones, which are considered in this study.

As a result of this analysis, several of the conserved regions described previously were found to be arranged in exactly the same order within hsp70 sequences. These sequence elements were used as anchor points for the alignment of sequences of hsp70 with the previously aligned α -ATPase and chaperonin sequences [6]. For illustration purposes, the alignment includes only two sequences for each protein family (Figure 2). Interestingly, conserved regions along chaperonin sequences are located in a continuous way, spanning approx. 70% of their entire length and concentrated towards the last two-thirds of the protein sequence. In contrast, a big insertion (229 aa), representing the central one-third of hsp70 sequence, interrupts the alignment of chaperonins and hsp70 (Figure 2, insert). The alignment is interrupted in α -ATPases by three insertions (90aa, 7aa and 19aa, Figure 2, insert, and see Figure 5 in [6]). Remarkably, the position of the second insertion in α -subunits matches that of the insertion in hsp70 (Figure 2).

The comparison of aligned sequences reveals 23 invariant and up to 74 identical plus highly conserved residues between chaperonins and hsp70. There are 34 invariant and 84 identical plus highly conserved residues in α -ATPases and hsp70. In both cases, most conserved residues are found clustered in previously defined regions (I to XI in Figure 2) of high sequence identity between chaperonins and α -ATPases [6]. The alignment reinforces sequence identity in the three protein families in another three regions that have not been defined previously (III*, VIII* and X* in Figure 2). Overall, conserved regions in each protein family represent a significant percentage of the total sequence length of the proteins (spanning approx. 80% of chaperonin and α -

ATPase and approx. 65% of 70 kDa chaperone sequences). When compared with the chaperonins, the alignment reveals that amino acid conservation within each region varies between hsp70 and α -ATPases (Figure 2). However, certain regions (IV, VIII, IX in Figure 2) showed a remarkably high degree of sequence conservation in the three protein families.

Sequence similarity between the three families was statistically evaluated with the Profile analysis [39], using a GCG software package (Genetics Computer Corporation, Inc., Madison, WI, U.S.A., 1993). Alignment of the sequences was performed using PILEUP. Position-specific sequence profiles were calculated from the alignment of the three protein families using PROFILEMAKE. Profiles were derived from the alignment of five sequences for each protein family (see the Experimental section), avoiding those sequence regions that contain big insertions (Figure 2 and Table 1). The profiles obtained were used to search the SwissProtein database [release 29.0 (6/94)] with PROFILESEARCH. The database search was carried out with both the alignment provided by PILEUP and with that adjusted manually according to Figure 2. The results obtained with both type of searches provided similar results, i.e. the protein families identified were the same, although both the Z scores and the number of sequences obtained were higher in those searches derived from the manually adjusted profile.

Table 1 illustrates that profiles generated from the three protein families, spanning from regions I to VIII of the alignment (searches A and B), provided a very high number of sequences of each family. Most importantly, the sequences provided fell in to the range of statistically significant Z scores ($Z > 5$), and none of the searches provided a representative number of sequences from other protein families from the database (Table 1). These results strongly suggest that the alignment presented for conserved regions I to VIII is not the result of fortuitous similarity, but rather of significant sequence similarity between the three protein

Figure 3 Comparison of the secondary and tertiary structures of hsc70 and GroEL chaperonin

(A) Alignment of secondary structure elements of hsc (HSC70) [14,15] and GroEL (GROEL) chaperonin [16] according to Figure 2. The numbers indicate the amino acid position in each sequence. The roman numerals above the alignment identify the regions of sequence similarity. Dots indicate gaps introduced for optimizing the alignment. A capital I within a circle (Ⓢ) indicates the position of an insertion. The closed circle (●) identifies a relevant glutamate (E) residue (see text). Red boxes, green arrows and black lines represent α -helices, β -strands and coils respectively. Filled boxes and arrows and thick lines illustrate conserved secondary structure elements. (B) An α -carbon trace of the crystallographic structures of the N-terminal ATPase fragment of hsc70 (HSC 70) [14] and the subunit of GroEL chaperonin [16]. Atomic coordinates were obtained from the Brookhaven Protein Data Bank and were processed by the RasMol 2.5 program. The colours identify conserved sequence regions in each structure: I, blue; II, red; III, green; III*, yellow; IV, dark yellow; V, magenta; VI, light green; VII, brown; VIII, violet; VIII*, purple; IX, red orange; X, flesh; X*, black and XI, pink. The numbers indicate the position of certain residues for orientation. The domains (apical, intermediate and equatorial) of GroEL and subdomains (IA, IB, IIA and IIB) of hsc70 are labelled according to previous publications [14,16].

families under consideration. In other words, they are homologous proteins for that part of their sequence. However, it should be noted that search C in Table 1, corresponding to sequence regions VIII* to XI, did not provide a reliable number of sequences of the chaperonin family, although it provided a high number of α -ATPase and hsp70 sequences. This result might indicate that either there is no common ancestry for that sequence region between chaperonins and the other two protein families or, alternatively, that a different evolutionary pressure and/or ancestry of the chaperonin family has resulted in the fading of a common amino acid signature between the three protein families during divergence.

To test the second possibility, position-specific sequence profiles were derived from that part of the alignment of the chaperonins and each of the other two protein families. The results obtained (Table 1) indicate that chaperonins also share significant sequence similarity in the C-terminal portion of their sequences with both the α -ATPases (search D in Table 1) and hsp70 members (search E in Table 1). In this situation, it is reasonable to suggest that sequence similarity between the three protein families spans along regions I to XI of their sequence. Given the long stretch of sequence similarity (> 400 residues), representing a high percentage of the sequence length in each protein, it seems implausible that such similarity could have arisen by convergence. On the other hand, the lack of a common signature between the chaperonin and the other two protein families in the C-terminal part of their sequences (search C in Table 1) suggests that chaperonins have either a higher mutational rate and/or more ancient origin, whereas α -ATPases and 70 kDa chaperones could be more closely related. Both situations are compatible with the observation that the overall sequence similarity is only 50 % among chaperonins [6,40] and 70 % for both α -ATPase [6,41] and hsp70 members [42–44].

In conclusion, the conservation of a very large sequence region (Figure 2; Table 1), together with the immunological cross-reactivity [6] and the analogous functional roles of these protein families [8,9,11–13], leads us to support that chaperonins, hsp70s and α -subunits are homologous proteins; in other words, most of the sequence of these three protein families derived from a common ancestral protein. The continuous arrangement of the conserved regions in chaperonins, and the observation that sequence conservation in α -ATPases and hsp70 is interrupted by insertion of non-homologous stretches (Figure 2), leads us to hypothesize that the latter two protein families could have evolved from an ancestral chaperonin. In fact, analysis of the alignment shown in Figure 2 with the DISTANCES program supports a phylogenetic scheme for the diversification of hsp70 and ATPases from a common ancestor after the diversification of the chaperonin family (results not shown). As discussed above, the lack of a common signature in the C-terminal region between chaperonins and the other two protein families (Table 1), as well as the genomic organization of mammalian chaperonins in a single exon [45], and the fact that mammalian genes for ATPase subunits [46,47] and hsp70 members [48,49] are organized in multiple exons, also argue in favour of an ancient origin of the chaperonins.

It is expected that the insertions in each family define, in addition to those built up in the conserved chaperonin module, functional specificities that characterize the protein family. For instance, the insertion between regions VI and VII in α -ATPases contains the P-loop, the so-called Walker phosphate-binding motif [50], which might explain why molecular chaperones not containing this insertion are weak ATPases [11]. The large insertion between regions VIII and VIII* in hsp70 is the sequence region that defines the substrate-binding domain of this protein

family [15,44,51]. Interestingly, it has been shown that 60 kDa and 70 kDa chaperones reveal different substrate-binding specificities [52]. However, as previously hypothesized, conserved sequence regions might be involved in conferring, at least in part, the activities shared by the three protein families; ATP binding/hydrolysis and chaperone function [53]. In this regard, it is remarkable to find that all of the proteins have the so-called 'ATPase signature' of V- and F-type ATPases (under region IX in Figure 1) [41,54–56], the Mg^{2+} -binding motif of F-type ATPases (region VII in Figure 1) [57] and the essential glutamic residue involved in the chaperone function of the proteins [58,59] (marked with a closed circle between regions VII and VIII in Figs. 1 and 2).

The problem now is that we face a situation in which three partially homologous protein families have little resemblance in their two- and three-dimensional structures [14–16,57]. For illustrative purposes, a comparison of the secondary (Figure 3A) and tertiary (Figure 3B) structures of hsc70 and GroEL chaperonin is presented. As can be seen, conservation of secondary structural elements between hsc70 and the chaperonin only surpasses sequence resemblance in the N-terminal region of their sequences (Figure 3A). In fact, the residues falling in the same structural category is 39 % in this sequence region (I to VI). It is remarkable to observe that in other regions in which the primary structure is fairly well conserved (VII, VIII, VIII*), the secondary structure elements show no resemblance, albeit for certain coils (VIII, IX) (Figure 3A). Interestingly, the conserved sequence regions that lack structural resemblance limit the boundaries of the large non-homologous insertion present in hsc70 (Figure 2). A similar picture is obtained when the tertiary structure of the proteins is compared (Figure 3B). The orientation of conserved sequence regions and the overall folding of the domains are different in each protein family. However, it is remarkable to note that some domains or subdomains of the proteins are organized and limited by the same conserved sequence regions (Figure 3B). For instance, regions III*, IV and V in subdomain IB of hsc70 and in the apical domain of GroEL; I, II in subdomain IA of hsc70 and in the intermediate domain of GroEL; III connecting, and VI and VII limiting subdomains IA and IB in hsc70 and the apical and intermediate domains in GroEL respectively. Recently, the three dimensional structure of the F1-ATPase complex from bovine mitochondria has been reported [57]. Interestingly, and in spite of the existence of significant sequence [6] and functional [8,9] similarities between the chaperonins and the α -ATPase protein families, these two ATP-binding proteins reveal no obvious structural resemblance [16,57]. Altogether, these findings illustrate that analogous proteins, showing significant sequence identity (Figure 2, Table 1), adopt different secondary and tertiary structures (Figure 3).

In principle, this observation seems to contradict the accepted dogma that conservation of protein tertiary structure always surpasses sequence resemblance. Nowadays, it is believed that extant proteins have evolved from a limited number of ancestral proteins [60]. The driving force of protein evolution is random mutagenesis of the amino-acid sequence. Natural selection, the other phenomenon in evolution, modulates the rate of protein divergence by imposing constraints on changes in the amino acid sequence and, thus, it appears that the major restriction of natural selection is to maintain the overall structure of the protein. However, it is also likely that the divergence of proteins early in evolution and the appearance of insertions and/or deletions within a given family (see Figure 2) could act as triggers for significant changes in the overall protein structure. In this situation, it is reasonable to suggest that the maintenance of protein scaffolds is not always the selective pressure that governs

the divergence of proteins during evolution. Therefore, it can be inferred that clusters of similar residues might encounter different structural solutions to perform similar functions. In this regard, it should be noted that both the detailed structure [61,62] and the relative orientation of the domains [61] have been shown already to vary within homologous proteins. Moreover, it has been demonstrated that non-local interactions can determine the secondary structure of peptide sequences [63]. In other words, tertiary interactions play a dominant role in determining the tendency of amino acid residues to adopt a given secondary structure [63]. An additional example in this regard is that provided by the structure of the reverse transcriptase of human immunodeficiency virus type-1 [64–66]. In this example, the quaternary structure of the holoenzyme reveals that subunit p51, with similar sequence to subunit p66, but lacking a C-terminal peptide, has a different tertiary structure to that of p66 [64–66].

In conclusion, we provide evidence supporting the view that chaperonins, hsp70 chaperones and α -ATPase protein families have evolved from a common ancestral protein. The identification of this evolutionary relatedness between extant ATPases provides a useful reference tool for the analysis of protein structure and evolution of the nucleotide-binding protein families.

We are indebted to Drs. L. Blanco and J. Satrústegui for valuable comments and critical reading of the manuscript. Dr. A. Valencia (CNB, Madrid) is acknowledged for helpful discussion. We thank M. López de Heredia (CBMSO, Madrid) for help in handling the 3-dimensional structures. The technical assistance of M. Chamorro and the secretarial assistance of Mrs. D. Jelenic is also gratefully acknowledged. A.I.F. was the recipient of a pre-doctoral fellowship from Comunidad de Madrid. The work was supported by a grant from Dirección General de Investigación Científica y Técnica PB94-0159, Spain. The CBMSO receives a grant from Fundación Ramón Areces.

REFERENCES

- Blundell, T. L., Sibanda, B. L., Sternberg, M. J. and Thornton, J. M. (1987) *Nature* (London) **326**, 347–352
- Thornton, J. M., Flores, T. P., Jones, D. T. and Swindells, M. B. (1991) *Nature* (London) **354**, 105–106
- Rost, B., Schneider, R. and Sander, C. (1993) *Trends Biochem. Sci.* **18**, 120–123
- Gribskov, M., McLachlan, A. D. and Eisenberg, D. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4355–4358
- Luis, A. M., Alconada, A. and Cuezva, J. M. (1990) *J. Biol. Chem.* **265**, 7713–7716
- Alconada, A., Flores, A. I., Blanco, L. and Cuezva, J. M. (1994) *J. Biol. Chem.* **269**, 13670–13679
- San Martín, C., Flores, A. I. and Cuezva, J. M. (1995) *J. Cell. Biochem.* **59**, 235–245
- Avni, A., Avital, S. and Gromet Elhanan, Z. (1991) *J. Biol. Chem.* **266**, 7317–7320
- Yuan, H. and Douglas, M. G. (1992) *J. Biol. Chem.* **267**, 14697–14702
- Chen, X. J. and Clark-Walker, G. D. (1995) *EMBO J.* **14**, 3277–3286
- Ellis, R. J. and van der Vies, S. M. (1991) *Annu. Rev. Biochem.* **60**, 321–347
- Gething, M. J. and Sambrook, J. (1992) *Nature* (London) **355**, 33–45
- Hendrick, J. P. and Hartl, F. U. (1993) *Annu. Rev. Biochem.* **62**, 349–384
- Flaherty, K. M., DeLuca Flaherty, C. and McKay, D. B. (1990) *Nature* (London) **346**, 623–628
- Morshauer, R. C., Wang, H., Flynn, G. C. and Zuiderweg, E. R. P. (1995) *Biochemistry* **34**, 6261–6266
- Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. and Sigler, P. B. (1994) *Nature* (London) **371**, 578–586
- Brändén, C.-I. (1990) *Nature* (London) **346**, 607–608
- Santarén, J. F., Alconada, A. and Cuezva, J. M. (1993) *J. Biochem.* **113**, 129–131
- Görg, A., Postel, W. and Günter, S. (1988) *Electrophoresis* **9**, 531–546
- Ansorge, W. (1985) *J. Biochem. Biophys. Methods* **11**, 13–20
- Chamberlain, J. P. (1979) *Anal. Biochem.* **98**, 132–135
- Reading, D. S., Hallberg, R. L. and Myers, A. M. (1989) *Nature* (London) **337**, 655–659
- Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. and Ellis, R. J. (1988) *Nature* (London) **333**, 330–334
- Thole, J. E. R., Keulen, W. J., Kolk, A. H. J., Groothuis, D. G., Berwald, L. G., Tiesjema, R. H. and van Embden, J. D. A. (1987) *Infect. Immunol.* **55**, 1466–1475
- Jindal, S., Dudani, A. K., Singh, B., Harley, C. B. and Gupta, R. S. (1989) *Mol. Cell. Biol.* **9**, 2279–2283
- Takeda, M., Chen, W. J., Saltzgeber, J. and Douglas, M. G. (1986) *J. Biol. Chem.* **261**, 15126–15133
- Braun, C. J. and Levings, III, C. S. (1985) *Plant Physiol.* **79**, 571–577
- Walker, J. E., Fearnley, I. M., Gay, N. J., Gibson, B. W., Northrop, F. D., Powell, S. J., Runswick, M. J., Saraste, M. and Tybulewicz, V. L. J. (1985) *J. Mol. Biol.* **184**, 677–701
- Lee, J. H., Garboczi, D. N., Thomas, P. J. and Pedersen, P. L. (1990) *J. Biol. Chem.* **265**, 4664–4669
- Gay, N. J. and Walker, J. E. (1981) *Nucleic Acids Res.* **9**, 2187–2194
- Drabent, B., Genthe, A. and Benecke, B. J. (1986) *Nucleic Acids Res.* **14**, 8933–8948
- Marimoto, R. I., Hunt, C., Huang, S.-Y., Berg, K. L. and Banerji, S. S. (1986) *J. Biol. Chem.* **261**, 12692–12699
- Bardwell, J. and Craig, E. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 848–852
- Powell, M. J. and Watts, F. Z. (1990) *Gene* **95**, 105–110
- Scherer, P. E., Krieg, U. C., Hwang, S. T., Vestweber, D. and Schatz, G. (1990) *EMBO J.* **9**, 4315–4322
- Welch, W. J. (1990) in *Stress Proteins in Biology and Medicine* (Morimoto, R. I., Tissiere, A. and Georgopoulos, C., eds.), pp. 223–278, Cold Spring Harbor Laboratory Press, New York
- Gutierrez, J. A. and Guerriero, Jr., V. (1995) *Biochem. J.* **305**, 197–203
- Rassow, J., Voos, W. and Pfanner, N. (1995) *Trends Cell Biol.* **5**, 207–212
- Gribskov, M., Luethy, R. and Eisenberg, D. (1989) *Methods Enzymol.* **183**, 146–159
- Gupta, R. S., Picketts, D. J. and Ahmad, S. (1989) *Biochem. Biophys. Res. Commun.* **163**, 780–787
- Nelson, N. (1992) *Biochim. Biophys. Acta* **1100**, 109–124
- Rochester, D. E., Winer, J. A. and Shah, D. M. (1986) *EMBO J.* **5**, 451–458
- Wadhwa, R., Kaul, S. C., Ikawa, Y. and Sugimoto, Y. (1993) *J. Biol. Chem.* **268**, 6615–6621
- Bhattacharyya, T., Karnezis, A. N., Murphy, S. P., Hoang, T., Freeman, B. C., Phillips, B. and Morimoto, R. I. (1995) *J. Biol. Chem.* **270**, 1705–1710
- Venner, T. J., Singh, B. and Gupta, R. S. (1990) *DNA Cell Biol.* **9**, 545–552
- Ohta, S., Tomura, H., Matsuda, K. and Kagawa, Y. (1988) *J. Biol. Chem.* **263**, 11257–11262
- Pierce, D. J., Jordan, E. M. and Breen, G. A. M. (1992) *Biochim. Biophys. Acta* **1132**, 265–275
- Sorger, P. K. and Pelham, H. R. (1987) *EMBO J.* **6**, 993–998
- Michikawa, Y., Baba, T., Arai, Y., Sakakura, T. and Kusakabe, M. (1993) *FEBS Lett.* **336**, 27–33
- Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982) *EMBO J.* **1**, 945–951
- Wang, T.-F., Chang, J.-H. and Wang, C. (1993) *J. Biol. Chem.* **268**, 26049–26051
- Hartl, F. U. and Martin, J. (1995) *Curr. Opin. Struct. Biol.* **5**, 92–102
- Alconada, A. and Cuezva, J. M. (1993) *Trends Biochem. Sci.* **18**, 81–82
- Gogarten, J. P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E. J., Bowman, B. J., Manolson, M. F., Poole, R. J., Date, D., Oshima, T. et al. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6661–6665
- Südhoff, T. C., Fried, V. A., Stone, D. K., Johnston, P. A. and Xie, X.-S. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6067–6071
- Futai, M., Noumi, T. and Maeda, M. (1989) *Annu. Rev. Biochem.* **58**, 111–136
- Abrahams, J. P., Leslie, A. G. W., Lutter, R. and Walker, J. E. (1994) *Nature* (London) **370**, 621–628
- Flaherty, K. M., Wilbanks, S. M., DeLuca Flaherty, C. and McKay, D. B. (1994) *J. Biol. Chem.* **269**, 12899–12907
- Buchberger, A., Valencia, A., McMacken, R., Sander, C. and Bukau, B. (1994) *EMBO J.* **13**, 1687–1695
- Chothia, C. (1992) *Nature* (London) **357**, 543–544
- Overington, J. P. (1992) *Curr. Opin. Struct. Biol.* **2**, 394–401
- Muñoz, V., Blanco, F. J. and Serrano, L. (1995) *Protein Sci.* **4**, 1577–1586
- Minor, D. L. and Kim, P. S. (1996) *Nature* (London) **380**, 730–734
- Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, Jr., A. D., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P. et al. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6320–6324
- Arnold, E., Jacobo-Molina, A., Nanni, R. G., Williams, R. L., Lu, X., Ding, J., Clark, Jr., A. D., Zhang, A., Ferris, A. L., Clark, P. et al. (1992) *Nature* (London) **357**, 85–89
- Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. and Steitz, T. A. (1992) *Science* **256**, 1783–1790