ELSEVIER

Biophysical Chemistry 71 (1998) 99–111

Biophysical Chemistry

Solvent accessibility analysis on the mutants of Hsc70 ATPase fragment

T.S. Kumarevel [a], M. Michael Gromiha [b], M.N. Ponnuswamy [a,*]

[a] *Department of Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai - 600 025, India*
[b] *International Centre for Genetic Engineering and Biotechnology, Area Science Park, Padriciano 99, 34012 Trieste, Italy*

Received 30 September 1997; revised 4 December 1997; accepted 5 December 1997

Abstract

Molecular chaperones are the cellular proteins which mediate the correct folding of other polypeptides. The concept of 'solvent accessibility' is one of the most powerful tools to understand the structure and stability of protein molecules. The hydrophobic variation of amino acid residues due to point mutations at many active sites of chaperone protein Hsc70 using solvent accessibility analysis is carried out. The numerical indices for several properties of amino acid residues, such as, reduction in accessibility, preference of amino acid residues in interior and surface parts, transfer free energy and the preference of amino acid residues to change their positions (buried/exposed) due to amino acid substitutions for Hsc70 and its mutants were set up. The accessibility of amino acid residues varies much between native and mutant proteins whereas there is no major changes on their conformations. The conformational stability for Hsc70 and its mutants were established and the computed hydrophobic free energy change is around 10 kcal/mol due to single amino acid substitution.

Keywords: Chaperone; Conformational stability; Free energy; HSC70; Hydrophobicity; Mutants; Solvent accessibility

1. Introduction

Classic in vitro studies show that many unfolded polypeptides are able to fold to the native state spontaneously, indicating that the polypeptide chain itself contains all the information necessary to specify its three-dimensional conformation [1]. Proteins that directly influence the processes by which other polypeptides attain their native structures within cell are termed as molecular chaperones [2], in view of their essential role in assisting the folding of polypeptides in vivo rather than misfolding or aggregation [3–6]. The Hsp70s and their partner proteins can stabilise the newly synthesised polypeptides until the segments necessary for the folding are available; the chaperonins are large cylindrical complexes that promote protein folding in the sequestered environment of their central cavity.

The chaperonins have an essential function in promoting the ATP-dependent folding of proteins, both under normal growth conditions and stress. There are two chaperonin subgroups; namely (i) members of the GroEL(HSP60) family, and (ii) the chaperonins of the TRiC family. GroEL type chaperonins are present in eubacteria, mitochondria and chloroplasts and are made of two stacked seven-

* Corresponding author.

0301-4622/98/$19.00
PII S0301-4622(97)00137-3

membered rings of 60K. They cooperate with a smaller protein cofactor, GroES in the *Escherichia Coli*, which is a single heptameric ring of 10K subunits. These proteins are stress inducible, except for the chloroplast homologue. The TRiC(TCP-1 ring complex) chaperonins in archaebacteria and the eukaryotic cytosol are eight- or nine-membered double rings containing a family of 55K subunits with homology to the TCP-1 tail complex protein from mouse. Only the archaebacterial members of this family are stress-inducible. Recently, the crystal structures of the chaperonin proteins GroEL and GroES have been determined in atomic detail [7–10] and functions of these proteins have been studied by electron microscopy [11,12].

The Hsp70s are a family of highly conserved ATPase of relative molecular mass of about 70 K found in prokaryotic and in most of the eukaryotic cells. This Hsp70s play an important role in protein metabolism under both stress and non-stress conditions, including de novo protein folding and translocation process on both sides of the mitochondrial membrane [13]. The versatility of the Hsp70s results from their basic function which is to bind and release hydrophobic segments of an unfolded polypeptide chain in an ATP hydrolytic reaction cycle [14,15]. Whereas the binding results in the stabilisation of the unfolded state, controlled release may allow progression along the folding pathway.

The three-dimensional structures of Hsp70 proteins are known in N-terminal ATPase domain of 45K [16]. The various Hsp70s have the binding specificity for hydrophobic peptides, their preference for certain amino acid patterns may vary from one compartment to another [15,17–19]. The endoplasmic reticulum (ER) homologue, Bip, binds with micromolar affinity to 7–8 residue peptides, corresponds to the segments of a polypeptide chain that would be exposed in the unfolded state but are buried within the hydrophobic core of the folded protein [15,17]. This is the ability of Hsp70 to interact with a wide range of unfolded polypeptides while ignoring the native states of these proteins.

The crystal structure of the 44K fragment of HSC70 shows several residues in the active site that might participate in ATP hydrolysis [16]. This was also studied with the ATPase core of clathrin uncoating protein [20] which has been demonstrated as a members of the Hsp70 family. All the mutant proteins were crystallized under the conditions almost similar to those used for the wild type protein. The structures of the two point mutations at the active site, at residue positions Asp10, Glu175, Asp199, Thr204, Asp206 were solved by X-ray crystallography and their coordinates are available in Protein Data Bank of Brookhaven National Laboratory [21,22].

The concept of accessible surface area (ASA) with an analytic procedure to realize it from the crystal structure has helped to compute the hydrophobic contribution to the stability of the folded states of proteins and to characterize them in more quantitative terms [23–28]. The ASA is defined as the protein surface that is in contact with solvent. It is measured by the set of points occupied by the center of a water molecule that is being rolled along the van der Waals surface of the protein molecule. Recently, several investigations have been done on accessible surface area of proteins and nucleic acids to compute the hydrophobic free energy and their stabilities [29–38].

The aim of this work is to analyse the effect of mutations on the hydrophobic variation of amino acid residues in Hsc70 and its mutants based on solvent accessibility study. The solvent accessible surface area of all atoms and residues in native and mutant proteins is computed and the available information are used to derive parameters for several properties such as, reduction in accessibility, preference of residues for surface or interior part, transfer free energy, etc. The results obtained are compared with those observed in globular proteins and the importance of mutant sites are explored.

2. Materials and methods

For our analysis, we have used the crystal structures of the native (3HSC) and mutant proteins (two point mutations at the five active sites, D10N(1NGH), D10S(1NGG), E175Q(1NGA), E175S(1NGB), D199N(1NGF), D199S(1NGE), T204E(1ATS), T204V(1ATR), D206N(1NGD) and D206S(1NGC)) available from the Protein Data Bank of Brookhaven National Laboratory [21,22].

2.1. Computation of accessible surface area and reduction in accessibility

The ASA of all atoms in Hsc70 and its mutants was computed using the procedure of Lee and Richards [23] and by the program ACCESS developed by Richmond and Richards [39].

We have grouped the atoms into five sets, namely, carbon (C), neutral N/O, charged nitrogen (N^+), charged oxygen (O^-) and sulfur (S). The average value of ASA for the native and mutants was calculated by dividing the total ASA by the number of atoms present in each group. The ASA of each amino acid residue was computed by adding the ASA of their respective atoms, as well as the average ASA for the 20 amino acid residues.

The reduction in accessibility was done using the expression of Rose et al. [29],

$$R_A = \frac{A_0 - <A>}{A_0} \quad (1)$$

where A_0 and $<A>$ represent, respectively, the accessible area in the unfolded (extended) and folded state of the protein. The A_0 values were taken from the work of Rose et al. [29].

2.2. Computation of residue preferences in interior and surface parts of Hsc70

The ASA of the residues in all considered proteins was computed as explained in the previous section. Then each residue was classified as on the surface or in the interior using the criteria that the residues with more than 5% accessibility are considered to be in the surface and the ones with less than 5% accessibility are considered to be in the interior part [25,33,40]. The solvent accessibility (in %) is defined as the solvent accessible surface area of the residue in its native protein, computed by ACCESS [39] divided by the solvent accessible surface area of the residue in an extended tripeptide Gly-X-Gly conformation [29]. The preference of each residue in molecular chaperones was obtained using the expression,

$$R_{\text{in}}(R_{\text{out}})^i = \frac{\text{No. of residues of type } i \text{ present in the interior (surface) part}}{\text{Total number of residues in the interior (surface) part}} \quad (2)$$

2.3. Computation of surrounding residues

The residues in a protein molecule are represented by their α-carbon atoms. Using the α-carbon coordinates, a volume of 8 Å is fixed around the residues at the mutant site and the residues occurring in this volume are identified. It is evident that the influence of each residue over the surrounding medium extends effectively only up to 8 Å [41,42]. The number of surrounding residues at the mutant site is computed using the formula,

$$N_j = \Sigma n_{i,j} \quad (3)$$

where $n_{i,j}$ is the number of surrounding residues of type i around the jth residue (mutant site residue) of the protein.

2.4. Computation of conformational stability

We followed the method of Ponnuswamy and Gromiha [37] to compute the free energy terms in the folded and unfolded states of molecular and their conformational stabilities. The conformational stability is given by

$$\Delta G = G_{\text{F}} - G_{\text{U}} \quad (4)$$

$$G_{\text{F}} = G_{\text{hy}} + G_{\text{hb}} + G_{\text{vw}} + G_{\text{el}} + G_{\text{ss}} \quad (5)$$

$$G_{\text{U}} = G_{\text{en}} + G_{\text{ne}} \quad (6)$$

where F and U represent folded and unfolded states. Hy, hb, vw, el, ss, en and ne represent hydrophobic, hydrogen bonding, van der Waals, electrostatic, disulfide, entropic and nonentropic, respectively. The hydrophobic free energy, G_{hy} was computed as

$$G_{\text{hy}} = \Sigma_i \Delta\sigma_i \left[A_i(\text{folded}) - A_i(\text{unfolded}) \right] \quad (7)$$

where A_i(folded) and A_i(unfolded) are solvent accessibilities of the folded and unfolded states of the molecule and $\Delta\sigma_i$ are the atomic solvation parame-

ters. The unfolded state solvent accessibilities were taken from Rose et al. [29] and the atomic solvation parameters were taken from Ponnuswamy and Gromiha [37]. G_{el} was taken to be the sum of contributions from ion-pairs and charge-helix dipoles, a surface ion-pair contributing 1 kcal/mol [43,44] and a buried ion-pair contributing 3 kcal/mol [45,46], and a helix-charge dipole contributing 1.6 kcal/mol [47,48]. G_{hb} was taken to be the sum of contributions from hydrogen bonds, each hydrogen bond contributing 1 kcal/mol [49,50]. G_{ss} was taken to be the sum of contributions from disulfides, each such link contributing 2.3 kcal/mol [51]. The entropic factor G_{en} was taken as 1.2 kcal/mol per residue [52] and half the amount of G_{hb} of the folded state [53] was added as the non-entropic term to represent the actual G_U value for the molecule. G_{vw} was calculated using the linear equation [37]

$$G_{vw} = 8.885 + 0.1413\,N \quad (8)$$

where N is the number of residues in the protein.

The computation of all the free energy terms and the approximations used to compute the conformational stability are discussed in an earlier article [37].

2.5. Computation of transfer free energy

We have followed the method of Miller et al. [31] to compute the transfer free energy using the expression,

$$\Delta G_t = -RT \ln f, \text{ where } RT = 0.596\,\text{kcal/mol at } 27^\circ \quad (9)$$

where R is the universal gas constant and T is the

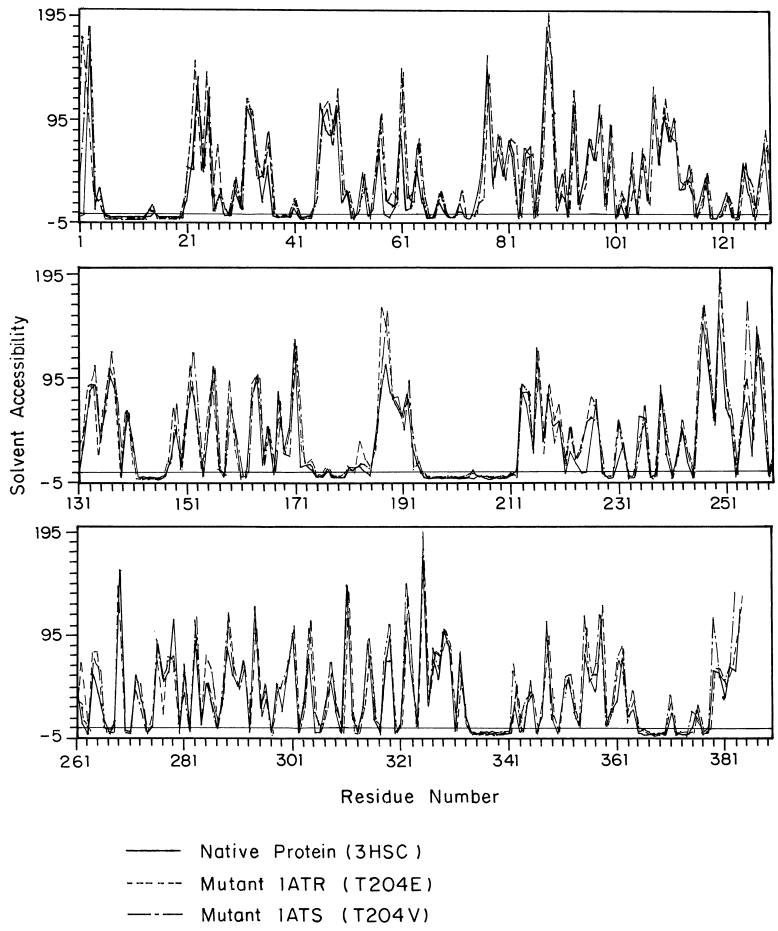

Fig. 1. The accessibility profile for the native and two mutant proteins.

temperature in Kelvin, f is the partition coefficient and is calculated using the equation,

$$f = (N_s/\Sigma N_s)/(N_b/\Sigma N_b) \quad (10)$$

where N_s and N_b are, respectively, the number of surface and buried residues of a given type of amino acid; ΣN_s and ΣN_b are the total number of surface and buried residues.

3. Results and discussion

3.1. Solvent accessibility of amino acid residues in Hsc70 and its mutants

We have analysed the variation of solvent accessibility for the substitution of amino acid residues in Hsc70. The accessibility profile for the native(3HSC) chaperone protein and two mutant proteins 1ATR(T204E) and 1ATS(T204V) are shown in Fig. 1. This profile is the plot of the accessibility values of the residues computed using the program ACCESS against their sequence numbers.

Fig. 1 shows that in the native protein(3HSC), the residues Asp10, Glu175, Asp199, Thr204 and Asp206 at mutant sites are completely buried (i.e., zero accessibility). It is interesting to note that the segments 6–20, 142–147, 195–212, 334–341, 365–370 are fully buried and the segments 21–25, 32–36, 45–51, 75–82, 84–91, 93–100, 106–115, 132–138, 148–153, 168–174, 186–194, 213–220, 246–252, 254–258, 288–292, 321–330, 351–367, 379–384 are exposed.

Among them, the segments 6–20 and 195–212 lie in the most interior part due to the formation of hydrophobic core around these segments. The segment 351–367 is surrounded by several charged residues, forming ion-pairs and prefers to be in the most exposed region. Also, it is to be noted that the residues 88Lys, 3Lys, 325Lys, 250Lys, 269Arg and 77Arg are highly exposed with more ASA (high peaks in 3HSC) and are positively charged.

The accessibility profiles obtained by changing Thr204 to Val(1ATR) and Thr204 to Glu(1ATS) are shown in Fig. 1. Interestingly, the residues 69Asp, 86Asp, 121Ser, 182Ala, 268Glu, 312Gly, 324Ala and 375Val in 1ATR(T204V) change their positions from exposed to buried. The residues 14Thr, 15Tyr, 26Val, 28Ile, 37Thr, 53Asp, 58Gln, 59Val, 130Ile, 185Leu, 204Thr, 221Ser, 225Asp, 228Leu, 275Ser, 280Ala, 284Ile, 287Leu, 305Leu, and 378Ala in 1ATS(T204E) also change their positions from buried to exposed. In 1ATR the residues 312Gly, 375Val and 86Asp dislike the surface considerably by reducing their ASAs from 14.70 to 1.10, 6.40 to 0.40 and 6.50 to 0.80 Å^2 respectively. Similarly in 1ATS, the residues 2Ser (not found in native protein), 225Asp, 305Leu, 58Gln, 59Val and 14Thr like to interact with other polar residues by increasing their ASA from 0 to 131.7, 5 to 53, 0.6 to 20.6, 2.9 to 18.9, 0.6 to 12.4, 0 to 8.6 Å^2, respectively.

Table 1
The ASA of the mutant sites of native and mutant proteins

Protein code	Accessible surface area (Å^2)[a]				
	Site 10	Site 175	Site 199	Site 204	Site 206
3HSC Native	390.4(52)	172.6(76)	311.4(80)	263.0(58)	314.6(84)
1NGH (D10N)	221.2(56)	202.8(90)	395.4(94)	332.2(44)	395.5(62)
1NGG (D10S)	406.4(46)	295.7(65)	352.7(84)	333.9(84)	333.1(94)
1NGB (E175N)	221.0(57)	114.9(94)	353.5(94)	331.7(43)	369.4(60)
1NGA (E175S)	213.8(66)	157.0(87)	370.0(93)	250.1(40)	217.9(38)
1NGF (D199N)	359.4(50)	196.8(78)	295.2(57)	308.7(60)	314.4(85)
1NGE (D199S)	221.0(55)	130.3(89)	333.6(94)	290.8(45)	345.2(52)
1ATR (T204V)	586.2(51)	250.0(71)	257.4(90)	488.7(82)	388.4(69)
1ATS (T204E)	541.7(64)	165.0(99)	204.8(88)	369.0(80)	285.1(66)
1NGD (D206N)	201.2(56)	94.7(91)	346.9(94)	267.1(42)	337.9(68)
1NGC (D206S)	305.8(59)	180.4(92)	377.3(94)	284.3(46)	275.9(54)

[a] Number of atoms are given in parenthesis.

Table 2
The total ASA (Å^2) in interior and surface parts

Protein code	< 5 Å^2 (interior)			> 5 Å^2 (surface)		
	ASA$^a_{tot}$	N^b	ASA$_{ave}$	ASA$^a_{tot}$	N^b	ASA$_{ave}$
3HSC	80.30	140	0.57	12,927.9	242	53.42
1NGH	60.60	132	0.46	13,954.9	246	56.73
1NGG	81.10	136	0.59	12,123.8	243	49.89
1NGB	64.90	135	0.48	13,136.6	243	54.06
1NGA	71.20	136	0.52	13,108.5	242	54.17
1NGF	109.10	145	0.75	12,959.4	237	54.68
1NGE	65.00	143	0.45	12,448.5	235	52.97
1ATR	90.80	131	0.69	14,555.8	252	57.76
1ATS	79.40	125	0.64	14,417.9	257	56.10
1NGD	56.80	134	0.42	13,366.2	244	54.78
1NGC	85.80	139	0.62	13,514.6	239	56.55
Average			0.56			54.67

[a] Total solvent accessibility.
[b] Total number of residues.

The ASA of the mutant sites of native and mutant proteins are given in Table 1.

This table shows that the five mutant sites in all mutant proteins varies with respect to native protein. When site 10 in native protein is considered, the ASA is 390.4 Å^2 from 52 atoms (1 O^-, 1 S, 2 N^+, 32 C and 16 O/N) whereas the ASA increased dramatically in 1ATR and 1ATS but decreased in 1NGD. The replacement of Asp10 by Asn shows a decrease of 169.2 Å^2 and by Ser indicates an increase of 16.0 Å^2 in their ASA values.

The total and average ASA values of all proteins are given in Table 2. We grouped the residues into two categories in such a way that ASA is less than or greater than 5%.

The table reveals that about 36% of residues are in buried part and 64% of residues are in exposed part with an average ASA of 0.56 and 54.6 Å^2

Table 3
The ASA of different types of atoms in native and mutant sites of molecular chaperones

Protein code	ASA (Å^2)[a]					
	Site	O^-	S	N^+	C	O/N
3HSC	10	0.80(1)	0.00(1)	18.40(2)	232.20(32)	139.00(16)
	175	19.80(1)	0.00(0)	10.10(1)	76.00(51)	66.70(23)
	199	6.30(1)	0.00(0)	0.00(0)	221.20(51)	83.90(28)
	204	6.30(2)	0.00(0)	16.90(1)	151.20(36)	88.60(19)
	206	6.30(2)	0.00(1)	16.90(1)	156.50(49)	134.90(31)
1NGH	10	0.00(0)	0.00(0)	0.00(0)	101.60(39)	119.60(17)
1NGG	10	0.00(0)	0.00(0)	34.90(1)	276.90(32)	94.60(14)
1NGB	175	0.00(1)	0.00(1)	9.70(1)	60.10(56)	45.10(35)
1NGA	175	0.00(1)	0.00(1)	6.30(1)	79.60(55)	71.10(29)
1NGF	199	0.40(1)	0.00(0)	0.00(0)	173.30(53)	121.50(30)
1NGE	199	24.90(2)	0.00(1)	22.50(1)	142.20(54)	144.00(36)
1ATR	204	13.00(1)	0.00(0)	0.00(0)	307.80(51)	167.90(30)
1ATS	204	0.00(0)	0.00(0)	0.00(0)	261.70(51)	107.30(29)
1NGD	206	36.20(3)	0.00(0)	14.60(1)	165.60(40)	121.50(24)
1NGC	206	23.00(3)	0.00(0)	15.90(1)	168.20(33)	68.80(17)

[a] Number of atoms are given in parenthesis.

respectively. Due to single amino acid substitution at D199N (1NGF) and T204E (1ATS), 15 residues came outside and the average surface area changed considerably (increased 28.7 Å^2 in 1NGF and decreased 23.5 Å^2 in 1NGD).

The results obtained by us for Hsc70 and mutants are comparable with the results of similar studies made on globular proteins. About 31% of residues are in buried part and 69% of residues are in exposed part with the average ASA of 0.70 and 58.85 Å^2, respectively [31].

3.2. ASA of all atom types at the mutant sites of Hsc70 chaperone

The ASA of five atom types at the mutant sites of native and mutant proteins are given in Table 3. From this table, it is observed that 1NGH and 1NGG adopt different trends. In these cases, Asn (1NGH) and Ser (1NGG) are present in the place of Asp10 (3HSC) and no major differences were observed in ASAs of O^- and S atoms. The number of atoms that contribute to the ASAs are zero whereas for 3HSC it is one. Also we note that the exposed sulfur atoms are from the residue Met and not from Cys since they are buried without forming any disulfide bond.

For the N^+ atom type of the native protein, the ASA is observed to be 18.40 (2) Å^2. In 1NGH, it is fully buried and in 1NGG it is 34.90 (1) Å^2. The ASA of carbon atom in native protein is 232.2 (32) and that in mutant 1NGH is 101.6 (39), lowered by 130.6 Å^2 though it has more number of atoms. But in 1NGG, interestingly higher ASAs is observed for the same number of atoms to that of native.

The 16 O/N atoms share the ASA of 139.0 Å^2 in native site but decreases with an additional atom in 1NGH and 1NGG (ASAs and number of contributions). We observed such interesting results in mutants of all active sites (D10, E175, D199, T204, and D206) in all the considered chaperone proteins (Table 3).

3.3. ASA of Hsc70 and its mutants in their atomic level

The computed average ASA values for the five different types of atoms (O^-, S, N^+, C and N/O) are given in Table 4. The average ASA of these five types of atoms in Hsc70s are 16.13, 8.14, 27.95, 3.78 and 4.30 Å^2, respectively and the average (%) contribution are 5.98, 0.48, 11.44, 51.19, 30.91%. The deviation is considerably more in the case of charged atoms.

The average and percentage contribution of ASA for these five different atom types in globular proteins are, respectively, 19.92, 3.59, 30.68, 4.55, 4.65

Table 4
The average ASA of different types of atoms in 3HSC its mutants

Protein code	ASA (Å^2)[a]				
	O^-	S	N^+	C	N/O
3HSC	14.58(52)	7.23(8)	24.63(57)	3.61(1857)	4.14(985)
1NGH	17.42(50)	7.60(8)	29.96(55)	3.85(1839)	4.47(976)
1NGG	12.61(48)	6.86(8)	26.57(56)	3.44(1841)	3.81(975)
1NGB	16.24(50)	9.83(8)	26.36(55)	3.69(1839)	4.18(976)
1NGA	14.31(50)	7.76(8)	27.96(56)	3.70(1844)	4.20(978)
1NGF	15.16(51)	6.26(8)	25.67(57)	3.61(1857)	4.13(985)
1NGE	14.90(50)	6.59(8)	25.23(55)	3.56(1838)	3.87(975)
1ATR	18.70(52)	8.77(8)	31.43(57)	4.58(1861)	5.55(987)
1ATS	17.69(53)	10.55(8)	33.22(56)	3.84(1855)	4.53(986)
1NGD	16.81(50)	10.26(8)	26.28(55)	3.82(1839)	4.13(985)
1NGC	15.68(50)	7.76(8)	24.47(55)	3.88(1838)	4.33(975)
Average[b]	16.13(1.68)	8.14(1.43)	27.95(2.76)	3.78(0.28)	4.30(0.44)
% contribution	5.98	0.48	11.44	51.19	30.91

[a] The number of atoms are given in brackets.
[b] Standard deviations are given in parenthesis in the row average.

Å^2 and 5.32, 0.49, 9.87, 54.68, 29.63%, a feature commonly observed in Hsc70s.

3.4. Effect of mutation at different sites

The effect of mutation at different sites has been analysed from the change in solvent accessible surface area of all residues due to amino acid substitutions. We have examined all the residues, which remain in the same position, or the ones prefer to move from surface to interior and vice versa. The results are given in Table 5.

The table shows that due to the substitution of Asn and Ser at Asp10, a total of 20 and 24 residues moved from their positions. In 1NGH, 14 residues came out from the interior to the surface out of 20 and 6 residues prefer to move in. They prefer some local interactions like hydrophobic, electrostatic, etc. In 1NGG, among the 24 residues, 14 residues like to move to the surface and 10 residues prefer to interact with the interior residues. More changes are observed in 1ATR and 1ATS, where a total of 25 and 20 residues change their ASA, respectively.

The residue preference for buried to surface and surface to buried are given in Table 6. The residues Leu, Ile, Ser, Asp, Ala and Val prefer to move from buried to exposed surface and the residues Asp, Ala, Phe, Glu, Leu and Asn prefer to be buried inside from the surface. It is noteworthy that the residues Ala, Asp and Leu change their positions very easily to adopt the environment due to the amino acid substitutions.

We have grouped the residues into charged polar (Lys, His, Arg, Glu, Gln, Asp, Asn, Ser, Thr and Pro) and non-polar (Ala, Val, Leu, Ile, Met, Phe, Tyr, Trp, and Gly) and computed the number of residues of each type which changes their positions. The results are given in Table 7.

It is interesting to note that some polar residues moved from surface to protein interior and many non-polar residues moved to the surface due to the

Table 5
The behaviour of amino acid residues with point mutations

Protein code	Buried to surface	Surface to buried
1NGH (D10N)	15Tyr, 3Asp, 58Gln, 59Val, 85Leu, 211Ile, 221Ser, 225Asp, 228Leu, 275Ser, 280Ala, 284Ile, 305Leu, 354Phe, 378Ala	69Asp, 86Asp, 165Ala, 182Ala, 184Gly, 354Phe
1NGG (D10S)	26Val, 28Ile, 53Asp, 58Gln, 59Val, 71Lys, 92Phe, 185Leu, 221Ser, 225Asp, 280Ala, 305Leu, 365Pro, 378Ala	86Asp, 121Ser, 141Asn, 174Asn, 184Gly, 194Asn, 232Asp, 266Ala, 307Ala, 354Phe
1NGB (E175Q)	6Ala, 58Gln, 59Val, 154Gln, 185Leu, 212Ile, 221Ser,225Asp, 275Ser, 280Ala, 284Ile, 305Leu, 378Ala	69Asp, 72Arg, 86Asp, 94Val, 173Ile, 268Glu, 354Phe, 359Leu
1NGA (E175S)	28Ile, 58Gln, 59Val, 185Leu, 221Ser, 225Asp, 228Leu, 275Ser, 284Ile, 305Leu, 378Ala	45Thr, 86Asp, 94Val, 174Asn, 182Ala, 268Glu, 354Phe
1NGF (D199N)	59Val, 185Leu, 221Ser, 225Asp, 275Ser, 284Ile, 305Leu	41Tyr, 62Asn, 72Arg, 94Val, 148Ala, 150Phe,266Ala, 268Glu, 316Pro, 321Leu, 354Phe, 359Leu
1NGE (D199S)	37Thr, 58Gln, 59Val, 225Asp, 228Leu, 275Ser, 284Ile, 378Ala	3Lys, 69Asp, 72Arg, 86Asp, 90Trp, 157Ala, 173Ile, 177Thr, 184Gly, 194Asn, 268Glu, 354Phe
1ATR (T204V)	2Ser(not found in 3HSC), 15Tyr, 26Val, 37Thr, 53Asp, 58Gln, 59Val, 92Phe, 185Leu, 212Ile, 221Ser, 225Asp, 228Leu, 275Ser, 284Ile, 287Leu, 305Leu, 366Pro	69Asp, 86Asp, 121Ser, 182Ala, 268Glu, 312Gly, 324Ala, 375Val
1ATS (T204E)	2Ser(not found in 3HSC), 14Thr, 15Tyr, 6Val, 28Ile, 37Thr, 53Asp, 58Gln, 59Val, 130Ile, 185Leu, 204Thr, 221Ser, 225Asp, 228Leu, 275Ser, 280Ala, 284Ile, 287Leu, 305Leu, 378Ala	75Gly, 86Asp, 121Ser, 266Ala, 359Leu
1NGD (D206N)	15Tyr, 58Gln, 59Val, 225Asp 228Leu, 275Ser, 284Ile, 305Leu, 378Ala	268Glu, 321Leu, 354Phe
1NGC (D206S)	37Thr, 58Gln, 59Val, 185Leu, 225Asp, 275Ser, 284Ile, 378Ala	69Asp, 72Arg, 150Phe, 182Ala, 268Glu, 321Leu, 354Phe

Table 6
The physicochemical properties of the 20 amino acids/residues in molecular chaperones[a]

Residue	Parameter											
	$Pref_{bs}$	$Pref_{sb}$	$< R_{Ac} >$	$< R_{Ag} >$ [b]	R^c_{ing}	R^c_{outg}	R_{inc}	R_{outc}	G_c	G^c_g	ASA_{hel}	ASA_{str}
Ala	10.74	15.38	0.86	0.76	11.0	7.9	14.3	7.0	0.43	0.20	14.7(26)	4.4(4)
Asp	11.57	16.66	0.72	0.65	2.2	7.4	5.0	9.1	−0.38	−0.72	41.6(10)	24.8(8)
Cys			1.00	0.90	5.4	1.8	1.4			0.67	0.0(1)	(0.0(1)
Glu		8.97	0.69	0.63	1.0	6.2	1.4	8.7	−0.92	−1.09	32.8(11)	74.4(6)
Phe	1.65	12.82	0.95	0.91	7.7	2.5	7.1	11.6	0.58	0.67	3.4(6)	11.4(7)
Gly		6.41	0.76	0.74	9.7	8.8	9.3	6.6	0.29	0.06	0.0(4)	1.8(3)
His			0.59	0.79	2.4	2.2		2.5		0.04	36.6(2)	0.0(0)
Ile	13.22	2.56	0.90	0.89	10.5	3.0	11.4	4.1	0.51	0.74	32.1(6)	16.2(14)
Lys	0.83	1.28	0.62	0.52	0.3	8.9	1.4	11.2	−1.21	−2.00	57.9(13)	72.6(4)
Leu	19.83	7.69	0.94	0.88	12.8	4.3	10.7	4.1	0.43	0.65	15.9(10)	8.5(7)
Met			0.76	0.86	3.0	0.9	1.4	1.6	−0.07	0.71	18.1(4)	107.9(1)
Asn		7.69	0.77	0.68	2.0	6.3	0.7	7.0	−0.96	−0.69	41.0(4)	43.6(5)
Pro	1.65	1.28	0.91	0.64	2.2	4.7	4.3	2.1	0.44	−0.44	6.0(5)	0.1(2)
Gln	8.26		0.73	0.65	1.3	4.5	1.4	4.1	−0.98	−0.74	47.9(7)	72.3(3)
Arg		5.13	0.69	0.65	0.4	4.0		9.1	−2.08	−1.34	71.0(15)	88.0(4)
Ser	13.22	3.85	0.71	0.67	5.0	8.9	5.0	5.4	−0.16	−0.34	38.4(6)	24.3(7)
Thr	4.96	2.56	0.79	0.71	4.6	7.1	7.9	6.6	0.09	−0.26	32.5(8)	21.6(8)
Val	10.74	5.13	0.94	0.86	12.7	4.6	16.4	3.3	0.90	0.61	10.2(9)	6.8(19)
Trp		1.28	0.94	0.88	2.7	1.3		0.4	−1.02	0.45	0.0(0)	0.0(0)
Tyr	3.30	1.28	0.82	0.82	3.3	4.8	0.7	3.7	−2.55	−0.22	55.3(2)	42.3(4)
					1346	4040	140	242				
					25.68%	74.32%	36.65%	63.35%				

[a] $Pref_{bs}$ and $Pref_{sb}$ are the preference of residues to move from buried to surface and surface to buried, respectively; $< R_{Ac} >$ and $< R_{Ag} >$ are solvent accessible reduction ratio for chaperones and globular proteins, respectively; R_{ing}, R_{outg}, R_{inc} and R_{outc}, are respectively, the preference of residues in interior and surface parts of globular and chaperone proteins; G_c and G_g are transfer free energy for chaperones and globular proteins; ASA_{hel} and ASA_{str} are accessible surface area of residues in helical and strand segments.
G_c and G_g are in kcal and the rest are dimensionless quantities.
[b] Data from Ponnuswamy (1993) [50].
[c] Data from Miller et al. (1987) [31,32].

mutations. In 1NGG, 8 non-polar residues moved from inner part to surface and seven polar residues buried inside the protein.

The computed surrounding residues within 8 Å of each mutant site in native and mutant proteins are given in Table 8.

Table 7
The number of polar and non-polar residues moved from interior to surface and vice versa

Name	Interior to surface			Surface to interior		
	Total	Polar	Non-polar	Total	Polar	Non-polar
1NGH	14	5	9	6	2	4
1NGG	14	6	8	10	6	4
1NGB	13	5	8	8	4	4
1NGA	11	4	7	7	4	3
1NGF	7	3	4	12	4	8
1NGE	8	4	4	12	7	5
1ATR	17	7	10	8	4	4
1ATS	20	8	12	5	2	3
1NGD	9	3	6	3	1	2
1NGC	8	4	4	7	3	4

Table 8
The surrounding residues within 8 Å of mutant sites in native and mutant proteins

Proteins	Site	Surrounding residues
3HSC(Native)	D10	G8, I9, L11, G12, S16, C17, V18, V144, T145, V146, E175, V369, A370, A373
	E175	D10, T145, V146, P147, A148, I173, N174, P176, T177, A178, A179, V369
	D199	I197, F198, L200, G201, F205, D206, V207, L336, V337, G338, T341
	T204	T13, L200, G201, G202, G203, F205, D206, G224, D225, T226, H227, L228, G229
	D206	I197, F198, D199, L200, T204, F205, V207, S208, T222, A223, G224, D225, T226
1NGH	10, 175	No change (with respect to Native)
	199	T204(Additional)
	204	D199(Additional)
	206	No change
1NGG	10, 175, 199, 204, 206	No change
1NGB	10	E175 (Missed)
	175	D10 (Missed)
	199	T204 (Additional)
	204	D199 (Additional)
	206	No change
1NGA	10, 175	No change
	199	T204 (Additional)
	204	D199 (Additional)
	206	No change
1NGF	10, 175	No change
	199	T204 (Additional)
	204	D199 (Additional), G224 (Missed)
	206	No change
1NGE	10, 175	No change
	199	T204 (Additional)
	204	D199 (Additional), G224 (Missed)
1ATR	10, 175, 199	No change
	204	G224 (Missed)
	206	No change
1ATS	10	E175 (Missed)
	175	D10 (Missed)
	199	No change
	204	T13 (Missed)
	206	No change
1NGD	10, 175, 199, 204, 206	No change
1NGC	10, 175,	No change
	199	T204 (Additional)
	204	D199 (Additional), G224 (Missed)
	206	No change

It is observed that more than 10 residues are present around each mutant site of native protein. A survey on the displacement of residues around each mutant site due to amino acid substitutions reveals no major changes in each site. Most of the residues at the mutant site keep the neighbouring and surrounding residues with them and in some site alone, the change is with one or two residues.

Though the behaviour of the surrounding residues changes in their accessibility, it is important to point out that the residue clusters are maintaining particular positions without moving much from the central residue.

3.5. Reduction in accessibility in Hsc70 ATPase fragments

The reduction in accessibility computed for all the 20 amino acid residues in molecular chaperones and of globular proteins are given in Table 6.

We note that the hydrophobic residues Ala, Gly, Val, Leu, Ile, Met, Phe, Trp, Tyr and Cys and the polar residues Gly, Ser, Thr, and His are buried to an extent of 88.7% and 71%, respectively in Hsc70 proteins.

Incidentally, the residue Cys buried to 100% in Hsc70s followed by Phe, Leu, Val and Trp. In globular proteins, Cys is buried to 90% and other residues Phe (91%), Ile (89%) and Leu and Trp (88%) follow almost the same level [50].

The highly exposed residues in chaperones is His followed by Lys, Arg, Ser, and Asp whereas in globular proteins the order is Lys, Glu, Pro, Gln, Arg and Asp. A good correlation ($r = 0.995$) is obtained between the surrounding hydrophobicity and the reduction in accessibility of amino acid residues in chaperone Hsc70 proteins, a similar trend was noticed in globular proteins also [54].

3.6. Surface and interior seeking amino acid residues in the chaperone Hsc70

The percentage of amino acid composition in the interior (R_{in}) (less than 5% accessibility) and their surface (R_{out}) (residues with more than 5%) parts of molecular chaperones is given in Table 6.

We note that 37% of residues in 3HSC are in buried part and 63% of residues are in surface part. The hydrophobic residues Val, Ala, Ile and Leu prefer to be in interior and the charged residues Lys, Asp, Glu and Arg prefer to be in surface, a normal trend observed in globular proteins [31]. When we consider the overall residue preferences a different kind of trend is observed for Ala i.e., only 2.8% is buried in globular proteins whereas 5.2% residues are buried in Hsc70.

The computed transfer free energy (Table 6) shows that Trp has negative free energy, indicating its preference to be outside the protein whereas Pro prefers to be interior of the protein Hsc70. The correlation of transfer free energy between Hsc70 and globular protein [31] is estimated to be 0.64. The correlation obtained with the transfer free energies of Hsc70 and its mutants with other globular protein scales are rather very poor. The computed values for the selected scales by Nozaki–Tanford–Jones [55], Kyte–Doolittle [56], Fauchere–Pliska [57], Eisenberg–McLachlan [30], Ponnuswamy [50], Ponnuswamy–Gromiha [58] are -0.22, 0.65, 0.21,0.20, 0.12, and 0.22, respectively. When compared to the globular proteins, it seems the chaperones need minimum free energy to assume the folded state from the unfolded state. This may be the reason that the chaperones fold quickly and helps others to attain their native conformation rather than misfolding or aggregation.

The ASAs calculated for 20 different amino acid residues in helices and strands are given in Table 6. In Hsc70s, the residues Ala, Arg, Lys, Glu, Asp and Leu prefer the helix part whereas in globular pro-

Table 9
The free energy contribution of chaperone protein 3HSC and its mutants[a]

Name	N	I_{tot}	I_b	I_s	G_{hy}	G_{el}	G_{hb}	G_{vw}	G_{ss}	G_U	G_f	ΔG
3HSC	382	21	2	19	252.86	42.60	257.86	62.86	0.0	587.33	616.18	28.85
1NGH	378	14	2	12	259.37	35.60	261.94	62.30	0.0	584.57	619.21	34.64
1NGG	379	14	1	13	252.63	33.60	262.67	62.44	0.0	586.14	611.34	25.20
1NGB	378	13	1	12	254.57	32.60	262.94	62.30	0.0	585.07	612.41	27.34
1NGA	378	15	1	14	253.88	34.60	260.94	62.30	0.0	584.07	611.72	27.65
1NGF	382	18	2	16	255.53	39.60	260.86	62.86	0.0	588.83	618.85	30.02
1NGE	378	19	1	18	253.25	38.60	256.94	62.30	0.0	582.07	611.09	29.02
1ATR	383	14	2	12	264.57	35.60	265.59	63.00	0.0	592.40	628.76	36.37
1ATS	382	18	2	16	263.71	39.60	260.86	62.86	0.0	588.83	627.03	38.20
1NGD	378	18	1	17	252.28	37.60	257.94	62.30	0.0	582.57	610.12	27.55
1NGC	378	20	2	18	250.12	41.60	255.94	62.30	0.0	581.57	609.96	28.39

[a] N, Total number of residues.
I_b, I_s, and I_{tot} are the buried, surface and total number of ion pairs, respectively.
G_{hy}, hydrophobic free energy; G_{el}, electrostatic free energy; G_{hb}, hydrogen bonding free energy; G_{vw}, van der Waals free energy; G_{ss}, disulfide free energy; G_u, unfolded energy; G_f, folded state free energy; ΔG, stability of the protein.

teins, the most preferred residues are Ala, Leu, Glu, Lys and Val [59,60]. Cys, Phe and Gly are buried in helices while the residues Arg, Lys and Tyr are exposed.

The residues Val, Ile, Asp and Thr are the highly preferred in strands of Hsc70 similar to globular proteins [59,60]. Ala, Cys, His and Pro are buried and the residues Met(107.9), Arg(88.0) and Glu(74.4) are highly exposed to the solvent in strands.

In our analysis, we noticed that 121 residues moved from buried to surface and 78 residues from surface to buried and 48% of the residues are from the secondary structures, alpha helices and beta strands.

The various free energy terms responsible for the conformational stability of all the proteins considered were computed by the method of Ponnuswamy and Gromiha [37] and the results are presented in Table 9. Due to a single substitution, it is to be noted that the hydrophobic free energy changes by more than 10 kcal/mol in the case of 1ATR and 1ATS. In 1NGA, 1NGB, 1NGG, 1NGH and 1ATR the electrostatic free energies lowered its value from that of native and this may be due to the loss of ion pairs. The stability of the mutant proteins also varies to the level of 10 kcal/mol due to the changes in ASA.

Further, the table shows that the hydrophobic free energy acquires 52.6% contribution and other factors such as electrostatic, hydrogen bonding and van der Waals interactions contribute 7.69%, 26.87% and 12.86%, respectively. The percentage contributions from hydrophobic, van der Waals and electrostatic interactions are slightly higher than the globular protein values (hydrophobic, 50.80%; hydrogen bonding, 27.1%; Van der Waals, 14.6%; electrostatic, 6.4%; and disulfide; 1.1%; Ponnuswamy and Gromiha [37]). It is interesting to note that the conformational stability is in the range of 25 to 40 kcal/mol observed for some globular proteins [37,61].

4. Conclusions

The solvent accessibility studies on the mutants of Hsc70 proteins provide useful information about the types of interactions and other essential features to understand their folding and stability. The accessible surface area of several residues change due to amino acid substitution and the clusters of residues around the central residue maintains their positions to be together, not moving apart. The variation of ASA increases the hydrophobic free energy to a level of more than 10 kcal/mol. The residues Ala, Asp and Leu shift their positions easily to adopt with the local environment with secondary interactions. Although the physicochemical properties of amino acid residues like reduction in accessibility, preference of residues in interior and surface part, etc., are similar to those of globular proteins the transfer free energy differs much.

Acknowledgements

T.S.K. thanks the Council of Scientific and Industrial Research, Government of India for the award of Senior Research Fellowship. M.M.G. is a recipient of an ICGEB fellowship.

References

[1] C.B. Anfinsen, Science 181 (1973) 223.
[2] R.J. Ellis, Nature 328 (1987) 378.
[3] M.J. Gething, J. Sambrook, Nature 355 (1992) 33.
[4] J.P. Hendrick, F.U. Hartl, Ann. Rev. Biochem. 62 (1993) 349.
[5] F.U. Hartl, R. Hlodan, T. Langer, Trends Biochem. Sci. 19 (1994) 20.
[6] F.U. Hartl, Nature 381 (1996) 571.
[7] K. Braig, Z. Otwinoski, R. Hedge, D.C. Boisvert, A. Joachimiak, A.L. Harwich, P.B. Sigler, Nature 371 (1994) 578.
[8] D.C. Boisvert, J. Wang, Z. Otwinowski, A.L. Horwich, P.B. Sigler, Nat. Struct. Biol. 3 (1996) 170.
[9] J.F. Hunt, A.J. Weaver, S.J. Landry, L. Gierasch, J. Deisenhofer, Nature 379 (1996) 37.
[10] S.C. Mande, V. Mehra, B. Bloom, W.G.J. Hol, Science 271 (1996) 203.
[11] T. Langer, C. Lu, H. Echols, J. Flanagan, M.K. Hayer, F.U. Hartl, Nature 35 (1992) 683.
[12] K. Braig, M. Simon, F. Furuya, J.F. Hainfeld, A.L. Horwich, Proc. Natl. Acad. Sci. USA 90 (1993) 3978.
[13] W. Neupert, F.U. Hartl, E.A. Craig, N. Pfanner, Cell 63 (1990) 447.
[14] G.C. Flynn, T.G. Chappell, J.E. Rothman, Science 248 (1989) 385.
[15] G.C. Flynn, J. Pohl, M.T. Flocco, J.E. Rothman, Nature 353 (1991) 726.

[16] K.M. Flaherty, C. DeLuca-Flaherty, D.B. McKay, Nature 346 (1990) 623.
[17] S. Blond-Eliguindi, S.E. Cwirla, W.J. Dower, R.J. Lipshutz, S.R. Sprang et al., Cell 75 (1993) 717.
[18] A. Gragerov, M.E. Gottesman, J. Mol. Biol. 24 (1994) 133.
[19] T.M. Takenaka, S.M. Leung, S.J. Mcandrew, J.P. Brown, L.E. Hightower, J. Biol. Chem. 270 (1995) 19839.
[20] T.G. Chappell, B.B. Konforti, S.L. Schmid, J. E Rothman, J. Biol. Chem. 262 (1987) 746.
[21] F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, J. Mol. Biol. 102 (1977) 535.
[22] Protein Data Bank, Brookhaven National Laboratory, Upton, New York, 1995.
[23] B. Lee, F.M. Richards, J. Mol. Biol. 55 (1971) 379.
[24] C. Chothia, Nature (London) 248 (1974) 338.
[25] C. Chothia, J. Mol. Biol. 105 (1976) 1.
[26] F.M. Richards, J. Mol. Biol. 82 (1974) 1.
[27] F.M. Richards, Annu. Rev. Biophys. Bioeng. 6 (1977) 151.
[28] T.J. Richmond, J. Mol. Biol. 178 (1984) 63.
[29] G.D. Rose, A.R. Geselowitz, G.J. Lesser, R.H. Lee, M.H. Zehfus, Science 229 (1985) 834.
[30] D. Eisenberg, A.D. McLachlan, Nature (London) 319 (1986) 199.
[31] S. Miller, J. Janin, A.M. Lesk, C. Chothia, J. Mol. Biol. 196 (1987) 641.
[32] S. Miller, A.M. Lesk, J. Janin, C. Chothia, Nature (London) 328 (1987) 834.
[33] J. Janin, S. Miller, C. Chothia, J. Mol. Biol. 204 (1988) 155.
[34] J. Janin, Bull. Inst. Pasteur 86 (1988) 21.
[35] G. Raghunathan, R.L. Jernigam, K.L. Ting, A.J. Sarai, Biomol. Struct. Dynam. 8 (1990) 187.
[36] P.K. Ponnuswamy, in: P. Balaraman, S. Ramaseshan (Eds.), Molecular Conformation and Biological Interactions, 1991, pp. 337.
[37] P.K. Ponnuswamy, M.M. Gromiha, J. Theor. Biol. 166 (1994) 63.
[38] P.K. Ponnuswamy, M.M. Gromiha, J. Theor. Biol. 169 (1994) 419.
[39] T.J. Richmond, F.M. Richards, J. Mol. Biol. 199 (1978) 537.
[40] M.M. Gromiha, P.K. Ponnuswamy, Int. J. Pept. Protein Res. 48 (1996) 452.
[41] P. Manavalan, P.K. Ponnuswamy, Arch. Biochem. Biophys. 184 (1977) 476.
[42] P. Manavalan, P.K. Ponnuswamy, Nature 275 (1978) 673.
[43] M.F. Perutz, H. Raidt, Nature 255 (1975) 256.
[44] L.R. Brown, A. DeMarco, A. Richatz, G. Wagner, K. Wuthrich, Eur. J. Biochem. 88 (1978) 87.
[45] A.R. Fersht, J. Mol. Biol. 64 (1972) 497.
[46] D.J. Barlow, J.M. Thornton, J. Mol. Biol. 168 (1983) 867.
[47] D. Sali, M. Bycroft, A.R. Fersht, Nature 335 (1988) 740.
[48] H. Nicholson, W.J. Bectal, B.W. Matthews, Nature 336 (1988) 651.
[49] A. Ben-Naim, J. Phys. Chem. 95 (1991) 1437.
[50] P.K. Ponnuswamy, Prog. Biophys. Mol. Biol. 59 (1993) 57.
[51] J.M. Thornton, J. Mol. Biol. 151 (1981) 261.
[52] W. Kauzmann, Sulfur in Proteins, Academic Press, New York, 1959.
[53] F.H. Stillinger, Science 209 (1980) 451.
[54] P.K. Ponnuswamy, M. Prabakaran, P. Manavalan, Biochem. Biophys. Acta 623 (1980) 301.
[55] D.D. Jones, J. Theor. Biol. 50 (1975) 167.
[56] J. Kyte, R.F. Doolittle, J. Mol. Biol. 157 (1982) 105.
[57] J.L. Fauchere, V. Pliska, Eur. J. Med. Chem. Chim. Ther. 18 (1983) 366.
[58] P.K. Ponnuswamy, M.M. Gromiha, Int. J. Pept. Protein Res. 42 (1993) 326.
[59] G.D. Fasman, in: G.D. Fasman (Ed.), Prediction of Protein Structures and Principles of Protein Conformation, Plenum, New York, 1989, pp. 193–316.
[60] M.M. Gromiha, P.K. Ponnuswamy, Int. J. Pept. Protein Res. 45 (1995) 225.
[61] T. Ooi, M. Oobatake, J. Biochem. 104 (1988) 440.