FEBS Letters 445 (1999) 126–130 FEBS 21582

Domain organization of flagellar hook protein from *Salmonella typhimurium*

Hatsuho Uedaira[a,b], Hisayuki Morii[a], Miyuki Ishimura[a], Hisaaki Taniguchi[c], Keiichi Namba[d], Ferenc Vonderviszt[d,e,*]

[a]*National Institute of Bioscience and Human-Technology, 1-1 Higashi, Tsukuba 305-8566, Japan*
[b]*Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba 305-0074, Japan*
[c]*Institute for Comprehensive Medical Science, Fujita Health University, Kutsukake-cho, Toyoake 470-1192, Japan*
[d]*International Institute for Advanced Research, Matsushita Electric Industrial Co., Ltd., 3-4 Hikaridai, Seika 619-0237, Japan*
[e]*Department of Physics, University of Veszprém, Egyetem u. 10, H-8200 Veszprém, Hungary*

Received 14 January 1999

Abstract Hook forms a universal joint, which mediates the torque of the flagellar motor to the outer helical filaments. Domain organization of hook protein from *Salmonella typhimurium* was investigated by exploring thermal denaturation properties of its proteolytic fragments. The most stable part of hook protein involves residues 148 to 355 and consists of two domains, as revealed by deconvolution analysis of the calorimetric melting profiles. Residues 72–147 and 356–370 form another domain, while the terminal regions of the molecule, residues 1–71 and 371–403, avoid a compact tertiary structure in the monomeric state. These folding domains were assigned to the morphological domains of hook subunits known from EM image reconstructions, revealing the overall folding of hook protein in its filamentous state.

Key words: Bacterial flagellum; Hook; Domain structure; Calorimetry; *Salmonella typhimurium*

1. Introduction

The helical filament of the bacterial flagellum is connected to a complex molecular motor located within the cell envelope through a highly curved tubular structure, called the hook ([1,2] and references therein). The hook plays the role of a flexible coupling that can transmit the torque, generated by the motor, to the filament in a wide range of directions. The hook is constructed from subunits of a single protein (FlgE) by a self-assembly process [3]. The molecular mass of hook protein varies considerably from species to species.

The polypeptide chain of *Salmonella* hook protein contains 403 amino acid residues [4]. Both terminal parts of hook protein are disordered in solution, involving about 70 NH_2-terminal and 30 COOH-terminal residues [5]. Upon polymerization, the disordered terminal regions fold into a predominantly α-helical domain [6]. It has been demonstrated that polymerization ability is lost completely upon removal of both disordered terminal regions of hook protein [5]. Analysis of the calorimetric melting profile of monomeric hook protein suggested that its compact part is constructed from three domains, which are β-sheeted structures as revealed by far-UV CD studies and secondary structure prediction [6].

In this study, we identified folding domains of hook protein by studying thermal unfolding properties of its major proteolytic fragments. It was found that in spite of its sequential dissimilarity, domain organization of hook protein shows many similarities to that of flagellin, which is the major component of flagellar filaments. We suggest that hook subunits and flagellin subunits also have a similar folding pattern in their filamentous state. Based on this assumption, the morphological domains of hook subunits, found earlier by EM image reconstruction [7], were assigned to the sequence.

2. Materials and methods

2.1. Preparation of hook protein and its fragments

Hook protein from *S. typhimurium* strain SJW880 was isolated and purified as described previously [6]. The H32 and H22 fragments of hook protein were purified from tryptic digestion mixtures by FPLC ion exchange chromatography [5]. Isolated fragments were stored at −20°C. The concentration of H32 and H22 was determined from absorption measurements at 280 nm. The extinction coefficients of $E^{1\%}_{280} = 7.32$ and $E^{1\%}_{280} = 9.42$ were used for H32 and H22, respectively, which were calculated from the known aromatic amino acid content of the molecule [8].

NH_2-terminal sequence analysis of the purified H22 fragment was performed using a gas-phase protein sequencer (Applied Biosystems model 477A) equipped with a PTH analyzer (Applied Biosystems model 130A). Then the COOH-terminal position of the H22 fragment was identified by electrospray mass spectrometry using a Sciex API-III mass spectrometer (PE Sciex, Ontario, Canada) [9].

2.2. Calorimetry

Differential scanning calorimetry (DSC) measurements were performed with a Microcal MCS adiabatic scanning microcalorimeter. All calorimetric measurements were performed in 10 mM sodium phosphate buffer solution (pH 7.0) at sample concentrations between 0.6 and 1.3 mg/ml and a scanning rate of 1°C/min. In the calculations of molar thermodynamic quantities, the molecular weights used were 31.3 and 21.6 kDa for H32 and H22, respectively, calculated from the known amino acid sequence. The heat capacity functions of hook fragments were analyzed by the double deconvolution procedure of Kidokoro and Wada [10], assuming that the native state heat capacities are quadratic functions of the temperature and the heat capacity changes between the native and intermediate states are constants. The obtained thermodynamic parameters were further adjusted by a nonlinear least-square fitting method [11] using the SALS program [12].

2.3. Fluorescence measurements

Thermal unfolding transitions of H32 and H22 were also monitored by fluorescence spectroscopy. Measurements were done in 10 mM phosphate buffer solutions (pH 7.0) at a heating rate of 1°C/min. Fluorescence measurements were made with a Shimadzu RF-5000 spectrofluorometer equipped with a thermoregulated cell holder. The intrinsic fluorescence intensities of the samples were monitored as a function of temperature using an excitation and emission wavelength of 295 nm and 350 nm, respectively. Protein concentrations were in the range of 6–13 μg/ml.

*Corresponding author. Department of Physics, University of Veszprém, Egyeten u. 10, H-8200 Veszprém, Hungary.
Fax: (36)-88-422103. E-mail: von007@almos.vein.hu

PII: S0014-5793(99)00110-6

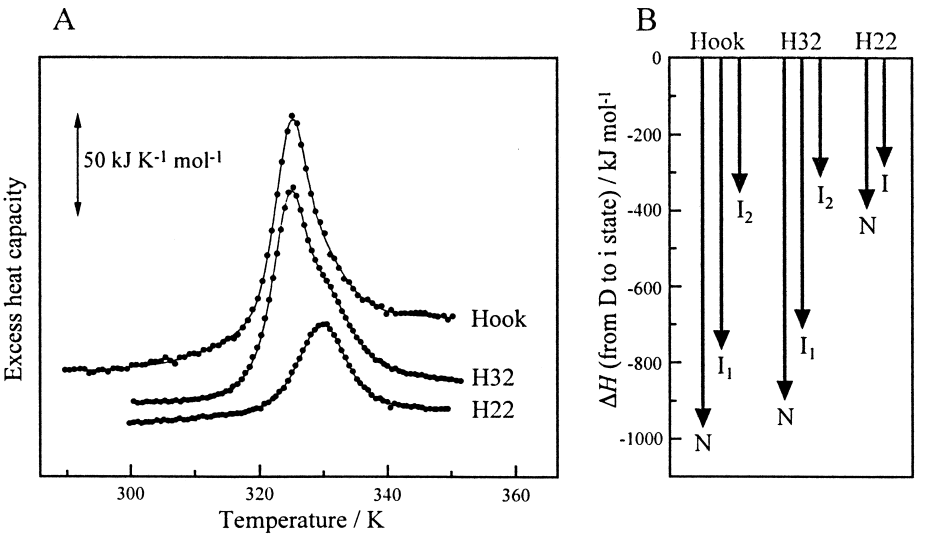

Fig. 1. A: Temperature dependence of the molar heat capacity of H32 and H22 fragments of hook protein. The molar heat capacity curve of intact hook protein is also displayed for comparison. The filled circles represent every 3rd collected data points. The continuous lines represent the calculated transition curves assuming a four-state transition for hook protein and H32, while H22 was approximated by a three-state transition model. B: Diagram of enthalpy differences for the native and intermediate states of hook protein, H32 and H22 relative to the denatured state at 320 K. Measurements were performed in 10 mM phosphate buffer (pH 7.0), at protein concentrations of 30 μM and 41 μM for H32 and H22, respectively. Scanning rate was 1°C/min.

3. Results

3.1. Tryptic fragments of hook protein

Limited proteolysis of hook protein from *Salmonella* results in a quick degradation of the disordered terminal regions of the molecule, and yields an intermediate fragment of 32 kDa (H32). In our previous study, H32 was identified to contain residues 72 to 370 of hook protein [5]. Further tryptic treatment of H32 results in the production of a 22 kDa fragment (H22), which can be slowly digested into even smaller species [5]. H32 and H22 were purified from tryptic hydrolysates by FPLC anion exchange chromatography. NH_2-terminal sequencing revealed that H22 lacks an additional 76 NH_2-terminal residues, as compared to H32. The apparent molecular mass of H22 together with the distribution of the possible tryptic cleavage sites suggested that its COOH-terminal sequence position should be at Lys-370 or Lys-355. High resolution electrospray mass spectrometry was used to unambiguously identify H22. The obtained molecular mass of H22 (21.6 kDa) perfectly matched that of the 148–355 fragment.

3.2. Calorimetric experiments

The thermal unfolding properties of intact hook protein and polyhook filaments have been reported previously [6]. In an attempt to reveal domain organization of hook protein, its major proteolytic fragments, H32 and H22, were also studied by differential scanning calorimetry.

Melting profiles of H32 and H22 are presented in Fig. 1A, where the melting profile of intact hook protein is also given for comparison. H32 exhibits a complex endotherm having a shoulder on the high temperature side, while H22 shows a symmetrical peak at the shoulder position of H32. Sequential heating experiments show that, like intact hook protein, thermal denaturation of the fragments is also highly reversible, indicating that the fragments are composed of autonomous folding units.

Thermodynamic parameters characteristic for the heat denaturation process are shown in Table 1. The denaturation temperature as well as the total transitional enthalpy value of H32 are very similar to those observed previously for hook protein [6], suggesting that the terminal parts have no significant internal stability. The melting profile of H22 is shifted slightly to higher temperatures, and its transitional enthalpy, corresponding to the heat absorption peak area, is less than half of H32. The ratio of the calorimetric and the effective van't Hoff enthalpies is significantly larger than 1.0 in both cases; therefore, these calorimetric profiles reflect multistate transitions [13], suggesting that H32 and H22 are composed of multiple cooperative folding units, or domains. The double deconvolution method of Kidokoro and Wada [10] followed by a non-linear least-square fitting procedure [11,12] was applied to analyze the multistate thermal denaturation of H32 and H22. The results of this analysis are illustrated in Table 2. Thermal denaturation of H22 can be nicely described by a model assuming a single intermediate state (three-state model), which gives an excellent fit to the measured heat capacity curve (Fig. 1). Thus, our analysis indicates the existence of two domains (G2 and G3) in H22. In the case of H32, a three-state model resulted in a poor fit to the observed heat capacity curve. However, assumption of a four-state transitional mechanism results in a very good approximation to the measured heat capacity profile with an r.m.s. deviation of 0.49 μW comparable to the normal noise level of our equipment. Therefore, in addition to the two domains of H22, H32 ap-

Table 1
Thermodynamic characteristics of the thermal denaturation of H32 and H22 fragments of hook protein

	T_m (K)	ΔH^{cal} (kJ mol^{-1})	ΔH^{vH} (kJ mol^{-1})	$\Delta H^{cal}/\Delta H^{vH}$	$\Delta_d C_p$ (kJ K^{-1} mol^{-1})
H32	325.4	1032	318	3.3	20
H22	328.0	476	300	1.6	5

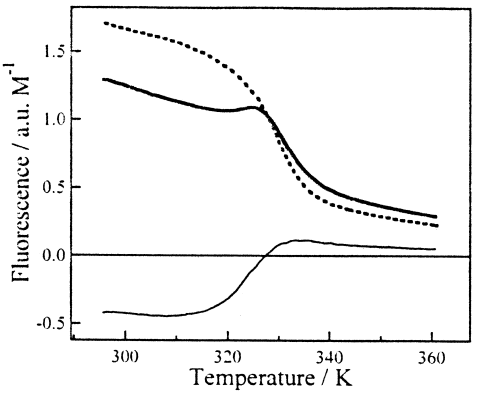

Fig. 2. Thermal denaturation of H32 (thick line) and H22 (dotted line) as followed by intrinsic fluorescence. The difference profile (thin line) shows a single transition around 325 K. Measurements were performed in 10 mM phosphate buffer (pH 7.0) solutions using an excitation and emission wavelength of 295 nm and 350 nm, respectively. Heating rate was 1°C/min.

pears to contain a domain (G1) discontinuous in the amino acid sequence, which is composed of segments Gly-72 to Lys-147 and Thr-356 to Lys-370.

Enthalpy values characteristic for the native and intermediate states of hook protein and its fragments relative to the denatured state at 320 K are compared in Fig. 1B. It is clearly seen that unfolding intermediates of hook protein and H32 are very similar concerning their characteristic enthalpy values. Together with the refolding ability of H32, this convincingly demonstrates that domains of H32 are in their native-like states; they are not significantly altered upon removal of the terminal regions. On the other hand, one of the domains of H22 seems to undergo partial destabilization as a result of the proteolytic treatment. Its stabilization enthalpy is significantly reduced, which may reflect the loss of interactions with the G1 domain of H32. The other domain of H22 exhibits a transition temperature and a denaturation enthalpy close to those obtained for the most stable domain of hook protein or H32.

3.3. Fluorescence measurements

Calorimetric experiments revealed that the isolated H22 fragment unfolds at higher temperatures than intact hook protein or H32. The fact that H22 exhibits a melting transition at the shoulder position of H32 suggests that H22 is the most stable part of the molecule. To demonstrate that the domains of H22 constitute the most stable part within H32 (or hook protein) as well, the unfolding transitions were also monitored by measuring the internal fluorescence intensity as a function of temperature.

Hook protein from *Salmonella* contains 11 tyrosine and two tryptophan residues [4]. Both Trp residues (Trp-208 and Trp-334) together with seven Tyr residues are located in the H22 region. The third domain of H32 contains only two additional Tyr residues, while the remaining two tyrosines are in the disordered terminal regions. Thus, the near-UV spectral properties of hook protein or H32 are dominated by the H22 region.

Trp fluorescence of H32 and H22 was monitored as a function of temperature using an excitation wavelength of 295 nm and detecting fluorescence at 350 nm (Fig. 2). While H22 exhibited a monotonously decreasing fluorescence signal with an unfolding transition around 330 K, H32 produced a complex profile with a local maximum around 327 K and a sharp decrease in the region around 330 K. The difference spectrum, however, gives a single sigmoidal transition of an increasing fluorescence intensity around 325 K, indicating that the G1 domain, corresponding to regions of H32 not involved in H22, unfold first around 325 K, followed by a conformational transition in the H22 region. Although G1 does not contain any Trp residues, still its unfolding resulted in an increase of the fluorescence of tryptophans located in H22. This observation indicates that the presence of intact G1 gives rise to a quenching process decreasing the fluorescence yield

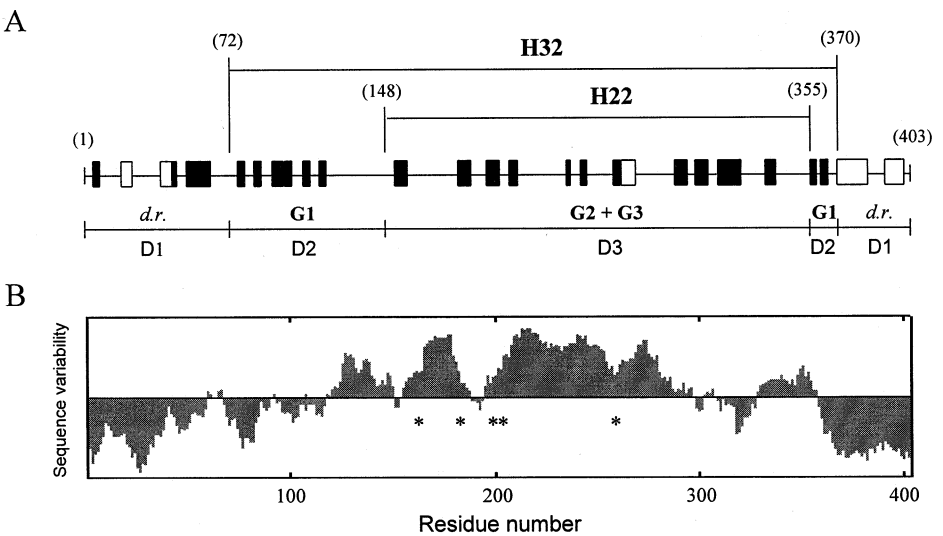

Fig. 3. A: Sequence positions of the proteolytic fragments (H32 and H22) of hook protein, positions of disordered regions (*d.r.*) and domains (G1–G3) identified by calorimetric analysis, predicted secondary structures (open box, α-helix; closed box, β-structure; [6]) and assignment of the morphological domains (D1 to D3) identified by EM image reconstructions [14]. B: Relative sequence variability profile of hook protein. Based on a multiple sequence alignment of the available hook sequences, sequence variability was calculated along the amino acid sequence of *Salmonella* hook protein by MaxHom [23] and plotted relative to its average value. Asterisks indicate the positions of insertions larger than 30 residues found in various hook sequences.

Table 2
Thermodynamic parameters for the melting transitions of hook protein fragments obtained by deconvolution analysis of the heat capacity curves

	Transitions								
	1			2			3		
	T_{01}[a] (K)	ΔH_{01}[b] (kJ mol^{-1})	ΔCp_{01}[c] (kJ K^{-1} mol^{-1})	T_{02} (K)	ΔH_{02} (kJ mol^{-1})	ΔCp_{02} (kJ K^{-1} mol^{-1})	T_{03} (K)	ΔH_{03} (kJ mol^{-1})	ΔCp_{03} (kJ K^{-1} mol^{-1})
H32	316.9 ± 0.4	164 ± 7	7 (fix)	321.7 ± 0.1	627 ± 7	23 (fix)	324.5 ± 0.1	1024 ± 8	27 (fix)
H22	329.0 ± 1.7	112 ± 8	0 (fix)	328.0 ± 0.2	432 ± 4	5.1 ± 0.2			

[a] T_{0i}, the equimolar fraction temperature of the native and ith states.
[b] ΔH_{0i}, the enthalpy difference of the ith state compared to the native state.
[c] ΔCp_{0i}, the heat capacity difference of the ith state compared to the native state.

of Trp fluorescence from G2 and G3. Thus, internal fluorescence measurements support the interpretation of calorimetric data that the domains of H22 correspond to the domains of H32 with the highest stability.

Similar conclusions can be obtained by monitoring thermal unfolding of H32 by CD spectroscopy in the near-UV region (data not shown). While hook protein or H32 gave a calorimetric melting peak with a transition temperature around 325 K, following their unfolding transition by CD in the near-UV region resulted in apparently a single transition around 328 K, the same transition temperature as obtained for H22. Since the majority of aromatic residues are situated in the H22 portion, these observations also indicate that H22 is the most stable part of hook protein.

4. Discussion

Heat capacity profiles of hook protein and its H32 fragment, deprived of the terminal regions, are very similar. Both of them can be nicely fitted by a four-state transition model with almost identical ΔH_{0i} and ΔCp_{0i} values ([6], this study). The terminal regions of hook protein appear to have a very small contribution to the melting transition indicating that these regions have no significant internal stability. Thus, calorimetric experiments reconfirm the conclusion, obtained previously from proteolytic experiments and NMR studies [5], that the terminal regions of hook protein are disordered in solution. They have no compact ordered structure and are in extensive contact with the surrounding media. The compact part of monomeric hook protein, the H32 portion, contains three cooperative folding units or domains. Analysis of the H22 fragment, comprising residues of 148 to 355, suggests that two domains are involved in H22 while the third domain of H32 is a discontinuous one, composed of residues 72 to 147 and 356 to 370 (Fig. 3A). H22 corresponds to the thermodynamically most stable part of hook protein.

Like hook is constructed from multiple copies of hook protein, the outer helical filaments of bacterial flagella are also composed of a single kind of protein, flagellin [1]. Both proteins exhibit very similar assembly characteristics: they pack in almost identical lattices [7,14,15] and have the ability to self-assemble [3,16]. Although there is no significant sequential homology between flagellin and hook protein [4], the Prop-Search algorithm [17], which is devised to recognize distant structural relationships not detected by common sequence alignment programs, strongly suggests that hook protein and flagellin are related to each other (F.V. unpublished results). Our present results also illuminate their close relationship: their domain structure looks very similar. Flagellin also has disordered terminal regions [18,19], whereas its compact portion (F40) is similarly constructed from three domains [20]. The highly conserved terminal portions of both molecule contain heptad repeats of hydrophobic amino acids and exhibit coiled-coil-forming propensities [4,6,21]. However, while flagellin contains several extended helical bundle-forming segments even at the conserved ends of its compact central part, hook protein exhibits coiled-coil-forming preferences only in its disordered terminal regions.

How do the domains of hook protein contribute to hook formation? We suggest that the hook subunit is folded essentially in the same way as has been revealed for flagellin subunits in the filament [1,21,22]. This assumption is consistent with all of the available biochemical and structural data, however, we can not claim that there are no other possible interpretations as well.

An EM analysis by Morgan et al. [7] revealed that the hook subunit is composed of three morphological domains: the rod-shaped inner domain (D1), the middle spherical domain (D2) and the outer ellipsoidal domain (D3). The volume fractions of these structural domains were estimated to be 0.28, 0.28 and 0.44 for D1, D2 and D3, respectively, based on the data of Morgan et al. [7]. These values correspond to molecular masses of 12 kDa, 12 kDa and 19 kDa, respectively, assuming a uniform density distribution. We assigned the folding domains found by calorimetric analysis (G1–G3) to the morphological domains (D1–D3) as follows (Fig. 3A):

Comparison of the available hook sequences from various bacterial species shows that the NH_2-terminal half of H22, residues 148 to 260, is exceptionally variable (Fig. 3B), containing several large insertions and deletions. This strongly suggests that the amino-terminal domain of H22 is exposed on the surface of hook. Indeed, the EM structure of *C. crescentus* hook, which is constructed from a hook protein containing an almost 150 residue long insertion in the variable region, shows an extra subdomain on the outer surface [7]. Since the estimated mass of D3 (19 kDa) is close to the molecular mass of H22 of 21.6 kDa, we assume that the D3 structural domain consists of the two strongly interacting domains of H22, namely G2 and G3.

Then D2 can be assigned to the G1 domain of H32, comprising residues 72–147 and 356–370, whose estimated molecular mass of 12 kDa fits well to that (10.5 kDa) calculated from the amino acid sequence. This assignment is reasonable, because G1 is closely associated with H22, together forming the compact H32 portion.

In contrast to the β-structural central portion, the highly

conserved disordered terminal regions of hook protein become stabilized upon hook formation into a predominantly α-helical domain [5], which seems to contain bundles of α-helices as suggested by coiled-coil prediction [6]. The disordered terminal regions of flagellin also fold up into axially oriented bundles of α-helices that gives rise to the innermost, rod-shaped domain seen in density maps [14,22,24]. Thus, it is plausible to assume that the innermost, rod-like domain of the hook subunit (D1) is formed by the terminal regions [7]. The disordered terminal regions altogether contain about 105 residues, which almost perfectly matches the estimated size of D1.

Acknowledgements: We thank Drs. F. Oosawa and T. Nitta for support and encouragement. This work was partially supported by Special Coordination Funds of the Science and Technology Agency of Japan to K.N., and the Hungarian OTKA T025468 grant to F.V.

References

[1] Namba, K. and Vonderviszt, F. (1997) Quart. Rev. Biophys. 30, 1–65.
[2] DeRosier, D.J. (1995) Curr. Opin. Struct. Biol. 5, 187–193.
[3] Kato, S., Aizawa, S.-I. and Asakura, S. (1982) J. Mol. Biol. 161, 551–560.
[4] Homma, M., DeRosier, D.J. and Macnab, R.M. (1990) J. Mol. Biol. 213, 819–832.
[5] Vonderviszt, F., Ishima, R., Akasaka, K. and Aizawa, S.-I. (1992) J. Mol. Biol. 226, 575–579.
[6] Vonderviszt, F., Závodszky, P., Ishimura, M., Uedaira, H. and Namba, K. (1995) J. Mol. Biol. 251, 520–532.
[7] Morgan, D.G., Macnab, R.M., Francis, N.R. and DeRosier, D.J. (1993) J. Mol. Biol. 229, 79–84.
[8] Gill, S.C. and von Hippel, P.H. (1989) Anal. Biochem. 182, 319–326.
[9] Taniguchi, H., Manenti, S., Suzuki, M. and Titani, K. (1994) J. Biol. Chem. 269, 18299–19302.
[10] Kidokoro, S. and Wada, A. (1987) Biopolymers 26, 213–229.
[11] Kidokoro, S., Uedaira, H. and Wada, A. (1988) Biopolymers 27, 271–297.
[12] Nakagawa, T. and Oyanagi, Y. (1980) in: K. Matusita (Ed.), Recent Developments in Statistical Inference and Data Analysis, North-Holland, Amsterdam, pp. 221–225.
[13] Privalov, P.L. (1982) Adv. Protein Chem. 35, 1–104.
[14] Morgan, D.G., Owen, C., Melanson, L.A. and DeRosier, D.J. (1995) J. Mol. Biol. 249, 88–110.
[15] Mimori, Y., Yamashita, I., Murata, K., Fujiyoshi, Y., Yonekura, K., Toyoshima, C. and Namba, K. (1995) J. Mol. Biol. 249, 69–87.
[16] Asakura, S. (1970) Adv. Biophys. (Japan) 1, 99–155.
[17] Hobohm, U. and Sander, C. (1995) J. Mol. Biol. 251, 390–399.
[18] Vonderviszt, F., Kanto, S., Aizawa, S.-I. and Namba, K. (1989) J. Mol. Biol. 209, 127–133.
[19] Kostyukova, A.S., Pyatibratov, M.G., Filimonov, V.V. and Fedorov, O.V. (1988) FEBS Lett. 241, 141–144.
[20] Vonderviszt, F., Uedaira, H., Kidokoro, S. and Namba, K. (1990) J. Mol. Biol. 214, 97–104.
[21] Yamashita, I., Vonderviszt, F., Mimori, Y., Suzuki, H., Oosawa, K. and Namba, K. (1995) J. Mol. Biol. 253, 547–558.
[22] Mimori-Kiyosue, Y., Vonderviszt, F. and Namba, K. (1997) J. Mol. Biol. 270, 222–237.
[23] Sander, C. and Schneider, R. (1991) Proteins 9, 56–68.
[24] Yamashita, I., Hasegawa, K., Suzuki, H., Vonderviszt, F., Mimori-Kiyosue, Y. and Namba, K. (1998) Nature Struct. Biol. 5, 125–132.