Folding behavior of a backbone-reversed protein: Reversible polyproline type II to β-sheet thermal transitions in retro-GroES multimers with GroES-like features

Shubbir Ahmed a,1, Anshuman Shukla a,1, Purnananda Guptasarma a,b,*

* Protein Science and Engineering Division, Institute of Microbial Technology (IMTECH), Chandigarh 160036, India
** Adjunct Faculty in Biology, Indian Institute of Science Education and Research (IISER Mohali), Chandigarh 160026, India

1 These authors contributed equally to the work.

1. Introduction

The structure of any naturally-occurring globular protein is determined by its amino acid sequence. The amino acid sequence has a definite directionality, in that the polypeptide backbone running through the sequence has an N-terminus and a C-terminus. If one were to retain the sequence of side-chains presented by the polypeptide backbone of a naturally-occurring protein, while reversing the direction in which the backbone runs through its sequence (e.g., by synthesizing a new gene encoding a recombinant protein with the exact reverse sequence of the original protein), what would be the likely structural-biochemical consequences of such a reversal? The new sequence would, of course, without doubt constitute an entirely new protein, but would it be a completely unrelated protein, despite the retention of the exact same sequence of side-chains? The possible structural–biochemical consequences of effecting such transformations, known as retro-transformations, have previously been discussed in the literature through a combination of structural–biochemical reasoning and molecular modeling [1–4]. Experimentally too, several groups have synthesized and studied backbone-reversed proteins, and the findings may be summarized as follows: (a) Serrano et al. created the retro forms of an all-beta SH3 domain from alpha-spectrin, a beta-alpha domain of streptococcal protein G, and an all-alpha domain of staphylococcal protein A, and reported that these remain largely unfolded [4], (b) Fassina and Ruvo created the retro form of an all-beta protein, rubredoxin, and found that it fails to fold and bind iron [5], (c) Thomas, Grutter, Gutte and coworkers created the retro form of the all-alpha leucine zipper protein, GCN4, and found that it folds into an extremely well-folded coiled-coil structure like GCN4, amenable to crystallization and structure determination [6,7], (d) Merrifield et al. synthesized the retro form of mellitin, a membrane pore forming protein, and found that it is able to form pores in membranes just like mellitin, with a level of activity comparable to that of mellitin [8], (e) Our group previously created retro forms of two all-beta proteins, retro-HSP12.6, and retro-CspA, and found that while retro-HSP12.6 folds into a thermal unfolding-resistant multimer displaying dramatic, irreversible structural consolidation upon heating [9], retro-CspA folds poorly and deposits readily into amyloid nanofibrils [10]. In summary, therefore, there are now several instances each of the folding of the retro forms of some all-alpha proteins, as well as some all-beta proteins. At the same time, there are also some reported instances in which certain retro forms of both all-alpha and all-beta proteins were reported to be unable to fold. With the question of the folding behavior of retro-proteins thus still remaining open, we attempted a retro-transformation on a well known beta strand-based...
protein, the co-chaperone GroES of *Escherichia coli* [11], to examine how this protein behaves upon backbone reversal. The reasons for selecting this protein were primarily that it is known to have a fairly high content of beta-strand structure, and also that it has a size (10 kDa) that makes it amenable to backbone reversal through new gene synthesis. We have argued earlier that there is a higher probability for beta-strand-based proteins to fold into beta-strand-based forms, and possibly even topologically mirror-imaged forms, following backbone reversal [12]; although alpha helical proteins might also be expected to have a tendency to fold into alpha helical forms after backbone reversal, as evidenced by the behavior of retro-GCN4 [6,7] the probability of their folding to adopt mirror-imaged forms is lower [1,2]. When one synthesizes a novel protein, there is always the hope that the protein will fold and be amenable to crystallographic structure determination. In the case of retro-GroES, there was the hope that we might be able to compare structure formation with that of normal GroES, which would be a first for any beta strand-based protein. We show here that retro-GroES folds and assembles into highly soluble and stable trimers and pentamers which are unfolded by chemical denaturants but show reversible heat-induced structural consolidation involving a structural transformation of PPII-like structures into β-sheet structures. Unfortunately, despite its high solubility, we could not find any conditions under which retro-GroES would crystallize, thus keeping us from examining whether the structure involves any topological mirror-imaging whatsoever. Retro-GroES also failed to bind GroEL, which was expected because binding functions involve the formation of surfaces exquisitely adapted to interact in specific ways and there is no reason to believe that a backbone-reversed protein would form the same surfaces as its parent protein, irrespective of whether it fails to fold, aggregates, or folds into a structure with native-like, or mirror-imaged topology.

### 2. Materials and methods

#### 2.1. Molecular genetic manipulations and design of constructs

From the sequence of the gene encoding the parent protein, GroES, the sequence of a novel gene encoding retro-GroES was created simply by reversing the order of the occurrence of the codons used by the parent gene. The DNA encoding retro-GroES was synthesized through a combination of contract synthesis and our own molecular genetic manipulations to derive constructs encoding retro-GroES with a choice of restriction sites flanking the sequence for future insertion into the vector, pQE30, for expression in fusion with a 6xHis affinity tag. The final construct encoded the backbone-reversed form of the 97 residues-long parent sequence flanked by N-terminal extension of 12 residues. The extension consisted of a 10 residues-long affinity tag (MRGSHHHHHHH) and a further two residues contributed by the cloning site (GS). Mutations replacing Y27, F31 and I73 by tryptophan (W) were introduced to create three separate single tryptophan-containing mutants of retro-GroES, with the objective of using the tryptophans as probes of structure formation though examination of changes in the wavelength of maximal fluorescence and accessibility to the quencher, acrylamide. The rationale for picking the locations in the sequence of retro-GroES of residues to be mutated into tryptophan was partially that they be occupied in the sequence of the parent protein, GroES, by an aromatic residue, or by a residue with a highly hydrophobic side-chain, like isoleucine. It may be noted that in the parent sequence of native GroES, the residue P67 which corresponds to residue F31 in retro-GroES is a part of the identified hydrophobic core of GroES [12]. The other two residues, Y27 and I73, were selected essentially because of their aromatic/hydrophobic character. After synthesis, the sequences of the genes encoding retro-GroES and all three single tryptophan mutants were confirmed through automated DNA sequencing.

#### 2.2. Expression, purification and folding

Plasmids containing the genes were transformed into the *E. coli* host M15pREP4 for protein expression. The sequences of GroES and retro-GroES are shown in Table 1. The expression of retro-GroES (and all of its mutants) was of the order of 12–15 mg/liter of LB medium. For expression, cells were grown overnight and a 1% secondary inoculum was added to an appropriate volume of LB. Cells were then induced with 1 mM IPTG at an OD of 0.6 and harvested 4 h after induction. Harvested cells were either suspended in lysis buffer containing urea (8 M Urea; 0.1 M NaH2PO4; 0.01 M Tris–Cl, pH 8.0), for purification under denaturing conditions, or in lysis buffer lacking urea but containing imidazole (0.05 M NaH2PO4; 0.3 M NaCl; 0.1 M imidazole; pH 8.0) for purification under non-denaturing conditions, and lysed through sonication. The lysate was centrifuged at 18000 × g for 1 h and the supernatant thus obtained was loaded onto a Ni-NTA column. For denaturing purification, the loading was done in the presence of urea.

### Table 1

<table>
<thead>
<tr>
<th>Amino acid sequences of <em>E. coli</em> GroES and retro-GroES</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli GroES (97 residues):</td>
</tr>
<tr>
<td>E. coli GroES (97 residues):</td>
</tr>
<tr>
<td>N-NINRPHLRDV  IVKRVKEVRK  SAGCIYLLGCVLVKEDQI</td>
</tr>
<tr>
<td>GEVKLPLDNK  VGDINFGDK  YGCYKEDQI</td>
</tr>
<tr>
<td>N-NIRPHLRDV  IVKRVKEVRK  SAGCIYLLGCVLVKEDQI</td>
</tr>
<tr>
<td>GEVKLPLDNK  VGDINFGDK  YGCYKEDQI</td>
</tr>
<tr>
<td>Retro-GroES (109 residues; including His tag):</td>
</tr>
<tr>
<td>N-NIRPHLRDV  IVKRVKEVRK  SAGCIYLLGCVLVKEDQI</td>
</tr>
<tr>
<td>GEVKLPLDNK  VGDINFGDK  YGCYKEDQI</td>
</tr>
<tr>
<td>LPKVECNELI  RNCQVAYVEG  RTSKAAASG</td>
</tr>
<tr>
<td>HPRINM-C</td>
</tr>
</tbody>
</table>

The residues shown in bold letters in retro-GroES show the sites of direct mutagenesis in the tryptophan-substituted variants.

Washing (8 M Urea; 0.1 M NaH2PO4; 0.01 M Tris–Cl, pH 8.0) and elution (8 M Urea; 0.1 M NaH2PO4; 0.01 M Tris–Cl, pH 5.5; and pH 4.5) were also done in similar denaturing conditions with dialysis of the eluted protein against 20 mM Tris, or 50 mM Tris, being performed to remove urea. Alternatively, for purification under non-denaturing conditions, protein was loaded onto the Ni-NTA column in the presence of 10 mM imidazole. Washing (0.05 M NaH2PO4; 0.3 M NaCl; 20 mM imidazole, pH 8.0) and elution (0.05 M NaH2PO4; 0.3 M NaCl; 250 mM imidazole, pH 8.0) were also done in similar non-denaturing conditions, with the concentration of imidazole increased in the elution buffer to replace His residues bound to Ni-NTA. Dialysis of the eluted protein was done against 20 mM Tris, to remove imidazole. After dialysis, concentration was done through a centrifugal Centrisart concentrator (Sartorius; with a 5 kDa membrane molecular weight cut-off), up to the point at which the protein started to show some visible lack of solubility. At this point the sample was taken out and centrifuged, and the supernatant taken for an estimation of concentration of soluble protein. SDS-PAGE analysis of the purified protein showed a pure protein band of the anticipated mobility, except in some preparations where a minor amount of degraded product was also seen.

#### 2.2. Protein parameters

Protein concentrations were estimated by absorption measurement at 280 nm on a Varian Cary 50 Bio UV-visible spectrophotometer, using predicted molar extinction coefficients of 1280 M⁻¹ cm⁻¹ for retro-GroES, and 5690, 6970 and 6970 M⁻¹ cm⁻¹ respectively, for Y27W, F31W and I73W retro-GroES. Inclusive of the affinity tag, the calculated molecular weights of the four forms above were: 11784.91, 11808.05, 11823.95 and 11857.96 Da, respectively. The isoelectric point of all forms was 6.40.

#### 2.3. Mass spectrometry

Mass spectrometry was carried out on an ABI Voyager DE STR MALDI-TOF mass spectrometer, with samples ionization assisted by the matrix, α-cyano, 6-hydroxy cinnamic acid (CHCA).

#### 2.4. Fluorescence spectroscopy

Fluorescence spectroscopic studies were carried out on a Jasco J-810 spectrophotolameter attached with the FMD-4275 steady-state fluorescence emission scanning accessory. For these studies, protein of concentrations ranging between 0.1–0.5 mg/ml was used, corresponding to optical densities of ~0.2 at 280 nm. An excitation wavelength of 280 nm was used with bandpass of 2–10 nm and emission was scanned between 200 and 420 nm with a bandpass of 2–10 nm, depending on protein concentration. For examining ANS fluorescence, excitation was carried out at 350 nm, and emission was scanned between 400 and 550 nm, with a bandpass of 10 nm.

#### 2.5. FTIR spectroscopy

FTIR spectra were collected on a Perkin-Elmer Spectrum BX instrument using evaporation-deposition of the protein sample between CaF₂ windows on a standard demountable cell prior to absorption measurements.

#### 2.6. Circular dichroism spectroscopy

Far-UV CD data was collected on a Jasco J-810 spectropolarimeter. The concentration of the protein sample used varied from ~0.15–0.30 mg/ml. In all cases a cuvette of path length of 0.1 cm was used and spectra were scanned from 250 nm to ~195 nm at a rate of 100 nm/min for each protein, with averaging of 5–10 scans performed as necessary. For collecting CD spectral data at high temperatures, the peltier controlled cuvette holder attachment of the spectropolarimeter was used, with 8 mm spacers for heat transfer to 0.2 mm cuvettes. Contributions of the buffer to the spectra were electronically subtracted and for each spectrum, mean residual ellipticity (MRE) was calculated and plotted.

#### 2.7. Gel-filtration chromatography

Gel-filtration chromatography was performed on a SMART chromatography system (Pharmacia), using analytical Superdex-75 column (bed volume 2.4 ml; void volume
(0.8 ml; fractionation range 3000–70,000 Da; exclusion limit (1,00,000 Da) at a flow rate of 0.1 ml/min, through use of 0.05 ml protein samples of concentration of approximately 1.0–1.5 mg/ml. Separately, for different NaCl concentrations a Bio-Sil column from Bio-Rad was used.

2.8. Examination of GroEL binding

Retro-GroES extracted from overexpressing E. coli under non-denaturing conditions was immobilized on a 2 ml Ni-NTA column through its 6xHis affinity tag. A cytoplasmic extract (volume 10 ml) of GroE-overproducing E. coli strain DH5α/pKY206, made by lysing cells harvested from a 2 l culture, was then passed through this column under gravity flow. Following this, imidazole was used to elute protein from this column and examine the same through electrophoresis on 15% SDS-PAGE gels. The rationale of the experiment was that the column-bound retro-GroES in trimeric form (in 50 mM Tris) would bind to GroEL if it were at all capable of doing so, and retain GroEL on the column, such that subsequent elution of retro-GroES and electrophoretic analysis would reveal the presence of GroEL in gels.

3. Results

No problems were faced in the expression of retro-GroES and its mutants in E. coli. All protein constructs were amenable to affinity purification under both denaturing, and non-denaturing, conditions. Retro-GroES was not found to deposit into inclusion bodies, but found to be extremely soluble and amenable to purification under non-denaturing conditions. Given this, our rationale for anyway purifying and studying retro-GroES under both denaturing and non-denaturing conditions was to examine whether there are any differences in the results obtained through the two approaches, i.e., whether de novo folding, and refolding, elicit the same results or different results. It may be noted that in the case of a backbone-reversed protein, the sequence in which the amino acid side-chains emerge from the ribosome is
effectively reversed, i.e., the side-chains which were the last in the sequence are now the first to emerge in the nascent polypeptide. Given that nature hasn’t had a chance to optimize the rates and outcomes of sequence-folding relationships involving such proteins, we imagined that there was a distinct possibility that different outcomes (owing e.g., to kinetics dominating the de novo folding process, and thermodynamics dominating the refolding process, over experimental timescales) could follow from (i) producing the sequence backwards on a ribosome, and (ii) allowing the sequence to be completely unfolded and given a chance to refold with the entire chain present. Interestingly, identical structural-biochemical data were obtained for protein that had been refolded through dialysis to remove denaturant (urea) following denaturing affinity purification, and for the same protein purified directly from E. coli lysates through non-denaturing affinity purification in the absence of urea. All protein constructs were soluble up to concentration levels of 5.0–5.5 mg/ml in aqueous buffers of pH 7.0–8.0.

The SDS-PAGE profiles of the elution of retro-GroES and its three tryptophan-substituted mutants from Ni-NTA columns following denaturing IMAC purification, are shown in Fig. 1A (for retro-GroES), and in Figs. S1A, S2A and S3A of the supplemental data section (respectively, for the three tryptophan-substituted mutants, Y27W, F31W, and I73W retro-GroES). Similarly, MALDI-TOF mass spectrometric data establishing the recovery of polypeptides of the correct size after expression is shown in Fig. 1B (for retro-GroES), and Figs. S1B, S2B and S3B of the supplemental data section (for the three mutants, in the same order as alluded to above), establishing the absence of any proteolytic processing of the protein and its tryptophan-substituted constructs, either within E. coli cells during expression, or during purification. Although the presence of very negligible amounts of possibly degraded (truncated and somewhat shorter) retro-GroES were also seen in some preparations, in particular, with one of the mutants, studies were carried out with the best preparations, keeping in mind that the biophysical and biochemical data would be dominated by the species of the right characteristics.

For samples refolded through removal of denaturant, gel-filtration chromatograms acquired using Superdex-75 microcolumns (Amersham-Pharmacia) equilibrated with either 20 mM or 50 mM Tris, pH 8.0, are shown in Fig. 1C (for retro-GroES), and Figs. S1C, S2C and S3C of the Supplementary data section (for the three mutants), establishing the ionic strength-dependent formation of trimeric or pentameric quaternary structural assemblies by retro-GroES and its mutants, in identical manners. In Fig. 1C, it is seen that the molecule is a pentamer in the presence of 20 mM Tris, but a trimer in 50 mM Tris (column calibration data is shown in Fig. S4 of the Supplementary data section). This intriguing modulation of quaternary structural status by changes in buffer strength was also seen using another column (Fig. S5) which, furthermore, showed that parallel changes occur up to 400 mM NaCl as seen with increase in Tris concentration to 50 mM. With increase in buffer ionic strength, there is a gradual shift of elution signaling that species exist in a rapid and dynamic equilibrium with low timescales of inter-conversion relative to the timescale of chromatography.

Fluorescence emission spectroscopic data in the range of 280–350 nm or 300–420 nm is shown in Fig. 1D (for retro-GroES), and Figs. S1D, S2D and S3D of the supplemental data section (for the three mutants), establishing both the tryptophan-less status of retro-GroES
itself, and the presence of tryptophan in the mutants. Far-UV CD spectra collected at 25 °C are shown in Fig. 1E (for retro-GroES), and Figs. S1E, S2E and S3E of the supplemental data section (for the three mutants), establishing the formation of β2 type beta secondary structure, typically characterized by a low intensity negative ellipticity band in the region of 215–235 nm and a more intense negative band with a minimum in the region of 200–205 nm. The origins of such CD spectra and their similarity to the spectra of GroES are detailed in the discussions section. Meanwhile, it may be noted that in Fig. 1E there appears to be greater structural content in retro-GroES in the presence of 20 mM Tris than in the presence of 50 mM Tris, and this could owe to the stabilization of certain regions of the chain into ordered structures through further quaternary structural interactions as the chain is pentameric in 20 mM Tris and trimeric in 50 mM Tris.

Fourier transform Infra-Red (FTIR) data for retro-GroES is shown in Fig. 1F, providing further confirmation by a technique other than circular dichroism that the protein is dominated by beta-strand structure mixed with PPII structure, typically characterized by peaks in the FTIR spectrum in the region of 1620–1635 cm⁻¹.

The conformational behavior of retro-GroES upon raising of temperature is extremely interesting. As Fig. 2A shows, raising of temperature to 40 °C, 60 °C and 95 °C leads progressively to increase in the intensity of the negative 215–235 nm band with an accompanying decrease in the intensity of the negative 200–205 nm band. These changes are, however, completely reversible as shown by the CD spectrum taken after cooling solutions back to 25 °C; the spectrum of the cooled protein is identical in shape and intensity to the spectrum of the protein at 25 °C prior to heating (seen in Fig. 1E; the spectrum is not shown again in Fig. 2A due to the high visual density of the data already present in the figure). Notably, as Fig. 2A shows, there is also a shift in the negative band position from 200 nm to higher wavelengths as temperature is increased. Another feature worth noting is that there is an isodichroic point at 210 nm which is a clear indication of a structural alteration involving a two-state transition. This behavior has been seen previously by other workers with both small peptides and certain small proteins (detailed in the Discussion and conclusions section); it is understood to owe to a heating-induced conversion of PPII structure to β-sheet structure.

Essentially similar structure-consolidation behavior with heating was observed with each of the three tryptophan-substituted retro-GroES mutants. The far-UV CD spectra of Y27W, F31W, and I73W retro-GroES, respectively, at 25 °C (prior to heating), at 95 °C and following cooling back to 25 °C, are shown in Fig. S6 (Supplementary data). In each case, there are comparable changes in the CD spectra with heating and cooling.

Importantly, in contrast to what is observed upon heating, upon exposure of the protein to increasing concentrations of the denaturant, guanidinium hydrochloride (Gdn–HCl), the 215–235 nm band in the far-UV CD spectrum of retro-GroES is seen to reduce progressively in intensity (Fig. 2B), as would be expected for a protein undergoing unfolding. A similar reduction in intensity of this band can also be seen, in Fig. 2C, upon exposure to increasing concentrations of the denaturant, urea, although the changes are less pronounced in urea for equivalent concentrations of the two denaturants, with unfolding also progressing to a lower degree. Plotting of changes in the mean residue ellipticity values at 220, 222 and 225 nm, during unfolding by Gdn–HCl and urea are shown in Figs. S7A and S7B, respectively, of the supplemental section.

The changes in hydrodynamic volume undergone by retro-GroES upon exposure to increasing concentrations of Gdn–HCl were also monitored in terms of changes in the molecule’s elution volume during gel filtration. This data can be seen in Fig. 2D, in which there is both an initial increase in elution volume (suggestive of a reduction in hydrodynamic volume) from 0.0 M to 0.25 M Gdn–HCl, followed subsequently by a progressive series of reductions in elution volume (suggestive of increase in hydrodynamic volume) from 0.25 M to 6.0 M Gdn–HCl. This behavior is compatible with an initial dissociation of subunits from a homo-pentameric form (in 20 mM Tris) to a form that is either an expanded monomer, or an equilibrium population of monomers and dimers, followed by the unfolding of the monomer. Similar gel-filtration behavioral changes detected during unfolding of the mutant, F31W retro-GroES, are shown in Fig. S7C of the data in the supplemental section; in this case, in order to emphasize that the behavior seen is independent of whether the initial population is pentameric, or trimeric, we show the data for an initial dissociation from a trimeric form (in 50 mM Tris) to a form that is either an

Fig. 3. Panel A: Gel-filtration chromatograms of refolded (open square) and de novo folded (open circle) retro-GroES in 20 mM Tris buffer (pH 8.0), showing that both adopt quaternary structures of the size of about five subunits. Panel B: Far-UV CD spectra of refolded (open square) and de novo folded (open circle) retro-GroES in 20 mM Tris buffer (pH 8.0), showing that both adopt similar secondary structure. Panel C: Increase in the negative MRE values of de novo folded retro-GroES at 222 nm as a function of increasing temperature (open square), and as a function of decreasing temperature (open circle), showing that the increase in the 222 nm signal with heating is completely reversible. For comparison with refolded retro-GroES, see Fig. 4A.
expanded monomer or an equilibrium of monomers and dimers, which is followed by the unfolding of the monomer.

The above data was for protein refolded from 8M urea, following denaturing Ni-NTA affinity purification. Below, we present data for protein directly purified from *E. coli* without exposure to denaturants. We purified retro-GroES from *E. coli* lysates under non-denaturing conditions to see whether the protein is in a soluble and folded form within the *E. coli* cell after biosynthesis, and also for certain other reasons already alluded to in the first paragraph of this section. For want of a more appropriate name, we refer to this form of retro-GroES as *de novo* folded retro-GroES, to distinguish it from the refolded retro-GroES for which data was presented in all earlier figures.

Fig. 3A plots the gel-filtration elution chromatogram of the *de novo* folded, and refolded, forms of retro-GroES in 20 mM Tris demonstrating that the elutions overlap exactly for pentameric forms adopted by proteins purified under both conditions. Similarly, Fig. 3B plots the far-UV CD spectra of *de novo* folded, and refolded, retro-GroES, showing that there is no significant difference in their secondary structural contents. The structural changes seen to accompany heating are less marked with *de novo* folded retro-GroES (Fig. 3C) than with refolded retro-GroES (Fig. 4A), in plots of changes in mean residue ellipticity (MRE) signals of the two forms measured at 222 nm as a function of temperature. With the *de novo* folded protein (Fig. 3C), the intensification of the negative ellipticity signal at 222 nm upon heating is of smaller magnitude, and can be seen to saturate at about 60 °C with no further increase observed upon heating up to 95 °C. In contrast, with the refolded protein (Fig. 4A) the changes continue up until a temperature of almost 90 °C before saturation is observed. Notably,

**Fig. 4.** Panels A–D: Increase in the negative MRE values of refolded retro-GroES at 222 nm with increasing (open square) and decreasing (open circle) temperature, in the absence of Gdn–HCl (Panel A) and in the presence of 1.0 M Gdn–HCl (Panel B), 2.0 M Gdn–HCl (Panel C) and 4.0 M Gdn–HCl (Panel D), showing that the PPII to beta-strand transition marked by gain of signal at 222 nm with increasing temperature occurs even in the presence of Gdn–HCl (up to 4.0 M), despite the fact that Gdn–HCl on its own causes loss of signal at 222 nm (Fig. 2B). Panel E: Far-UV CD spectra of retro-GroES in the presence of 4.0 M Gdn–HCl at 25 °C (open triangle), at 98 °C (open square), and upon cooling back to 25 °C (open circle), showing that the protein reversibly adopts a significantly greater level of beta-strand structure in 4.0 M Gdn–HCl at 98 °C (with a near-doubling of the signal strength at 222 nm). Panel F: Fluorescence emission of ANS from a control 10 μM solution (open circles) and from an equivalent ANS solution containing 0.6 mg/ml retro-GroES (open squares), showing that there is no increase in ANS fluorescence indicative of any significant surface hydrophobicity in retro-GroES.
the changes seen are entirely reversible with both the de novo folded and refolded forms. Together with the remaining data comparing the two forms, this establishes that the form adopted by the protein through de novo folding is identical in all respects to that obtained through refolding from urea (with the fact of urea-mediated unfolding already having been established).

As already demonstrated, retro-GroES shows a decrease in the intensity of the CD signal at 222 nm with increase in the concentration of Gdn–HCl, but an increase in signal with increase in temperature. We decided to see what happens when the two unfolding influences are combined, i.e., whether such a combining results in dominance of the chemical unfolding effect or the structural consolidation effect. Changes in CD signal strength at 222 nm with heating and cooling were monitored at four different Gdn–HCl concentrations. The results for Gdn–HCl concentrations of 0.0 M, 1.0 M, 2.0 M and 4.0 M, respectively, are shown in Fig. 4A–D from which it may be seen that the presence of the denaturant merely results in an off-setting of the starting and ending MRE values, without any profound effect on the heat-induced structural consolidation. Astonishingly, the reversible increase in intensity is observed even in the presence of 4.0 M Gdn–HCl.

To illustrate how this change is reflected in the far-UV CD spectrum of refolded retro-GroES, Fig. 4E shows CD spectra of the protein in the presence of 4.0 M Gdn–HCl at 25 °C (prior to heating), at 95 °C and then at 25 °C again, following cooling of the solution. The two spectra taken at 25 °C overlap to a remarkable extent and are markedly different from the spectrum at 95 °C, with clear evidence of increased beta-strand content at 95 °C even in the presence of 4M Gdn–HCl. The intensification of the 215–235 nm band observed in the absence of Gdn–HCl is seen to an even greater extent in the presence of 4.0 M Gdn–HCl.

Fig. 4F shows that retro-GroES is not bound by the fluorescent dye, ANS (8-anilino-1, napththalene sulfonic acid) since there is no intensification of the fluorescence of the dye in the presence of the protein.

To further explore retro-GroES, the residue tryptophan was introduced as a spectroscopic probe of structure into different locations in the protein, since tryptophan’s wavelength of maximal fluorescence emission, and its quenchability by quenchers such as acrylamide, both depend acutely on the extent of the burial of the residue. Substitution mutations incorporating tryptophan were made at three sites in the protein, as already mentioned (see Materials and methods for details of rationale), to create Y27W, F31W, and I73W retro-GroES. Locations for the introduction of tryptophan were based essentially on the locations of the single phenylalanine (F31), the single tyrosine (Y27W), and one of the several isoleucines in the protein, all hydrophobic analogs of tryptophan. The hope was that one or more of these would yield a tryptophan residue buried within the (yet unknown) structure of retro-GroES. Data demonstrating that there is no marked effect on the structure and structural-biochemical characteristics of retro-GroES through Trp-substitutions, has already been presented in Figs. S1–S3, S6 and S7 of the Supplementary data section. Now we summarize the characterization of the burial, or exposure, of the Trp in the three mutants. Fig. S8A–S8C of the Supplementary data section show that all mutants emit maximally between 348–352 nm at 25 °C, suggesting that there is only marginal burial of the residue in any of the mutants (it may be noted a tryptophan residue fully exposed to the aqueous solvent is expected to emit maximally at 352–353 nm). This conclusion regarding marginal burial is supported by the fact that all mutants require \( F_0 \) values in excess of 4.0 to achieve full quenching by the fluorescence quencher, acrylamide. Notably, the quenching profiles of the three mutants are neither identical in shape, nor superimposable (Fig. S8D of the Supplemental section) demonstrating that they differ in the accessibility of the tryptophans at the three different locations. Of the three mutants, Y27W retro-GroES has the most exposed tryptophan residue. F31W and I73W retro-GroES have tryptophans with comparable accessibilities that are less exposed, and more buried. However, the fact that all of them reach \( F_0 \) values higher than 4.0 with 200 mM acrylamide suggests that they are all marginally buried.

Our conclusion regarding Y27W having the most exposed tryptophan residue is supported by the fact that it shows the longest wavelength of maximal emission at 352 nm (Fig. S8A in the Supplementary data section), unlike I73W and F31W which show emission maxima at 351, and 348 nm, respectively (Fig. S8B and S8C in the Supplementary data section). Upon heating, the emission of F31W (already at 348 nm) shows no changes. However, the emission of I73W (at 351 nm) changes to 348 nm. In contrast, the emission of Y27W shifts from 352 nm to 348–349 nm upon heating; a blue-shifting of emission compatible with the structural consolidation observed also through CD spectroscopy.

Retro-GroES extracted from overexpressing E. coli under non-denaturing conditions was found to be incapable of binding to GroEL, using immobilization experiments in which retro-GroES bound to a Ni-NTA column was allowed to interact with the cytoplasmic extract of a GroE-overproducing strain, following which SDS-PAGE gel electrophoresis failed to reveal any elution of GroEL bound to retro-GroES eluted from the column under non-denaturing conditions using imidazole as a histidine substitute.

4. Discussion and conclusions

About 31% of the residues in the co-chaperone, GroES, of the GroEL-GroES system is made up of beta strand-based structure, with the remaining sections of the molecule adopting beta turn structure or lacking regular secondary structure altogether [13]. About 16 residues out of the 97 residues in the main chain of GroES are known to adopt a structure when GroES binds to GroEL [14]. The far-UV CD spectrum of the molecule is typical of the \( \beta_{hs} \) type of far-UV CD spectrum [14,15]. Proteins with such spectra are now understood to consist of a combination of beta-strand structure and polyproline type II (PP2) structure [16]. Such proteins also display some non-cooperativity in regard to their folding, and unfolding, in a manner dependent on the content of PP2 structure [17]. Consequently, the spectra are characterized by two bands with negative mean residue ellipticity. One of these bands, which owes to the PPII structure, has a clear minimum in the region of 200–210 nm; the minimum is usually closer to 200 nm than to 210 nm, because of which proteins with such spectra were earlier mostly assumed to consist of a large random coil content. The other band has a shoulder in the region of 220–230 nm. This band owes to the beta-strand structure (ordinarily manifested as a negative band with a minimum between 216–218 nm). The exact positions of the minima, of course, depend on the relative contents of the two types of structure, as well as on whether there is any random coil structure present. PPII structure is commonly seen in alanine-rich and other peptides [18,19], as well as in numerous proteins that exist in an energetically and structurally preorganized state prior to interactions with a designated cellular partner [20]. In recent years, it has been reported from studies of certain small peptides and proteins that PPII structures transform into beta strand-based structures with increase in temperature [18,21]. Thus, in a protein with mixed sheet and PPII structural contents, the effects of heating are likely to be somewhat unpredictable, depending on which of the two dominate. If the beta structure starts to unravel, as a result of the weakening of the hydrogen bonding interactions stabilizing that structure, and without any balancing effect of the conversion of PPII structures to beta strands, a protein may be expected to unfold upon heating. If, however, the beta-strand interactions are strong enough to survive the initial increase in temperature without concomitant unfolding, then the conversion of PPII structure to beta-strand structure with heating would, of course, increase the net numbers of interactions stabilizing the protein and manifest as heating-induced consolidation of structure.

Very intriguingly, in this study we have found that retro-GroES also adopts a structure with a \( \beta_{hs} \) type of far-UV CD spectrum which is similar to that of GroES in both qualitative and quantitative terms. The shapes of the spectra are similar. The positions of the two negative
bands and their mean residue ellipticities too are similar [14,15]. Like GroES, retro-GroES has a negative band at 203 nm with a signal strength of ~10,000 deg cm^2 dmol^{-1}, and a second negative band with a shoulder at 222–225 nm. Both GroES and retro-GroES associate into multimers. One further interesting parallel exists between GroES and retro-GroES. It is known that GroES unfolds in a non-cooperative manner upon exposure to increasing concentrations of Gdn–HCl [15]. In this paper, we show that a similar non-cooperative unfolding is seen for retro-GroES, in regard to unfolding by both Gdn–HCl and urea. It is worth noting that non-cooperative unfolding is a known phenomenon for beta strand-based proteins [17,22–25]. This is primarily thought to be due to the existence of stabilizing hydrophobic interactions either between sheets, or amongst residues on the same face of a sheet, which fail to be weakened or destroyed in concert with hydrogen bonds stabilizing the sheet, i.e., these two types of interactions do not always cooperate as they are destroyed, and show some autonomy from each other. When either class of interactions is independently weakened (especially when hydrogen bonding is weakened by a denaturant like Gdn–HCl), the rate at which unfolding equilibrium is effectively reached is considerably slowed down by the survival of the other class of interactions (hydrophobic interactions) leading to apparently non-cooperative unfolding. Non-cooperativity of structural changes is seen also in cold denaturation of beta-rich proteins for the same reasons.

As regards differences between GroES and retro-GroES, GroES is well known to form heptamers both in the GroEL-bound and GroEL-unbound states. Retro-GroES, in contrast, adopts structures that appear to be either trimeric or pentameric based on gel-filtration studies. Further, GroES is known to undergo unfolding by denaturants, as well as by heat. In contrast, although retro-GroES undergoes unfolding by denaturants, it shows structural consolidation upon heating. The changes seen in the spectra are entirely similar to those seen in other proteins and peptides in which there is a conversion of PPII structure to beta-strand structure with heating, with an isodichroic point at around 210 nm [21]. In retro-GroES too (Fig. 2A), there is a clear isodichroic point at 210 nm around which the intensities of the two negative bands rise and fall in inverse proportionality with each other. Importantly also, the heat-induced structural changes are entirely reversible, supporting our conclusion that they involve a two-state transition characterized by a reversible population inversion around an isodichroic point. Finally, of course, the two proteins differ in respect of interactions with GroEL. As already discussed in the introduction section, this is expected; notably, even if some interaction were seen, it could involve the sections of GroEL used in the binding of sub-strate proteins rather than those involved in the binding of the co-chaperonin, GroES, and so further exploration of this issue must await high-resolution NMR-based structural biochemistry or crystallographic studies.

The most important message from this study is the following: With the exception of the heat-induced changes seen in structure, and the lack of unfolding of retro-GroES by heat, structure-forming behavior in the backbone-reversed retro-GroES is very similar to that shown by the parent polypeptide, GroES. Given the profoundness of the chemical transformations involving backbone reversal, it is an astonishing fact that retro-GroES folds and assembles into ordered forms characterized by structure at the secondary, tertiary and quaternary structural levels which display not just unfolding and folding behavior but full reversibility of denaturant- and heat-induced structural alterations. In this sense, the behavior of retro-GroES which folds into a structure with a ββ type of CD spectrum like GroES is similar to that of retro-GCN4-p1, which has been reported to fold into a helical structure like GCN4-p1. The high solubility of retro-GroES and the lack of aggregation at high concentrations, and upon heating, suggest that the protein lacks significant exposed hydrophobic surfaces. We admit that these studies have not provided any fundamental insights into why certain retro-proteins display features akin to those shown by their parent, native proteins, while other retro-proteins do not. We hope, however, that these results will someday become incorporated into a larger picture concerning structural-biochemical behavior upon backbone reversal of beta-strand protein structures.

Acknowledgements

SA and AS thank the Council of Scientific and Industrial Research (CSIR), New Delhi, for the research fellowships. PG thanks the CSIR, the Indian National Science Academy (INSA), New Delhi, and the Department of Biotechnology (DBT), Govt. of India, for the research grants to study protein folding and aggregation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbapap.2008.02.009.

References