Biochimie 92 (2010) 491–498

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Equilibrium and kinetics of the unfolding and refolding of *Escherichia coli* Malate Synthase G monitored by circular dichroism and fluorescence spectroscopy☆

Aditi Maheshwari [a,1], Vikash K. Verma [a,1], Tapan K. Chaudhuri [a,b,*]

[a] Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India
[b] School of Biological Sciences, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

ARTICLE INFO

Article history:
Received 5 October 2009
Accepted 11 January 2010
Available online 4 February 2010

Keywords:
Unfolding and refolding of large multi-domain protein Malate Synthase G
Kinetics of refolding and unfolding process of MSG
Mechanism of protein folding and unfolding

ABSTRACT

The equilibrium and kinetics studies of an 82 kDa large monomeric *Escherichia coli* protein Malate Synthase G (MSG) was investigated by far and near-UV CD, intrinsic tryptophan fluorescence and extrinsic fluorescence spectroscopy. We find that despite of its large size, folding is reversible, *in vitro*. Equilibrium unfolding process of MSG exhibited three-state transition thus, indicating the presence of at least a stable equilibrium intermediate. Thermodynamic parameters suggest this intermediate resembles the unfolded state. However, the equilibrium intermediate exhibits pronounced secondary structure as measured by far-UV CD, partial tertiary structure as delineated by near-UV CD, compactness (m value) and exposed hydrophobic surface area as assessed by ANS binding, typically depicting a molten globule state. The stopped-flow kinetic data provide clear evidence for the presence of a burst phase during the refolding pathway due to the formation of an early Intermediate, within the dead time of the instrument. Refolding from 4 M to various lower concentrations until 0.4 M of GdnHCl follow biphasic kinetics at lower concentrations of GdnHCl (<0.8 M), whereas monophasic kinetics at concentrations above 1.5 M. Also, rollover in the refolding and unfolding limbs of chevron plot verifies the presence of a fast kinetic intermediate at lower concentration of GdnHCl. Based upon the above observations we hereby propose the folding pathway of a large multi-domain protein Malate Synthase G.

1. Introduction

Understanding the folding mechanism of proteins is a real challenge and has been an unresolved issue since long [1]. Folding pathway of many small proteins have been explored that usually follow two-state mechanism: N ↔ U [2–8]. However, large proteins (>70 kDa) that constitute more than 70% of total proteins in the cell have not been subjected to such studies due to their complex structure and difficulty in characterization. Large proteins composed of multiple domains often refold inefficiently because of the formation of partially folded intermediates (unfolded conformers) that tend to self-associate (aggregate) into disordered complexes, driven by hydrophobic forces and inter-chain hydrogen bonding [9]. Individual domains of multi-domain proteins doesn't always fold independently. There are several factors that effect the folding of multi-domain proteins such as the nature of domain interface and influence of neighboring domains. There are instances, where a domain exerts control over the folding of neighboring domain by virtue of interfacial contacts and incase the exposed interface of an isolated domain has a large hydrophobic surface, this can lead to aggregation [10]. This aggregation process irreversibly removes protein from their productive folding pathways. It is desirable to explore protein-folding pathway of such large proteins for which characterization of intermediates is an important step.

Malate Synthase G is a large 82 kDa monomeric protein consisting of four domains. It's one of the two enzymes working in glyoxylate pathway that bypasses the decarboxylation steps of Citric acid cycle by converting acetate to succinate. Malate Synthase G catalyzes the condensation of glyoxylate and acetyl-coenzyme A and hydrolyse the intermediate to yield malate and coenzyme A to replenish the pool of citric-acid-cycle intermediates [11].

The structure of this protein is known and it has been elucidated not just by X-Ray crystallography [12] but also by NMR spectroscopy [13]. For the first time, we present *in vitro* equilibrium and kinetic studies of Malate Synthase G (MSG) with the purpose to define the

☆ This work was supported by financial assistance from MHRD division, Govt. of India, Department of Science and Technology Govt. of India and IRD, IIT Delhi.

* Corresponding author at: School of Biological Sciences, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India. Tel.: +91 11 2659 1012; fax: +91 11 2658 2282.
E-mail address: tapan@dbeb.iitd.ac.in (T.K. Chaudhuri).

[1] These authors contributed equally.

0300-9084/$ – see front matter © 2010 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.biochi.2010.01.009

conformational changes occurring during the folding–unfolding pathway.

2. Materials and methods

2.1. Expression and purification of Malate Synthase G (MSG)

Recombinant, monomeric MSG was isolated and purified from BL21-(DE3) strain of *Escherichia coli*, carrying the pCS19 plasmid that encodes MSG preceded by six histidine codons, under LacZ promoter. Purification was done through a single step process using Ni-NTA chelating column, HiTrap HP (Pharmacia, Piscataway, NJ, USA) in AKTA FPLC. The final protein was obtained after dialyzing it against buffer (sodium cacodylate buffer, 50 mM NaCl, pH 7.4) and further concentrating it using amicon tubes (from Millipore, Billerica, MA, U.S.A.) with 50 kDa molecular weight cut off membrane.

2.2. Enzymatic activity of MSG

Enzymatic activity of MSG was determined by the method previously designed by Ornston and Ornston [18], with necessary modifications, in Bekman coulter DU800 spectrophotometer (USA). Each assay mixture contained 10 μl of 1 M $MgCl_2$, 16 μl of 0.01 M Acetyl CoA, 10 μl of 0.1 M glyoxylate, and 100 μl of 1 M tris buffer (pH 8.0) in a total volume of 900 μl. The reaction was initiated by adding protein sample to the above mixture such as the final concentration of protein was 1 μM. The reaction was stopped after 10 min by adding 2 ml of 6 M GdnHCl and 100 μl of 10 mM DTNB for color development. Absorbance was measured at 412 nm after another 10 min of incubation at RT. For the above activity assay 3 ml cuvette with a light path of 1 cm was used. Each sample was corrected for contributions from buffer solutions to get rid of basal activity, in absence of protein.

2.3. Equilibrium unfolding studies by intrinsic tryptophan fluorescence spectroscopy

For equilibrium unfolding studies, fluorescence spectra were recorded on Perkinelmer life sciences luminescence spectrometer LS 55 (Perkein-Elmer, USA), using an optical cuvette of path length 1 cm. The protein samples (final concentration 1 μM) were prepared in 10 mM sodium cacodylate, 50 mM NaCl at pH 7.4 and 25 °C and incubated for 2 h to equilibrate at appropriate final concentration of GdnHCl/Urea. Each spectrum was corrected for contributions from buffer solutions, containing increasing concentration of GdnHCl (0–5 M)/Urea (0–6 M). The samples were excited at 295 nm and emission spectra were recorded between 310 and 400 nm with excitation and emission slit width 5 nm each.

2.4. Equilibrium unfolding–refolding studies by circular dichroism spectroscopy

CD spectra were recorded on JASCO J-810 CD polarimeter (JASCO, Japan), flashed with nitrogen gas, using an optical cuvette of path length 2 mm for far-UV CD and 1 cm for near-UV CD, respectively. The protein samples (final concentration 1 μM and 12 μM, respectively for far and near-UV CD) were prepared in 10 mM sodium cacodylate, 50 mM NaCl at pH 7.4 and 25 °C and incubated for 2 h to equilibrate at appropriate final concentration of GdnHCl. Refolding was achieved by series of dilutions of the denatured protein with refolding buffer, 10 mM sodium cacodylate, 50 mM NaCl at pH 7.4 and 25 °C for 2 h to equilibrate. Each spectrum was corrected for contributions from buffer solutions, containing increasing concentration of GdnHCl (0–5 M). The emission spectra were recorded between 250 and 205 nm for far-UV CD and 310–250 nm for near-UV CD spectroscopy (Fig. 2).

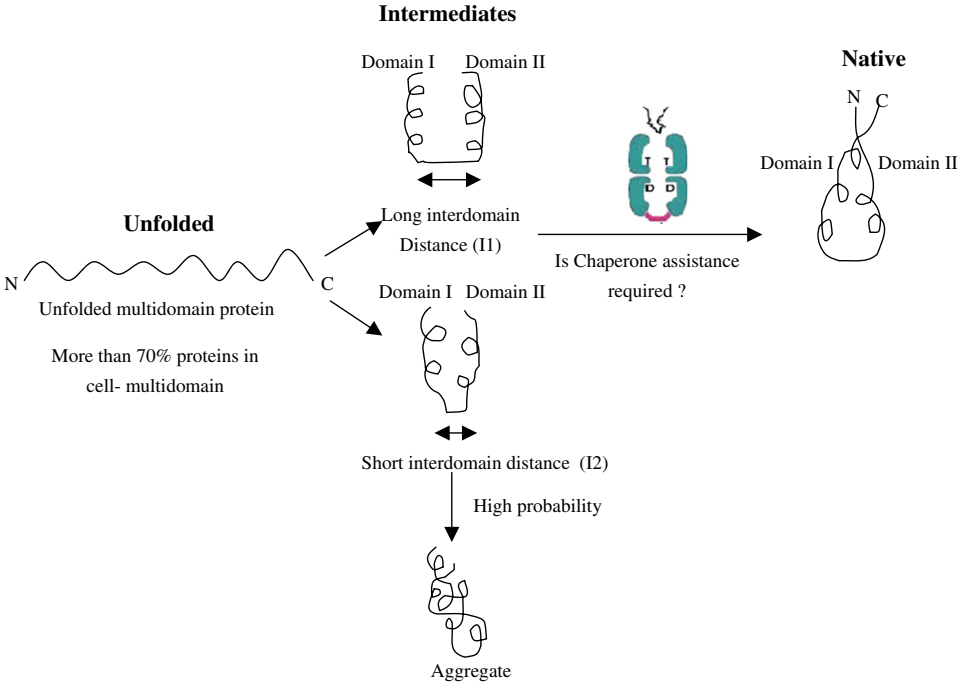

Fig. 1. Proposed schematic presentation for multi-domain large protein folding. Unfolded protein undergoes refolding/folding process through the formation of intermediate species (I1, I2, etc.) and leads to the folded state. If the inter-domain distance is high (I1), there will be weak inter-domain interaction and the domains can fold independently [14]. However, if the inter-domain distance is smaller (I2) there are possibilities for co-operative folding of the domains [15,16]. There is also high probability for aggregation during the refolding processes of multi-domain proteins, when there exist strong inter-domain interaction. The final step in the folding pathway, of multi-domain proteins, involve the slow rearrangement among the partially or completely folded domains or there are possibilities for chaperone assistance if the intermediates are susceptible to aggregation process [17].

2.5. Equilibrium unfolding studies by extrinsic fluorescence spectroscopy using ANS as external fluorophore

ANS spectra were recorded on Perkinelmer life sciences, luminescence spectrometer LS 55, using an optical cuvette of path length 1 cm. The protein samples (final concentration 1 μM) were prepared in 10 mM sodium cacodylate, 50 mM NaCl at pH 7.4, 25 °C and incubated for 2 h to equilibrate at appropriate final concentration of GdnHCl and 20 μM ANS. Each spectrum was corrected for contributions from buffer solutions, containing increasing concentration of GdnHCl (0–5 M) and 20 μM ANS. The samples were excited at 370 nm and emission spectra were recorded between 410 and 600 nm with excitation and emission slit width 5 nm each.

2.6. Thermodynamic calculations

Global fitting was performed, in which the transition curves measured at different wavelengths – 467 nm, 222 nm and 340 nm were fitted simultaneously, with the assumption that native (N), intermediate (I) and unfolded (U) state are the only three states sufficient for interpreting the unfolding transition of the protein. The best fit values of the equilibrium unfolding (thermodynamic) parameters thus obtained; by the method of non-linear least squares (three-state model) [3,19] are summarized in Table 1.

According to the three-state model, equilibrium unfolding intermediate is stably populated during the unfolding transition from N to U. The unfolding reaction is thus represented by the following scheme:

$$N \overset{K_{ni}}{\rightleftharpoons} I \overset{K_{iu}}{\rightleftharpoons} U$$

where K_{ni} and K_{iu} are the equilibrium constant of the unfolding transition from N to I and I to U, respectively and their dependence on the concentration is given in [20].

$$K_{ni} = \exp[-(\Delta G_{ni}H_2O - m_{ni}[C])/RT]$$

$$K_{iu} = \exp[-(\Delta G_{iu}H_2O - m_{iu}[C])/RT]$$

ΔG_{ni} H_2O and ΔG_{iu} H_2O are the free energy differences between N → I and I → U state, respectively. m_{ni} and m_{iu} are the cooperativity index of the two above mentioned transitions. [C] corresponds to the GdnHCl concentration.

2.7. Kinetic studies by stopped-flow fluorescence spectroscopy

The unfolding and refolding kinetics of MSG were monitored by following the spectral changes in intrinsic tryptophan fluorescence at 340 nm, in a stopped-flow apparatus-Applied photophysics RX 2000 (Surrey, UK) spectrophotometer, using an optical cuvette of path length 1 mm. All kinetic measurements were made in the presence of 10 mM sodium cacodylate, 50 mM NaCl at pH 7.4 and 25 °C. Drive syringes of 250 μl and 2.5 μl were used to inject protein/denatured-protein and denaturant/buffer respectively, resulting in an 11-fold dilution of the protein solution. Dead time of the instrument is 20 ms. The samples were excited at 295 nm and sampling was done at an interval of 0.5 s upto 100 s, except first data point that was collected after 20 ms.

The unfolding kinetic data was analyzed by non-linear least squares fitting to the equation:

$$A(t) = \sum A_i \exp(-k_i t) + A_{infinity}$$

where A(t) is the total signal, $A_{infinity}$ is the signal at infinite time, A_i is the signal attained during each (fast or slow) phase, k_i is the rate constant for that phase and t stands for time. The refolding kinetic data was analyzed by non-linear least squares fitting to the equation:

$$A(t) = \sum A_i(1 - \exp(-k_i t)) + A_0$$

where A_o is the signal at 0 s. Rest of the notations is the same as for unfolding kinetics.

All the chemicals used were of analytical grade.

3. Results

3.1. Malate Synthase G exhibits three-state unfolding at equilibrium

The results obtained for GdnHCl-induced conformational transitions of MSG, studied by intrinsic tryptophan fluorescence, far-UV CD, near-UV CD and ANS fluorescence are shown in Fig. 2 (A–D), respectively. All 4 probes show evidence of two unfolding transitions with an intermediate getting populated at lower concentration of GdnHCl (~0.4 M). The equilibrium refolding of 4 M GdnHCl denatured MSG was monitored by changes in far-UV CD signal (as shown in Fig. 2B). The reversible nature of the reaction suggests for accumulation of at least one stable intermediate at the concentration of ~0.4 M GdnHCl and not a salt binding artifact.

Significantly lower value of ΔG_{iu}, free energy change from I to U state, than ΔG_{ni}, free energy change from N to I state, shows that the intermediate is energetically closer to the fully unfolded state rather than the native state. 'm' value is an important reaction coordinate that provides a measure of the solvent accessibility and consequently average compactness of intermediates. The above idea is further supported by higher value of m_{ni} than m_{iu} showing that more surface area is available to the solvent from N to I state, consequently intermediate is closer to unfolded state. It is important to note that if the measured m-value is significantly below the value expected for the size of the analyzed protein, this is a good hint that the two-state assumption is no longer correct, and that partially folded states are populated at equilibrium in addition to U and N. Urea-induced equilibrium unfolding studies were performed in presence of different NaCl concentrations. An increase in fluorescence intensity was observed at 0.4–0.5 M concentration of urea (data not shown). This confirms that the formation of stable intermediate state during GdnHCl-induced equilibrium unfolding was not due to intrinsic behavior of the salt denaturant (GdnHCl), since the similar intermediate was also observed during the urea-induced equilibrium unfolding in the presence of different NaCl concentrations (0 M, 0.05 M, 0.5 M). Further the data remain unchanged over a range of MSG concentrations (1–10 μM), which makes the population of oligomeric states of the protein unlikely in the course of our experiments. In order to verify the midpoint of I–U transition, equilibrium data in the range 0.4–3 M, from Fig. 2B and D, were fitted and the C_m (midpoint concentration value of

Table 1
Thermodynamic parameters of MSG.

Protein	ΔG_{nu} H_2O kcal/mol	m_{nu} kcal/mol M	Cm_{nu} M	ΔG_{ni} H_2O kcal/mol	m_{ni} kcal/mol M	Cm_{ni} M	ΔG_{iu} H_2O kcal/mol	m_{iu} kcal/mol M
Malate Synthase G	7.5 ± 0.2	15.9 ± 5.0	0.5 ± 0.1	5.9 ± 0.1	14.5 ± 5.0	0.4 ± 0.1	1.6 ± 0.2	1.5 ± 0.1

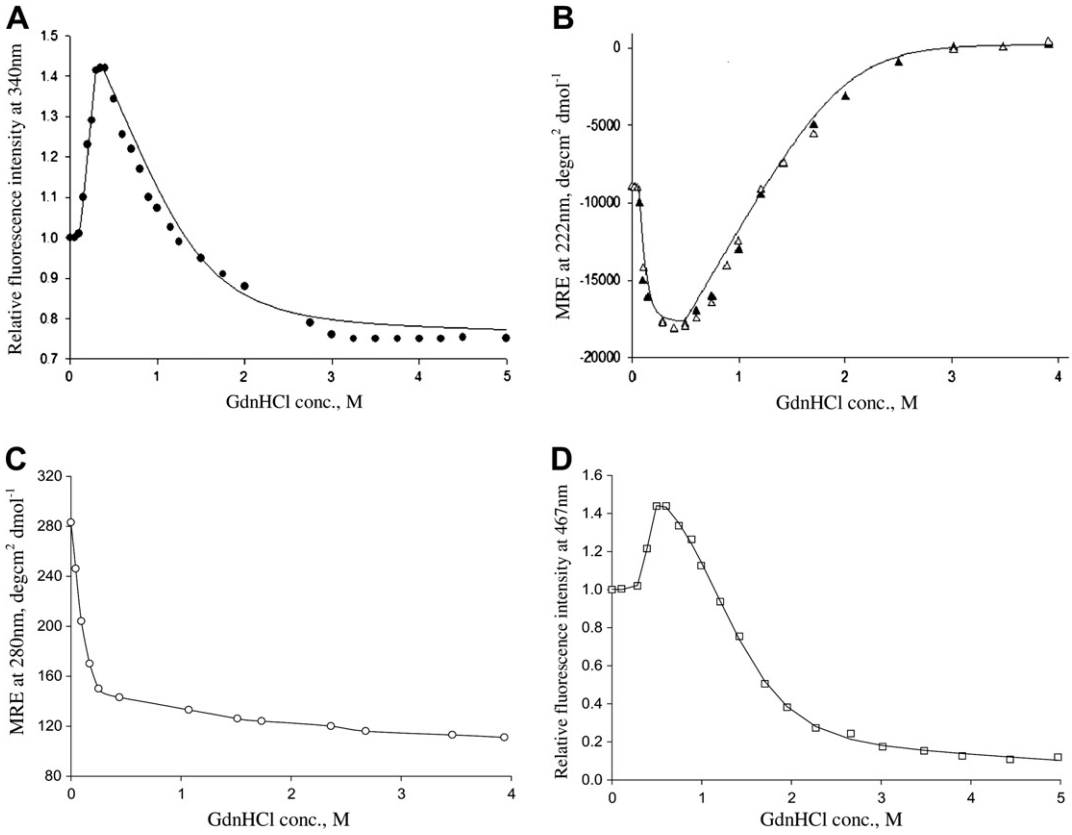

Fig. 2. Effect of GdnHCl on equilibrium unfolding of Malate Synthase G. Continuous line depicts the theoretical curve obtained on three-state fitting. A) (●) Change in tryptophan fluorescence at 340 nm for 1 μM of native protein (pH 7.4, 25 °C), denatured by increasing concentration of GdnHCl (0–5 M). B) Change in Ellipticity at 222 nm for 1 μM of native protein denatured by increasing conc. of GdnHCl (0–5 M). (▲) and (△) signify equilibrium refolding and unfolding at various GdnHCl concentrations, respectively. C) (○) Change in ellipticity at 280 nm ($[\theta]_{280}$) for 12 μM of native protein, denatured by increasing concentration of GdnHCl concentration (0–5 M) and D) (□) Change in tryptophan fluorescence at 467 nm for 1 μm of native protein and 20 μM of ANS, denatured by increasing concentration of GdnHCl (0–5 M). Global fitting was performed, in which the transition curves measured at different wavelengths 467 nm, 222 nm and 340 nm were fitted simultaneously. The best fit values of the equilibrium unfolding (thermodynamic) parameters thus obtained are summarized in Table 1.

GdnHCl for the transition from I to U) values were obtained at around 1.5 M (data not shown).

3.2. The intermediate state of MSG is a molten globule state

The spectroscopic features of the intermediate state (as compared to native and unfolded state) populated at 0.4 M GdnHCl are illustrated in Fig. 3(A–C). It is quite evident from Fig. 3A that the intrinsic fluorescence intensity of the I state increases with respect to the native and denatured state and the fluorescence maxima got shifted from 344 nm (N) – 347 nm (I) – 350 nm (U). Over the transition from N → I, intrinsic tryptophan fluorescence intensity increases until ~0.4 M concentration of GdnHCl, as shown in Fig. 2A. It is likely that at this concentration, intermediate (I) is in such a conformation where tryptophans are now exposed to the outside, but not to the solvent completely. Drop in intensity over the transition from I → U, indicates a completely unfolded structure, where tryptophan are quenched externally by the solvent molecules.

Increase in the ellipticity (by 2-fold) was observed at 0.4 M GdnHCl (as shown in Fig. 3B). The far-UV CD spectra of intermediate (I) indicate that the secondary structure is still intact (infact more pronounced). Since the CD ellipticity in far-UV region is mainly due to the helical structure, it appears that the non-native helical structure is more populated in the intermediate state. This might be consequential of β → α transition in the protein. Disordered proteins, including denatured species, have a strong negative signal at 200 nm but a weak negative or positive signal at 222 nm [21]. A comparison of $[\theta]_{280}$ values between the native, intermediate and unfolded states (Fig. 3C) show a significant change (almost half) in the MRE value from native to intermediate but a slight change in MRE value from intermediate to unfolded state. This means that intermediate state has lost a part of its native tertiary interactions involving aromatic amino acid residues and tight packing of side chains. The comparison suggests that the stable folding intermediate that populates around 0.4 M GdnHCl concentrations carries characteristic features of a molten globule state.

It was observed that enzymatic activity of MSG dropped sharply until GdnHCl concentration 0.5 M, as shown in Fig. 4. The results go in parallel with the near-UV CD observations; where ~20% of tertiary structure and ~20% of enzymatic activity is left at 0.5 M GdnHCl. This might indicate sequential unfolding of the domains, where the domain containing active site is most susceptible to denaturation and most of the structure and activity loss occurs at lower concentration of GdnHCl (<0.5 M). The residual activity left beyond 1 M is within error limits.

The complete reversibility of the unfolding process is required to validate the thermodynamic parameters obtained from curve fitting. Therefore, activity regain on refolding by dilution from 4 M to 0.5 M GdnHCl was monitored, part of Fig. 4.

Equilibrium studies convey a global picture about the conformational changes in the protein and nothing much about fast kinetics. Therefore, further stopped-flow kinetic experiments were

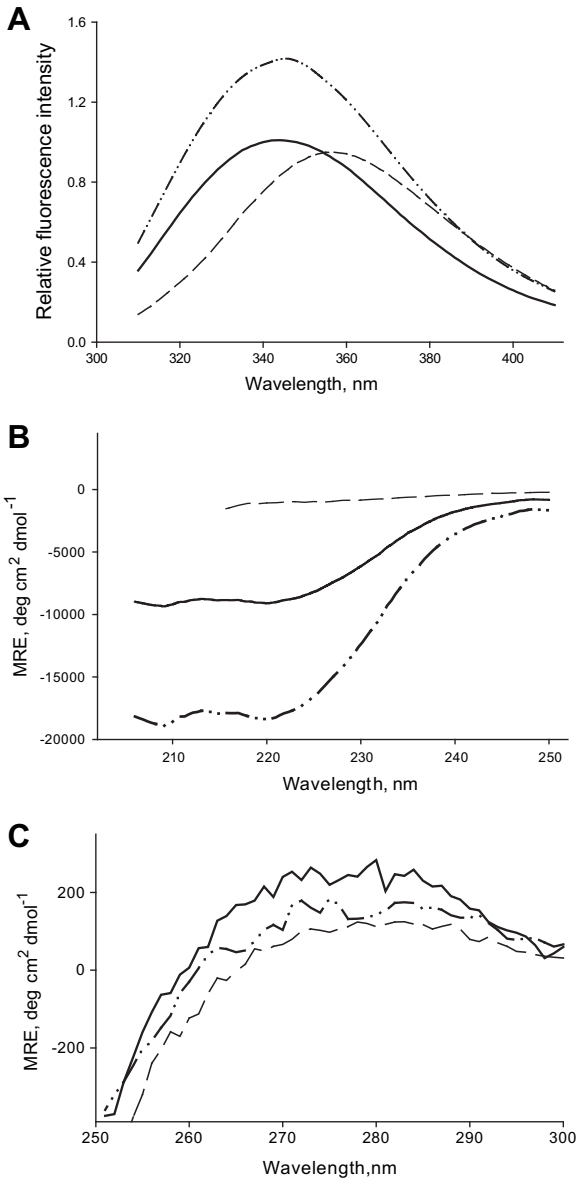

Fig. 3. Spectra for 1 μM of protein sample at different GdnHCl concentrations – 0 M (solid line), 0.4 M (dash-dotted line) and 4 M (dashed line), from various spectroscopic techniques. A) Intrinsic tryptophan fluorescence spectra, B) far-UV CD spectra and C) near-UV CD spectra.

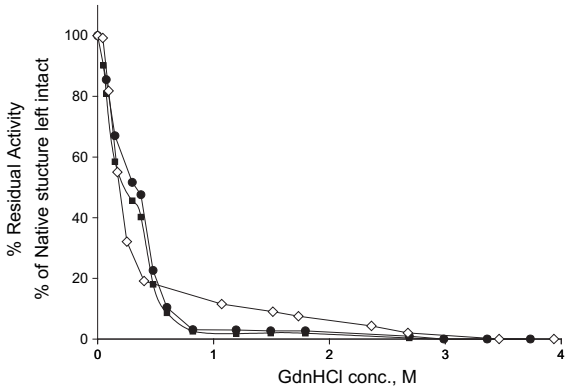

Fig. 4. Double plot-percentage of native tertiary structure from near-UV CD (◇) and Residual activity (◆, on unfolding; ■, on refolding) Vs GdnHCl concentration. Activity assay was performed as per the protocol given by Ornston and Ornston [18]. It is observed that at ~0.5 M GdnHCl conc., 80% activity and tertiary structure is lost. Final protein concentration was 1 μM during the assay and the basal activity (in absence of protein) was subtracted from the activity attained at each GdnHCl concentration.

performed, before putting forward a model of folding mechanism of Malate Synthase G.

3.3. Kinetic studies

As with any complex process, such as folding of multi-domain protein, time-resolved data are essential for elucidating the mechanism [22]. For large proteins it is predicted that initial state from where the refolding is initiated might be heterogeneous ensembles. In order to understand this problem, best approach would be to perform kinetic studies in stopped-flow mode.

Kinetic curves of MSG unfolding and refolding from the native (0 M) and unfolded state (4 M), which were monitored by intrinsic tryptophan fluorescence at various GdnHCl concentrations, are shown in Figs. 5 and 6, respectively. The observations from unfolding kinetic measurement indicate that substantial part of the unfolding of MSG occurs within the dead time of the stopped-flow instrument (20 ms), since burst phase becomes more pronounced as the stability of the folding intermediate decreases at higher concentration of GdnHCl (beyond 3 M) (Fig. 5a). The burst phase during unfolding may have been caused by rapid initial unzippering of tertiary structure of the protein. However, the time range under which these tertiary structures unfold is exactly not known. Until 1.5 M of GdnHCl concentration, transition curves fit well into monophasic kinetic equation beyond which data fits well in biphasic kinetic equation, as evident from the residual plot 5b.

At 4 M of GdnHCl concentration, non-linear least squares fitting gave an $A_{infinity}$ of 0.71 ± 0.00, A_1 of 0.14 ± 0.01, A_2 of 0.04 ± 0.01, k_1 of 1.23 ± 0.02 s^{-1} and k_2 of 0.06 ± 0.01 s^{-1}.

The refolding reaction was initiated by concentration jumps from 4 M (unfolded state) to various GdnHCl concentrations (4–0.4 M) at same temperature and pH. No prominent kinetics were observed above 1 M of GdnHCl, but became evident below this concentration. The fastest phase that disappeared within few milliseconds time was inferred only from an abrupt increase in fluorescence intensity at 340 nm (burst phase) and the second relatively slower phase was marked by slow increase in fluorescence intensity. The burst phase becomes clear and pronounced at lower concentrations (beyond 1 M), where more than 50% signal disappears within the burst phase (Fig. 6a). The fluorescence amplitude values recovered in the burst phase at final GdnHCl concentrations below 0.55 M show little dependence on GdnHCl concentration. Until 1.5 M the curves fit well in single exponential curve but beyond this concentration, the curves fit best in double exponential curve, as can be seen from the plots of the residuals in 6b.

At 0.4 M GdnHCl concentration, non-linear least squares fitting gave an A_o of 1.33 ± 0.08, A_1 of 0.09 ± 0.01, A_2 of 0.06 ± 0.00, k_1 of 0.98 ± 0.24 s^{-1} and k_2 of 0.05 ± 0.04 s^{-1}.

This refolding kinetic data is also quite in agreement with our observation from equilibrium refolding data, which shows about 50% activity regain at 0.4 M GdnHCl concentration.

Another factor that makes the picture clearer and strongly suggests for the presence of a kinetic intermediate is the roll over in both unfolding and refolding arm of chevron plot (as shown in Fig. 7). Curvature in the limbs of chevron plots may result from changes in the position of the transition state with the concentration of denaturant, i.e., Hammond behavior [23]. Roll over in refolding limb particularly, suggest of an early intermediate formation at low concentration of GdnHCl [24].

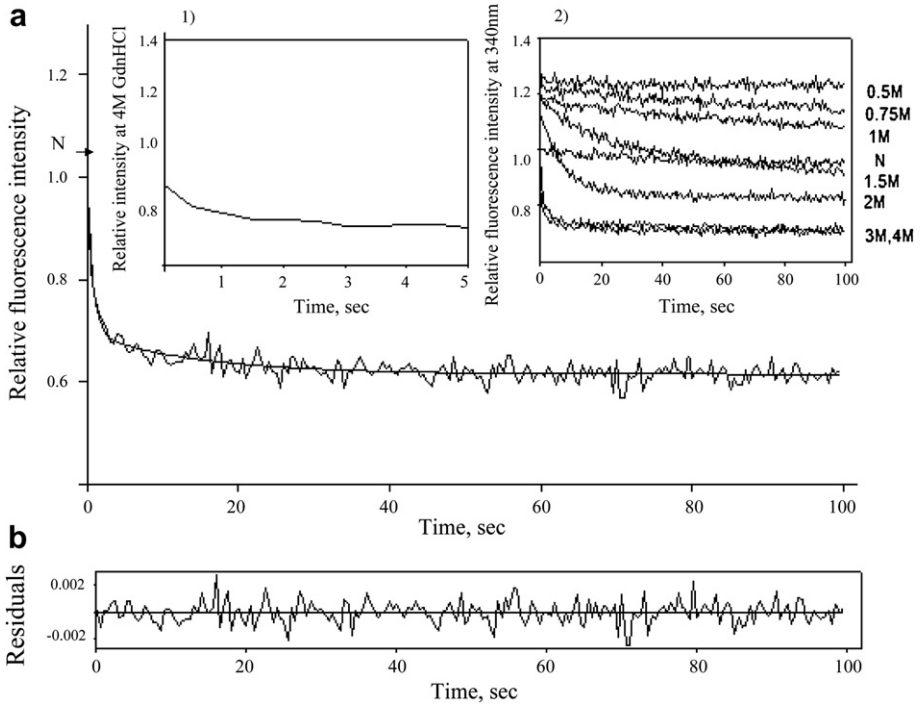

Fig. 5. Kinetic unfolding of MSG measured by change in fluorescence intensity at 340 nm. a) The unfolding kinetic trace of MSG at 4 M GdnHCl concentration was well fitted to double exponential curve (thick solid line). N denotes the fluorescence value of MSG in the native state obtained by extrapolation of the baseline of the N state. Inset 1 shows a log plot for first 5 s of kinetic phase. The final concentration of GdnHCl after mixing is indicated for each curve (in inset 2). b) The lower panel shows the residual between the experimental points and the theoretical curves obtained at 4 M GdnHCl conc.

The midpoint for folding according to the kinetic chevron plot (Fig. 7) is close to 1.6 M, where k_{obs} reach its minimum. This is in agreement with the values reported from the equilibrium data where the highest value for C_m is around 1.5 M for the midpoint of the I–U transition.

To see the possibility of any aggregate formation, which is quite usual for large proteins, kinetic measurements were repeated at few GdnHCl concentrations (0.4 M, 0.6 M, 1 M), for protein concentrations 1 μM, 5 μM and 10 μM. The rate constants did not show any appreciable protein concentration dependence, indicating that

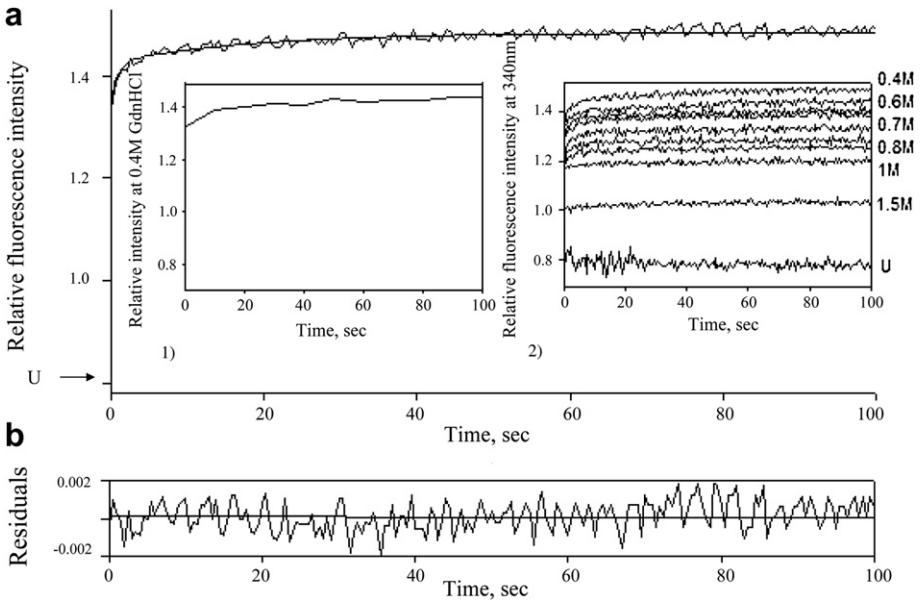

Fig. 6. Refolding kinetics trace of MSG measured by change in fluorescence intensity at 340 nm. a) The kinetic refolding trace of MSG at 0.4 M GdnHCl was well fitted to double exponential (thick solid line). U denotes the fluorescence value of MSG in the unfolded state obtained by extrapolation of the baseline of the U state. Inset 1 shows a log plot for first 5 s of kinetic phase. The final GdnHCl concentration is mentioned at the side of each curve (in inset 2). b) The lower panel shows the residual between the experimental points and the theoretical curves obtained at 0.4 M GdnHCl concentrations.

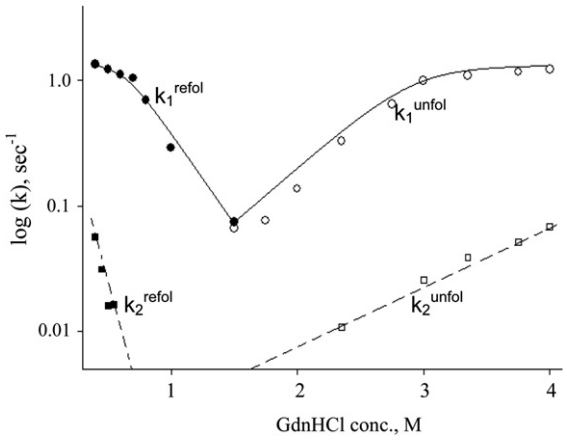

Fig. 7. Chevron plot—GdnHCl concentration dependence of kinetic rate constants. (○) and (□) represent the rate constants for unfolding and (●) and (■) represent the rate constants for refolding, of fast and slow phases, respectively.

under the experimental conditions, the folding reaction doesn't involve any aggregates.

Furthermore, since these experiments have been performed under near physiological conditions, it may be inferred that these *in vitro* results hold true for *in vivo* also.

4. Discussion

The function of kinetic and equilibrium intermediates in protein folding remains divisive. Do they represent any signature of distinct pathway or adventitious traps in funnel like landscape? Characterization of intermediates during the folding pathway of proteins is necessary to resolve this issue.

All the equilibrium unfolding results including the thermodynamic parameters are in congruence and clearly indicate towards one stable equilibrium intermediate. This intermediate has all the characteristics of a molten globule state, pronounced secondary structure, little or no tertiary structure, higher hydrophobic nature and compactness. A comparison of the properties of the native and unfolded states with the molten globule state has been summarized in Table 2.

Kinetic experiments give an opportunity to study the transition states that have their structure modified by change in conditions like change in GdnHCl concentration [25]. The refolding of MSG (Fig. 6) is characterized by the formation of an early intermediate state. However, the present results reveal that the intermediate state is formed within the dead time of the stopped-flow apparatus. The dead time of 20 ms restricted us from positioning the burst phase intermediate and we further need to employ a stopped-flow apparatus with lesser dead time or continuous flow apparatus, to investigate the intermediate state. Based on our results, it seems that the kinetic intermediate is nearly similar to the equilibrium intermediate. Since, this burst phase intermediate forms within the dead time, it indicate towards a substantial conformational recovery with most of the local secondary and long range tertiary interactions taking place during this time. The subsequent folding after the burst phase is a biphasic process leading to a final state similar to equilibrium transition state at the lower concentration of GdnHCl (<0.5 M). The presence of pronounced burst phase along with the curvature observed in folding arm of chevron plot (Fig. 7) strongly suggests towards the formation of a kinetic intermediate.

Unlike unfolding, refolding kinetics does not show the slow phase. The reason behind fast refolding kinetics could be the sudden dilution of the denaturant, resulting in the collapse of the extended peptide chain mainly by non-local hydrophobic interactions and by formation of hydrogen bonds between neighboring residues to form ordered secondary structural elements being followed by slower rearrangements of side chain packing in the subsequent stages [26,27]. Secondly, the reason for absence of slow phase during refolding might be that the native state could not be achieved, due to dilution limitation posed by the stopped-flow instrument. Therefore, refolding kinetics what we observed is only the transition of U → I. I → N phase is simply missing, which might have been a slow phase involving final inter-domain rearrangements. Also, the proteins with higher helical propensity follow hierarchical mechanism of folding where the formation of secondary structure follows the formation of hydrophobic core. In multi-domain proteins like ours, this collapse may lead to the formation of individual domains, before assembly, to form the native structure [28].

Domains may fold independently and to different extent. Generally, the most stable domain folds first, giving rise to kinetic intermediate with properties resembling the corresponding equilibrium molten globule intermediate and after the folding of all domains is done; they form specific docking interactions to reach an active three dimensional conformation. This step could be the rate-determining step where all the domains come closer and rearrange amongst themselves to achieve the final structure. A possible schematic presentation for the large protein folding, developed from the literature data [11,12,14–17] as well as from our experimental findings, has been given in Fig. 1.

Based on the observations in this paper (along with few hypothesis), we propose the following model of folding mechanism of MSG.

$$U \overset{fast}{\rightleftharpoons} I_{MG} \overset{K_{in}}{\rightleftharpoons} I_D \overset{K_{rd}}{\rightleftharpoons} N$$

where U is the unfolded state, I_{MG} is the molten globule state, I_D is the intermediate predicted to be formed when all the domains are in position to dock together and N is the finally achieved native state. k_{in} is the kinetic rate constant at molten-globule stage and k_{rd} is the predicted rate constant for rate-determining step.

Demonstrating whether the intermediate is on-pathway or off-pathway is not simple and requires further detailed structural studies or thorough kinetic analysis of the intermediate. Therefore, at this point nothing can be commented on the nature of the intermediate.

Folding studies on multi-domain proteins is quite a difficult task. One can't ignore the inter-domain effects and effects of the geometry of domain interfaces on overall folding of protein. For detail investigation of the effects of neighboring domains on each other, one has to study both the thermodynamics and kinetics of

Table 2
Comparison of conformational properties of MSG in different states.

State	λ_{max} nm	Relative fluorescence intensity at 340 nm	CD ($[\theta]_{222}$) deg cm^2 dmol^{-1}	CD ($[\theta]_{280}$) deg cm^2 dmol^{-1}	% Activity	Relative fluorescence intensity at 467 nm
Native state	343	1.0	−8946.1	283.1	100	1.0
Equilibrium intermediate	346	1.4	−18100.0	145.2	50	1.2
Unfolded state	357	0.7	100.0	111.0	0	0.2

the domain in isolation as well as in multi-domain proteins. But, the thermodynamics and kinetics of multi-domain can be complex to analyze. Also, in multi-domain proteins it is difficult to define the limit, where one domain begins and other ends. Furthermore, there can be difficulties in expression of soluble form of an individual domain, where there are extensive inter-domain interfaces. In such cases, expression of individual domains can lead to aggregate formation. Due to these limitations not much folding studies has been done on multi-domain protein.

Our next set of experiments would include folding studies of individual domains and its effect on overall folding of this protein.

Acknowledgement

The authors thankfully acknowledge the gift of Malate Synthase G plasmid from Dr. Laura Baldoma, University of Barcelona, Spain. We are grateful to Dr. J. Udgaonkar (National Centre for Biological Sciences, Bangalore (NCBS)) for invaluable assistance and advices.

References

[1] C.B. Anfinsen, Principles that govern the folding of protein chains. Science 181 (1973) 223–230.
[2] M. Jamin, R.L. Baldwin, Refolding and unfolding kinetics of the equilibrium unfolding intermediate of apomyoglobin. Nature 3 (1996) 613–619.
[3] T.K. Chaudhuri, M. Arai, T.P. Terada, T. Ikura, K. Kuwajima, Equilibriums and kinetic studies on folding of the authentic and recombinant forms of human – lactalbumin by circular dichroism spectroscopy. Biochemistry 39 (2000) 15643–15651.
[4] E.N. Baryshnikova, B.S. Melnik, A.V. Finkelsteinx, G.V. Semisotnov, V. E. Bychkova, Three-state protein folding: experimental determination of free-energy profile. Protein Sci. 14 (2005) 2658–2667.
[5] M. Arai, K. Kuwajima, Rapid formation of a molten globule intermediate in refolding of – lactalbumin. Folding Des. 1 (1996) 275–285.
[6] B.W. Matthews, Structural and genetic analysis of protein stability. Annu. Rev. Biochem. 62 (1993) 139–160.
[7] T. Schindler, M. Herder, M.A. Marahiel, F.X. Schmid, Extremely rapid protein folding in the absence of intermediates. Nat. Struct. Biol. 2 (1995) 663–673.
[8] K. Teilum, K. Maki, B.B. Kragelund, F.M. Poulsen, H. Roder, Early kinetic intermediate in the folding of Acyl-Coenzyme A binding protein detected by fluorescence labeling and ultrarapid mixing. Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 9807–9812.
[9] W.A. Houry, D. Frishman, C. Eckerskorn, F. Lottspeich, F.U. Hartl, Identification of in vivo substrates of the chaperonin GroEL. Nature 402 (1999) 147–154.
[10] I.E. Sanchez, M. Morillas, E. Zobeley, T. Kiefhaber, R. Glockshuber, Fast folding of the two domains semliki forest virus capsid protein explains co-translational proteolytic activity. J. Mol. Biol. 338 (2004) 159–167.
[11] J.M. Berg, J.L. Tymoczko, L. Stryer, The citric acid cycle, in: Biochemistry, fifth ed. WH Freemam & Co., New York, 2002, p. 484.
[12] B.R. Howard, J.M. Endrizzi, J.M. Remington, Crystal structure of *Escherichia coli* malate synthase G complexed with magnesium and glyoxylate at 2.0 A resolution: mechanistic implications. Biochemistry 39 (2000) 3156–3168.
[13] V. Tugarinov, W.-Y. Choy, V.Y. Orekhov, L.E. Kay, Solution NMR-derived global fold of a monomeric 82-kDa enzyme. Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 622–627.
[14] J.H. Han, S. Batey, A.A. Nickson, S.A. Teichmann, J. Clarke, The folding and evolution of multidomain proteins. Nat. Rev. Mol. Cell. Biol. 8 (2007) 319–330.
[15] P. Privalov, Intermediate states in protein folding. J. Mol. Biol. 258 (1996) 707–725.
[16] E. Sudharshan, A.G. Rao, Involvement of cysteine residues and domain interactions in the reversible unfolding of lipoxygenase-I. J. Biol. Chem. 274 (1999) 35351–35358.
[17] T.K. Chaudhuri, G.W. Farr, W.A. Fenton, S. Rospert, A.L. Horwich, GroEL/GroES-mediated folding of a protein too large to be encapsulated. Cell 107 (2001) 235–246.
[18] L.N. Ornston, M.K. Ornston, Regulation of glyoxylate metabolism in *Escherichia coli* K-12. J. Bacteriol. 98 (1969) 1098–1108.
[19] M. Mizuguchi, M. Arai, K. Nitta, K. Kuwajima, Equilibrium and kinetics of the folding of equine lysozyme studied by circular dichroism spectroscopy. J. Mol. Biol. 283 (1998) 265–277.
[20] C.N. Pace, Determination and analysis of urea and guanidine hydrochloride denaturation curves. Methods Enzymol. 131 (1986) 266–280.
[21] G.D. Fasman, Determination of protein secondary structure, in: Circular Dichroism and the Conformational Analysis of Biomolecules. Plenum Press, New York, 1996, pp. 69–104.
[22] H. Roder, K. Maki, H. Cheng, M.C.R. Shastry, Rapid mixing methods for exploring the kinetics of protein folding. Methods 34 (2004) 15–27.
[23] A.R. Fersht, Structure and Mechanism in Protein Science. WH Freeman, New York, 1998.
[24] M.J. Parker, S. Marqusee, The cooperativity of burst phase reactions explored. J. Mol. Biol. 293 (1999) 1195–1210.
[25] A.R. Fersht, Characterizing transition states in protein folding: an essential step in the puzzle. Curr. Opin. Struct. Biol. 7 (1997) 3–9.
[26] K.A. Dill, D.O.V. Alonso, K. Hutchinson, Thermal stabilities of globular proteins. Biochemistry 28 (1989) 5439–5449.
[27] K.A. Dill, Dominant forces in protein folding. Biochemistry 29 (1990) 7133–7156.
[28] A.R. Fersht, Nucleation mechanism in protein folding. Curr. Opin. Struct. Biol. 5 (1995) 79–84.