Elucidation of the Molecular Mechanisms of Protein Folding

Department of Life and Coordination-Complex Molecular Science
Division of Biomolecular Functions

KUWAJIMA, Kunihiro	Professor
MAKABE, Koki	Assistant Professor
NAKAMURA, Takashi	IMS Research Assistant Professor
CHEN, Jin	OIIB Research Assistant Professor
TAKENAKA, Toshio	Post-Doctoral Fellow
MIZUKI, Hiroko	Technical Fellow
TANAKA, Kei	Secretary

Kuwajima group is studying mechanisms of *in vitro* protein folding and mechanisms of molecular chaperone function. Our goals are to elucidate the physical principles by which a protein organizes its specific native structure from the amino acid sequence. In this year, we studied the fibrillogenic propensity of the GrEL apical domain and the sequential four-state folding/unfolding of goat α-lactalbumin and its N-terminal variants.

1. Fibrillogenic Propensity of the GroEL Apical Domain: A Janus-Faced Minichaperone

The chaperonin GroEL plays an essential role in promoting protein folding and in protecting against misfolding and aggregation in the cellular environment. In this study, we report that both GroEL and its isolated apical domain form amyloid-like fibrils under physiological conditions, and that the fibrillation of the apical domain is accelerated under acidic conditions. We also found, however, that despite its fibrillation propensity, the apical domain exhibits a pronounced inhibitory effect on the fibril growth of β_2-microglobulin. The analysis of the primary amino-acid sequence by programs, PASTA, TANGO and Zyggregator, which predict aggregation-prone sequences, indicates that the most aggregation-prone sequence is located in residues 260–280, which is coincident with the substrate protein-binding site in the chaperonin GroEL. Therefore, there is a close relationship between the fibrillogenic propensity and the substrate binding of GroEL. Furthermore, the analysis of 1003 sequences of the chaperonin family proteins by the Zyggregator program has shown that the aggregation-prone sequence is present in the substrate-binding site, indicating that the close relationship between the fibrillogenic propensity and the substrate binding is a general property of the chaperonin family.

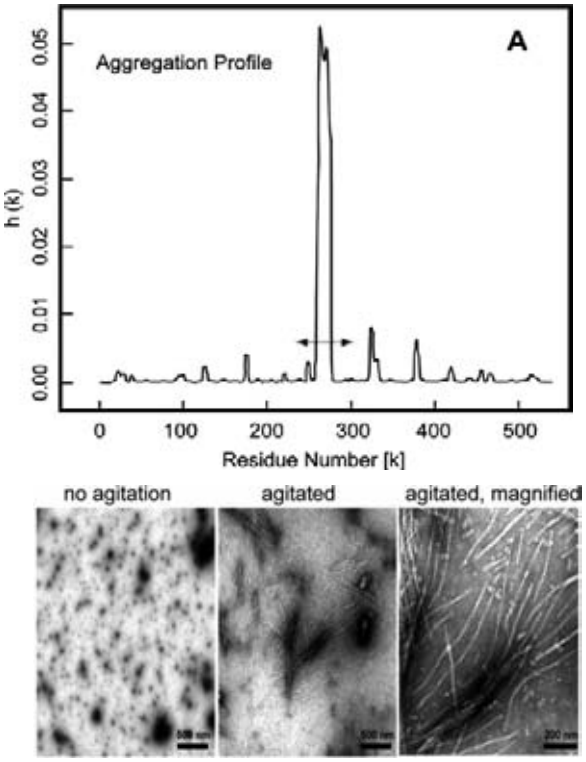

Figure 1. The aggregation profile of GroEL obtained by the analysis with PASTA (A), and the transmission electron microscopic observations of the isolated apical domain at pH 7.0 and 37 °C at 0.1 M NaCl in the presence of 0.5 mM SDS, where the amyloid-like fibrils of the apical domain were formed when the solution was agitated.

2. Sequential Four-State Folding/Unfolding of Goat α-Lactalbumin and Its N-Terminal Variants

Equilibria and kinetics of folding/unfolding of goat α-lactalbumin (GLA) and its two N-terminal variants were studied by circular dichroism spectroscopy. The two variants were wild-type recombinant and Glu1-deletion (E1M) variants expressed in *Escherichia coli*. The presence of an extra methionine at the N terminus in recombinant GLA destabilized the protein by 2 kcal/mol, while the stability was recovered in the E1M variant in which Glu1 was replaced by Met1. Kinetic folding/unfolding reactions of the proteins, induced by stopped-flow concentration jumps of guanidine hydrochloride, indicated the presence of a burst-phase in refolding, and gave chevron plots with significant curvatures in both the folding and unfolding limbs. The folding-limb curvature was interpreted in terms of accumulation of the burst-phase intermediate (I). However, there was no burst phase observed in the unfolding kinetics to interpret the unfolding-limb curvature. We thus assumed a sequential four-state mechanism, in which the folding from the burst-phase intermediate takes place via two transition states separated by a high-energy intermediate (J). We estimated changes in the free energies of the burst-phase intermediate I and two transition states (‡1 and ‡2), caused by the N-terminal variations and also by the presence of stabilizing calcium ions. The Φ values at the N terminus and at the Ca^{2+}-binding site thus obtained increased successively during folding, demonstrating the validity of the sequential mechanism. The stability and the folding behavior of the E1M variant were essentially identical to those of the authentic protein, allowing us to use this variant as a pseudo-wild-type GLA in future studies.

Residue No. / Protein	0	1	2	3	4	5	6	7	8	9	10
Authentic GLA		E	Q	L	T	K	C	E	V	F	Q
Recombinant GLA	M	E	Q	L	T	K	C	E	V	F	Q
E1M variant		M	Q	L	T	K	C	E	V	F	Q

Figure 2. The amino acid sequences of the N-terminal residues of authentic GLA, recombinant GLA, and the E1M variant.

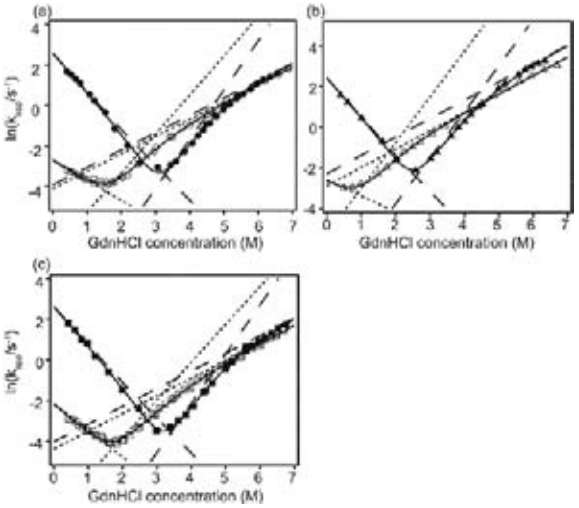

Figure 3. Chevron plots of authentic GLA (a), recombinant GLA (b), and the E1M variant (c) in the holo (filled symbols) and apo (open symbols) forms at pH 7.0 and 25 °C. The solid lines are theoretical curves fitted by the sequential four-state model. The broken lines and dotted lines represent the GdnHCl dependence of logarithms of microscopic rate constants ($\ln(k_1)$, $\ln(k_{-1}k_{-2}/k_2)$ and $\ln(k_{-2})$) of folding and unfolding kinetics; the first two are rate-limited by the transition state 1 (‡1), and the last one by the transitions state 2 (‡2).

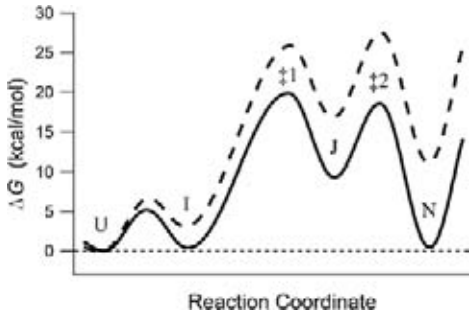

Figure 4. Schematic free-energy profiles of unfolding of GLA under a weakly unfolding condition (solid line) and under a strongly unfolding condition (broken line). Under the weakly unfolding condition, the transition state is located at ‡1, while it is located at ‡2 under the strongly unfolding condition at a high GdnHCl concentration (>5.5*M*). The free-energy profiles shown correspond to those for authentic holo GLA at 3.3*M* GdnHCl (solid line) and at 6.5*M* GdnHCl (broken line). A hypothetical intermediate (J) located between ‡1 and ‡2 is metastable under all conditions, that is, it is higher in free energy than U, I, and N.

References

1) J. Chen, H. Yagi, P. Sormanni, M. Vendruscolo, K. Makabe, T. Nakamura, Y. Goto and K. Kuwajima, *FEBS Lett.* **586**, 1120–1127 (2012).

2) K. Tomoyori, T. Nakamura, K. Makabe, K. Maki, K. Saeki and K. Kuwajima, *Proteins; Struct., Funct., Bioinf.* **80**, 2191–2206 (2012).