

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 30 日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2013～2016

課題番号：25440075

研究課題名(和文)速度データの相関解析に基づく非二状態蛋白質のフォールディング機構研究

研究課題名(英文)Folding mechanism of non-two-state proteins based on correlation analyses of kinetic folding data

研究代表者

桑島 邦博(Kuwajima, Kunihiro)

東京大学・大学院理学系研究科(理学部)・名誉教授

研究者番号：70091444

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：球状蛋白質のフォールディング速度過程に関する文献調査を行い、蛋白質フォールディングのデータベースを構築した。データベースの総エントリー数は141(非二状態蛋白質59,二状態蛋白質82)である。このデータベースから、非二状態蛋白質と二状態蛋白質を、それぞれ別々に取りだし、フォールディング速度定数の対数($\ln k_f$)と、蛋白質の立体構造特性との間の相関を解析した。立体構造特性として、研究代表者等が提案している非局所的接触クラスターの数の対数($\ln N_c$)を用いることにより、非二状態蛋白質と二状態蛋白質のフォールディング分子機構の違いを明確に示すことが出来た。

研究成果の概要(英文)：We carried out literature searching on the folding kinetics of globular proteins, and constructed a protein-folding database. The total entry number of the database is 141, consisting of 59 non-two-state proteins and 82 two-state proteins. We took the non-two-state proteins and the two-state proteins separately from the database, and investigated correlations between the logarithmic rate constant ($\ln k_f$) of folding and the structural properties of proteins. We show that using the logarithmic number of nonlocal contact clusters, a structural property proposed by us, we can clearly describe differences in the molecular mechanisms of folding between the non-two-state and the two-state proteins.

研究分野：生物物理学

キーワード：フォールディング 速度論 データベース

1. 研究開始当初の背景

(1) 蛋白質のフォールディング問題は、人工蛋白質の *de novo* 設計やアルツハイマー病などのフォールディング病治療法の開発にも関係した生物物理学における重要な問題である。

(2) 二状態フォールディングを示す小さな単一ドメイン蛋白質のフォールディング速度定数 k_f が主鎖構造のトポロジカルな複雑さを示す構造特性 (Relative contact order 等) に強く相関することが知られ、蛋白質のフォールディング機構は主鎖構造の複雑さによって決まるのだという理解が定着しつつある。しかし、中間体を示す、非二状態蛋白質のフォールディングがどのような分子機構によって決まるのかは実は未だ全く明らかではなかった。

(3) 過去数年間の間に、蛋白質のフォールディング速度に関する多くの実験データが蓄積され、これらのデータをまとめた、新たな、蛋白質のフォールディング速度に関するデータベースを構築する必要がある。

2. 研究の目的

(1) 蛋白質フォールディングデータベース構築：最新の蛋白質フォールディング速度に関するデータも含めた、最新の蛋白質フォールディングデータベースを構築する。

(2) 非二状態蛋白質のフォールディング機構：上で作成されるデータベースを基にして、 k_f と様々な構造特性との間の相関解析を行う。その際、二状態蛋白質と非二状態蛋白質を、それぞれ、別々に解析し、非二状態蛋白質に見出される特徴を明らかにする。非二状態蛋白質のフォールディング分子機構の解明に寄与する。

3. 研究の方法

(1) フォールディングデータベース構築の方法：PubMed 等で、球状蛋白質のフォールディング速度過程に関する実験の論文を検索し、フォールディング速度に関するデータを収集した。その際、単一ドメインの球状蛋白質で蛋白質データバンク (PDB) に立体構造が登録されているもののみを対象とし、複数ドメインや PDB 未登録の蛋白質は除外した。ジスルフィド結合や共有結合で繋がった補欠分子族はフォールディング速度に影響すると考えられるので、これらを含む蛋白質も除外した。また、アミノ酸残基数 34 未満の小ペプチドも除外した。収集されたデータと、既存の蛋白質フォールディングデータベース (Garbuzynskiy 等[1]及び Wagaman 等[2]) との比較を行い、最新でエントリー数最大のデータベースを構築した。既存のデータベースには二重登録や k_f 値の間違いなどがあったので、これらも修正した。データベース中の

エントリー項目は、蛋白質名、PDB コード、PDB 中の蛋白質長、実際の蛋白質長、構造クラス (α , β , $\alpha\beta$, $\alpha+\beta$)、実験条件 (pH と温度)、 k_f の対数 ($\ln k_f$)、フォールディング型 (二状態または非二状態) の 8 項目である。

(2) ($\ln k_f$) と構造特性との間の相関解析：上で作成されたデータベースを基に、($\ln k_f$) と、様々な構造特性との間の相関解析を行った。既存の構造特性である、Relative contact order (RCO) [3], Absolute contact order (ACO) [4], Long-range order (LRO) [5], Chain topology parameter (CTP) [6], Total contact distance (TCD) [7]に加え、研究代表者等が独自に提案している、非局所的接触クラスター (Nonlocal contact cluster, NLCC) の数の対数 ($\ln N_c$) と ($\ln k_f$) との間の相関解析を行った。相関解析では、各構造特性を決める各パラメータ値を、交差検証を用いて最適化した上で、($\ln k_f$) と各構造特性との間の線形回帰分析から相関係数を求めた。

(3) NLCC の定義とパラメータ最適化：本研究の NLCC は、以前、研究代表者等が提案した構造特性に改良を加えたものであり[8]、二つの残基の間の接触強度 γ_c は、残基間の (水素原子を除いた) 原子-原子ペア (AA pair) 間の距離 ($d_{m,n}$) (m と n は、それぞれの残基に属する原子の順番を表す) のロジスティック関数の和として、次式で与えられる (図 1)。

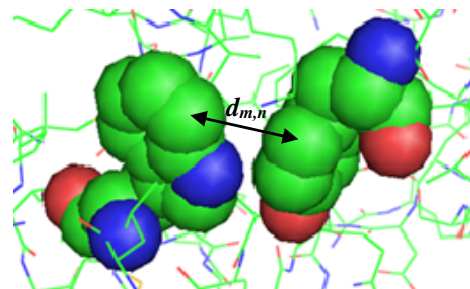


図 1: 二つの残基の原子-原子ペア間の距離

$$\gamma_c = \sum_{m,n}^{\text{AA pairs}} \frac{1}{1 + \exp[\alpha(d_{m,n} - d_{\text{cut}})]} \quad (1)$$

ここで、 d_{cut} はカットオフ距離、 α は関数の傾斜強度を表す。接触強度カットオフを g とし、 $\gamma_c \geq g$ のとき二つの残基は互いに空間的に接触していると定義する。さらに、空間的に接触している、二つの残基 (残基 i と j) が、配列に沿ってカットオフ配列距離 L 以上離れているとき (i.e., $|j - i| \geq L$)、残基 i と j は非局所的接触をしていると定義する。非局所的接触にある二つの残基ペア (i, j) と (k, l) が互いに接近しているとき (i.e., $|k - i| + |l - j| \leq 4$)、二つの残基ペアは同じ NLCC に属すると定義する。ただし、NLCC には少なくとも三つ以上の接触残基ペアがなくてはならない。これらの定義を用いることにより、蛋白質のコンタ

クトマップ上に NLCC を描き、その数 N_c を求めることが出来る (図2)。(1)式を用いる方が、 d_{cut} で不連続に接触を定義するよりも、より良い ($\ln k_f$) と ($\ln N_c$) との相関が観測された。

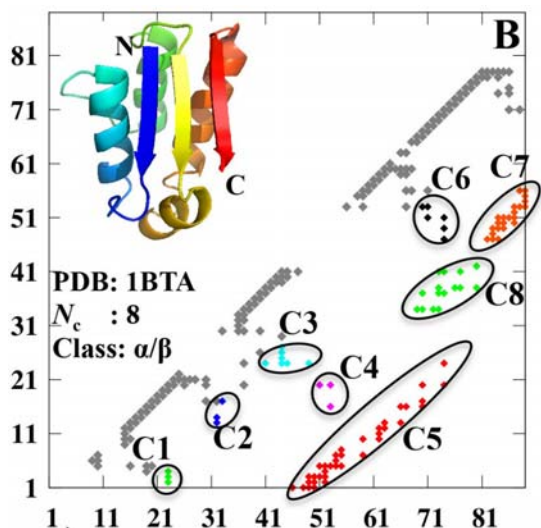


図2：バルスターのコンタクトマップ。線で囲まれた領域がパラメータ最適化後のNLCCであり、 $N_c=8$ である。

本研究では、($\ln k_f$) と ($\ln N_c$) との間の相関が最大となるように四つのパラメータ (L , a , d_{cut} , g) の値を、交差検証を用いて最適化した。作成されたデータベースから、非二状態蛋白質と二状態蛋白質を、それぞれ別々に取り出し、それぞれについて、パラメータの最適化を行った。既存の構造特性 (RCO, ACO, LRO, CTP, TCD) についても同様にパラメータの最適化を行った。

4. 研究成果

(1) データベースの構築：上記の方法に従って蛋白質フォールディングデータベースを構築した。総エントリー数は141であり、二状態蛋白質 82, 非二状態蛋白質 59である。既存のデータベースにはない、新規なエントリー数が 32 あり、本データベースが最新で最大のものとなる。構造クラス別に分類すると、 α 型が 35 (二状態 25, 非二状態 10), α/β 型が 15 (二状態 1, 非二状態 14), $\alpha+\beta$ 型が 41 (二状態 23, 非二状態 18), β 型が 50 (二状態 33, 非二状態 17)であった。今後、本データベースが、蛋白質の構造やフォールディングの理論的な研究や計算機シミュレーションに活用されることが期待される。完成されたデータベースは、共同研究者である Jooyoung Lee 教授 (韓国高等科学院) の研究室ホームページ上にアップロードされる (<http://lee.kias.re.kr/~protein/wiki/doku.php?id=pkhd:view>)。

(2) 構造特性のパラメータの最適化：データベースから非二状態蛋白質と二状態蛋白質を別々に取り出し、それぞれについて、($\ln k_f$)

との相関が最大となるように、構造特性のパラメータの最適化を行なった。NLCCの四つのパラメータについて、最適化した結果が表1である。表1より、非二状態蛋白質と二状態蛋白質では、最適化パラメータの値が明確に異なっており、これは両者の間で、フォールディング分子機構に違いがあるためと考えられる。非二状態蛋白質では、一次配列上十分離れた ($L=16$) 残基同士の、空間的に接近した ($d_{cut}=4.4 \text{ \AA}$) 相互作用が支配的である。一方、二状態蛋白質では一次構造上余り離れていない ($L=6$) 残基同士の、空間的にはもっと広い範囲に及ぶ ($d_{cut}=8.2 \text{ \AA}$) 相互作用も支配的である (表1)。

表1：NLCCの最適化パラメータ値

	L	a (\AA^{-1})	d_{cut} (\AA)	g	相関係数
非二状態	16	2.2	4.4	1.2	-0.879
二状態	6	1.8	8.2	0.6	-0.719

そこで、表1のパラメータ値を使い、代表的な非二状態蛋白質4種について、それらのコンタクトマップ上にNLCCを描いてみた (図3)。同様に、二状態蛋白質4種についてもNLCCを描いてみた (図4)。

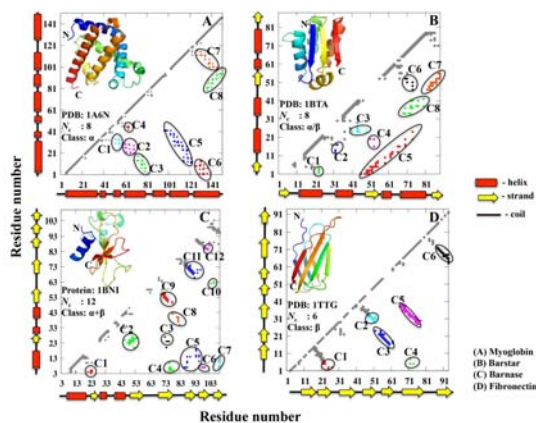


図3：非二状態蛋白質 (A, ミオグロビン; B, バルスター; C, バルナーゼ; D, フィブネクチン・タイプ III モジュール) のコンタクトマップとNLCC

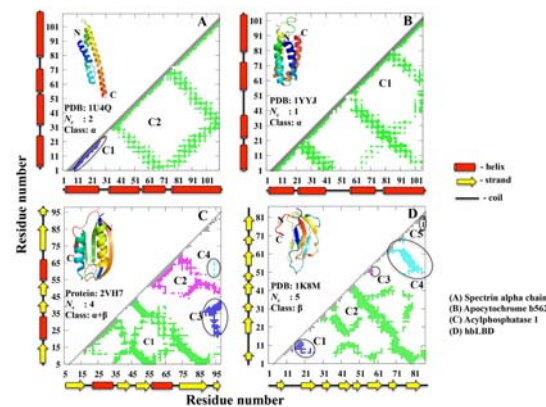


図4：二状態蛋白質 (A, α スペクトリン R15 ドメイン; B, アポシトクローム b_{562} ; C, アクリルフォスファターゼ 1; D, hbLBD) のコンタクトマップとNLCC

図3と図4の比較から、非二状態蛋白質と二状態蛋白質のフォールディング分子機構の違いが明らかに見て取れる。非二状態蛋白質では、コンタクトマップ上に多くのNLCCが描かれた(図3)。この場合、NLCCは二次構造要素(α ヘリックス、 β ヘアピン等)の相互作用領域に対応すると見なされるので、図3の結果は、非二状態蛋白質のフォールディングが、二次構造の形成された中間体を通して進む、所謂、「枠組みモデル」に従っていることを示す。一方、二状態蛋白質では、NLCCの数が少なく、分子全体が一つのNLCCであるか(図4B)、分子の半分以上が一つの大きなNLCCとなり、それに小さなNLCCが少数加わっている場合が多い(図4A, C, D)。このような図4の結果は、二状態蛋白質のフォールディングが、悉無律的な「核形成凝縮モデル」に従っていることを示す。このように、NLCCという構造特性を用いることにより、非二状態蛋白質と二状態蛋白質のフォールディング分子機構の違いを明確に示すことができた。また、蛋白質の立体構造から得られる($\ln N_c$)値からフォールディング速度定数 k_f を予測する方法を確立できた。

(3) ($\ln N_c$)と既存の構造特性との比較：本研究で提案された($\ln N_c$)と既存の構造特性(RCO, ACO, LRO, CTP, TCD)との間で、実験的に観測された($\ln k_f$)との相関がどのように違うかを調べてみた。

非二状態蛋白質に対する比較の結果が図5である。($\ln N_c$)が、他の構造特性に比べ明らかにより良い相関を示しており、NLCCや($\ln N_c$)が非二状態蛋白質のフォールディング速度過程を記述する上で、極めて良好な構造特性であることが明らかとなった。

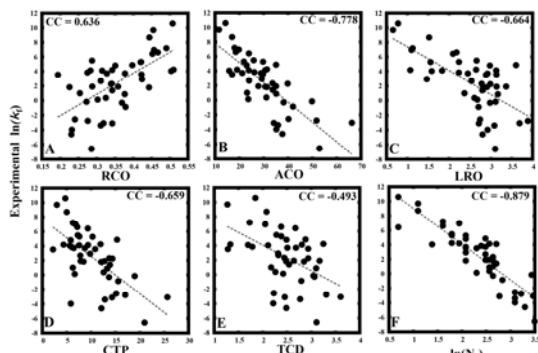


図5：非二状態蛋白質の各構造特性と($\ln k_f$)との関係

一方、二状態蛋白質においては、($\ln N_c$)と($\ln k_f$)との相関係数は-0.719と十分有意ではあるが、LRO(相関係数=-0.769)やTCD(相関係数=-0.723)よりは相関が弱いことが分かった(図6)。NLCCが接触残基ペアのクラスター(構造要素間の接触領域)を表しているのに対し、LROとTCDは、いずれも、一次配列上離れた残基間の接触ペアの数や接触している残基間の長さにより定義される

構造特性である[5, 7]。二状態蛋白質では、構造要素間の相互作用よりも、残基間の相互作用の方がより直接的にフォールディング速度過程に関わっていることを示している。

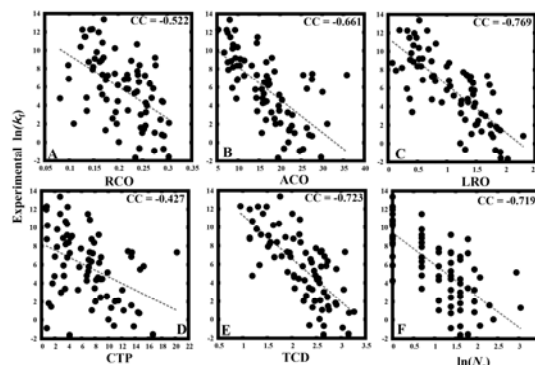


図6：二状態蛋白質の各構造特性と($\ln k_f$)との関係

<引用文献>

- Garbuzynskiy, S.O., Ivankov, D.N., Bogatyreva, N.S. & Finkelstein, A.V. (2013). Golden triangle for folding rates of globular proteins. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 147-150.
- Wagaman, A.S., Coburn, A., Brand-Thomas, I., Dash, B. & Jaswal, S.S. (2014). A comprehensive database of verified experimental data on protein folding kinetics. *Protein Sci.* **23**, 1808-1812.
- Plaxco, K.W., Simons, K.T. & Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985-994.
- Kamagata, K., Arai, M. & Kuwajima, K. (2004). Unification of the Folding Mechanisms of Non-two-state and Two-state Proteins. *J. Mol. Biol.* **339**, 951-965.
- Gromiha, M.M. & Selvaraj, S. (2001). Comparison between long-range interactions and contact order in determining the folding rate of two-state proteins: application of long-range order to folding rate prediction. *J. Mol. Biol.* **310**, 27-32.
- Nölting, B., Schälke, W., Hampel, P., Grundig, F., Gantert, S., Sips, N., Bandlow, W. & Qi, P.X. (2003). Structural determinants of the rate of protein folding. *J. Theor. Biol.* **223**, 299-307.
- Zhou, H. & Zhou, Y. (2002). Folding rate prediction using total contact distance. *Biophys. J.* **82**, 458-463.
- Kamagata, K. & Kuwajima, K. (2006). Surprisingly high correlation between early and late stages in non-two-state protein folding. *J. Mol. Biol.* **357**, 1647-1654.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 8件)

- ① Goyal, M., Chaudhuri, T.K., Kuwajima, K. (2014). Irreversible denaturation of maltodextrin glucosidase studied by differential scanning calorimetry, circular dichroism, and turbidity measurements. *PLoS One* **9**, e115877, 査読有, DOI: 10.1371/journal.pone.0115877.
- ② Nakagawa, A., Moriya, K., Arita, M., Yamamoto, Y., Kitamura, K., Ishiguro, N., Kanzaki, T., Oka, T., Makabe, K., Kuwajima, K., Yohda, M. (2014). Dissection of the ATP-dependent conformational change cycle of a group II chaperonin. *J. Mol. Biol.* **426**, 447–459, 査読有, DOI: 10.1016/j.jmb.2013.09.034.
- ③ Ohmae, E., Miyashita, Y., Tate, S., Gekko, K., Kitazawa, S., Kitahara, R., Kuwajima, K. (2013). Solvent environments significantly affect the enzymatic function of *Escherichia coli* dihydrofolate reductase: comparison of wild-type protein and active-site mutant D27E. *Biochim. Biophys. Acta* **1834**, 2782–2794, 査読有, DOI: 10.1016/j.bbapap.2013.09.024.
- ④ Sekiguchi, H., Nakagawa, A., Moriya, K., Makabe, K., Ichiyangi, K., Nozawa, S., Sato, T., Adachi, S., Kuwajima, K., Yohda, M., Sasaki, Y.C. (2013). ATP dependent rotational motion of group II chaperonin observed by X-ray single molecule tracking. *PLoS One* **8**, e64176, 査読有, DOI: 10.1371/journal.pone.0064176
- ⑤ Yagi-Utsumi, M., Kunihara, T., Nakamura, T., Uekusa, Y., Makabe, K., Kuwajima, K., Kato, K. (2013). NMR characterization of the interaction of GroEL with amyloid β as a model ligand. *FEBS Lett.* **587**, 1605–1609, 査読有, DOI: 10.1016/j.febslet.2013.04.007.
- ⑥ Chandak, M.S., Nakamura, T., Makabe, K., Takenaka, T., Mukaiyama, A., Chaudhuri, T.K., Kato, K., Kuwajima, K. (2013). The H/D-exchange kinetics of the *Escherichia coli* co-chaperonin GroES studied by 2D NMR and DMSO-quenched exchange methods. *J. Mol. Biol.* **425**, 2541–2560, 査読有, DOI: 10.1016/j.jmb.2013.04.008.
- ⑦ Nakamura, T., Aizawa, T., Kariya, R., Okada, S., Demura, M., Kawano, K., Makabe, K., Kuwajima, K. (2013). Molecular mechanisms of the cytotoxicity of human α -lactalbumin made lethal to tumor cells (HAMLET) and other protein-oleic acid complexes. *J. Biol. Chem.* **288**, 14408–14416, 査読有, DOI: 10.1074/jbc.M112.437889.
- ⑧ Chandak, M.S., Nakamura, T., Takenaka, T., Chaudhuri, T.K., Yagi-Utsumi, M., Chen, J., Kato, K., Kuwajima, K. (2013). The use of spin desalting columns in DMSO-quenched H/D-exchange NMR experiments, *Protein Sci.* **22**, 486–491, 査読有, DOI: 10.1002/pro.2221.
- [学会発表] (計 14 件)
- ① Makabe, K., Kuwajima, K. Folding mechanism of *Borrelia burgdorferi* Osp A, a dumbbell-shaped two-domain protein, The 3rd Korean-Polish Conference on Protein Folding: Theoretical and Experimental Approaches, 2017年2月5–9日, Jeongsun-kun (Korea)
- ② Manavalan, B., Kuwajima, K., Lee, J. Relationship between the folding rate and structure-based parameters of globular proteins, 第54回日本生物物理学会年会, 2016年11月25–27日, つくば国際会議場 (茨城県つくば市)
- ③ Kuwajima, K., Manavalan, B., Lee, J. The problem of protein folding, The 16th KIAS Conference on Protein Structure, 2016年9月22–24日, Seoul (Korea)
- ④ Kuwajima, K., Manavalan, B., Lee, J. The problem of protein folding, The 2nd Polish-Korean Conference on Protein Folding: Theoretical and Experimental Approaches, 2016年5月28日–6月1日, Gdansk (Poland)
- ⑤ Manavalan, B., Kuwajima, K., Lee, J. Relationship between the folding rate and structure-based parameters of globular proteins, 第16回日本蛋白質科学会年会, 2016年6月7–9日, 福岡国際会議場 (福岡県福岡市)
- ⑥ Kuwajima, K., Chandak, M.S., Yagi-Utsumi, M., Yamaguchi, T., Nakamura, T., Kato, K. The H/D-exchange kinetics of a fully unfolded protein studied by the DMSO-quenched and 2D NMR techniques, The 8th Japan-Korea Seminar on Biomolecular Sciences: Experiments and Simulations, 2016年2月15–17日, 岡崎コンファレンスセンター (愛知県岡崎市)
- ⑦ Kuwajima, K., Manavalan, B., Lee, J. The problem of protein folding, The 14th FAOBMB Congress, 2015年11月27–30日, Hyderabad (India)
- ⑧ Kuwajima, K. Formation of the chaperonin complex studied by 2D NMR and isothermal titration calorimetry, The 15th KIAS Conference on Protein Structure, 2015年9月17–19日, Seoul (Korea)
- ⑨ Kuwajima, K. Anti-tumor complexes formed by oleic acid and molten globule intermediates of proteins, The 1st Korean-Polish Conference on Protein Folding: Theoretical and Experimental Approaches, 2015年5月24–28日, Seoul (Korea)

- ⑩ Manavalan, B., Kuwajima, K. Lee, J. A contact-based topological quantity and its relationships with the folding kinetics of non-two-state and two-state proteins, The 7th Japan–Korea Seminar on Biomolecular Sciences: Experiments and Simulations, 2014年11月26–28日, Seoul (Korea)
- ⑪ Kuwajima, K. The problem of protein folding and its relationship to bioscience, The Second International Symposium on Protein Folding and Dynamics, 2014年11月5–7日, Bangalore (India)
- ⑫ Manavalan, B., Kuwajima, K. Lee, J. A contact-based topological quantity and its relationships with the folding kinetics of non-two-state and two-state proteins, The 14th KIAS Conference on Protein Structure, 2014年9月18–20日, Seoul (Korea)
- ⑬ Kuwajima, K. The problem of protein folding and its relationship to bioscience, The 4th Asia Pacific Protein Association Conference, 2014年5月17–20日, Jeju (Korea)
- ⑭ Kuwajima, K. The problem of protein folding and non-folding, International Symposium "Dynamic Ordering of Biomolecular Systems for Creation of Integrated Functions", 2014年1月11–12日, キャンパスプラザ京都 (京都府京都市)

〔図書〕 (計 1 件)

Kuwajima, K. (第12章を分担執筆), Springer Japan, Molecular Science of Fluctuations toward Biological Functions, 2016, 総ページ数 270 (245-270)

〔産業財産権〕

○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕

本研究で作成のデータベース

<http://lee.kias.re.kr/~protein/wiki/doku.php?id=pkhd:view>

6. 研究組織

(1)研究代表者

桑島 邦博 (KUWAJIMA, Kunihiro)

東京大学・大学院理学系研究科・名誉教授

研究者番号：70091444

(2)研究分担者

()

(3)連携研究者

()

(4)研究協力者

LEE, Jooyoung (李 柱榮)

韓国高等科学院・教授

Professor, Korea Institute for Advanced Study

MANAVALAN, Balachandran

韓国高等科学院・研究員

Research Fellow, Korea Institute for Advanced Study