

令和 2 年 6 月 8 日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2016～2019

課題番号：16K07314

研究課題名(和文) 速度データと天然構造の相関解析による蛋白質フォールディングの遷移状態の研究

研究課題名(英文) Studies on the transition state of protein folding by means of correlation analysis between the folding rate and the native structure

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) 3,700,000円

研究成果の概要(和文)：本研究では、(1) 温度補正を取り入れ、標準温度(25℃)におけるフォールディング速度定数をまとめたデータベースを構築し、インターネット上に公開した。(2) フォールディング反応の遷移状態における構造形成度を表す値と蛋白質の構造的特性との間の相関解析を行なった。型蛋白質16種の343箇所における値を調べ、構造特性と遷移状態における臨界構造との関係を明らかにした。(3) フォールディング速度定数とさまざまな構造特性との間の関係について調べた学術文献を調査し、見かけ上異なって見える、二状態と非二状態フォールディング反応が、本質的に同一の物理原理に従っていることを明らかにした。

研究成果の学術的意義や社会的意義

蛋白質フォールディング機構の研究は、蛋白質の間違ったフォールディングによって引き起こされる、ヒトの様々なフォールディング病(アルツハイマー病、白内障、クロイツフェルト・ヤコブ病、パーキンソン病、II型糖尿病など多数)などの原因解明や治療法の構築の応用研究に役立つと期待される、基盤的な研究である。本研究で構築されたデータベースは、蛋白質フォールディングの計算機シミュレーションや理論研究の研究者によって、広く活用されることが期待される。本研究で実施された、値やフォールディング速度と構造特性との間の相関解析は、本データベースを活用するための事例研究として用いることも出来る。

研究成果の概要(英文)：In this study, (1) a standardized protein folding database, in which the folding rate constant of all listed proteins are calculated at the standard temperature (25°C), has been constructed and opened on the internet homepage. (2) The correlations between the values, which represent the degree of structural formation in the transition state of folding, and different structure-based properties of proteins were analyzed, and the relationships between the structure-based properties and the critical structure in the transition state were discussed. (3) Academic literatures concerning the relationships between the folding rate constant and various structure-based properties were investigated, and it has been found that the two-state and the non-two-state folding reactions of proteins apparently seem very different, but are based on essentially the same physical principles.

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 遷移状態

様式 C-19, F-19-1, Z-19 (共通)

## 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) 蛋白質フォールディング速度と構造パラメータとの間の相関解析の研究は、近年、大変盛んになされており、"K-Fold"等の「蛋白質フォールディング速度予測ツール」も、現在8種以上がウェブ上に公開されている [1]。しかし、その基礎となっているフォールディング速度データは、個々の研究者により、様々な実験条件下で得られたものであり、特に測定温度は、低いところは5°C、高いところでは75°Cのデータがそのまま使われているのが現状である。

(3) また、現在まで、多くの球状蛋白質のフォールディング反応遷移状態の臨界構造に関して、変異誘発  $\Phi$  値解析が行われ、蛋白質の遷移状態の臨界構造に関する実験データが蓄積されている [2]。しかし、個々のアミノ酸残基の  $\Phi$  値と蛋白質の構造特性との間の関係は明らかではない。実験的に得られた  $\Phi$  値は誤差範囲内であり、蛋白質構造に関する情報はほとんど含まれていないとする極論もある [3]。

(4) 単ドメインの球状蛋白質には、天然状態 (N) と構造の完全にほどけた状態 (U) の二状態のみを示す、二状態蛋白質と、モルテン・グロビュール (MG) 状態などの中間状態 (I) を示す、非二状態蛋白質があるが、これら両者の関係は明確ではない。I状態は、蛋白質の会合などによりもたらされた、間違っただけの状態であり、真のフォールディング中間体ではないと主張する研究者もいる [4]。

## 2. 研究の目的

本研究では、以下の三点に焦点を当て、蛋白質フォールディングの分子機構を明らかにすることを目的とする。

(1) 温度補正を取り入れた、より質の高い、蛋白質フォールディング速度データベースを構築する。

(2) 速度論的実験から得られた  $\Phi$  値データと蛋白質構造パラメータとの間の相関解析を行い、 $\Phi$  値が立体構造をどの様に反映しているかを明らかにする。

(3) 二状態蛋白質と非二状態蛋白質のそれぞれについて、フォールディング速度と蛋白質の構造特性との相関を研究した学術論文を調査し、二状態蛋白質と非二状態蛋白質との関係を明らかにする。

## 3. 研究の方法

(1) 蛋白質フォールディング・データベースへの温度補正の導入：28年度までの基盤研究(C)「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」において、球状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した。このデータベースの中から、Eyring プロットが定量的に報告されている蛋白質のデータを選び、二状態蛋白質と非二状態蛋白質の、それぞれについて、フォールディング及びアンフォールディング速度定数の温度補正に必要な最適パラメータ値を求め、温度補正を取り入れた。

(2)  $\alpha$ ヘリックス型蛋白質の  $\Phi$  値と立体構造特性との間の相関解析：二状態転移を示す  $\alpha$ ヘリックス型タンパク質のフォールディング速度過程に関する学術論文を調査し、フォールディング反応遷移状態における、各アミノ酸残基の  $\Phi$  値とその残基の立体構造特性との相関を解析した。

(3) 二状態蛋白質と非二状態蛋白質との間の関係：1998年から2018年の間に公表された、蛋白質フォールディング速度と蛋白質構造特性との相関について研究した、学術論文30報を詳細に検討し、二状態蛋白質と非二状態蛋白質との間で、構造特性との相関において、どのような類似点と相違点があるかを検討した。

## 4. 研究成果

(1) 蛋白質フォールディング・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速度定数の対数 ( $\ln(k_f)$ ) とアンフォールディング速度の対数 ( $\ln(k_u)$ ) に温度補正を導入し、変性剤濃度ゼロで25°Cにおける、 $\ln(k_f)$  及び  $\ln(k_u)$  の値を掲載したフォールディング・データベースを構築した [5]。データベースのエントリー数は、二状態蛋白質89種、非二状態蛋白質52種の計141種であり、現在までの蛋白質フォールディング・データベースでは最大となる。データベースはインターネット (<http://lee.kias.re.kr/~bala/PFDB>) 上に公開されているので、広く、

蛋白質フォールディングのシミュレーションや理論研究に活用されることが期待される。

温度補正の妥当性は、以下のように確認した。14種の蛋白質については、 $k_f$  が 25°C と 25°C から離れた温度の両点において測定されており、また、11種の蛋白質については、 $k_u$  もこれら二つの温度で測定されている（これらの蛋白質は、上記の温度補正最適パラメータの推定に用いられた蛋白質とは異なる）。これらを、テストセットとして用い、上記で得られた最適パラメータを用いて、正しく温度補正できるかを検証した結果、 $\ln(k_f)$  と  $\ln(k_u)$  のいずれも、正しく温度補正されていることが分かった。

図1は、インターネット上に公開されている、データベース PFDB (Protein Folding DataBase) ホームページの画面である。各蛋白質について、最左列から順に以下の情報が掲載されている：(i) 蛋白質の簡略名；(ii) PDB コード；(iii) 構造クラス ( $\alpha$ ,  $\beta$ ,  $\alpha/\beta$ ,  $\alpha+\beta$ )；(iv) SCOP 分類に基づくフォールド；(v) PDB に登録されている構造の残基長；(vi) 実験に使われた蛋白質の残基長；(vii) 実験条件 (pH と温度)；(viii) フォールディングの型 (二状態 (2S) または非二状態 (N2S))；(ix) 論文中に報告されている  $\ln(k_f)$ ；(x) 温度補正後の  $\ln(k_f)$ ；(xi) 中間体形成速度定数の対数  $\ln(k_i)$  (非二状態蛋白質についてのみ)；(xii) 論文中に報告されている  $\ln(k_u)$ ；(xiii) 温度補正後の  $\ln(k_u)$ ；(xiv) Tanford  $\beta$  ( $\beta_T$ )；(xv) 参考として AG データセット [6, 7] の値；(xvi) 補記。

**PFDB: A standardized protein folding database with temperature correction**

		HOME		N2S		2S		DOWNLOAD DATASET		CONTACT										
No.	Protein short name	PDB	Class	Fold	Present dataset										AG dataset		Comments			
					$L_{pts}$	$L$	pH	Temp (°C)	Folding type	$\ln(k_f)$	$\ln(k_f)$ (25°C)	$\ln(k_u)$	$\ln(k_u)$ (25°C)	$\beta_T$	pH	Temp (°C)		Folding type	$\ln(k_f)$	
1	Apomyoglobin (Whako) [1]	1A6N	$\alpha$	Globin-like	151	153	6.2	5	N2S	1.1	4.5	NA	-3.8	5.5	0.72	-	-	-	-	
2	Pti1 [2]	1AU7 (103-160)	$\alpha$	DNARNA-binding 3-helical bundle	58	63	5.5	25.0	N2S	9.7		12.6	5.5		0.74	-	-	-	-	
3	4-helix bundle protein PRB [3]	1AUR (Chain B: 202-2115)	$\alpha$	Four-helical up-and-down bundle	94	95	7.5	10	N2S	5.4	6.7	NA	-5.2	-1.0	0.83	-	-	-	6.0	Both the AG and our datasets adopted the same reference [3]. The ACPro and our datasets reported the same $\ln(k_f)$ value, but the Garbuzovskiy dataset reported a different value.
4	IM7 [4]	1AY1 (1-86)	$\alpha$	HIV-1 gp41 fragments	86	86	7.0	10	N2S	5.7	6.9	8.0	-0.84	3.0	0.90	-	25	2S	7.2	The $\ln(k_f)$ value reported in the AG dataset was based on the 2S model [5]. However, the N2S nature of this protein is well established [4], so that our reported value is based on the N2S model.
5	Apomyoglobin (Horse) [6]	1DWR (1-152)	$\alpha$	Globin-like	152	153	6.0	26	N2S	2.9	2.9	5.3	NA		NA	NA	NA	NA	NA	The value of $\ln(k_f)$ was taken from the rate constant of the fast phase of the bi-phasic refolding kinetics reported.
6	Fingerted Homodomain [7]	1ENH	$\alpha$	DNARNA-binding 3-helical bundle	54	54	5.7	25.0	N2S	10.6		NA	7.6		0.83	-	-	-	-	
7	FF domain from human HYPA/FP1 [8]	1UZZ (3-71)	$\alpha$	3-Helical bundle	69	71	5.7	25	N2S	8.0		9.9	3.4		0.91	-	10	-	7.6	The same experimental group reported the $\ln(k_f)$ values at 25°C [8] and at 10°C [9]. The ACPro and our datasets reported the value at 25°C, but the Garbuzovskiy dataset reported the value at 10°C.

図1：データベース PFDB ホームページ (<http://lcc.kias.re.kr/~bala/PFDB>) の画面 (一部)

(2)  $\Phi$  値と立体構造特性との間の相関解析：16種の $\alpha$ ヘリックス型二状態蛋白質 (PDB codes: 1E41, 1IDY, 1IMQ, 1LMB, 1SS1, 1ST7, 1U4Q\_R15, 1U4Q\_R16, 1U4Q\_R17, 1VII, 1W4E, 1W4J, 1YYJ, 2LLH, 2WQG, 2WXC) の計 343 個の  $\Phi$  値データとを対象に、 $\Phi$  値と立体構造特性との相関を調べた。立体構造特性としては、(i) AGADIR に基づくヘリックス傾向値を各蛋白質について 0 と 1 の間に規格化した値 (NHP), (ii) 相対溶媒接触表面積 (RSA), (iii) アミノ酸配列に沿って長距離の残基間の原子-原子接触密度 (LCD), (iv) 中距離の残基間原子-原子接触密度 (MCD), (v) 短距離の残基間原子-原子接触密度 (SCD), について調べた。その結果、NHP と RSA に対しては、

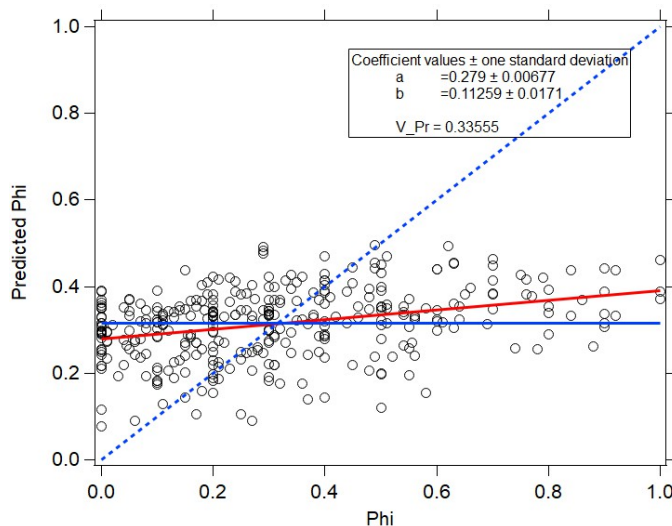


図2：多重回帰分析による理論的な  $\Phi$  値 (縦軸) と実験で得られた  $\Phi$  値 (横軸) との相関

相関係数  $r = 0.15 \sim 0.20$ , LCDに対しては  $r = -0.25$  となり, 5つ全ての構造特性の線形結合を仮定した多重回帰分析の理論値と  $\Phi$  値との相関は  $r = 0.34$  であった (図2)。また,  $t$  検定の  $t$  値は  $t = 6.57$  であり,  $\Phi$  値はこれらの構造特性 (特に, NHP, RSA, LCD) と, 大きくはないが, 有意に相関していることが分かった。したがって, これは, ヘリックス傾向値と溶媒接触度が大きくて, 一次構造上離れている残基と余り接触していない残基が, より大きな  $\Phi$  値を取る傾向にあることを示している。

(3) 二状態蛋白質と非二状態蛋白質との間の関係: 二状態蛋白質と非二状態蛋白質の双方について, フォールディング速度定数の対数 ( $\ln(k_f)$ ) と以下の構造特性との相関が調べられている: (i) Absolute contact order (ACO); (ii) chain length; (iii) long-range order; (iv) the number of sequence-distant native pairs; (v)  $n$ -order contact distance; (vi) geometric contact number; (vii) inter-residue interaction parameter; (viii) average topological information; (ix) logarithmic ACO; (x) 鎖長と構造クラスにより決まるアルゴリズムPREFURによる予測値; (xi) effective chain length (これらの構造特性の詳細については拙者の総説[8]を参照)。これらの構造特性は, いずれも, 主鎖構造によって決まり, 主鎖構造が, 一次構造上離れた残基間の相互作用によって決まる, 複雑な構造であればあるほど,  $\ln(k_f)$  は低下する。二状態蛋白質と非二状態蛋白質の, いずれも, これらの構造特性に対して類似した依存性を示すので, 二状態フォールディングと非二状態フォールディングの物理機構は本質的に同じであることが期待される。非二状態蛋白質の中間体形成速度定数の対数 ( $\ln(k_i)$ ) と二状態蛋白質の  $\ln(k_f)$  との比較から, 以下の結論がなされた。(i) MG状態などの中間体を示す階層的な非二状態フォールディングが蛋白質フォールディングの一般的な機構である。(ii) 二状態フォールディングは非二状態フォールディングが単純化した結果であるが, それには二つのシナリオが考えられる。つまり, シナリオ1は, 中間体 (I) から天然状態 (N) に至る過程の自由エネルギー障壁が低下したことにより, UからIを蓄積せず直接Nに至るようになったというシナリオであり, シナリオ2は, IからNへの障壁の低下はないが, Iの不安定化により, 見かけ上二状態転移を示すシナリオである。図3は, これらのシナリオに基づく, 蛋白質フォールディング反応の自由エネルギー・プロフィールを示す[8]。

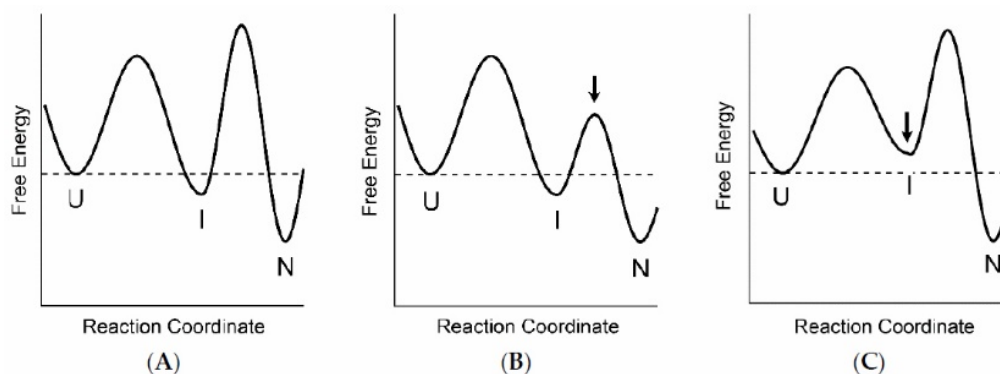


図3: 蛋白質フォールディングの自由エネルギー・プロフィール。(A) 階層的な非二状態フォールディング; (B) シナリオ1に基づく二状態フォールディング; (C) シナリオ2に基づく二状態フォールディング。

#### 参考文献

1. C. C. Chang, J. Song, B. T. Tey, R. N. Ramanan, Towards more accurate prediction of protein folding rates: A review of the existing Web-based bioinformatics approaches. *Brief. Bioinform.* **16**, 314-324 (2015).
2. A. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (W. H. Freeman Co., New York, 1998), pp. 631 (桑島他訳「タンパク質の構造と機構」医学出版).
3. A. N. Naganathan, V. Munoz, Insights into protein folding mechanisms from large scale analysis of mutational effects. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8611-8616 (2010).
4. B. A. Krantz, L. Mayne, J. Rumbley, S. W. Englander, T. R. Sosnick, Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *J Mol Biol* **324**, 359-371 (2002).
5. B. Manavalan, K. Kuwajima, J. Lee, PFDB: A standardized protein folding database with temperature correction. *Sci Rep* **9**, 1588 (2019).
6. A. S. Wagaman, A. Coburn, I. Brand-Thomas, B. Dash, S. S. Jaswal, A comprehensive database of verified experimental data on protein folding kinetics. *Protein Sci.* **23**, 1808-1812 (2014).
7. S. O. Garbuzynskiy, D. N. Ivankov, N. S. Bogatyreva, A. V. Finkelstein, Golden triangle for folding rates of globular proteins. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 147-150 (2013).
8. K. Kuwajima, The Molten Globule, and Two-State vs. Non-Two-State Folding of Globular Proteins. *Biomolecules* **10** (2020).

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 6件/うち国際共著 5件/うちオープンアクセス 3件）

1. 著者名 Shimizu Masahiro, Kajikawa Yukihiro, Kuwajima Kunihiro, Dobson Christopher M., Okamoto Yuko	4. 巻 87
2. 論文標題 Determination of the structural ensemble of the molten globule state of a protein by computer simulations	5. 発行年 2019年
3. 雑誌名 Proteins: Structure, Function, and Bioinformatics	6. 最初と最後の頁 635 ~ 645
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/prot.25688	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Manavalan Balachandran, Kuwajima Kunihiro, Lee Jooyoung	4. 巻 9
2. 論文標題 PFDB: A standardized protein folding database with temperature correction	5. 発行年 2019年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 1588- (9 pages)
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-018-36992-y	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Makabe Koki, Nakamura Takashi, Dhar Debanjan, Ikura Teikichi, Koide Shohei, Kuwajima Kunihiro	4. 巻 430
2. 論文標題 An Overlapping Region between the Two Terminal Folding Units of the Outer Surface Protein A (OspA) Controls Its Folding Behavior	5. 発行年 2018年
3. 雑誌名 Journal of Molecular Biology	6. 最初と最後の頁 1799 ~ 1813
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.jmb.2018.04.025	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Takenaka Toshio, Nakamura Takashi, Yanaka Saeko, Yagi-Utsumi Maho, Chandak Mahesh S., Takahashi Kazunobu, Paul Subhankar, Makabe Koki, Arai Munehito, Kato Koichi, Kuwajima Kunihiro	4. 巻 12
2. 論文標題 Formation of the chaperonin complex studied by 2D NMR spectroscopy	5. 発行年 2017年
3. 雑誌名 PLOS ONE	6. 最初と最後の頁 e0187022-
掲載論文のDOI (デジタルオブジェクト識別子) 10.1371/journal.pone.0187022	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Ku wajima Kunihiro	4. 巻 10
2. 論文標題 The Molten Globule, and Two-State vs. Non-Two-State Folding of Globular Proteins	5. 発行年 2020年
3. 雑誌名 Biomolecules	6. 最初と最後の頁 407- (17 pages)
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/biom10030407	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Cheng Qianyi, Joung InSuk, Lee Juyong, Ku wajima Kunihiro, Lee Jooyoung	4. 巻 15
2. 論文標題 Exploring the Folding Mechanism of Small Proteins GB1 and LB1	5. 発行年 2019年
3. 雑誌名 Journal of Chemical Theory and Computation	6. 最初と最後の頁 3432 ~ 3449
掲載論文のDOI (デジタルオブジェクト識別子) 10.1021/acs.jctc.8b01163	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

[学会発表] 計10件 (うち招待講演 8件 / うち国際学会 8件)

1. 発表者名 Kunihiro Ku wajima
2. 発表標題 Are there any correlations between mutational values and structural properties in all two-state proteins?
3. 学会等名 The 18th KIAS Conference on Protein Structure and Function (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 真壁幸樹, 桑島邦博
2. 発表標題 An overlapping region between the two terminal folding units of outer surface protein A (OspA) controls its folding behavior
3. 学会等名 The 17th KIAS Conference on Protein Structure and Function (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 桑島邦博
2. 発表標題 The Problem of Protein Folding
3. 学会等名 IGER International Symposium on Science of Molecular Assembly and Biomolecular Systems 2017 (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 Manavalan, B., Kuwajima, K. & Lee, J.
2. 発表標題 Relationship between the folding rate and structure-based parameters of globular proteins
3. 学会等名 第16回日本蛋白質科学会年会
4. 発表年 2016年

1. 発表者名 Kuwajima, K., Manavalan, B. & Lee, J.
2. 発表標題 The problem of protein folding
3. 学会等名 The 2nd Polish-Korean Conference on Protein Folding: Theoretical and Experimental Approaches (招待講演) (国際学会)
4. 発表年 2016年

1. 発表者名 Kuwajima, K., Manavalan, B. & Lee, J.
2. 発表標題 The problem of protein folding
3. 学会等名 The 16th KIAS Conference on Protein Structure (招待講演) (国際学会)
4. 発表年 2016年

1. 発表者名 Manavalan, B., Kuwajima, K. & Lee, J.
2. 発表標題 Relationship between the folding rate and structure-based parameters of globular proteins
3. 学会等名 第54回日本生物物理学会年会
4. 発表年 2016年

1. 発表者名 Makabe, K. & Kuwajima, K.
2. 発表標題 Folding mechanism of <i>Borrelia burgdorferi</i> Osp A, a dumbbell-shaped two-domain protein
3. 学会等名 The 3rd Korean-Polish Conference on Protein Folding: Theoretical and Experimental Approaches (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 Kuwajima, K.
2. 発表標題 The problem of protein folding
3. 学会等名 The 10th Toyota RIKEN International Workshop on Science of Life Phenomena (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Kuwajima, K, Yagi-Utsumi, M., Yanaka, S. & Kato, K.
2. 発表標題 The H/D-exchange kinetics of unfolded ubiquitin in 6 M GdmCl studied by the DMSO-quenched 2D NMR techniques
3. 学会等名 The 19th KIAS Conference on Protein Structure and Function (招待講演) (国際学会)
4. 発表年 2019年



〔図書〕 計0件

〔産業財産権〕

〔その他〕

球状蛋白質フォールディング速度のデータベース  
<http://lee.kias.re.kr/~bala/PFDB/>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	李 柱榮  (Lee Jooyoung)	韓国高等科学院・Computational Science・Professor	