



Large-scale analysis of macromolecular crowding effects on protein aggregation using a reconstituted cell-free translation system

Tatsuya Niwa^{1†}, Ryota Sugimoto^{1†}, Lisa Watanabe¹, Shugo Nakamura², Takuya Ueda³ and Hideki Taguchi^{1*}

¹ Department of Biomolecular Engineering, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan, ² Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan, ³ Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan

OPEN ACCESS

Edited by:

Salvador Ventura,
Universitat Autònoma de Barcelona,
Spain

Reviewed by:

Dong-Woo Lee,
Kyungpook National University, South
Korea

Pierre Genevieux,
Centre National de la Recherche
Scientifique, France

*Correspondence:

Hideki Taguchi
taguchi@bio.titech.ac.jp

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 26 June 2015

Accepted: 25 September 2015

Published: 08 October 2015

Citation:

Niwa T, Sugimoto R, Watanabe L,
Nakamura S, Ueda T and Taguchi H
(2015) Large-scale analysis
of macromolecular crowding effects
on protein aggregation using
a reconstituted cell-free translation
system. *Front. Microbiol.* 6:1113.
doi: 10.3389/fmicb.2015.01113

Proteins must fold into their native structures in the crowded cellular environment, to perform their functions. Although such macromolecular crowding has been considered to affect the folding properties of proteins, large-scale experimental data have so far been lacking. Here, we individually translated 142 *Escherichia coli* cytoplasmic proteins using a reconstituted cell-free translation system in the presence of macromolecular crowding reagents (MCRs), Ficoll 70 or dextran 70, and evaluated the aggregation propensities of 142 proteins. The results showed that the MCR effects varied depending on the proteins, although the degree of these effects was modest. Statistical analyses suggested that structural parameters were involved in the effects of the MCRs. Our dataset provides a valuable resource to understand protein folding and aggregation inside cells.

Keywords: protein aggregation, protein folding, cell-free translation system, macromolecular crowding, large-scale analysis

Most proteins must properly fold into their native tertiary structures, defined by the primary amino acid sequences, to perform their functions (Anfinsen, 1973; Dobson, 2003). However, protein folding is a highly complicated physicochemical process, and many proteins require the aid of molecular chaperones to fold into their correct structures, both *in vitro* and *in vivo* (Tyedmers et al., 2010; Hartl et al., 2011). Misfolded proteins often form protein aggregates, which leads to the loss of protein function and sometimes cause toxic effects in the cells (Tyedmers et al., 2010).

To clarify the principles of the protein aggregation and the properties associated with it, we conducted a comprehensive analysis of protein aggregation under the completely chaperone-free condition by using a *Escherichia coli* reconstituted cell-free translation system (Niwa et al., 2009). In this analysis, thousands of bacterial proteins were expressed separately, and their aggregation propensities were evaluated by using a centrifugation-based method. Statistical analyses revealed significant insights concerning protein aggregation (Niwa et al., 2009).

In the previous analysis, the aggregation propensity was evaluated in a diluted solution, in which the protein concentration was at most 1–2 mg/mL (Shimizu et al., 2001, 2005; Niwa et al., 2009).

However, the intracellular environment is much more crowded with macromolecules such as proteins and nucleic acids, and such an environment has been thought to affect the protein folding properties and the aggregation propensity (Zhou et al., 2008; Zhou, 2013). The effect of macromolecular crowding on protein folding and aggregation has been studied extensively, from both theoretical and experimental viewpoints, for decades (Zhou et al., 2008; Elcock, 2010; Gershenson and Gierasch, 2011; Zhou, 2013), and some studies suggested that the macromolecular crowding effects increase the intermolecular interactions mainly by its excluded volume effect, and hence facilitate the aggregation of some proteins (van den Berg et al., 1999; Munishkina et al., 2004). In contrast, other studies predicted that the crowding effects increase the stability of the native state and tend to bias proteins toward the native structure, although the effect on the stability was suggested to be modest (Cheung et al., 2005; Christiansen et al., 2010; Hong and Gierasch, 2010; Mittal and Best, 2010; Wang et al., 2010; Gershenson and Gierasch, 2011). However, in either case, these studies were limited to the experiments with a small number of model substrates or theoretical approaches.

To gain insight into the effects of macromolecular crowding on protein folding and aggregation and confirm these theories, we performed a large-scale analysis of the macromolecular crowding effects with a variety of proteins, by attempting the “*in vitro* proteome” approach reported previously (Niwa et al., 2009, 2012). By using a reconstituted cell-free translation system (Shimizu et al., 2001, 2005), we can easily evaluate the macromolecular crowding effects during the translation reaction for various kinds of proteins. In this analysis, we chose two macromolecular crowding reagents (MCRs), Ficoll 70, and dextran 70, because both two MCRs are hydrophilic polysaccharide and expected to have low interaction in specific amino acid side chains. Hence, the effects of these two MCRs can be thought to be mainly attributed to its excluded volume effect without any significant inhibition of expression reactions by the cell-free translation system (Zhou et al., 2008). In fact, we tried to use polyethylene glycol (PEG) 3350 as another MCR, but we could not evaluate its effect because the presence of PEG 3350 almost entirely abolished the protein expression by the cell-free translation system.

The method for the evaluation followed the previous comprehensive analyses of protein aggregation (see Materials and Methods). The measurement error of the solubility in the presence of the MCRs was about $\pm 10\%$, which is nearly equal to that in the absence of MCRs, as reported previously (Niwa et al., 2009).

We performed this experiment for 150 *E. coli* proteins under three conditions: no addition of MCRs, Ficoll-added, and dextran-added conditions. These 150 proteins were chosen at random among the proteins that were annotated as cytoplasmic proteins and whose aggregation propensities were evaluated in the previous comprehensive analysis (Niwa et al., 2009). Among the tested proteins, 142 proteins were quantified under the three conditions. All obtained data are shown in Supplementary Table S1 in the dataset, which is available at

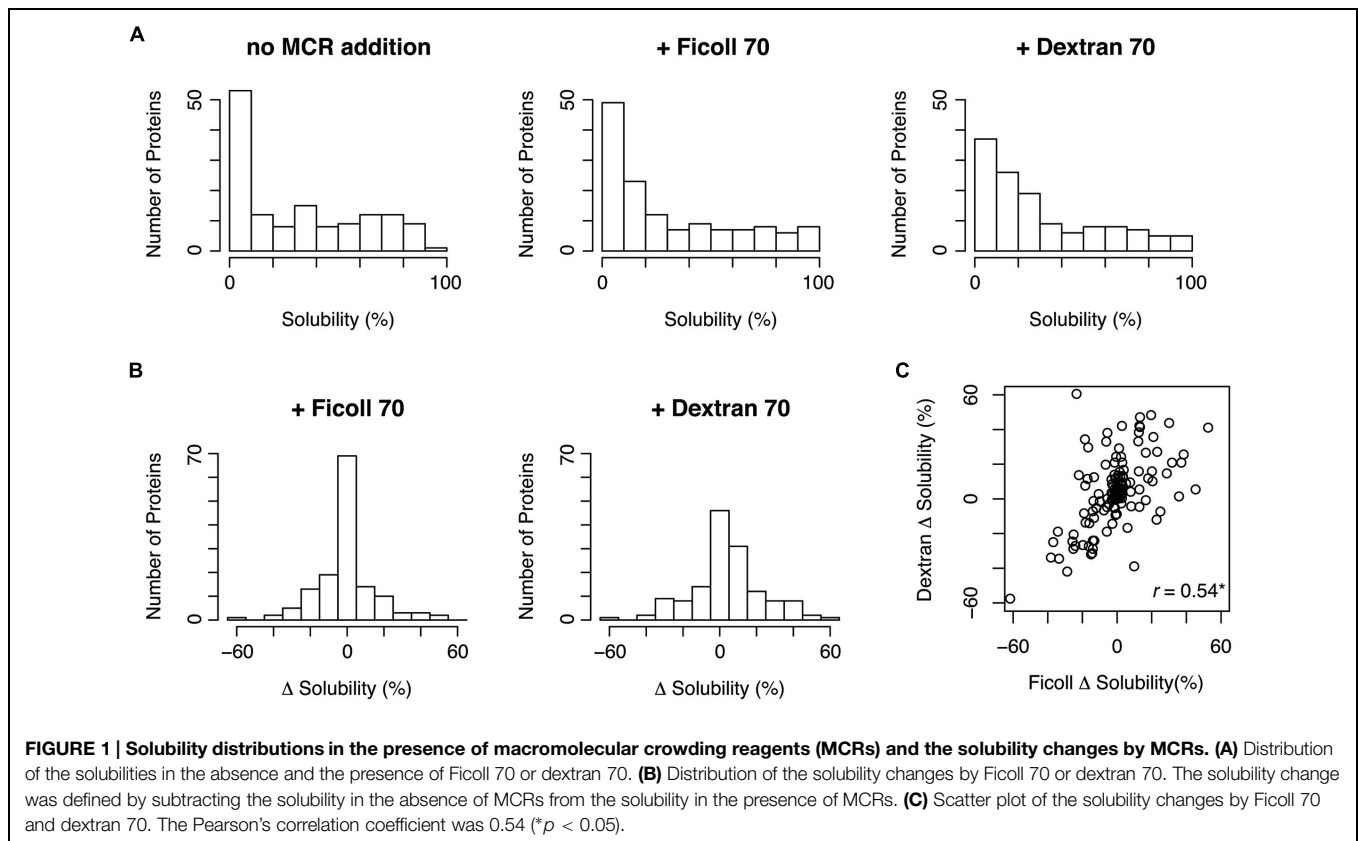
figshare repository¹. The distributions of the solubilities under the Ficoll- and dextran-added conditions were similar to that in the absence of MCRs (Figure 1A). This result suggested that Ficoll and dextran do not exert strong effects on the overall aggregation propensity. However, the distribution of the solubility changes under the dextran-added conditions was slightly biased toward a higher level, suggesting that dextran tends to act to prevent protein aggregation (Figure 1B). Furthermore, the solubility changes by Ficoll or dextran were widely distributed between -50 and $+50\%$, suggesting that the MCRs could act both positively and negatively on aggregate formation. In other words, the degree or direction of the effect of the MCRs on the aggregation propensity depends on the properties of the proteins. Moreover, the solubility changes by Ficoll correlated well with those by dextran (Figure 1C), indicating the similar effects of the two MCRs on the aggregation propensity.

As expected, the addition of both two MCRs did not cause drastic changes in the synthetic yield of the cell-free translation system. Furthermore, the change of the synthetic yield did not show a significant correlation with the solubility changes, suggesting that the effects of the changes in the synthetic yield on the solubility changes by the MCRs were small.

To determine which properties were related to the effects of the MCRs, we compared the solubility changes by MCRs and the physicochemical properties, such as molecular weight and isoelectric point. Although the molecular weight did not correlate with the solubility change, the solubility change by dextran positively correlated with the isoelectric point (Figure 2A). Moreover, the net charge, calculated by the number of charged amino acid residues, also correlated with the solubility change by dextran. These results suggested that dextran tends to act as an aggregation inhibitor for positively charged proteins. Concerning the properties derived from the primary sequence, we compared the solubility changes with the content ratios of the four amino acid groups, classified according to their properties. However, no obvious correlation was observed in the ratio of negatively charged (Asp and Glu), positively charged (Lys, Arg, and His), aromatic (Phe, Tyr, and Trp), or hydrophobic (Val, Leu, and Ile) amino acids.

Most proteins adopt a unique tertiary structure defined by the amino acid sequence, and a wide variety of structures exist. To compare the structural properties of proteins with the effects of the MCRs, we used the structural classification of proteins (SCOP) database (Murzin et al., 1995). As reported previously, the SCOP fold seems to have a formidable influence on the aggregation propensity, and proteins with specific folds have a strong tendency to form aggregates (Niwa et al., 2009). We then extracted the proteins with aggregation-prone folds, and investigated the distribution of their solubility changes by the MCRs. All annotations of the SCOP folds for tested proteins are listed in Supplementary Table S2 in the dataset¹, and the aggregation-prone folds in the previous report (Niwa et al., 2009) were as follows; c37: P-loop containing nucleoside triphosphate hydrolases, a4: DNA/RNA-binding 3-helical bundle, c1: TIM

¹<http://dx.doi.org/10.6084/m9.figshare.1495333>



β/α -barrel, c3: FAD/NAD(P)-binding domain, c55: Ribonuclease H-like motif, and c94: Periplasmic binding protein-like II. The histograms of the solubility changes for the proteins with the aggregation-prone folds revealed strong biases toward lower solubility, indicating that the MCRs tended to enhance the aggregate formation of these proteins (Figure 2B).

To compare further structural features, we constructed structural models for 41 proteins, by using a template-based modeling method. The structural templates used for the modeling are listed in Supplementary Table S3 in the dataset². Comparisons of the radius of gyration and the surface area of the amino acid main chains showed that both parameters positively correlated with the solubility change by dextran (Figures 2C,D). In addition, we compared the solubility change by the MCRs and the relative contact order, which is considered to be related to protein folding (Plaxco et al., 1998). Although the contact order negatively correlated with the solubility change by dextran, the correlation between them was not statistically significant.

The data obtained from this study suggested that the macromolecular crowding effect enhances aggregate formation for some proteins and prevents it for others. Previous experimental and theoretical studies suggested that the macromolecular crowding effects are often quite complicated, and particularly difficult to understand quantitatively (Zhou et al., 2008; Elcock, 2010; Zhou, 2013). The results obtained here

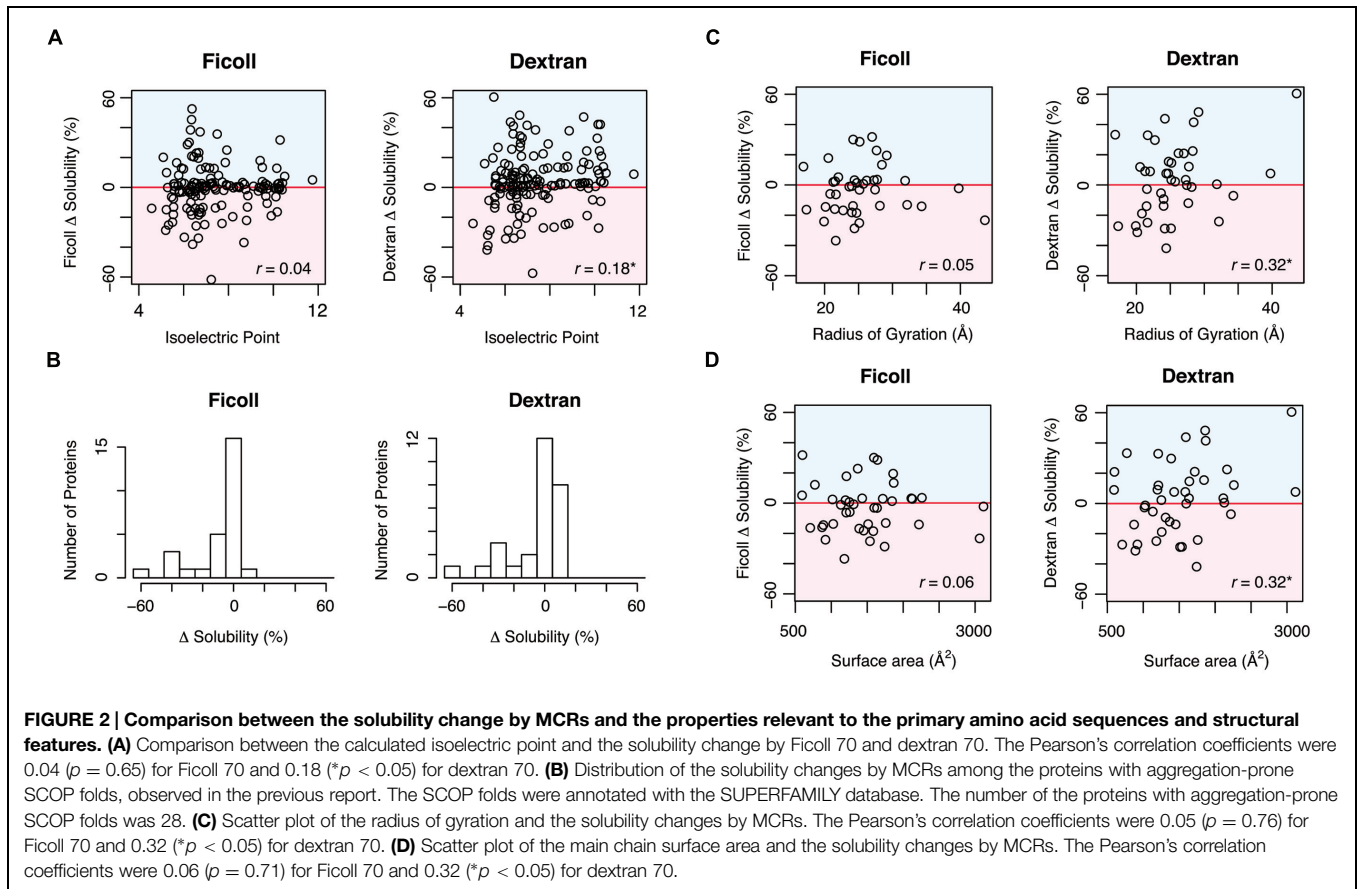
seem to reflect this complexity of the macromolecular crowding effects on protein folding and aggregation. In addition, some studies suggested that the effect of macromolecular crowding on protein folding is modest (Zhou et al., 2008; Mittal and Best, 2010; Zhou, 2013). Our data seem to be in agreement with these ideas, because the effects of macromolecular crowding were not strong, in comparison with the influences of molecular chaperones reported previously (Figure 1A; Niwa et al., 2012). Although our statistical analyses gave some insights for understanding the macromolecular crowding effects as described above, their influences on protein folding and aggregation are quite complicated and further detailed analysis are needed. Our dataset obtained from the “*in vitro* proteome” approach has great potential, as a valuable dataset that will contribute to further understanding of the effects of macromolecular crowding and protein folding inside cells.

MATERIALS AND METHODS

Method for the Evaluation of the Aggregation Propensity

The method for the evaluation of the aggregation propensity followed those used in previous comprehensive analysis (Niwa et al., 2009). The template DNA for expression by the cell-free translation system was amplified from an *E. coli* ORF library [ASKA library (Kitagawa et al., 2005; Riley et al., 2006)] by PCR,

²<http://dx.doi.org/10.6084/m9.figshare.1495333>



as described previously (Niwa et al., 2009). The transcription-translation-coupled expression was conducted by a reconstituted cell-free translation system [PURE system (Shimizu et al., 2001, 2005)] at 37°C for 1 h. For detection, L-[³⁵S]-methionine was added to the PURE system. Ficoll 70 (GE Healthcare) or dextran 70 (Sigma-Aldrich) was also included at the concentration of 80 mg/ml in the reaction, to evaluate the effects of MCRs. After the expression, an aliquot was withdrawn as the total fraction, and the remainder was centrifuged at 20,000 × *g* for 30 min. The total and supernatant fractions were separated by SDS-PAGE, and the band intensities were quantified by autoradiography (FLA7000 image analyzer and Multi Gauge software, Fujifilm). The ratio of the supernatant to the total protein was defined as the solubility, as referred to as the index of aggregation propensity.

Data Analysis

The molecular weight, amino acid content, and net charge were calculated from the amino acid sequences obtained from GenoBase³ (Kitagawa et al., 2005; Riley et al., 2006). Estimation of pI values was conducted with a web tool⁴ (Sillero and Maldonado, 2006). The SCOP (Murzin et al., 1995) classification was obtained from the dataset distributed

by GenoBase. The SCOP fold annotation in GenoBase was based on the SUPERFAMILY database (Madera et al., 2004). The SCOP folds annotated as aggregation-prone folds were as follows; c37: P-loop containing nucleoside triphosphate hydrolases, a4: DNA/RNA-binding 3-helical bundle, c1: TIM β/α-barrel, c3: FAD/NAD(P)-binding domain, c55: Ribonuclease H-like motif, and c94: Periplasmic binding protein-like II (Niwa et al., 2009). The modeled structures were obtained from the database by Zhang's group⁵ or modeled by the MODELER program⁶ (Eswar et al., 2006). Among the 41 modeled structures, 28 were selected from Zhang's database with the following criteria: >80% template identity, >80% template coverage, and >0.7 TM-score to the template. The remaining 13 structures were modeled by MODELER with the template PDBs determined by a PSI-BLAST search with the following criteria: >80% template identity and >80% template coverage. The radius of gyration and the relative contact order were calculated by using in-house developed scripts. Surface area was calculated with the NACCESS software⁷ (Hubbard and Thornton, 1993). All statistical tests were conducted with the R software⁸.

³<http://zhanglab.cmb.med.umich.edu/QUARK/ecoli2/>

⁶<https://salilab.org/modeller/>

⁷<http://www.bioinf.manchester.ac.uk/naccess/>

⁸<http://www.r-project.org/>

³<http://ecoli.naist.jp/GB/>

⁴<http://isoelectric.ovh.org/>

ACKNOWLEDGMENTS

This work was supported in part by JSPS KAKENHI Grant Numbers 22870010 and 25840045 (to T. N.), MEXT KAKENHI Grant Numbers 19058002 and 26116002 (to H. T.), and the Platform Project for Supporting in Drug Discovery and Life Science Research (Platform for Drug Discovery, Informatics, and Structural Life Science) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the

Japan Agency for Medical Research and Development (AMED) (to S. N. and T. U.)

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.01113>

REFERENCES

- Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223–230. doi: 10.1126/science.181.4096.223
- Cheung, M. S., Klimov, D., and Thirumalai, D. (2005). Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* 102, 4753–4758. doi: 10.1073/pnas.0409630102
- Christiansen, A., Wang, Q., Samiotakis, A., Cheung, M. S., and Wittung-Stafshede, P. (2010). Factors defining effects of macromolecular crowding on protein stability: an in vitro/in silico case study using cytochrome c. *Biochemistry* 49, 6519–6530. doi: 10.1021/bi100578x
- Dobson, C. M. (2003). Protein folding and misfolding. *Nature* 426, 884–890. doi: 10.1038/nature02261
- Elcock, A. H. (2010). Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr. Opin. Struct. Biol.* 20, 196–206. doi: 10.1016/j.sbi.2010.01.008
- Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., et al. (2006). Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics* Chap. 5, Unit 5.6. doi: 10.1002/0471250953.bi0506s15
- Gershenson, A., and Gierasch, L. M. (2011). Protein folding in the cell: challenges and progress. *Curr. Opin. Struct. Biol.* 21, 32–41. doi: 10.1016/j.sbi.2010.11.001
- Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature* 475, 324–332. doi: 10.1038/nature10317
- Hong, J., and Gierasch, L. M. (2010). Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J. Am. Chem. Soc.* 132, 10445–10452. doi: 10.1021/ja103166y
- Hubbard, S. J. T., and Thornton, J. M. (1993). “NACCESS; Computer Program”. London: Department of Biochemistry and Molecular Biology, University College London.
- Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., et al. (2005). Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* 12, 291–299. doi: 10.1093/dnares/dsi012
- Madera, M., Vogel, C., Kummerfeld, S. K., Chothia, C., and Gough, J. (2004). The SUPERFAMILY database in 2004: additions and improvements. *Nucleic Acids Res.* 32, D235–D239. doi: 10.1093/nar/gkh117
- Mittal, J., and Best, R. B. (2010). Dependence of protein folding stability and dynamics on the density and composition of macromolecular crowders. *Biophys. J.* 98, 315–320. doi: 10.1016/j.bpj.2009.10.009
- Munishkina, L. A., Cooper, E. M., Uversky, V. N., and Fink, A. L. (2004). The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J. Mol. Recognit.* 17, 456–464. doi: 10.1002/jmr.699
- Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536–540. doi: 10.1016/S0022-2836(05)80134-2
- Niwa, T., Kanamori, T., Ueda, T., and Taguchi, H. (2012). Global analysis of chaperone effects using a reconstituted cell-free translation system. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8937–8942. doi: 10.1073/pnas.1201380109
- Niwa, T., Ying, B. W., Saito, K., Jin, W., Takada, S., Ueda, T., et al. (2009). Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of *Escherichia coli* proteins. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4201–4206. doi: 10.1073/pnas.0811922106
- Plaxco, K. W., Simons, K. T., and Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* 277, 985–994. doi: 10.1006/jmbi.1998.1645
- Riley, M., Abe, T., Arnaud, M. B., Berlyn, M. K., Blattner, F. R., Chaudhuri, R. R., et al. (2006). *Escherichia coli* K-12: a cooperatively developed annotation snapshot–2005. *Nucleic Acids Res.* 34, 1–9. doi: 10.1093/nar/gkj405
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., et al. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755. doi: 10.1038/90802
- Shimizu, Y., Kanamori, T., and Ueda, T. (2005). Protein synthesis by pure translation systems. *Methods* 36, 299–304. doi: 10.1016/j.jymeth.2005.04.006
- Sillero, A., and Maldonado, A. (2006). Isoelectric point determination of proteins and other macromolecules: oscillating method. *Comput. Biol. Med.* 36, 157–166. doi: 10.1016/j.compbiomed.2004.09.006
- Tyedmers, J., Mogk, A., and Bukau, B. (2010). Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell Biol.* 11, 777–788. doi: 10.1038/nrm2993
- van den Berg, B., Ellis, R. J., and Dobson, C. M. (1999). Effects of macromolecular crowding on protein folding and aggregation. *EMBO J.* 18, 6927–6933. doi: 10.1093/emboj/18.24.6927
- Wang, Y., He, H., and Li, S. (2010). Effect of Ficoll 70 on thermal stability and structure of creatine kinase. *Biochemistry (Mosc.)* 75, 648–654. doi: 10.1134/S0006297910050160
- Zhou, H. X. (2013). Influence of crowded cellular environments on protein folding, binding, and oligomerization: biological consequences and potentials of atomistic modeling. *FEBS Lett.* 587, 1053–1061. doi: 10.1016/j.febslet.2013.01.064
- Zhou, H. X., Rivas, G., and Minton, A. P. (2008). Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* 37, 375–397. doi: 10.1146/annurev.biophys.37.032807.125817

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Niwa, Sugimoto, Watanabe, Nakamura, Ueda and Taguchi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.