Amphiphilic structured PEG derivatives suppressing protein thermal aggregation at extremely low molecular ratio

Adam M. Wawro,[a,b] Takahiro Muraoka,[a,c] Wijak Yospanya,[a] Mayu Kawame,[a] Yudai Nakagawa,[a] Mihoko Ui,[d] Paweł M. Antonik,[e] Peter B. Crowley,[e] Kouhei Tsumoto,[b,f,g] and Kazushi Kinbara*[a,h]

[a] Adam M. Wawro, Takahiro Muraoka, Wijak Yospanya, Mayu Kawame, Yudai Nakagawa, Kazushi Kinbara

School of Life Science and Technology, Tokyo Institute of Technology, Yokohama, Japan, E-mail: kkinbara@titech.ac.jp

[b] Adam M. Wawro, Kouhei Tsumoto

Department of Bioengineering, School of Engineering, The University of Tokyo, Tokyo, Japan

[c] Takahiro Muraoka

PRESTO, Japan Science and Technology Agency, Saitama, Japan

[d] Mihoko Ui

Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai, Japan

[e] Paweł M. Antonik, Peter B. Crowley

School of Chemistry, National University of Ireland Galway, Galway, Ireland

[f] Kouhei Tsumoto

Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Tokyo, Japan

[g] Kouhei Tsumoto

Institute of Medical Science, The University of Tokyo

Tokyo, Japan

[h] Kazushi Kinbara

Living Systems Materialogy (LiSM) Research Group, International Research Frontiers Initiative (IRFI), Tokyo Institute of Technology, Yokohama, Japan

Abstract: Amphiphilic structured PEG derivatives consisting of octa(ethylene glycol) chains connected with aromatic vertices inhibited the thermally-induced lysozyme aggregation even when present at below 0.1 mM concentration. This concentration range is close to a 1:1 molar ratio with the additive and lysozyme, and was completely inaccessible for previously reported stabilizers. The possible mechanisms of the stabilizing actions revealed that the PEG-based amphiphiles do not in fact prevent aggregation at the molecular level, as assumed before, but rather prevent macroscopic precipitation of the denatured protein molecules and enable dissolution of them to the native, folded state at ambient temperature. The stabilizers do not interact with properly folded native lysozyme and therefore do not affect its natural catalytic properties, indicating a potential for practical use in protein-based therapeutics.

Proteins realize numerous biochemical roles in an extremely precise manner thanks to their intricate molecular structures. Linear polypeptide chains in aqueous environment organize spontaneously into three-dimensional self-assemblies—spatially organized dynamic frameworks, exhibiting unique binding and catalytic properties. However, their metastable structure also makes them prone to various forms of chemical and physical degradation.

Protein aggregation is arguably the most common process leading to the loss of the protein original properties.[1] Protein aggregation occurs naturally in vivo, leading to development of various diseases,[2] and in vitro, interfering with manufacturing and storage of therapeutic macromolecules.[3]

Undoubtedly suppression of protein aggregation both in vivo and in vitro is recognized as a vital yet challenging goal for basic and applied biosciences. A glimpse into nature's solution to this issue, as usual, reveals an attractive perspective for scientists. In cells, a class of molecular chaperone proteins called chaperonins is known to accommodate unfolded proteins in their hydrophobic cavities to protect the proteins from aggregation and promote ATP-driven refolding into stable folded structures, using spatially organized hydrophobic residues inside its cavity.[4] A single molecule of molecular chaperon complex stabilizes and refold one molecule of a guest protein. In sharp contrast, stabilization of proteins by artificial means is usually achieved by addition of large stoichiometric excess kosmotropic salts,[5] saccharides and related polyols,[6] polar amino acids,[7–10] polymers,[11–17] nanogels[18] and ionic liquids.[19–21] A common characteristic of these additives is that typically their high concentration (>100 mM for small molecules or 10 mg mL^{-1} for polymers) is required to achieve satisfactory results. In such cases, however, the high concentration of additives changes the general properties of the solution, including viscosity, ionic strength, refractive index and chemical compatibility. Moreover, additives may have a negative impact on the catalytic or binding properties of the stabilized proteins.[13,22]

Our group has been looking for chaperone-like synthetic polymers, capable of stabilizing proteins at sub-millimolar concentrations. In this work we report a new synthetic well-defined polymer, exhibiting highly effective stabilization of lysozyme against thermally induced aggregation even at 0.1 mM concentration, i.e. 250-fold lower than other known compounds.[23,24] The compound is an amphiphilic macrocycle constructed from discrete octa(ethylene glycol). Its stabilizing effect was clearly observed at an unprecedentedly low 1:1 molar ratio with respect to the protein. Here we discuss our unexpected findings and provide an insight into the mechanism behind this effect.

The concept of the macrocyclic amphiphilic PEGs reported here was inspired by two molecules investigated by our group earlier.[23, 24] Both compounds showed unquestioned anti-aggregation properties, despite different structural features. The new compounds reported herein, **1a** and **1b**, consist of a large, oligo(ethylene glycol)-based macrocyclic ring with three evenly distributed hydrophobic sections. The use of monodisperse octa(ethylene glycol)[25,26] provided the general hydrophilic–hydrophobic balance of the molecule, enabled maintaining predictable, well-defined and reproducible molecular structure, and ensured non-random distribution of the hydrophobic units within the ring. The vertices are constructed from 3,5-dihydroxybenzoic acid, derivatized into methyl ester **1a** or *N*-(2-hydroxyethyl)amide **1b** in order to investigate the effect of structural fine-tuning on the properties of the scaffold (Scheme 1A). In addition to the two triangle-shaped compounds, we synthesized their non-cyclic counterparts, **2a** and **2b**, featuring two detached tetra(ethylene glycol) chains instead of one octa(ethylene glycol) (Scheme 1B). This allowed us to investigate the effect of the cyclic structure alone, while maintaining nearly unaltered molecular weight and functional groups.

Stabilization properties were evaluated using hen egg white lysozyme. It is a small cationic protein exhibiting rapid aggregation above its melting point and good stability below, thus being widely considered as a model protein for studying thermal aggregation and folding events. Thermal deactivation of lysozyme was performed using a classic heat shock method, i.e., heating its solution without stirring for 20 min at 90 °C, followed by rapid cooling. Under these conditions, lysozyme in a buffered solution typically loses all the initial enzymatic activity (residual activity <5%), due to irreversible aggregation, followed by precipitation. In the presence of a stabilizer, however, the enzyme may remain solubilized and retain a significant part of its original catalytic activity once cooled back to the ambient temperature.

In the initial stage we investigated the relationship between the concentration of the discrete amphiphilic PEGs **1a**, **1b**, **2a** and **2b** and the residual activity of the enzyme after the heat shock. Quite surprisingly, all compounds exhibited excellent ability to maintain the enzymatic activity of lysozyme at 1 mM concentration (Figure 1A).

Strikingly, the most effective compound **1b** showed a significant stabilizing effect even when present at concentration as low as 0.03 mM, nearly in 1:1 molar ratio with lysozyme (Figures 1A, S4 and S5). Slightly above, at 0.1 mM, the sample with **1b** reached the plateau of concentration–activity relationship, suggesting that further increasing the additive concentration above this value no longer enhances the protein recovery. It was clearly shown by performing the lysozyme activity assay where all additives were used at equal, 3 mM concentration (Figure 1B). The stabilizing effect of all four compounds at that concentration was nearly identical, despite their structural differences; the residual activity of lysozyme remained at the level of ca. 65%. The proposed interpretation of the physical meaning of the plateau is discussed in the supporting information (pages S10-S11).

As expected, the molecular structure of the additives significantly affected their properties. In terms of stabilizing efficiency, defined as the concentration of an additive required to achieve a comparable level of residual catalytic activity, clearly more hydrophilic structures of **1b** and **2b** were advantageous over their more hydrophobic counterparts. Corresponding methyl esters **1a** and **2a** required 3–10-fold higher concentration to exhibit the same effect. That said, even the least efficient compound 2a exhibited maximum enzyme recovery when present at 1 mM concentration. In addition, we observed that the cyclic structure of the stabilizers was also beneficial, although the difference between compounds **1b** and **2b** is not as large as between **1a** and **2a**. One should notice, however,

that at concentrations around 0.03 mM the stoichiometric ratio between the additive and the protein is nearly 1:1. Assuming the stabilizing effect is based on non-covalent interactions between the additive and the protein, the concentration corresponding to the equimolar ratio is plausibly the lower limit in general, since below that the number of the stabilizing molecules is simply insufficient to interact with the overrepresented protein molecules.

The observation of efficient protein stabilization achieved at nearly equimolar concentrations of amphiphilic PEGs has a great importance for understanding the aggregation inhibition mechanism. Previous works hypothesized that the stabilization of lysozyme is achieved by interaction on the folded or partially unfolded protein surface.[23] Considering the hydrodynamic radius of folded lysozyme in phosphate buffer (1.85 nm)[27] and its ellipsoid-like shape determined by X-ray diffraction, solvent-accessible area of lysozyme can be roughly estimated to be 35–45 nm^2. Moreover, unfolded lysozyme molecules would have even larger surface area. At the same time, the approximate maximum surface area that could be physically covered by a single **1b** molecule is just 2 nm^2, which is approximately 20-fold smaller. With this in mind, we found it very unlikely that the stabilizing mechanism of amphiphilic PEGs is as suggested previously.

We were curious what happens to the protein molecules during the heat treatment at the molecular level, and what is the role of the amphiphilic PEGs in this process. If PEGs interacted with a specific area of folded or partially unfolded lysozyme molecule, we would expect it to be seen by protein NMR experiments. We therefore investigated it by 1H–15N HSQC NMR of a ^{15}N-labeled lysozyme solution in the presence or absence of **1b** (Figures 2 and S3).[28]

At 40 °C, well below the melting point of the protein (73 °C), no significant differences between the samples without (A) and with 10 eq. of **1b** (B) were observed; lysozyme was in its native state and no changes in the signal patterns suggesting binding of **1b** were noticed. Raising the temperature to 75 °C resulted in spectral broadening, consistent with the formation of high molecular weight assemblies, or aggregates. Interestingly, the same effect was observed equally in both solutions, regardless of the presence of **1b**. After cooling back to 40 °C, however, the spectra looked completely different. In the sample with no additive the signals hardly recovered, suggesting the irreversible aggregation of lysozyme and/or precipitation. In contrast, in the presence of **1b** the signals reappeared at their original positions, meaning that the protein molecules were mostly solubilized and refolded properly into their native structures. It is important to link the NMR observations with the sample appearance. The lysozyme only sample (Figure 2A) contained a ‘heavy’ precipitate after heating–cooling cycle, in contrast the sample (Figure 2B) in which **1b** was present was practically transparent. This observation and the high temperature NMR spectrum suggest that soluble protein‑**1b** mixed aggregates were reversibly formed at 75 °C.

The results of this experiment are in good agreement with the residual activity assay shown earlier. It is somewhat surprising though that 1b does not in fact prevent aggregation by stabilizing unfolded protein molecules; apparently it prevents irreversible precipitation of macroscopic aggregates but does not interfere with reversible formation of soluble aggregates, causing NMR signals to broaden significantly.

This unexpected hypothesis was further tested by investigation of dynamic light scattering (DLS) of a buffered solution of lysozyme and **1b**. Usually DLS data are presented as a distribution of particle sizes, i.e., as a plot of scattering intensity against scattering particle diameter. However, the particle

distribution is a highly processed data, obtained through fitting the raw input to a model and requires particles of relatively uniform sizes for reliable interpretation. We analyzed the correlation function instead, which allows for qualitative, but more certain conclusions (Figure 3). The correlation function provides very approximate information about the hydrodynamic diameter of particles, but more importantly it reveals when the scattering particles in solution form large aggregates, resulting in easily distinguished contributions from non-Brownian motions. At 20 °C lysozyme exists in solution as a monomer, a small particle diffusing quickly, yielding rapidly declining correlation function (Figure 3A). After heating the solution to 70 °C, the correlation function changes drastically; its value partially exceeds the theoretical limit of 1 and the function loses monotonicity, particularly in the long-correlation-time region. This suggests formation of large aggregates. After cooling back to 20 °C, the slope partially recovers its original shape, but the correlation coefficient still does not reach value of 0 and a part of the function is not monotonic, indicating presence of some remaining large particles. In the samples where **1b** was present (Figure 3B and 3C), the correlation curves at the initial state and at 70 °C are almost identical with the sample containing no additive (A), while after cooling the sample back to 20 °C a difference can be noticed, especially when the additive is present at higher concentration (Figure 3C).

The correlation coefficient value reaches zero quickly and remains monotonic, evidently suggesting very low concentration of large aggregates. The DLS experiment confirmed the presence of lysozyme aggregates at elevated temperatures in all samples, and together with the results of ^{1}H–^{15}N HSQC NMR allowed us to conclude that very likely **1b** enables formation of soluble lysozyme aggregates[29,30] in a reversible manner and inhibits irreversible precipitation. The chaperon-like properties of the investigated compounds, in particular compound 1b, are likely caused by certain interactions between molecules of the protein and the stabilizer. However, no interactions of 1b with folded protein, leading potentially to resonance peak shift, were revealed by the NMR experiment. Indeed, it would be highly desirable that the stabilizer interacts selectively with the unfolded protein, which have no catalytic properties anyway, but do not bind the properly folded native enzyme, as such interactions at ambient temperature could potentially negatively affect the enzymatic activity. We further investigated these hypothetical interactions by lysozyme unfolding in the presence or absence of 1b using differential scanning calorimetry (Figure 4A).

Addition of up to 0.5 mM of **1b** did not change lysozyme melting point and had only a minor effect on the thermal unfolding enthalpy ΔH. This suggests that there is no significant lysozyme–**1b** interaction at low temperatures. This conclusion was also supported by isothermal titration calorimetry analysis of lysozyme solution titrated with **1b** (Figure 4B). The heat flow upon addition of PEG was negligible and no significant exothermic peaks, typical for protein–small molecule binding, were observed. Another supporting observation was that the presence of 1b did not affect the enzymatic activity of lysozyme at ambient temperature, without thermal treatment.

This work has two major logical parts. In the first one, we designed and synthesized a series of discrete, octa(ethylene glycol)-based synthetic chaperons. We evaluated their properties using a lysozyme thermal aggregation model and discovered that one of the compounds can inhibit the lysozyme aggregation even when present at below 0.1 mM concentration. This concentration range was completely inaccessible for previously reported stabilizers, which were required at least 250-fold higher concentrations. We also showed that the overall hydrophobic–hydrophilic balance and molecular topology have an impact on the anti-aggregation capability of the amphiphiles. In the

second part we investigated the possible mechanisms of the stabilizing actions and revealed that the PEG-based amphiphiles do not in fact prevent aggregation at the molecular level, as assumed before, but rather prevent macroscopic precipitation and enable dissolution of the protein molecules to the native, folded state. We also demonstrated that the stabilizers do not interact with properly folded native lysozyme and therefore do not affect its natural catalytic properties.

Acknowledgement: The authors thank Suzukakedai Materials Analysis Division, Open Facility, Tokyo Institute of Technology, for ESI-TOF mass spectrometry measurements. This work was supported by MEXT Grant-in-Aid for Scientific Research on Innovative Areas “Molecular Engine” (JP18H05418 and JP18H05419 to K.K.) and JSPS Grant-in-Aid for Scientific Research B (JP19H02831 to K.K.).

References

[1] J. S. Lawrence, C. R. Middaugh, in Aggregation of Therapeutic Proteins, W. Wang, C. J. Roberts, Eds. (Wiley, 2010), pp. 1–62.

[2] Y. S. Eisele, C. Monteiro, C. Fearns, S. E. Encalada, R. L. Wiseman, E. T. Powers, J. W. Kelly, *Nat. Rev. Drug Discov.* **2015**, *14*, 759–780.

[3] M. Lebendiker, T. Danieli, *FEBS Lett.* **2014**, *588*, 236–246.

[4] H. Saibil, *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 630–642.

[5] J. W. Bye, R. J. Falconer, *Protein Sci.* **2013**, *22*, 1563–1570 (2013).

[6] C. Avanti, V. Saluja, E. L. P. Van Streun, H. W. Frijlink, W. L. J. Hinrichs, *PLoS One.* **2014**, *9*, 2–7 (2014).

[7] K. Shiraki, M. Kudou, S. Fujiwara, T. Imanaka, M. Takagi, *J. Biochem.* **2002**, *132*, 591–595.

[8] K. Shiraki, M. Kudou, S. Nishikori, H. Kitagawa, T. Imanaka, M. Takagi, Eur. *J. Biochem.* **2004**, *271*, 3242–3247.

[9] K. Shiraki, M. Kudou, R. Sakamoto, I. Yanagihara, M. Takagi, *Biotechnol. Prog.* **2005**, *21*, 640–643.

[10] R. Chang, M. Gruebele, D. E. Leckband, *Biomacromol*. **2022**, *23*, 4063−4073.

[11] R. C. Lee, F. Despa, L. Guo, P. Betala, A. Kuo, P. Thiyagarajan, *Ann. Biomed. Eng.* **2006**, *34*, 1190–1200.

[12] D. Mustafi, C. M. Smith, M. W. Makinen, R. C. Lee, *Biochim. Biophys. Acta - Gen. Subj.* **2008**, *1780*, 7–15.

[13] S. Ganguli, K. Yoshimoto, S. Tomita, H. Sakuma, T. Matsuoka, K. Shiraki, Y. Nagasaki, *J. Am. Chem. Soc.* **2009**, *131*, 6549–6553.

[14] M. J. Poellmann, T. R. Sosnick, S. C. Meredith, R. C. Lee, *Macromol. Biosci.*, **2016**, 1–9.

[15] N. Kameta, T. Matsuzawa, K. Yaoi, M. Masuda, *RSC Adv.* **2016**, *6*, 36744–36750.

[16] R. Rajan, N. Kumar, K. Matsumura, *Biomacromol.* **2022**, *23*, 487−496.

[17] A. Debas, K. Matsumura, R. Rajan, *Mol. Syst. Des. Eng.*, **2022**, *7*, 1327–1335.

[18] T. Nishimura, K. Akiyoshi, *Bioconjugate Chem.*, **2020**, *31*, 1259–1267.

[19] H. Zhao, *J. Chem. Technol. Biot.* **2016**, *91*, 25–50.

[20] T. A. Shmool, L. K. Martin, R. P. Matthews, J. P. Hallett, *JACS Au* **2022**, *2*, 2068−2080.

[21] P. Kushwaha, N. P. Prabhu, *New J. Chem.*, **2022**, *46*, 11082–11094.

[22] W. M. Neville, H. Eyring, *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 2417–2419.

[23] T. Muraoka, K. Adachi, M. Ui, S. Kawasaki, N. Sadhukhan, H. Obara, H. Tochio, M. Shirakawa, K. Kinbara, *Angew. Chem. Int. Ed.* **2013**, *52*, 2430–2434.

[24] N. Sadhukhan, T. Muraoka, M. Ui, S. Nagatoishi, K. Tsumoto, K. Kinbara, *Chem. Commun.* **2015**, *51*, 8457–8460.

[25] A. M. Wawro, T. Muraoka, K. Kinbara, *Polym. Chem.* **2016**, *7*, 2389–2394.

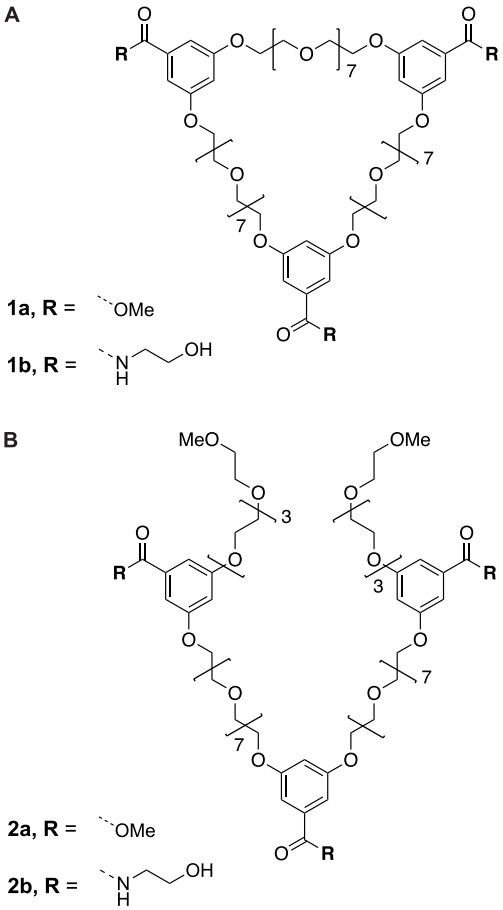
[26] A. M. Wawro, T. Muraoka, M. Kato, K. Kinbara, *Org. Chem. Front.* **2016**, *3*, 1524–1534 (2016).

[27] A. S. Parmar, M. Muschol, *Biophys. J.* **2009**, *97*, 590–598.

[28] Because the recombinant ^{15}N-labeled lysozyme was found to be more prone to thermal aggregation, a slightly higher molar ratio of 1b was used in this experiment (see SI for details).

[29] N. Golub, A. Meremyanin, K. Markossian, T. Eronina, N. Chebotareva, R. Asryants, V. Muronetsb, B. Kurganov, *FEBS Lett.* **2007**, *581*, 4223–4227.

[30] K. A. Markossian, I. K. Yudin, B. Kurganov, *Int. J. Mol. Sci.* **2009**, *10*, 1314–1345.



Scheme 1. Molecular structures of amphiphiles **1a**, **1b**, **2a** and **2b**.

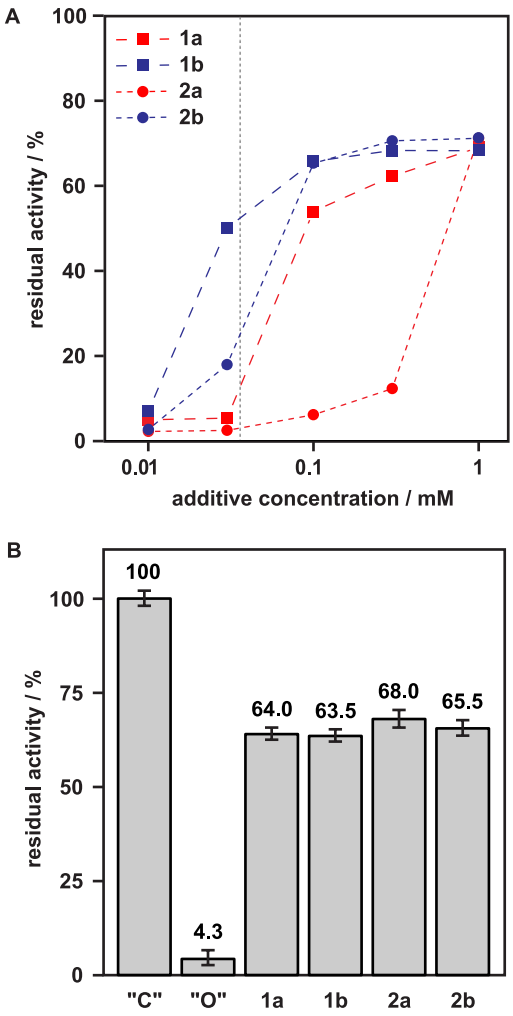


Figure 1. (A) Residual activity of thermally deactivated lysozyme (20 min, 90 °C) depending on the concentration of additives. Vertical dotted line marks the concentration of lysozyme (0.035 mM). (B) Residual activity of thermally deactivated lysozyme in the presence of 3 mM additives, no additive (“O”) and non-heated control (“C”). Each measurement was repeated three times.

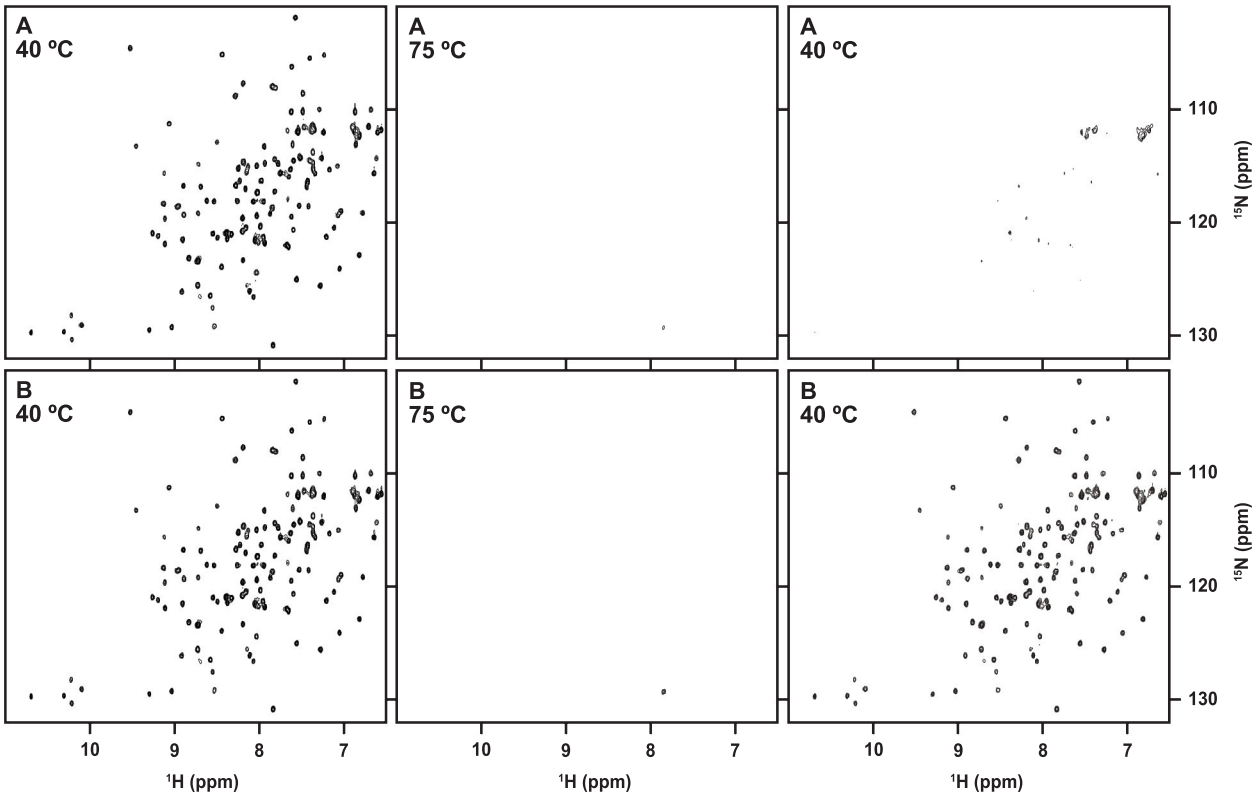


Figure 2. ^{1}H–^{15}N HSQC spectral regions of lysozyme in the presence of 0 (A) or 10 eq. (B) of compound **1b** at 40 °C (left), heated to 75 °C (middle) and then cooled back to 40 °C (right).

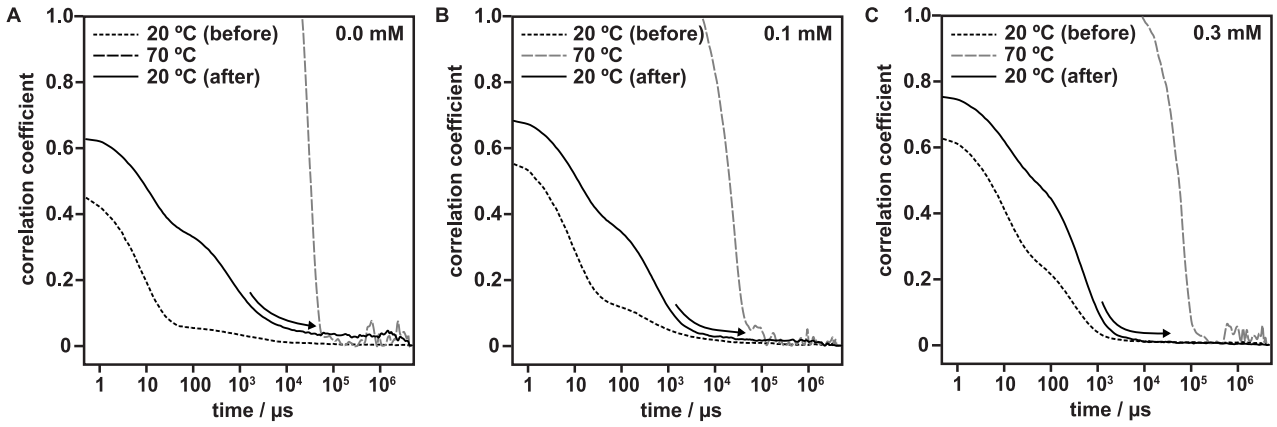


Figure 3. Correlation function curves obtained from DLS analysis of buffered 0.5 mg mL^{-1} (0.035 mM) solutions of lysozyme alone (A) or in the presence of compound **1b** at 0.1 mM (B) or 0.3 mM (C).

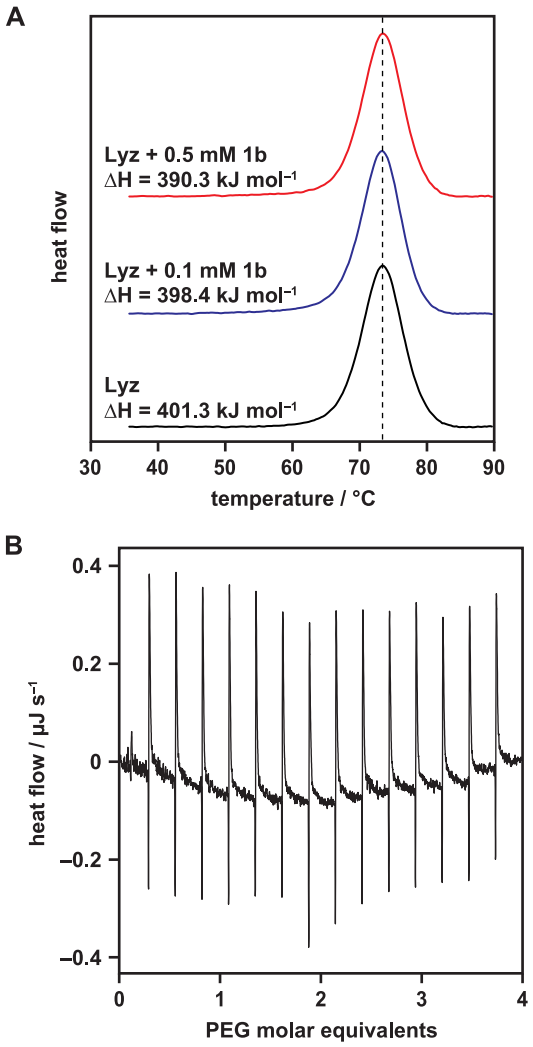


Figure 4. (A) Heat flow curves of lysozyme (0.035 mM) unfolding in the absence or presence of **1b**. Baselines were corrected and transition enthalpy was calculated using non-two-state model. (B) Heat flow curve of lysozyme solution (0.1 mM) titrated with **1b** (4.0 mM, 2 μL each injection) For details, see: Supporting Information, page S3.