


 Cite this: *RSC Adv.*, 2022, **12**, 26658

# Cysteine-based protein folding modulators for trapping intermediates and misfolded forms†

 Hayato Nishino,<sup>‡a</sup> Mai Kitamura,<sup>‡a</sup> Shunsuke Okada,<sup>a</sup> Ryosuke Miyake,<sup>Ⓜb</sup> Masaki Okumura<sup>c</sup> and Takahiro Muraoka<sup>Ⓜb</sup>\*<sup>ade</sup>

Folding is a key process to form functional conformations of proteins. Folding *via* on-pathway intermediates leads to the formation of native structures, while folding through off-pathways affords non-native and disease-causing forms. Trapping folding intermediates and misfolded forms is important for investigating folding mechanisms and disease-related biological properties of the misfolded proteins. We developed cysteine-containing dipeptides conjugated with amino acids possessing mono- and diamino-groups. In oxidative protein folding involving disulfide-bond formation, the addition of cysteine and oxidized glutathione readily promoted the folding to afford native forms. In contrast, despite the acceleration of disulfide-bond formation, non-native isomers formed in significantly increased yields upon the addition of the dipeptides. This study provides a molecular design of cysteine-based protein-folding modulators that afford proteins adopting non-native conformations through intermolecular disulfide-bond formation. Because of the intrinsic reversibility of the disulfide bonds upon redox reactions, the disulfide bond-based approach demonstrated here is expected to lead to the development of reversible methodologies for trapping transient and misfolded forms by intermolecular disulfide bond formation and restarting the folding processes of the trapped forms by subsequent cleavage of the intermolecular disulfide bonds.

 Received 30th June 2022  
 Accepted 12th September 2022

DOI: 10.1039/d2ra04044a

[rsc.li/rsc-advances](https://rsc.li/rsc-advances)

## Introduction

The folding process of a polypeptide chain influences the fate of the protein.<sup>1,2</sup> Folding *via* stepwise conversion among on-pathway intermediates allows for the formation of the native conformation of the protein with its biological activity.<sup>3–9</sup> Meanwhile, an off-pathway folding process forming non-native intermediates affords misfolded and even pathological isomers of the protein.<sup>10–12</sup> For instance, side effects of pharmacological administration for psychiatric disorders induce aggregation-prone proinsulin containing non-native disulfide bonds, leading to type-II diabetes.<sup>13</sup> To investigate the biological as well

as disease-related properties of a protein, it is necessary to develop methodologies of not only folding promotion to the native form but also folding-intermediates trapping.<sup>14</sup> In the oxidative protein folding reactions involving disulfide-bond formation, reactions of cysteine residues are one of the major factors that regulate the folding process.<sup>15,16</sup> Namely, disulfide-bond formation between the native cysteine-residue pairs allows for folding of the polypeptide chain to the native conformation, while folding through disulfide bonding between non-native cysteine-residue pairs results in the formation of misfolded isomers.

For the regulation of protein folding, chemical approaches are effective.<sup>17,18</sup> In the endoplasmic reticulum, enzymes such as protein disulfide isomerase promote disulfide-bond shuffling by cleavage and re-formation of disulfide bonds to encourage the structural conversion of substrate proteins into thermodynamically stable conformations, *i.e.*, the native structure.<sup>19</sup> Inspired by the enzymatic process, a variety of compounds such as thiols with increased acidity, glutathione derivatives, and selenols have been developed for promotion of protein folding.<sup>20–25</sup> On the other hand, methodologies that can trap folding intermediates and misfolded forms are limited to genetic mutations<sup>26–28</sup> and chemical treatments using iodoacetic acid and maleimide-appending compounds.<sup>29,30</sup> Here, we demonstrate that newly developed cysteine-based dipeptides, Cys-Dap and Cys-Tamp (Fig. 1), facilitated the oxidation of proteins and afforded

<sup>a</sup>Department of Applied Chemistry, Graduate School of Engineering, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, Japan. E-mail: muraoka@go.tuat.ac.jp

<sup>b</sup>Department of Chemistry and Biochemistry, Graduate School of Humanities and Sciences, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan

<sup>c</sup>Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, 6-3 Aramaki-Aza-Aoba, Aoba-ku, Sendai 980-8578, Japan

<sup>d</sup>Institute of Global Innovation Research, Tokyo University of Agriculture and Technology, 3-8-1 Harumi-cho, Fuchu, Tokyo 183-8538, Japan

<sup>e</sup>Kanagawa Institute of Industrial Science and Technology (KISTEC), Kanagawa, 243-0435, Japan

† Electronic supplementary information (ESI) available: Materials, instrumentation, synthesis, SDS-PAGE and reverse-phase HPLC analyses. See <https://doi.org/10.1039/d2ra04044a>

‡ These authors contributed equally to this work.



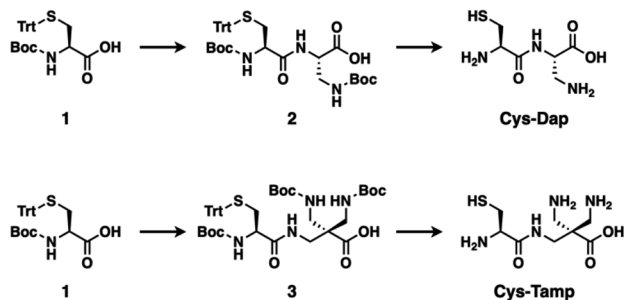


Fig. 1 Synthetic schemes of Cys-Dap and Cys-Tamp.

isomers with non-native conformations preferentially through intermolecular disulfide-bond formation.

## Results and discussion

### Synthesis of cysteine-containing dipeptides

The cysteine-containing dipeptides, Cys-Dap and Cys-Tamp, were synthesized from cysteine protected with Boc and trityl groups in a straightforward approach (Fig. 1). Boc- and trityl-protected cysteine **1** was coupled with Boc-protected (*S*)-diaminopropionic acid (Dap) to afford **2**, and subsequent deprotection of Boc and trityl groups provided Cys-Dap. Cys-Tamp was synthesized from compound **3** that was produced by coupling **1** and 3-amino-2,2-bis[[(*tert*-butoxycarbonyl)amino]methyl] propanoic acid (Boc-protected Tamp).<sup>31,32</sup>

### Effects of cysteine-containing dipeptides on disulfide-bond formation of RNase A

To investigate the effects of the cysteine-containing dipeptides on the regulation of oxidative protein folding, folding reactions of ribonuclease (RNase) A were analyzed. RNase A forms four disulfide bonds in its native form (N), and the reduced form (R) folds into N by consecutive disulfide-bond formation through 1SS, 2SS, and 3SS intermediates containing one, two, and three disulfide bonds, respectively (Fig. 2a).<sup>5,33–35</sup> As a side reaction, 4SS<sub>U</sub>, non-native forms with four disulfide bonds, can be produced from the 3SS intermediate.

Disulfide-bond formation of reduced RNase A was conducted by oxidized glutathione (GSSG) as an oxidant, and the reaction was monitored by electrophoresis. Treatment of the reaction mixture with malPEG-2000 (average  $M_n = 2000$ ), which reacts with free thiol groups on RNase A, allows for the separation of the folding intermediates by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). In the presence of GSSG, reduced RNase A transformed into 1SS, 2SS, and 3SS intermediates over the incubation time, and a band corresponding to 4SS (N + 4SS<sub>U</sub>) appeared after 30 min incubation (Fig. 3a). In the presence of the cysteine and GSSG mixture, a band corresponding to 4SS was observed after 10 min incubation (Fig. 3b). Interestingly, the disulfide-bond formation proceeded even faster in the presence of the cysteine-containing dipeptides and GSSG than in the other conditions. Particularly, in the presence of Cys-Tamp and GSSG, the band of 4SS appeared after 5 min

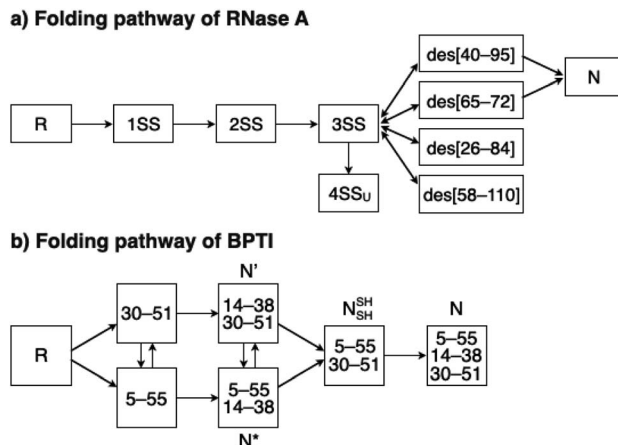


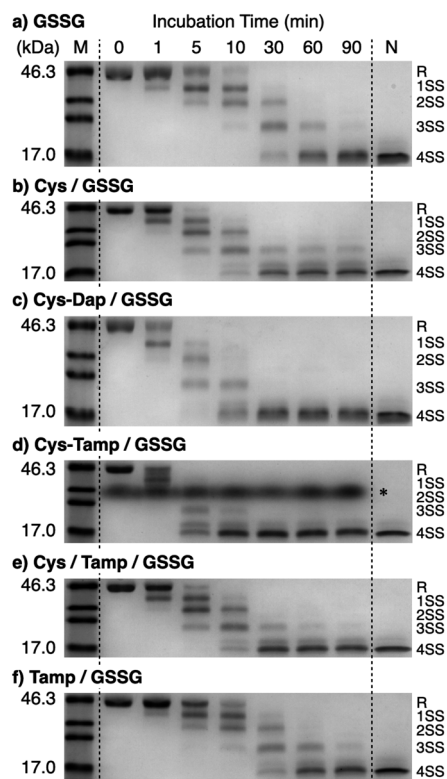
Fig. 2 Folding pathways of (a) RNase A and (b) BPTI. R and N denote reduced and native forms, respectively. In (a), 1SS, 2SS, 3SS: folding intermediates containing one, two, and three disulfide bonds, respectively; 4SS<sub>U</sub>: non-native forms with four disulfide bonds; des[M–N]: a folding intermediate with three native disulfide-bond pairings lacking the Cys<sub>M</sub>–Cys<sub>N</sub> disulfide bond. In (b) N', N\*, N<sup>SH</sup>: folding intermediates with two disulfide bonds at Cys14–Cys38 and Cys30–Cys51, Cys5–Cys55 and Cys14–Cys38, and Cys5–Cys55 and Cys30–Cys51, respectively.

incubation (Fig. 3d). A quantitative analysis of the fully oxidized RNase A formation based on the SDS-PAGE analyses shows that nearly 90% conversion from R was completed after 10 min incubation in the condition of Cys-Tamp and GSSG (Fig. 4a, blue squares). The condition of cysteine and GSSG required 90 min incubation for 90% conversion, while the oxidation proceeded more slowly in the presence of GSSG alone (Fig. 4a, black circles and orange diamonds). Importantly, the mixtures of cysteine, Tamp, and GSSG or Tamp and GSSG did not accelerate the oxidation compared to the Cys-Tamp and GSSG mixture (Fig. 3e and f). Therefore, the covalent conjugation of Tamp largely influenced the properties of the cysteine thiol group, enabling rapid disulfide-bond formation.

### Effects of cysteine-containing dipeptides on folding of RNase A to native and misfolded forms

The efficiency of oxidative refolding of RNase A to its native form was investigated by the recovery of enzymatic activity (Fig. 4b). In the presence of GSSG, RNase A showed an increase in the enzymatic activity during the initial 120 min incubation to reach a plateau of 12% recovery. The cysteine/GSSG system showed an enhanced recovery of the RNase A activity to 62% after 360 min incubation. Interestingly, the Cys-Dap/GSSG system resulted in a significant drop in the recovery yield to 12%, and the Cys-Tamp/GSSG system hardly showed any recovery of RNase A activity (<2%). These results indicate that coupling mono- and diamino-groups lowers the efficiency of cysteine to promote the protein refolding to the native form. As indicated in Fig. 4a, the Cys-Dap/GSSG and Cys-Tamp/GSSG systems facilitated the disulfide-bond formation of reduced RNase A more quickly than the cysteine/GSSG system. Therefore, it is likely that cysteine coupled with mono- and diamino-





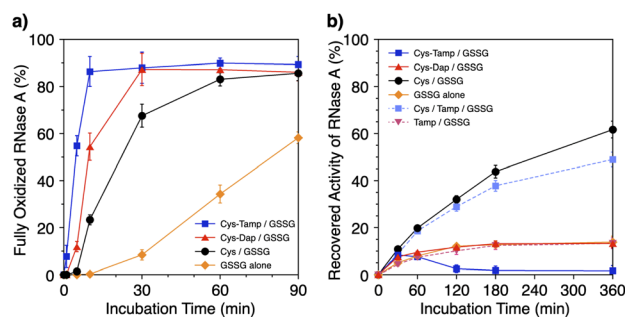
**Fig. 3** Time course analyses of RNase A oxidation by SDS-PAGE. SDS-PAGE gel images monitoring the oxidation of RNase A (8.0  $\mu$ M) in the presence of (a) GSSG, (b) cysteine and GSSG, (c) Cys-Dap and GSSG, (d) Cys-Tamp and GSSG, (e) cysteine, Tamp and GSSG, and (f) Tamp and GSSG (thiol compounds and Tamp: 1.0 mM; disulfide compounds: 0.20 mM) in a buffer (50 mM Tris-HCl, 300 mM NaCl, pH 7.5). The oxidation reactions were quenched with malPEG-2000 after 1, 5, 10, 30, 60, and 90 min incubations. The leftmost and rightmost lanes show the bands corresponding to the markers (M) and native RNase A (N), respectively. The bands marked with an asterisk in (d) were also observed under the same condition regardless of the absence of RNase A (see also Fig. S1 in ESI<sup>†</sup>). The bands marked with an asterisk in (d) are assigned to a complex of SDS and Cys-Tamp (see also Fig. S1 in ESI<sup>†</sup>). R represents reduced form, and 1SS, 2SS, 3SS, and 4SS represent species containing one, two, three, and four disulfide bonds, respectively. In these species, conjugates between RNase A and thiol-additives can be included.

groups produced non-native isomers of RNase A in the oxidation reactions driven by GSSG.

To investigate the effects of Tamp on the refolding reaction, enzymatic activity recoveries of RNase A in the Cys/Tamp/GSSG and Tamp/GSSG systems were measured. In the Tamp/GSSG system, reduced RNase A showed an analogous activity recovery profile compared to that in GSSG alone. Meanwhile, in the Cys/Tamp/GSSG system, reduced RNase A readily recovered its enzymatic activity up to 49% after 360 min incubation. These results suggest that simply adding Tamp to the reaction condition has limited influence on the refolding process of RNase A.

#### Effects of cysteine-containing dipeptides on BPTI folding

To gain a more general view of the effects of the cysteine-containing dipeptides on the protein folding reaction, we



**Fig. 4** (a) Time-course changes of the percentages of fully oxidized RNase A (8.0  $\mu$ M) and (b) recovered enzymatic activity of RNase A (8.0  $\mu$ M) in the presence of (blue squares) Cys-Tamp and GSSG, (red triangles) Cys-Dap and GSSG, (black circles) cysteine and GSSG, (orange diamonds) GSSG alone, (light blue squares) cysteine, Tamp and GSSG, and (pink inverted triangles) Tamp and GSSG (thiol compounds and Tamp: 1.0 mM; disulfide compounds: 0.20 mM) in a buffer (50 mM Tris-HCl, 300 mM NaCl, pH 7.5). The percentages of 4SS were quantified by SDS-PAGE analyses (original data: Fig. 3). The activity was evaluated by spectroscopic monitoring of the hydrolysis of cCMP to 3'-CMP at 30 °C. Error bars indicate the means  $\pm$  SEM of three independent experiments.

studied the folding reactions of bovine pancreatic trypsin inhibitor (BPTI). BPTI possesses three disulfide bonds, *i.e.*, C<sub>5</sub>-C<sub>55</sub>, C<sub>14</sub>-C<sub>38</sub> and C<sub>30</sub>-C<sub>51</sub>, and forms quasi-native intermediates such as N' and N\* with two disulfide bonds in the on-pathway folding process to afford the native structure (N, Fig. 2b).<sup>36</sup> Oxidative folding of reduced and denatured BPTI (R, 30  $\mu$ M) was monitored by reversed-phase high-performance liquid chromatography (HPLC). As we reported previously, R-BPTI showed spontaneous folding into N in the presence of GSSG as an oxidant.<sup>24</sup> The HPLC analysis indicated that the fraction of R disappeared after 30 min incubation and N-BPTI formed in 21% yield after 60 min. In the cysteine/GSSG system, the folding reaction of BPTI to its N form proceeded faster, where R disappeared within the initial 10 min incubation and the yield of N after 60 min was enhanced to 38% (Fig. 5a). By using Cys-Dap or Cys-Tamp as a thiol additive, the oxidation reaction of R-BPTI was further accelerated. Namely, in the Cys-Dap/GSSG and Cys-Tamp/GSSG systems, the fractions of R disappeared within 5 min. Interestingly, despite such rapid oxidation reactions, the yields of N were lower than that observed in the cysteine/GSSG system (Cys-Dap/GSSG: 30%; Cys-Tamp/GSSG: 18%). Furthermore, the HPLC profile changed slowly, particularly after 5 min incubation in the Cys-Tamp/GSSG system, and the HPLC trace after 60 min incubation was rather complicated with some fractions that usually did not appear (fractions marked with a red circle and bracket in Fig. 5c). These results indicate that R-BPTI transformed into multiple isomers by rapid disulfide-bond introduction and less efficient disulfide-bond shuffling in the Cys-Tamp/GSSG condition. The complicated HPLC profile indicates the formation of structural isomers of BPTI including not only on-pathway folding intermediates but also off-pathway species.<sup>37,38</sup> It should also be noted that, analogous to the case of RNase A folding, addition of Tamp hardly influenced the folding process of BPTI. The yields of N-BPTI after 60 min



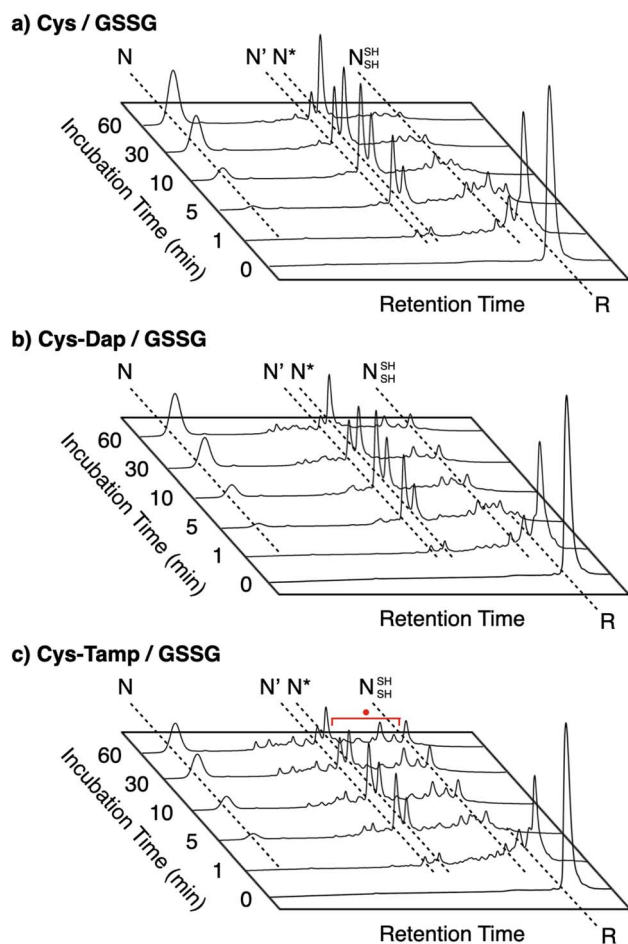


Fig. 5 Time-course reversed-phase HPLC traces of oxidative folding of BPTI (30  $\mu\text{M}$ ) in the presence of GSSG (0.20 mM) and (a) cysteine, (b) Cys-Dap, and (c) Cys-Tamp (1.0 mM) between 16 and 45 min retention time. N and R depict native and reduced forms of BPTI, respectively. The structural information of N', N\*, and N<sub>SH</sub><sup>SH</sup> is indicated in the caption of Fig. 2. Eluents: water (containing 0.05% TFA) and CH<sub>3</sub>CN (containing 0.05% TFA) with a linear gradient (water/CH<sub>3</sub>CN = 80/20 at 16 min to 65/35 at 45 min); flow rate: 1.0 mL min<sup>-1</sup>; detection wavelength: 229 nm.

incubation were 23% in the Tamp/GSSG system and 41% in the Cys/Tamp/GSSG system (Fig. S2<sup>†</sup>).

### Discussion of chemical properties of cysteine-containing dipeptides on the folding process

To investigate the reactivities of the thiol groups in the oxidative protein folding, chemical properties, *i.e.*, acidity ( $\text{pK}_a$ ) and redox potential ( $E^0$ ), of the cysteine-based dipeptides were evaluated. The properties of cysteine thiol group were reported as  $\text{pK}_a = 8.37$  and  $E^0 = -220$  mV.<sup>39</sup> The  $\text{pK}_a$  values were lowered by increasing the number of amino groups in the molecules:  $\text{pK}_a = 6.83 \pm 0.09$  (Cys-Dap) and  $6.67 \pm 0.09$  (Cys-Tamp).§ The  $E^0$

§  $\text{pK}_a$  values of cysteine in the presence of Dap and Tamp were  $8.36 \pm 0.03$  and  $8.27 \pm 0.04$ , respectively, indicating that the acidity of cysteine is hardly influenced by Dap or Tamp. Measurement condition: [cysteine] = [Dap] = [Tamp] = 125  $\mu\text{M}$ .

values also decreased to  $-256 \pm 8$  mV (Cys-Dap) and  $-281 \pm 1$  mV (Cys-Tamp). The decreased  $E^0$  values indicate that Cys-Dap and Cys-Tamp prefer the oxidized states to cysteine, so that, in the oxidative reaction by GSSG, intermolecular disulfide bond formation between a protein and a dipeptide would be preferential to the intramolecular disulfide bond formation between cysteine residues in a protein. This electrochemical property is likely the major reason of the faster disulfide bond formation of proteins observed in the presence of dipeptides. The increased acidity of Cys-Dap and Cys-Tamp relative to cysteine suggests enhanced nucleophilicity of the thiol units. Nucleophilic reactions of Cys-Dap and Cys-Tamp to cleave an intramolecular disulfide bond between cysteine residues in a protein can also provide protein–dipeptide conjugates with intermolecular disulfide bonds. Indeed, in the RNase A folding reactions, the formation of folding intermediates conjugated with the dipeptides was suggested by matrix-assisted laser desorption/ionization time-of-flight mass (MALDI-TOF MS) spectrometry. In the presence of cysteine and GSSG, a signal assigned to a one-to-one conjugate between RNase A and cysteine was observed after 1 min incubation (RNase A + Cys, red circle in Fig. 6a). After 180 min incubation, the signal of the conjugate could not be identified clearly due to significantly weakened intensity (Fig. 6c). In contrast, in the presence of Cys-Dap or Cys-Tamp, signals assigned to one-to-one and even one-to-two conjugates between RNase A and the dipeptides were clearly observed after 180 min incubation (orange circle and square in Fig. 6f and pink circle and square in Fig. 6i). Although quantitative comparisons are technically difficult based on the mass spectroscopic signal intensities, this result suggests relatively long lifetimes of the conjugates between RNase A and the dipeptides forming intermolecular disulfide bonds. Thus, the folding intermediates and non-native forms could be stabilized.

## Experimental

### Concentration determination of thiol compounds

A thiol compound dissolved in 10 mM HCl aq. was diluted in a buffer (50 mM Tris–HCl, 0.3 M NaCl, pH 7.5). The mixture was added to an aqueous solution of 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB). The concentration of the thiol compound was determined by the absorbance at 412 nm measured at 30 °C.<sup>40</sup>

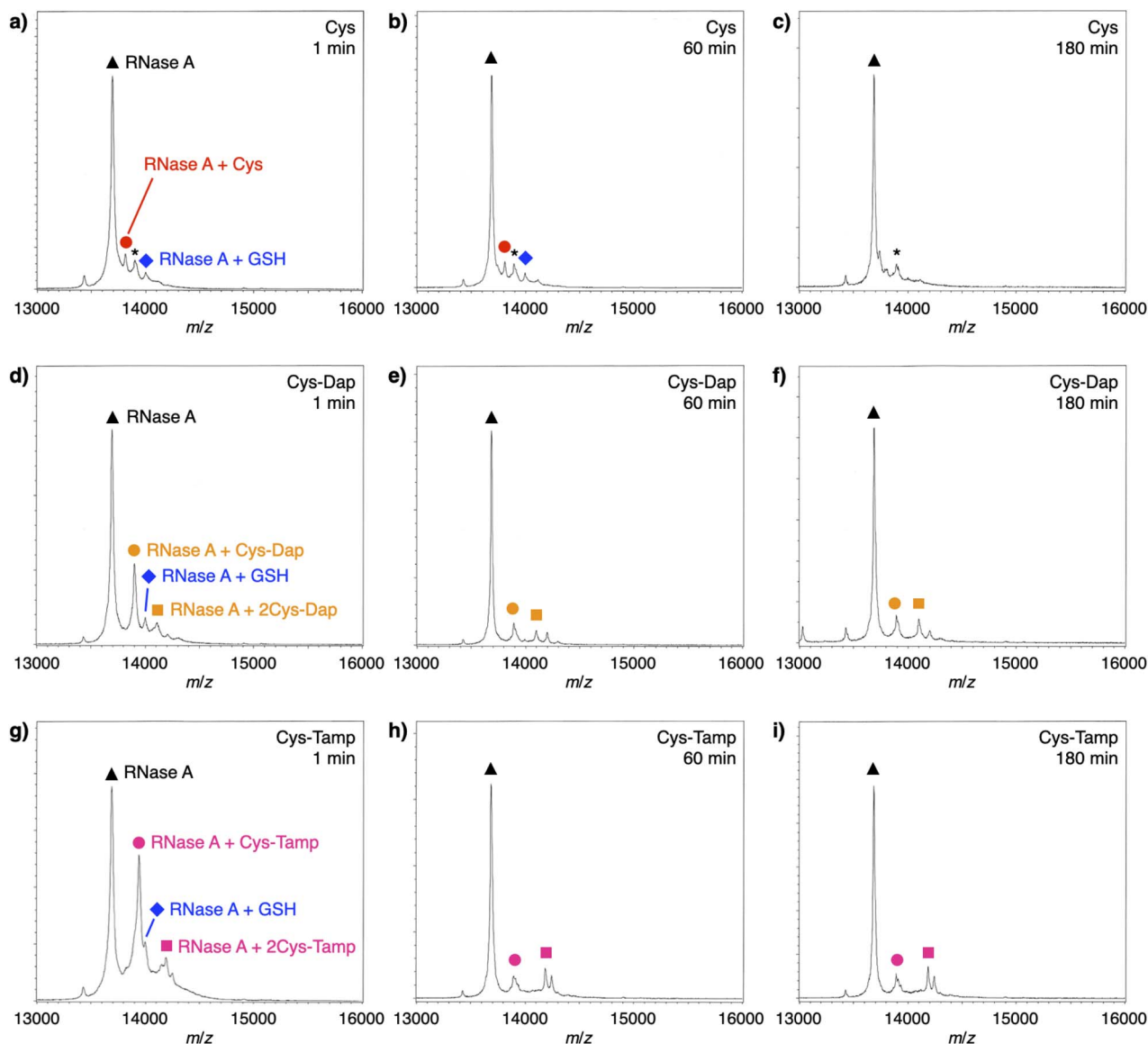
### Preparation of reduced and denatured RNase A

RNase A (from bovine pancreas) dissolved in a buffer (200 mM Tris–HCl, pH 8.6) containing 6.0 M guanidium chloride and 100 mM dithiothreitol (DTT) was incubated for 2 h at 25 °C. The mixture was dialyzed for 2 h at 25 °C with 10 mM HCl aq., and the dialysis was performed three times. The content of the secondary structure after the dialysis was characterized by the circular dichroism spectroscopic measurement (Fig. S3<sup>†</sup>).

### SDS-PAGE measurements of RNase A

A buffer (50 mM Tris–HCl, 0.3 M NaCl, pH 7.5) containing oxidized glutathione (GSSG, 0.20 mM) and a thiol compound (1.0 mM) was preincubated for 10 min at 30 °C.<sup>41</sup> Reduced and





**Fig. 6** Time-course MALDI-TOF MS analyses of RNase A (8.0  $\mu\text{M}$ ) in the presence of (a–c) cysteine and GSSG, (d–f) Cys-Dap and GSSG, and (g–i) Cys-Tamp and GSSG at 1, 60 and 180 min of incubation after the addition of the thiol and disulfide compounds (thiol compounds: 1.0 mM; disulfide compounds: 0.20 mM) in a buffer (50 mM Tris-HCl, 300 mM NaCl, pH 7.5). Matrix: sinapinic acid. Linear and positive modes. Signals marked with black triangles, blue diamonds, red circles, orange circles, orange squares, pink circles, pink squares, and asterisks correspond to RNase A, RNase A–GSH conjugate, RNase A–cysteine conjugate, RNase A–Cys-Dap conjugate, RNase A–(Cys-Dap)<sub>2</sub> conjugate, RNase A–Cys-Tamp conjugate, RNase A–(Cys-Tamp)<sub>2</sub> conjugate, and RNase A–sinapinic acid complex, respectively.

denatured RNase A (final concentration: 8.0  $\mu\text{M}$ ) was added to the buffer, and the mixture was incubated for 90 min at 30 °C. During the incubation, 16  $\mu\text{L}$  of the reaction mixture was collected at designated times (0, 1, 5, 10, 30, 60, and 90 min). The collected mixture was added to Laemmli's  $\times 4$  SDS-loading buffer<sup>42</sup> (16  $\mu\text{L}$ ) containing malPEG-2000 (10 mM) to quench disulfide bond formation of free thiol groups. Redox states of RNase A obtained by the above treatment were separated by non-reducing 14% SDS-PAGE using WIDE RANGE gel (Nacalai Tesque, Kyoto, Japan). Proteins were visualized by staining with Coomassie brilliant blue G-250. The band intensities were analyzed by a ChemiDoc Touch imaging system and Image Lab (Bio-Rad, Hercules, CA, USA).

### Enzymatic activity measurements of RNase A

A buffer (50 mM Tris-HCl, 0.3 M NaCl, pH 7.5) containing oxidized glutathione (GSSG, 0.20 mM) and a thiol compound (1.0 mM) was preincubated for 10 min at 30 °C. Reduced and denatured RNase A (final concentration: 8.0  $\mu\text{M}$ ) was added to the buffer, and the mixture was incubated for 360 min at 30 °C. During the incubation, 110  $\mu\text{L}$  of the reaction mixture was collected at designated times (30, 60, 120, 180, and 360 min). The collected mixture was added to a buffer (50 mM Tris-HCl, 0.3 M NaCl, pH 7.5, 330  $\mu\text{L}$ ) containing cCMP (final concentration: 0.80 mM). The mixture was poured into a quartz cuvette to monitor the change in absorption intensity at 284 nm at 30 °C for 4 min by a V-650 UV-vis spectrophotometer (JASCO, Tokyo,



Japan). Values represent means  $\pm$  SEM from three independent experiments.

### BPTI folding assay

Reduction and denaturation of BPTI (from bovine lung) was carried out as follows. BPTI (10 mg) was dissolved in 1 mL of 0.1 M Tris-HCl (pH 8.0) containing 30 mM DTT and 8 M urea and the solution was allowed to stand for 3 h at 50 °C. The reduced and denatured protein was purified by HPLC using an InertSustain C18 column ( $\phi$ 4.6  $\times$  250 mm, GL Sciences, Tokyo Japan), lyophilized and stored at -30 °C until use. Purified reduced/denatured BPTI was analyzed by MALDI-TOF MS, resulting in the confirmation of disulfide bonds reduction. For the assay, fully reduced and denatured BPTI (600  $\mu$ M) was dissolved in 6 M urea containing 0.05% trifluoroacetic acid. A mixture of GSSG and a thiol compound was dissolved in a buffer (50 mM Tris-HCl, 300 mM NaCl, pH 7.5), and the resulting mixture was preincubated for 10 min at 30 °C. To the preincubated mixture, the solution of reduced and denatured BPTI was added (final concentrations: BPTI: 30  $\mu$ M, urea: 300 mM, GSSG: 200  $\mu$ M, thiol compound: 1.0 mM, Tris-HCl: 50 mM, NaCl: 300 mM, pH 7.5). After incubation at 30 °C for pre-determined periods, the reaction was quenched by adding an equal volume of 1 M HCl aq., which was then analyzed by reversed-phase HPLC at a flow rate of 1.0 mL min<sup>-1</sup> monitoring at 229 nm with a linear gradient of elutions (solvent A: 0.05% trifluoroacetic acid in water; solvent B: 0.05% trifluoroacetic acid in acetonitrile; percentages of solvent A: 95% at 0 min, 80% at 15 min, 30% at 115 min).

### Determination of acid dissociation constant $pK_a$

Citrate buffers (pH 3.0, 4.0), phosphate buffers (pH 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0), and boric-acid buffers (pH 8.5, 9.0, 9.5, 10.0, 10.5, 11.0, 11.5, 12.0, 12.5) were prepared for the measurements. To a quartz cell were added a thiol compound (5.0 mM, 12.5  $\mu$ L) and a buffer of the desired pH (487.5  $\mu$ L) at 25 °C. After mixing the sample, the pH value and absorbance at 240 nm of the resultant solution were measured. The plot of the absorbance values at each pH was analyzed by KaleidaGraph (version 5.0.1) for curve fitting to evaluate  $pK_a$ .

### Determination of redox potential $E^{0'}$

The  $E^{0'}$  value of a thiol compound was determined by following the previously reported protocol.<sup>43</sup> A buffer (100 mM Tris-HCl, 1.0 mM EDTA, pH 7.0) was degassed by N<sub>2</sub> bubbling for more than 1 h prior to use. DTT<sup>red</sup> (60  $\mu$ M, 2.5 mL) in the buffer was added to a disulfide (60  $\mu$ M, 2.5 mL) in the buffer under N<sub>2</sub>, which was stirred at 25  $\pm$  0.1 °C for 24 h. To quench the reaction, an aliquot of the reaction mixture (1 mL) was added to 1 M HCl aq. (200  $\mu$ L), and the obtained sample solution was immediately analyzed by reversed-phase HPLC (Tosoh TSKgel ODS-100V column,  $\phi$ 4.6  $\times$  250 mm). The column was equilibrated with water containing 0.1% trifluoroacetic acid at a flow rate of 1.2 mL min<sup>-1</sup> prior to use. The RP-HPLC analysis was conducted with water containing 0.1% trifluoroacetic acid (eluent A) and CH<sub>3</sub>CN containing 0.1% trifluoroacetic acid

(eluent B) with a linear gradient (percentage of eluent B: 0–4% in 0–10 min, 4–7% in 10–20 min). The concentrations of the species at equilibrium were calculated from the observed peak areas and corresponding calibration curves.

The equilibrium constant  $K_{eq}$  of the reaction (eqn (1)), described as eqn (2), was determined by averaging seven individual experiments following the above procedure.



$$K_{eq} = \frac{[\text{Thiol}]^2[\text{DTT}^{\text{ox}}]}{[\text{Disulfide}][\text{DTT}^{\text{red}}]} \quad (2)$$

The redox potential  $E^{0'}$  was calculated by the Nernst equation (eqn (3)).

$$E^{0'} = E_{\text{DTT}}^{0'} + \frac{RT}{nF} \ln K_{eq} \quad (3)$$

where  $n$  is the number of transferred electrons ( $n = 2$ ),  $F$  is Faraday's constant (96 500 C mol<sup>-1</sup>),  $R$  is the universal gas constant (8.314 J K<sup>-1</sup> mol<sup>-1</sup>),  $T$  is the temperature (298 K), and  $E^{0'}_{\text{DTT}}$  is the redox potential of DTT (-327 mV).

## Conclusions

Protein misfolding is believed to be the primary cause of many common degenerative diseases. Therefore, trapping folding intermediates and misfolded forms is fundamentally important for the characterization of the disease-related biological properties of the misfolded proteins. Investigation of the folding intermediates is also important to understand the detailed mechanisms and pathways of protein folding. Currently, the methods to trap folding intermediates and misfolded forms are limited to essentially irreversible processes. In this study, we demonstrated rapid disulfide-bonding and stable formation of non-native isomers of RNase A and BPTI using cysteine-containing dipeptides, Cys-Dap and Cys-Tamp, in the presence of GSSG. Owing to the intrinsic reversibility of the disulfide bonds upon redox stimuli, the disulfide bond-based approach has great potential to develop novel reversible methodologies enabling the trapping of transient and misfolded forms of proteins by forming intermolecular disulfide-bonds and restarting the oxidative folding of the trapped forms by subsequent cleavage of the intermolecular disulfide bonds. The thiol-disulfide bond chemistry also relates to virus infection and immune suppression activities. ORF8 is an accessory protein of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that is proposed to interfere with immune responses.<sup>44,45</sup> In the large-scale dimeric structure, ORF8 contains three sets of intramolecular disulfide bonds per monomer and a single intermolecular disulfide bond connecting the monomers. Disulfide-bond cleaving and disturbing compounds such as Cys-Dap and Cys-Tamp would be useful to investigate the mechanisms of the unique molecular structure of ORF8 to fulfill the functions, which can also lead to a drug design. Thus, the molecular design reported in this study is an important step to develop methodologies for such reversible protein folding



manipulations as well as for analysis of virus infection activities.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

S. O. thanks the JSPS Research Fellowship for Young Scientists. We thank Mr Takaaki Toyoda for the experimental supports. This work was supported by research funds from a Grant-in-Aid for Transformative Research Areas (JP21H05096 to T. M., JP21H05095 to M. O.), JST CREST (JPMJCR19S4 to T. M.), JST FOREST (JPMJFR2122 to T. M.), Takeda Science Foundation (T. M.) and Asahi Glass Foundation (T. M.).

## Notes and references

- C. M. Dobson, *Philos. Trans. R. Soc. London, Ser. B*, 2001, **356**, 133.
- Y. E. Kim, M. S. Hipp, A. B. M. Hayer-Hartl and F. U. Hartl, *Annu. Rev. Biochem.*, 2013, **82**, 323.
- C. B. Anfinsen, *Science*, 1973, **181**, 223.
- T. E. Creighton, *Biol. Chem.*, 1997, **378**, 731.
- M. Narayan, E. Welker, W. J. Wedemeyer and H. A. Scheraga, *Acc. Chem. Res.*, 2000, **33**, 805.
- C. M. Dobson, *Nature*, 2003, **426**, 884.
- F. Baneyx and M. Mujacic, *Nat. Biotechnol.*, 2004, **22**, 1399.
- J. L. Arolas, F. X. Aviles, J.-Y. Chang and S. Ventura, *Trends Biochem. Sci.*, 2006, **31**, 292.
- K. Arai and M. Iwaoka, *Molecules*, 2021, **26**, 195.
- M. Bucciantini, E. Giannoni, F. Chiti, F. Baroni, L. Formigli, J. Zurdo, N. Taddei, G. Ramponi, C. M. Dobson and M. Stefani, *Nature*, 2002, **416**, 507.
- C. Haass and D. J. Selkoe, *Nat. Rev. Mol. Cell Biol.*, 2007, **8**, 101.
- D. Eisenberg and M. Jucker, *Cell*, 2012, **148**, 1188.
- S. Ninagawa, S. Tada, M. Okumura, K. Inoguchi, M. Kinoshita, S. Kanemura, K. Imami, H. Umezawa, T. Ishikawa, R. B. Mackin, S. Torii, Y. Ishihama, K. Inaba, T. Anazawa, T. Nagamine and K. Mori, *eLife*, 2020, **9**, e60970.
- T. E. Creighton, N. J. Darby and J. Kemmink, *FASEB J.*, 1996, **10**, 110.
- Y. Konishi, T. Ooi and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.*, 1982, **79**, 5734.
- D. P. Goldenberg, *Trends Biochem. Sci.*, 1992, **17**, 257.
- O. Hanpanich and A. Maruyama, *Biomaterials*, 2020, **254**, 120150.
- T. Nishimura and K. Akiyoshi, *Bioconjugate Chem.*, 2020, **31**, 1259.
- B. Wilkinson and H. F. Gilbert, *Biochim. Biophys. Acta, Proteins Proteomics*, 2004, **1699**, 35.
- W. J. Lees, *Curr. Opin. Chem. Biol.*, 2008, **12**, 740.
- M. Okumura, S. Shimamoto and Y. Hidaka, *FEBS J.*, 2012, **279**, 2283.
- J. C. Lukesh III, M. J. Palte and R. T. Raines, *J. Am. Chem. Soc.*, 2012, **134**, 4057.
- P. S. Reddy and N. Metanis, *Chem. Commun.*, 2016, **52**, 3336.
- S. Okada, M. Matsusaki, K. Arai, Y. Hidaka, K. Inaba, M. Okumura and T. Muraoka, *Chem. Commun.*, 2019, **55**, 759.
- S. Tsukagoshi, R. Mikami and K. Arai, *Chem.-Asian J.*, 2020, **15**, 2646.
- H. Feng, Z. Zhou and Y. Bai, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 5026.
- P. Neudecker, P. Robustelli, A. Cavalli, P. Walsh, P. Lundström, A. Zarrine-Afsar, S. Sharpe, M. Vendruscolo and L. E. Kay, *Science*, 2012, **336**, 362.
- M. Tinzl and D. Hilvert, *ChemBioChem*, 2021, **22**, 92.
- T. Zander, N. D. Phadke and J. C. A. Bardwell, *Methods Enzymol.*, 1998, **290**, 59.
- H. A. Scheraga, W. J. Wedemeyer and E. Welker, *Methods Enzymol.*, 2001, **341**, 189.
- R. Miyake, S. Tashiro, M. Shiro, K. Tanaka and M. Shionoya, *J. Am. Chem. Soc.*, 2008, **130**, 5646.
- R. Miyake, A. Ando, M. Ueno and T. Muraoka, *J. Am. Chem. Soc.*, 2019, **141**, 8675.
- A. Bierzynski, P. S. Kim and R. L. Baldwin, *Proc. Natl. Acad. Sci. U.S.A.*, 1982, **79**, 2470.
- D. M. Rothwarf and H. A. Scheraga, *Biochemistry*, 1993, **32**, 2680.
- M. Iwaoka, D. Juminaga and H. A. Scheraga, *Biochemistry*, 1998, **37**, 4490.
- J. S. Weissman and P. S. Kim, *Science*, 1991, **253**, 1386.
- R. Kojima, M. Okumura, S. Masui, S. Kanemura, M. Inoue, M. Saiki, H. Yamaguchi, T. Hikima, M. Suzuki, S. Akiyama and K. Inaba, *Structure*, 2014, **22**, 431.
- M. Okumura, S. Kanemura, M. Matsusaki, M. Kinoshita, T. Saio, D. Ito, C. Hirayama, H. Kumeta, M. Watabe, Y. Amagai, Y.-H. Lee, S. Akiyama and K. Inaba, *Structure*, 2021, **29**, 1357.
- P. C. Jocelyn, *Eur. J. Biochem.*, 1967, **2**, 327.
- G. L. Ellman, *Arch. Biochem. Biophys.*, 1959, **82**, 70.
- M. M. Lyles and H. F. Gilbert, *Biochemistry*, 1991, **30**, 613.
- U. K. Laemmli, *Nature*, 1970, **227**, 680.
- K. Arai, H. Ueno, Y. Asano, G. Chakrabarty, S. Shimodaira, G. Mugesh and M. Iwaoka, *ChemBioChem*, 2018, **19**, 207.
- T. G. Flower, C. Z. Buffalo, R. M. Hooy and J. H. Hurley, *Proc. Natl. Acad. Sci. U.S.A.*, 2021, **118**, e2021785118.
- B. E. Young, S.-W. Fong, Y.-H. Chan, T.-M. Mak, L. W. Ang, D. E. Anderson, C. Y.-P. Lee, S. N. Amrun, B. Lee, Y. S. Goh, Y. C. F. Su, W. E. Wei, S. Kalimuddin, L. Y. A. Chai, S. Pada, S. Y. Tan, L. Sun, P. Parthasarathy, Y. Y. C. Chen, T. Barkham, R. T. P. Lin, S. Maurer-Stroh, Y.-S. Leo, L.-F. Wang, L. Renia, V. J. Lee, G. J. D. Smith, D. C. Lye and L. F. P. Ng, *Lancet*, 2020, **396**, 603.

