# Protein Prosthesis: A Nonnatural Residue Accelerates Folding and Increases Stability

Ulrich Arnold, Matthew P. Hinderaker, Jens Köditz, Ralph Golbik, Renate Ulbrich-Hofmann, and Ronald T. Raines*

*Department of Biochemistry/Biotechnology, Martin-Luther University, 06114 Halle, Germany, and Department of Chemistry and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, USA*

**Synthesis of Fmoc-Asn(trt)-5,5-Dimethyl-L-Proline.** The synthesis of 5,5-dimethyl-L-proline is also reported in the text of Hinderaker, M. P.; Raines, R. T. *Protein Sci.* **2003**, *12*, 1188–1194.

Chemicals and solvents were from Aldrich (Milwaukee, WI). Reactions were monitored by thin-layer chromatography using TLC plates (AL SIL G/UV) from Whatman (Kent, UK) with visualization by illumination with ultraviolet light or staining with $I_2$. NMR spectra were obtained with Bruker AC-250, Bruker AC-300, and Varian UNITY-500 spectrometers. Mass spectra were obtained with electrospray ionization (ESI) techniques at the University of Wisconsin Biotechnology Center.

2-(3-Methyl-3-nitro-butyl)-[1,3]dioxolane was prepared as described previously.[S1] 2-(3-Methyl-3-nitro-butyl)-[1,3]dioxolane (68 g, 0.36 mol) was dissolved in methanol (0.25 L). Raney nickel was added to this solution, and the flask was filled with $H_2$(g) from a balloon. (CAUTION! Raney nickel is highly pyrophoric and will ignite methanol vapors if dry.) Periodically, the reaction was assayed for completion by $^{13}$C NMR spectroscopy. Completion times ranged from overnight to 10 days, depending on the scale and catalyst loading. Upon completion, the catalyst was carefully filtered away. The solvent was removed under reduced pressure to yield 3-[1,3]dioxolan-2-yl-1,1-dimethyl-propylamine as a pale yellow oil (56.5 g, 99%). $^1$H NMR ($CDCl_3$, 300 MHz) $\delta$: 4.84 (t, $J$ = 5 Hz, 1H), 3.82–3.98 (m, 4H), 1.65–1.75 (m, 2H), 1.42–1.51 (m, 2H), 1.10 (s, 6H). $^{13}$C NMR ($CDCl_3$, 75.5 MHz) 104.0, 64.1, 48.3, 38.1, 29.67, 28.5. MS (ESI) *m/z* 160.2 ($MH^+$).

3-[1,3]Dioxolan-2-yl-1,1-dimethyl-propylamine (15.3 g, 96 mmol) was dissolved in hot water (40 mL) and the pH lowered to ~3 with 2 N HCl. The resulting solution was heated at reflux for 30 min, and then made basic by the addition of 6 N KOH and extracted (4×) with chloroform. The organic layer was dried over $MgSO_4$(s), and the solvent was removed under reduced pressure to yield a black oil. The black oil was purified by distillation under reduced pressure to yield 2,2-dimethyl-3,4-dihydro-2H-pyrrole as a clear pungent oil (7.5 g, 80%). $^1$H NMR ($CDCl_3$, 300 MHz) $\delta$: 7.33 (broad s, 1H), 2.52 (t, $J$ = 7 Hz, 2H), 1.56 (t, $J$ = 7 Hz, 2H), 1.2 (s, 6H). $^{13}$C NMR ($CDCl_3$, 75.4 MHz) $\delta$ 163.1, 72.9, 36.7, 34.3, 28.5.

2,2-Dimethyl-3,4-dihydro-2H-pyrrole (12.4 g, 128 mmol) was dissolved in water (70 mL) at 0 ºC. KCN (16.3 g, 251 mmol) was added to the solution, and the pH was lowered from ~13.4 to ~6 over 2 h by the addition of 2 N HCl (128 ml, 256 mmol). The pH continued to rise, but was maintained near 6 by the addition of 2 N HCl. After 3 h at 0 °C, the solution was made basic by the addition of 2 N NaOH, and extracted (4×) with chloroform. The combined organic extract was dried with $MgSO_4$(s), and the solvent removed under reduced pressure. The residue was purified by flash chromatography (0.4 kg silica gel, ethyl acetate/hexanes (65:35 v/v)). Fractions containing 5,5-dimethyl-pyrrolidine-2-carbonitrile were pooled, and the solvent was removed under reduced pressure to yield 5,5-dimethyl-pyrrolidine-2-carbonitrile as a pale yellow oil

(7.47 g, 46%). $^1$H NMR ($CDCl_3$, 300 MHz) $\delta$: 4.07–4.12 (apparent dd, $J$ = 8, 5 Hz, 1H), 2.12–2.36 (m, 2H), 1.79–1.90 (m, 2H), 1.61–1.72 (m, 1H), 1.30 (s, 3H), 1.17 (s, 3H). $^{13}$C NMR ($CDCl_3$, 75.4 MHz) $\delta$ 122.0, 59.5, 46.5, 38.3, 30.9, 29.0, 28.8. $^{13}$C DEPT-135 $\delta$: 46.5, 29.0, 28.8 positive (CH or $CH_3$), 38.3, 30.9 negative ($CH_2$). MS (ESI) *m/z* 125.2 ($MH^+$).

5,5-Dimethyl-pyrrolidine-2-carbonitrile was hydrolyzed to form 5,5-dimethylproline as described previously.[S1] Racemic 5,5-dimethylproline was resolved with D-tartrate as described previously[S2] to yield 5,5-dimethyl-L-proline in greater than 97% ee. 5,5-Dimethyl-L-proline (170 mg, 1.19 mmol) was dissolved in dry DMF (18 mL). N-α-Fmoc-N-β-trityl-L-asparagine pentafluorophenyl ester (906 mg, 1.19 mmol) and N-hydroxybenzotriazole (165 mg, 1.19 mmol) were added to the solution, and the resulting mixture was stirred overnight at room temperature. The solvent was removed under reduced pressure, and the residue was purified by flash chromatography (20 g silica gel, 4% v/v methanol in methylene chloride). Fractions containing Fmoc-Asn(trt)-5,5-dimethyl-L-proline were pooled, and the solvent was removed under reduced pressure to yield Fmoc-Asn(trt)-5,5-dimethyl-L-proline as a white crystalline solid (710 mg, 83%). MS (ESI) *m/z* 722.6 ($MH^+$).

**Peptide Synthesis.** The sequence of the synthetic peptide was CAYKTTQANKHIIVACEGN–5,5-dimethyl-L-proline–YVPVHFDASV (which corresponds to residues 95–124). This peptide was synthesized on a 100-μmol scale with a Pioneer automated peptide synthesizer (Applied Biosystems) at the University of Wisconsin Biotechnology Center. A methylbenzhydrylamine polystryrene resin that had been functionalized with a 4-hydroxymethylphenoxy acid-labile linker and loaded with valine (which is the C-terminal residue of RNase A) was used for peptide synthesis. Appropriate Fmoc-protected monomers and the Fmoc-Asn(trt)-5,5-dimethyl-L-proline dimer were coupled by using cycles of *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate/diisopropylethyl amine (HATU) activation of the carboxylic acid group, followed by piperidine deprotection of the Fmoc group. After cleavage from the resin and side-chain deprotection, the peptide was purified by high-performance liquid chromatography on a C18 reverse-phase column. The monoisotopic mass of the purified peptide was in accord with expected values for $MH^+$, $MNa^+$, and $MK^+$: MS (MALDI) *m/z* 3303.9 ($MH^+$ [$C_{148}H_{227}N_{39}O_{43}S_2H$] = 3303.6), 3326.0 ($MNa^+$ [$C_{148}H_{227}N_{39}O_{43}S_2Na$] = 3325.6), 3343.9 ($MK^+$ [$C_{148}H_{227}N_{39}O_{43}S_2K$] = 3341.6). For reference, the monoisotopic mass of the peptide corresponding to wild-type residues 95–124 is lower by 28 units: MS (MALDI) *m/z* 3275.2 ($MH^+$ [$C_{146}H_{223}N_{39}O_{43}S_2H$] = 3275.6).[S3]

**Plasmid Construction.** Plasmid pTXB1, *E. coli* cells strain ER2566, and chitin beads were from New England Biolabs (Beverly, MA). Wild-type RNase A and thermolysin were from Sigma Chemical (St. Louis, MO). All other chemicals were of commercial reagent grade or better, and were used without further purification.

By use of standard techniques the DNA sequence encoding the 1–94 fragment of RNase A (from plasmid pBXR[S4]) was inserted into plasmid pTXB1 as described elsewhere.[S3] It is noteworthy that proteins produced by this method possess an additional methionine residue at their N-terminus: Met(–1).

**Production of Intein Fusion Protein.** *Escherichia coli* ER2566 cells that had been transformed with a plasmid that directs the expression of RNase A(1–94) fused to the *mxe*-intein were cultivated in Luria–Bertani medium containing ampicillin (100 μg/mL) with shaking at 37 °C until O.D. = 0.5 at 600 nm. Gene expression was then induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG; to 0.5 mM), and the cultures were grown for an

additional 3–4 h at 25 °C. The yield of harvested cells was ~2.3 g per liter of bacterial culture. Cells were stored at –20 °C until used.

**Production of Thioester-Tagged RNase A(1–94).** Thioester-tagged RNase A(1–94) was produced according to the instructions given by New England Biolabs with the modifications described previously:[S3] lysis and column buffer was 20 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS)–NaOH buffer (pH 6.8) containing NaCl (0.5 M), ethylenediaminetetraacetic acid (EDTA; 0.1 mM), and Triton X-100 (0.1% v/v) and cleavage buffer was 50 mM MOPS–NaOH buffer (pH 6.8) containing NaCl (0.5 M), 2-mercaptoethanesulfonic acid (MESNA; 50 mM), and EDTA (0.1 mM). The cleavage product, namely RNase A(1–94) thioester, was eluted from the resin after 16 h at 4 °C and subsequently precipitated by sodium deoxycholate (to 0.24 mM) and trichloroacetic acid (to 30 mM). This precipitate was washed with acetone, dissolved in 2 M guanidine–HCl (Gdn–HCl) to a concentration of approximately 5 mg/mL, and used in a ligation reaction.

**Ligation of Thioester-Tagged Protein and Peptide.** The ligation reaction was performed in 0.20 M Tris–HCl buffer (pH 8.0) containing Gdn–HCl (2.0 M), the thioester-tagged protein (5 mg/mL), and peptide (12.5 mg/mL; 40-fold molar excess). The pH was adjusted to ~8 with 1 M NaOH, and the ligation reaction was allowed to proceed at room temperature for at least 16 h.

**Protein Folding and Purification**. The ligation reaction mixture was diluted with 0.10 M Tris–HCl buffer (pH 8.0) to <0.1 mg of protein per milliliter and <0.1 M Gdn–HCl. Reduced glutathione (GSH) and oxidized glutathione (GSSG) were added with stirring at 4 °C to final concentrations of 1.0 and 0.2 mM, respectively. After adjusting the pH to ~8 with 1.0 M NaOH, the solution was transferred to room temperature and protein folding was allowed to proceed for at least 16 h.

Prior to application of the folding reaction to columns, any precipitate formed during folding was removed by centrifugation at 15,000*g* for 30 min. Purification employed hydrophobic interaction chromatography (Phenyl Sepharose™ 6 Fast Flow, Amersham Biosciences) and cation-exchange chromatography (HiTrap SP Sepharose™ cation-exchange resin, Amersham Biosciences) as described previously.[S3] Fractions (2 mL) were collected manually, and absorbance at 278 nm as well as the determination of ribonucleolytic activity were used to locate the protein.

Samples from throughout the purification procedure were analyzed by SDS–PAGE through a 15% (w/v) acrylamide gel. After purification, 25 mg of peptide yielded 0.8 mg of a semisynthetic RNase A with dmP as residue 114 (dmP114 RNase A) (40% yield based on thioester-tagged protein; 1% yield based on peptide).

**Mass Spectrometry.** After desalting the samples by use of μ-C18 ZipTip™ (Millipore), a mass spectrum of dmP114 RNase A was obtained with a Bruker Biflex III matrix-assisted laser desorption/ionization (MALDI) instrument (Figure S1). MS (MALDI) *m*/*z* 13,845 (expected: 13,841). For reference, the mass of semisynthetic wild-type RNase A is lower by approximately 28 units: MS (MALDI) *m*/*z* 13,815 (expected: 13,813).[S3]

**Thermal Denaturation.** The thermal stability of dmP114 RNase A was determined by monitoring the change in absorbance at 287 nm with temperature. Protein samples (0.2 mg/mL) were in 50 mM sodium phosphate buffer (pH 8.0) containing NaCl (25 mM). The thermal transition upon heating at 0.15 °C/min was monitored with a Cary Model 3 UV/VIS spectrophotometer (Varian). The value of $T_m$, which is the temperature at the midpoint of the thermal transition, was determined as described previously.[S3]

**Ribonucleolytic Activity.** Values of $k_{cat}/K_M$ for the enzymatic cleavage of a fluorogenic substrate, 6-carboxyfluorescein~dArU(dA)$_2$~6-carboxytetramethylrhodamine, were determined as described previously.[S5] Additionally, the residual activities of wild-type RNase A and the dmP114 variant were determined in the presence of 0–400 mM Gdn–HCl (Figure S2).

**Determination of Unfolding Rate Constants ($k_U$).** Values of $k_U$ of wild-type RNase A and dmP114 RNase A were determined by proteolysis with thermolysin as described previously.[S6,S7] In a typical experiment, 20 µL of thermolysin (2 mg/mL in 50 mM Tris–HCl buffer, pH 8.0, containing 10 mM $CaCl_2$) was added to 160 µL of 50 mM Tris–HCl buffer (pH 8.0), which had been pre-incubated in a thermostat (Lauda, accuracy ±0.1 °C). The unfolding reaction was started by addition of wild-type RNase A or dmP114 RNase A (20 µL of a 1.0 mg/mL solution). After distinct time intervals, 25 µL was taken and mixed rapidly with 8 µL of 50 mM EDTA. SDS–PAGE and densitometric evaluation of the RNase A bands were carried out as described previously.[S6] From the decrease in the peak areas of the intact RNase A band, which followed a first-order reaction, the rate constants of proteolysis were determined. Under the conditions applied these values correspond to the respective unfolding rate constants. The data are shown as an Eyring plot in Figure S3.

**Gdn–HCl-Induced Denaturation**. Fluorescence spectroscopy was used to determine the thermodynamic stability of dmP114 RNase A at 25 °C in 0.10 M MES–NaOH buffer (pH 6.0) containing NaCl (0.10 M). Fluorescence emission spectra of protein samples (100 µg/mL) with increasing concentration of Gdn–HCl were recorded from 290–350 nm in 1.0-nm increments on a Fluoro-Max-2 spectrometer (Yvon-Spex) (Figure S4A). The slit width was 5 nm for excitation at 278 nm and 5 nm for emission, and the cuvette path length was 1 cm. For the transition curve, the signal with maximum difference between native and unfolded protein (303 nm) was recorded and fitted to a two-state model by nonlinear regression (Figure S4B).[S8,S9]

**Stopped-Flow Refolding.** Wild-type RNase A and dmP114 RNase A (500 ng/mL) were incubated in 4.5 M Gdn–HCl in 0.10 M MES–NaOH buffer (pH 6.0) containing NaCl (0.10 M) at 5–25 °C. Refolding was initiated by mixing 1 volume of RNase A solution with 25 volumes of 0.10 M MES–NaOH buffer (pH 6.0) containing NaCl (0.10 M) and 6-carboxyfluorescein–dArU(dA)$_2$–6-carboxytetramethylwrhodamine (50 nM). The resulting Gdn–HCl concentration was 173 mM. The recovery of enzymatic activity was used as a measure for refolding of the enzymes. Refolding was followed by fluorescence spectroscopy using an Applied Photophysics BioSequential DX.17 MV stopped-flow spectrometer. Emission was recorded as integral fluorescence using a 515 nm cut-off filter with excitation at 490 nm. Typically, 8–10 traces were averaged. Data were fitted to the equation:

$$\text{fluorescence}(t) = Ae^{-k_F t} + mt + n \qquad \text{(S1)}$$

where fluorescence($t$) is the fluorescence signal at time $t$, $A$ is the signal amplitude, $k_F$ is the reactivation rate constant, $t$ is time, $m$ is the slope of the linear part, and $n$ is the signal off-set. This equation is sufficient to describe the refolding kinetics of wild-type RNase A and its dmP114 variant, and is similar in form to equations used previously for similar purposes.[S10,S11] The data are shown in Figure S5.

**References**

(S1) Bonnett, R.; Clark, V. M.; Giddey, A.; Todd, A. *J. Chem. Soc.* **1959**, 2087–2093.
(S2) An, S. S. A.; Lester, C. C.; Peng, J. -L.; Li, Y. -J.; Rothwarf, D. M.; Welker, E.; Thannhauser, T. W.; Zhang, L. S.; Tam, J. P.; Scheraga, H. A. *J. Am. Chem. Soc.* **1999**, *121*, 11558–11566.
(S3) Arnold, U.; Hinderaker, M. P.; Nilsson, B. L.; Huck, B. R.; Gellman, S. H.; Raines, R. T. *J. Am. Chem. Soc.* **2002**, *124*, 8522–8523.
(S4) delCardayré, S. B.; Ribó, M.; Yokel, E. M.; Quirk, D. J.; Rutter, W. J.; Raines, R. T. *Protein Eng.* **1995**, *8*, 261–273.
(S5) Kelemen, B. R.; Klink, T. A.; Behlke, M. A.; Eubanks, S. R.; Leland, P. A.; Raines, R. T. *Nucleic Acids Res.* **1999**, *27*, 3696–3701.
(S6) Arnold, U.; Ulbrich-Hofmann, R. *Biochemistry* **1997**, *36*, 2166–2172.
(S7) Arnold, U.; Ulbrich-Hofmann, R. *Eur. J. Biochem.* **2001**, *268*, 93–97.
(S8) Santoro, M. M.; Bolen, D. W. *Biochemistry* **1988**, *27*, 8063–8068.
(S9) Pace, C. N.; Shirley, B. A.; Thomson, J. A. In *Protein Structure*; Creighton, T. E., Ed.; IRL Press: New York, 1989, pp 311–330.
(S10) Koerber, S. C.; Fink, A. L. *Anal. Biochem.* **1987**, *165*, 75–87.
(S11) Liu, W.; Tsou, C. L. *Biochim. Biophys. Acta* **1987**, *916*, 465–473.

**Figure Legends**

***Figure S1.*** MALDI mass spectrum of dmP114 RNase A.

***Figure S2.*** Residual activity of wild-type RNase A (●) and dmP114 RNase A (△) as a function of the concentration of Gdn–HCl. Activity was determined at 25 °C with 6-carboxyfluorescein–dArU$(dA)_2$–6-carboxytetramethylrhodamine as substrate in 0.10 M MES–NaOH buffer (pH 6.0) containing NaCl (0.10 M).

***Figure S3.*** Eyring plot for the unfolding of wild-type RNase A (●) and dmP114 RNase A (△) at pH 8.0.

***Figure S4.*** Fluorescence properties at 25 °C of dmP114 RNase A as a function of the concentration of Gdn–HCl in 0.10 M MES–NaOH buffer (pH 6.0) containing NaCl (0.10 M). A. Fluorescence emission spectra. B. Transition curve determined from the change of the emission intensity at 303 nm (insert).

***Figure S5.*** Reactivation of unfolded dmP114 RNase A (top) and wild-type RNase A (bottom) at 5 °C. Upper panels: Lines indicate the fit of the raw data to eq S1. These data are also shown in Figure 2. Lower panels: Deviation of the raw data from the fit to eq S1.

***Figure S1***

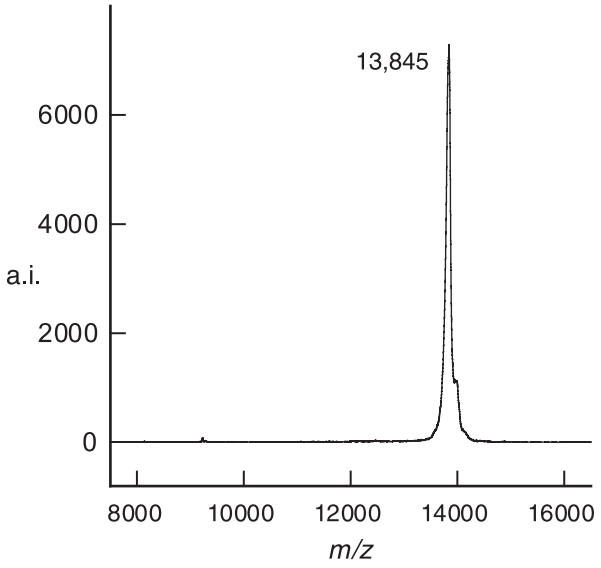


***Figure S2***

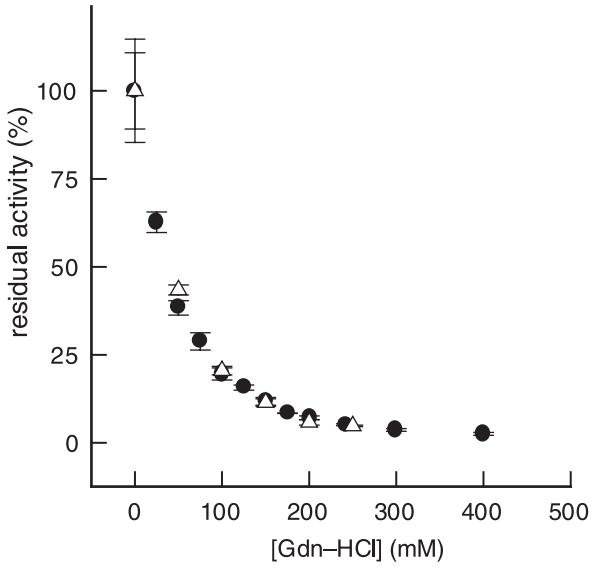

residual activity (%)
100
75
50
25
0
0
100
200
300
400
500
[Gdn–HCl] (mM)

***Figure S3***

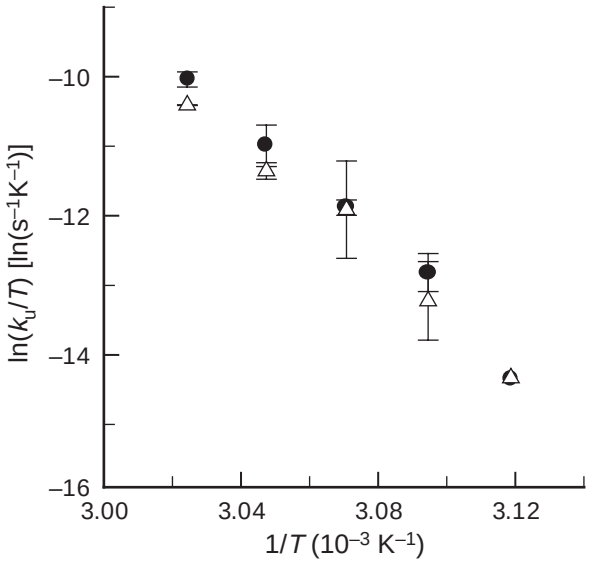


***Figure S4***

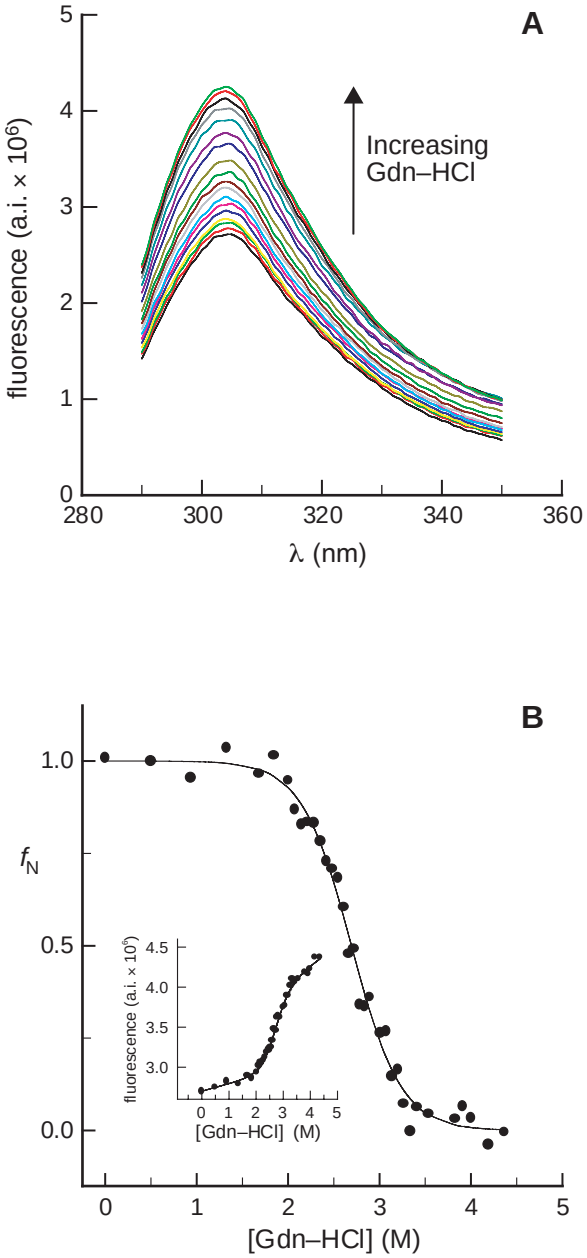
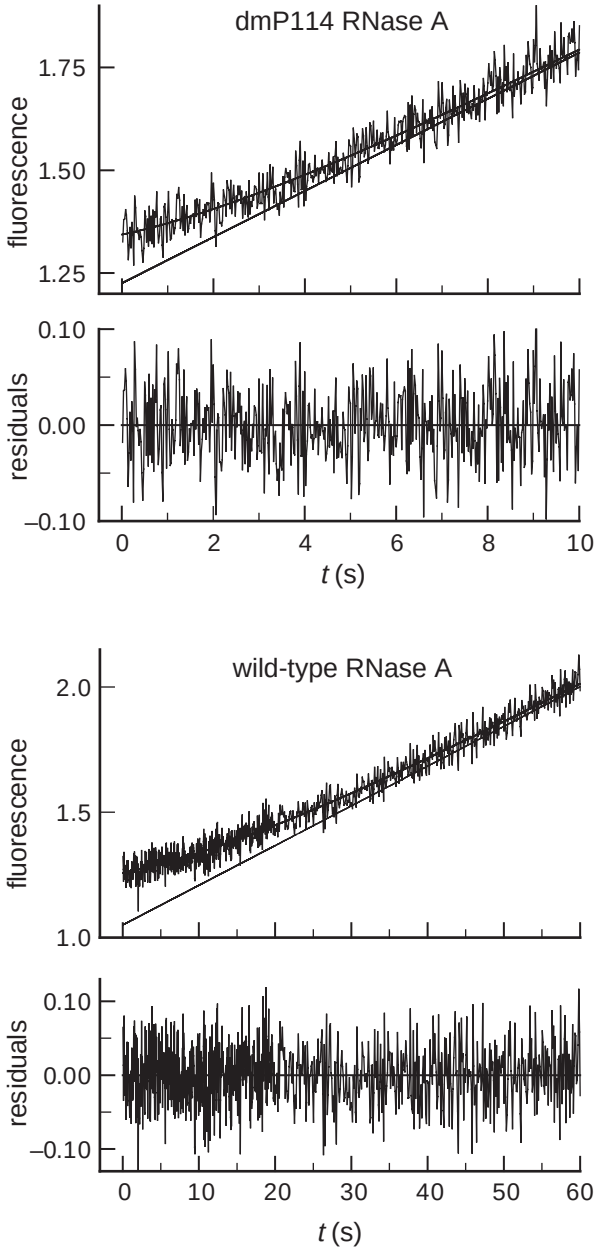

A
fluorescence (a.i. × 10^6)
λ (nm)
Increasing
Gdn–HCl
B
$f_N$
[Gdn–HCl] (M)
fluorescence (a.i. × 10^6)
[Gdn–HCl] (M)

***Figure S5***


dmP114 RNase A
fluorescence
1.75
1.50
1.25
residuals
0.10
0.00
–0.10
0
2
4
6
8
10
t (s)
wild-type RNase A
fluorescence
2.0
1.5
1.0
residuals
0.10
0.00
–0.10
0
10
20
30
40
50
60
t (s)