SUPPORTING MATERIAL

Dynamics of thermodynamically stable, kinetically trapped and inhibitor-bound states of pepsin

Derek R. Dee, Brenna Myers and Rickey Y. Yada

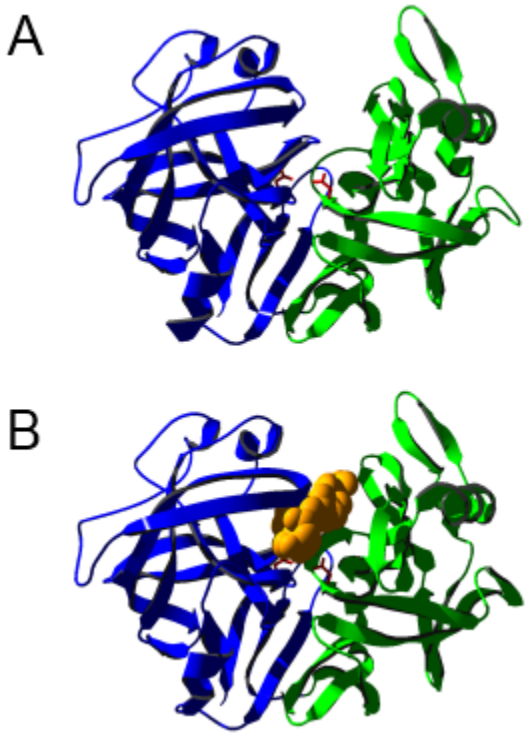

FIGURE S1 Crystal structures of (A) human pepsin (PDB: 1PSN) and (B) human pepsin-pepstatin (PDB: 1PSO). The N- and C-terminal domains are indicated in blue and green, respectively, while the catalytic D32 and D215 reside side-chains are shown in red.

QENS sample preparation

Commercial porcine pepsin was further purified using an ACTA FPLC (GE Healthcare, Chalfont St. Giles, UK) and either gel filtration or ion-exchange chromatography. Gel filtration was performed using a Superdex-75 gel filtration column (GE Healthcare, Chalfont St. Giles, UK) and a volatile buffer system consisting of 100 mM acetic acid titrated to pH 5.55 with ammonium hydroxide. 10 mg samples were loaded onto the column at a flow rate of 0.8 ml/min. Fractions were stored at 4°C and were then combined to obtain a uniform protein concentration, dispensed into 1 ml aliquots in 2 ml Eppendorf tubes, flash frozen by immersion in liquid nitrogen and were lyophilized. Ion-exchange chromatography was carried out using a MonoQ 10/100GL (GE Healthcare, Chalfont St. Giles, UK) ion-exchange column with buffers (A) 100

mM acetic acid/NH_4OH pH 5.5 and (B) 100 mM acetic acid/NH_4OH pH 5.5, 100 mM NaCl and a single-step gradient from 100% buffer A to 100% buffer B. Approximately 100 mg of pepsin was loaded onto the column in a volume of 8 ml with a constant flow rate of 3 ml/min. Each run yielded one 15 ml fraction, and different fractions were combined and subjected to buffer exchange, using an Amicon stirred cell (Millipore Corp., Bedford, MA), to reduce the NaCl content to below 0.2 mM. Samples were then flash frozen, lyophilized and stored at –20 ºC.

Ip was obtained by adding 200 µl of 20 mM NH_4OH to 4 mg of purified and lyophilized pepsin resulting in a final pH of 8. Rp was obtained from the Ip solution by adding 1.3 ml of 40 mM acetic acid/NH_4OH buffer of pH 5.3. The pH at each step was monitored by using pH indicator paper. Ip and Rp solutions were then flash frozen and lyophilized.

To replace exchangeable H atoms with D atoms prior to the neutron scattering experiments, samples were treated with three rounds of re-dilution with D_2O buffer, flash freezing and lyophilization. Np and Rp samples were diluted with 20 mM acetic acid/NH_4OH in D_2O, pD 5.5, while Ip samples were diluted with 60 mM NH_4OH in D_2O, with a final pD of 8. NpP samples were made in D_2O by dissolving purified Np in 40 mM acetic acid/ NH_4OH pD 5.5 and adding an aliquot of pepstatin/ethanol to give a ratio of 1:1.5 of pepsin:pepstatin. The NpP samples were then subjected to three rounds of H/D exchange by lyophilization.

Samples were diluted to a concentration of 50 mg/ml (Np, Rp and Ip) or 100 mg/ml (Np and NpP) in D_2O buffer consisting of 100 mM acetic acid-D4/NaOD with a final pD of 5.55 (Np, Rp and NpP) or 8 (Ip). The 50 mg/ml solutions were centrifuged at 10,000 rpm for 10 minutes and passed through a 0.22 µm filter and then loaded into an annular aluminium can with a 0.1 mm sample thickness and ~ 1 ml sample volume for DCS and an annular can (with a flanged inner can) of 0.5 mm sample thickness and a volume of ~ 2 ml for HFBS. The 100 mg/ml samples were centrifuged at 12,000 rpm for 10 minutes to remove the small amount of insoluble protein and were then loaded into an annular aluminium can with a thickness of 0.4 mm and volume of 3.5 ml for DCS and an annular can (with a 'DCS-type', flangeless inner can), with a thickness of 1 mm and volume of 4 ml for HFBS.

Quasielastic neutron scattering

Neutrons with wavelengths comparable to interatomic distances have energies proportional to molecular vibrational, rotational and translational energy levels and can be used to probe both the structure and dynamics of biomolecules. Scattered neutrons are characterized by an energy transfer, $\hbar\omega = E - E_0 = \hbar^2/2m\ (k_f^2 - k_i^2)$, and a momentum transfer, $\hbar Q = \hbar k_f - \hbar k_i$, where ω is the neutron frequency, Q is the scattering vector, $\hbar\omega$ is the energy transferred between the scattered neutron and the sample, $\hbar = h/2\pi$ where h is Planck's constant, and k_i and k_f are the incident and scattered wave vectors, respectively. Neutron scattering can be separated into coherent and incoherent contributions, giving rise to cross-correlation and self-correlation scattering functions, respectively. A feature particularly important to studies of dynamics is that H has a much larger incoherent scattering power than other atoms, including D. Thus, incoherent neutron scattering predominantly follows the motions of H atoms. Furthermore, neutron scattering experiments are usually performed in D_2O to minimize the background signal and samples are pre-equilibrated in D_2O buffer to replace all exchangeable H with D. As backbone H atoms generally exchange very quickly with D atoms, neutron scattering from protein solutions is dominated by the incoherent scattering from H atoms of side chain groups (1). QENS from proteins is largely incoherent scattering from H and analysis of the scattering data can be

simplified to include only the incoherent contribution. An analytical model convoluted with an experimentally determined resolution function is used to fit the measured data:

$$S_{measure}(Q,\omega) = \left[S_{theory}(Q,\omega) \otimes S_{resolution}(Q,\omega) \right]. \qquad (16)$$

The analytical scattering function, $S_{theory}(Q,\omega)$, contains terms for each of the elastic, quasielastic and inelastic contributions to the incoherent scattering (2):

$$S_{\text{theor}}(Q,\omega) = e^{-Q^2\langle u^2\rangle/3} \cdot \left[A_0(Q) \cdot \delta(\omega) + \sum_{i=1} A_{\text{i}}(Q) \cdot L_{\text{i}}(\Gamma_{\text{i}}, \omega) \right]. \qquad (17)$$

The elastic component is described by a delta function, $\delta(\omega)$, and the quasielastic components are represented as a sum of Lorentzian functions $L_{\text{i}}(\Gamma_{\text{i}}, \omega)$ of half-width at half maximum (HWHM), Γ_{i}, where Γ_{i} is inversely proportional to the correlation time of motion, τ. The other terms have the same meaning as described in the main text.

Global Motions

It is often assumed that, in a comparative QENS study focused on ps internal dynamics, the effect of slower, ns whole-particle diffusive motions can be neglected (3-6). For the TOF experiments described here, which measured motions faster than 100 ps, whole-particle motions would be characterized by energy transfers much smaller than that of the experimental resolution (i.e., less than 57 μeV), and would contribute mainly to the elastic scattering. However, the HFBS measurements could be sensitive to translational motions as correlation times up to 8 ns were measured. Whole-particle diffusive motions would be faster for the more compact states, Np and Rp, which were shown to have reduced radius of gyration values compared to Ip (7); therefore, these motions would contribute to an overestimation of the quasielastic broadening in Np compared to Ip. However, Ip showed a larger FWHM than Np (**Fig. 3B** from the main text), indicating that internal motions dominated the scattering. Furthermore, the finding that the FWHM tend towards non-zero values as $Q \rightarrow 0$ Å^{-1} supports that whole-particle diffusive motions (translational and rotational) can safely be neglected in a comparative analysis (5, 6, 8). This is due to the unconstrained versus constrained nature of external and internal motions: if the whole-particle position is fixed in space (i.e., the protein is not moving on a given timescale), the only motions detected are those due to internal diffusion and these motions give rise to finite FWHM towards small Q/large length scales because they are restricted in space (2). Additional considerations are that whole-particle diffusion decreases with the increased viscosity of D_2O versus H_2O solutions and with increasing protein concentration (6).

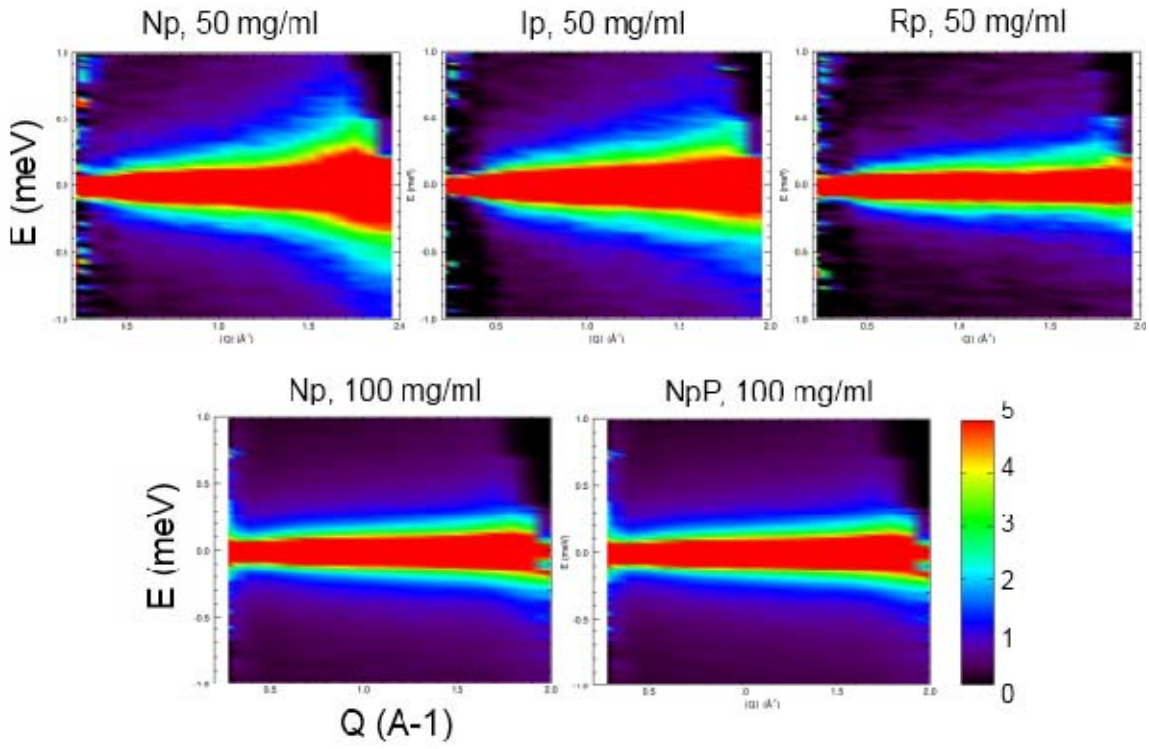

FIGURE S2 TOF-QENS scattering profiles after correction for detector efficiency and background and solvent scattering. The scattering spectra, $S_{protein}(Q, \omega)$, are shown vs. energy transfer, E (meV), and scattering vector, Q (Å^{-1}). Intensity coloring is according to the scale bar in the lower right panel.

QENS Curve Fitting

The reduced and corrected difference spectra were fit according to Eq. 3 in the Q-range 0.4 to 1.8 Å^{-1}, and in the energy range –0.9 to +0.9 meV, except for the Q = 1.8 Å^{-1} spectrum which was restricted to –0.9 to +0.7 meV. The data were adequately fit with a Gaussian, a Lorentzian and a flat background (**Fig. S3**), representing the elastic, quasielastic and inelastic scattering, respectively

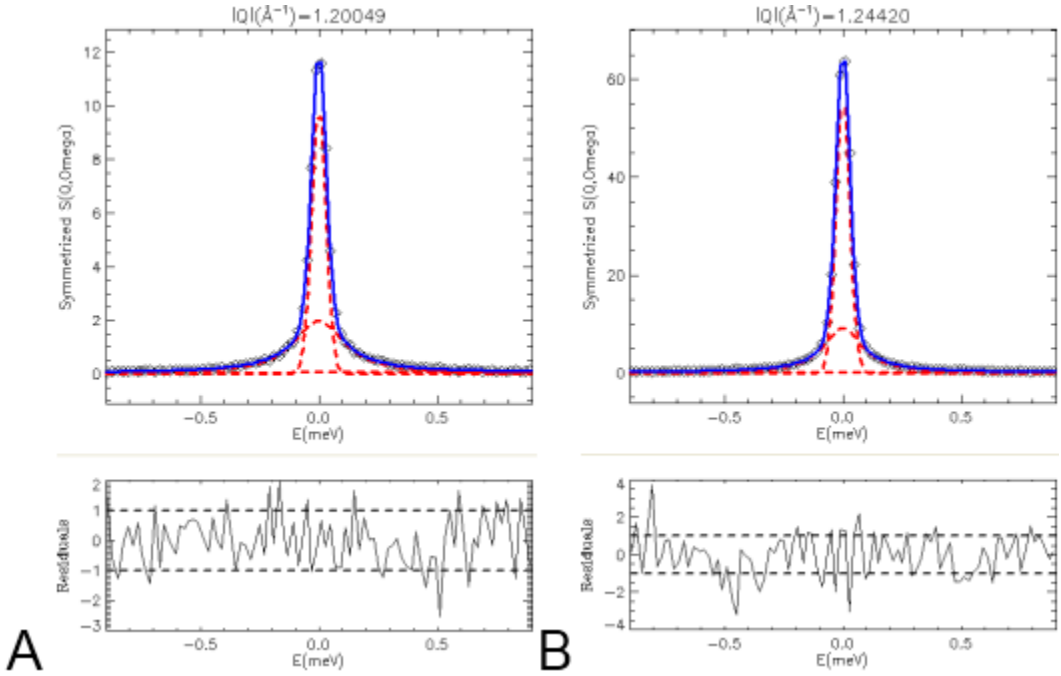

FIGURE S3 QENS curve-fitting procedure. The TOF spectra were fit using a flat background, a Gaussian and one Lorentzian. Examples of fitting are shown for Np at (A) 50 mg/ml and (B) 100 mg/ml.

A 'shoulder' in the EISF

For each of the pepsin samples measured, the EISF followed a general trend with Q in that the values decreased at low Q, began to plateau around 1.1 Å$^{-1} \leq Q \leq$ 1.3 Å^{-1}, and then continued to decrease towards high Q giving the appearance of a 'shoulder' in the EISF (**Fig. 2B** from the main text). This result contrasts the EISF trends reported for other proteins in solution (4, 5), in which the EISF tends toward a plateau and can be adequately accounted for by Eq. 10 or similar models. However, data on *Staphylococcal* nuclease in solution also resulted in a shoulder in the EISF around Q = 1.5 Å^{-1}, and a mixed analytical model, combining vibrational, diffusion in a spherical volume, and jump diffusion motions, was required to fit the EISF (9). Data from native dihydrofolate reductase (DHFR) (8) and aspartate transcarbamylase (10) also gave a shoulder in the EISF, although the data were fit with the DSV model only. No explanation of this feature was given in (8), while the authors in (10) proposed that the shoulder feature was due to a small

coherent scattering contribution. However, the coherent contribution is generally considered negligible (11) and a definitive explanation for the shoulder feature is currently lacking. Further examples of proteins that gave a shoulder in the EISF include yeast phosphoglycerate kinase (12), hydrated bacteriorhodopsin (13) and ribonuclease A (14), although no discussion of this feature was given in these reports.

Analysis of *Q*-binned data to compare Np and NpP

From the reduced TOF data (i.e., **Fig. S2**) it appeared that there were differences in the Np and NpP at 100 mg/ml scattering that may not have been characterized by fitting the *Q*-dependent spectra, particularly as the FWHM were fixed to similar values. Thus, for comparison, the *Q*-binned, single spectra of Np and NpP were also analyzed. The 913 DCS spectra for Np and NpP were reduced, buffer subtracted and combined into one spectrum covering the *Q*-range from 0.4 to 1.8 $Å^{-1}$ and fit according to Eq. 3. The data were thus averaged over all measured scattering angles, such that spatial information could not be obtained, but the counting statistics were greatly improved. The results, as shown in **Table S1**, indicated that Np was characterized by a relatively greater quasielastic component (smaller EISF) and diffusive motions with a faster averaged correlation time, ‹τ›.

Similarly, the angular independent, *Q*-binned HFBS spectra were also compared as the angular dependent analysis of the FWHM and EISF for Np and NpP did not reveal large differences between the samples. Of the 16 HFBS detectors, numbers 3 to 13 were grouped, corresponding to a *Q*-range of 0.47 to 1.51 $Å^{-1}$, and the resulting spectra were analyzed according to Eq. 3. Similarly to the *Q*-dependent analysis, the observed changes in dynamics were relatively small. However, Np was characterized by a larger FWHM than for NpP, and thus by a faster averaged correlation time, ‹τ› (**Table S2**).

TABLE S1 Fitting results from the *Q*-binned Np and NpP spectra, averaged over Q = 0.4 $Å^{-1}$ to 1.8 $Å^{-1}$*.

	Np	**NpP**
EISF	0.637 ± 0.002	0.656 ± 0.003
FWHM (meV)	0.195 ± 0.002	0.172 ± 0.002
‹τ› (ps)	42.5 ± 0.4	48.3 ± 0.5

*Uncertainties indicate fitting errors.

TABLE S2 Fitting results from the Q-binned HFBS data from Np and NpP, averaged over Q = 0.47 Å^{-1} to 1.51 Å^{-1}*.

	Np	**NpP**
EISF	0.130 ± 0.006	0.120 ± 0.006
FWHM (μeV)	8.061 ± 0.174	7.321 ± 0.166
‹τ› (ns)	1.03 ± 0.02	1.13 ± 0.03

*Uncertainties indicate fitting errors.

Differential scanning calorimetry

Calorimetry experiments were carried out with a MicroCal VP-DSC (MicroCal, Northampton, MA) using a scan rate of 1.5 °C/min. Protein concentrations ranged from 0.4 to 1.4 mg/ml, as measured by $A_{280\ nm}$, with $E^{0.1\%}$ = 1.529. Pepstatin solutions were made by dissolving 2 mg in 1 ml of ethanol and heating to 60 ºC. NpP samples were made by combining an aliquot of the pepstatin/ethanol solution with pepsin in 100 mM acetic acid/NaOH, pH 5.3, to give a molar ratio, pepsin:pepstatin, of 1:1.3. Samples were dialyzed overnight in the buffer used as the reference for DSC experiments. Data analysis was performed using MicroCal Origin v7.0 (OriginLab, Northampton, MA). Measurements were repeated three times for each sample. The DSC scans for Np and NpP are shown in **Fig. S4**.

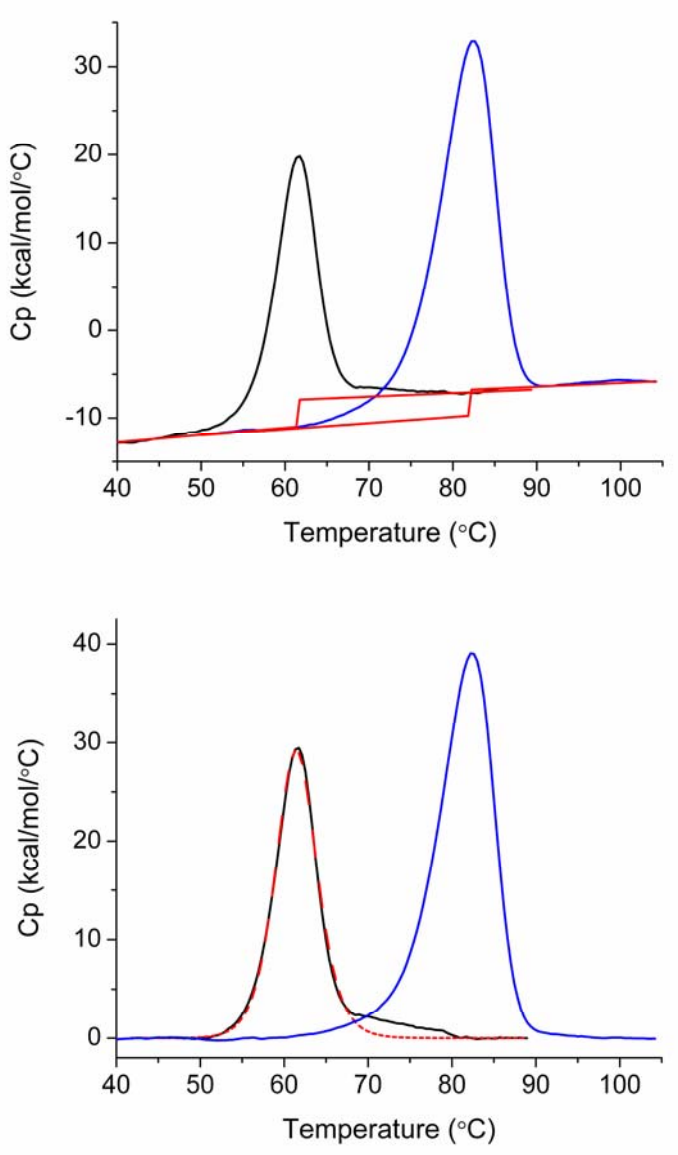

FIGURE S4 DSC measurements of Np (black) and NpP (blue) in 100 mM acetic acid/NaOH, pH 5.3 buffer. Top: buffer-subtracted, normalized and fit with a linear 'step-at-peak' baseline (red) to emphasize the change in heat capacity of unfolding, ΔC_p. Bottom: buffer-subtracted, normalized and with a cubic polynomial baseline subtracted. NpP was analyzed by direct integration of the data, whereas Np was fit according to a non-two state model (red, dashed curve) in order to avoid the post-T_m broadening that was attributed to the presence of Rp (7).

H/D-exchange monitored by FTIR

FT-IR spectra were recorded with a Bruker IFS66vs spectrometer with a DTGS detector from 1000 to 2000 cm^{-1} at 4 cm^{-1} resolution. Lyophilized protein samples were dissolved in D_2O buffer (100 mM acetic acid-D4/NaOD, pD 5.4), loaded between CaF_2 windows with a 56 µm spacer, and the measurements were initiated within four minutes. Sample temperature was

maintained at 25 °C using a circulating water bath. Measurements were repeated three times for each sample.

The exchange kinetics of backbone amide H with D upon transferring a protein into a deuterated solution is a function of the pH, temperature and the accessibility of the H to solvent (15). Thus, H located in the solvent-excluded hydrophobic core of a protein should exchange only upon becoming exposed due to conformational fluctuations of the protein and monitoring H/D exchange kinetics can give information on the slow timescale (> ms) fluctuations of a protein (4).

FTIR spectra were measured for Np, NpP and Rp upon transfer into D_2O buffer, in which the amide II band, around 1550 cm^{-1}, was expected to decrease and the amide II' band, around 1450 cm^{-1}, would increase upon H/D exchange. Although spectra were recorded over a period of 20 hours, only the first spectrum for each sample, at $t \sim 4$ minutes, is shown in **Fig. S5**. It can be seen that by the time of the first scan the amide II intensity of Np was already diminished considerably while that of NpP and Rp remained visible. Qualitatively this supports that Np is more flexible than either Rp or NpP on slower timescales. Combined with the findings from the QENS data, which indicated that Np was more flexible on a ps timescale, it appears that Np is overall more flexible than Rp.

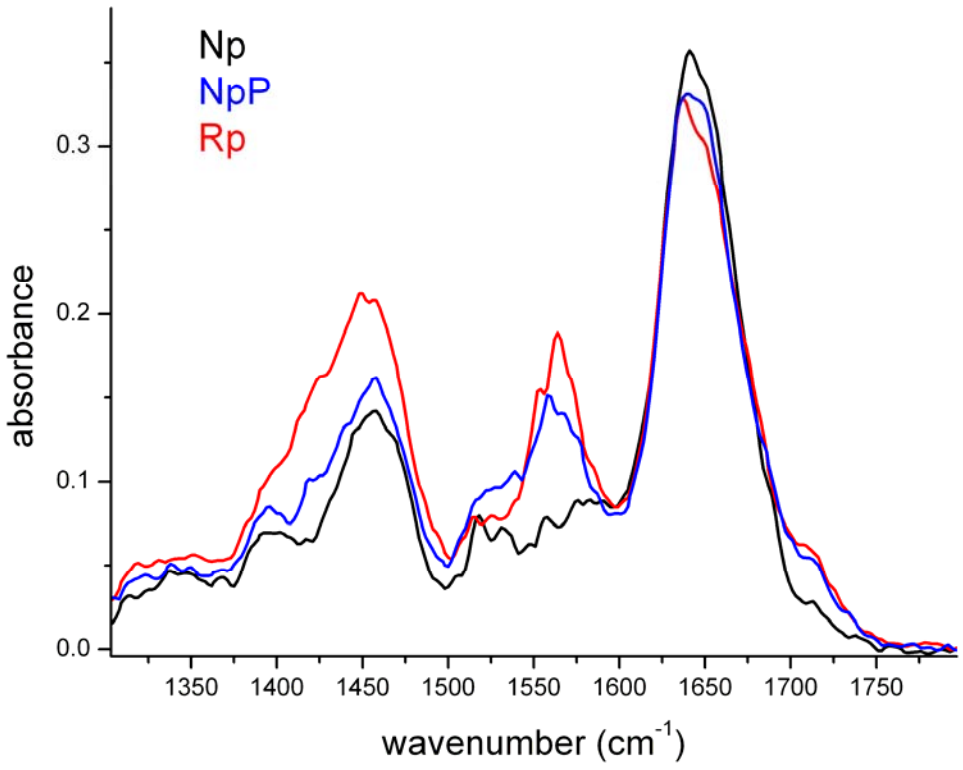

FIGURE S5 FTIR spectra for Np, Np-pepstatin (NpP) and Rp. Spectra were recorded within 4 minutes after dissolving the lyophilized proteins in D_2O buffer (100 mM acetic acid-D4/NaOD, pD 5.5). Spectra were normalized according to the absorbance at 1640 cm^{-1}. Representative spectra are shown.

REFERENCES

1. Gabel, F., D. Bicout, U. Lehnert, M. Tehei, M. Weik and G. Zaccai. 2002. Protein dynamics studied by neutron scattering. Quart. Rev. Biophys. 35:327-367.
2. Bée, M. 1988. Quasielastic Neutron Scattering: Principles and Applications in Solid State Chemistry, Biology and Materials Science. Adam Hilger, Bristol, England.
3. Fitter, J. 2003. A measure of conformational entropy change during thermal protein unfolding using neutron spectroscopy. Biophys. J. 84:3924-3930.
4. Fitter, J. and J. Heberle. 2000. Structural equilibrium fluctuations in mesophilic and thermophilic alpha-amylase. Biophys. J. 79:1629-1636.
5. Bu, Z., D.A. Neumann, S. Lee, C.M. Brown, D.M. Engelman and C.C. Han. 2000. A view of dynamics changes in the molten globule-native folding step by quasielastic neutron scattering. J. Mol. Biol. 301:525-536.
6. Tehei, M., R. Daniel and G. Zaccai. 2006. Fundamental and biotechnological applications of neutron scattering measurements for macromolecular dynamics. Eur. Biophys. J. 35:551-558.
7. Dee, D., J. Pencer, M. Nieh, S. Krueger, J. Katsaras and R.Y. Yada. 2006. Comparison of solution structures and stabilities of native, partially unfolded and partially refolded pepsin. Biochemistry. 45:13982-13992.
8. Tehei, M., J.C. Smith, C. Monk, J. Ollivier, M. Oettl, V. Kurkal, J.L. Finney and R.M. Daniel. 2006. Dynamics of immobilized and native *Escherichia coli* dihydrofolate reductase by quasielastic neutron scattering. Biophys. J. 90:1090-1097.
9. Kataoka, M., M. Ferrand, A.V. Goupil-Lamy, H. Kamikubo, J. Yunoki, T. Oka and J.C. Smith. 1999. Dynamical and structural modifications of *Staphylococcal* nuclease on C-terminal truncation. Physica B. 266:20-26.
10. Zanotti, J. M., G. Herve and M.C. Bellissent-Funel. 2006. Picosecond dynamics of T and R forms of aspartate transcarbamylase: a neutron scattering study. Biochim. Biophys. Acta. 1764:1527-1535.
11. Dunn, B. M., M. Lewitt and C. Pham. 1983. Inhibition of pepsin by analogues of pepsinogen-(1-12)-peptide with substitutions in the 4-7 sequence region. Biochem J. 209:355-362.
12. Receveur, V., P. Calmettes, J.C. Smith, M. Desmadril, G. Coddens and D. Durand. 1997. Picosecond dynamical changes on denaturation of yeast phosphoglycerate kinase revealed by quasielastic neutron scattering. Proteins. 28:380-387.
13. Fitter, J., R.E. Lechner, G. Buldt and N.A. Dencher. 1996. Internal molecular motions of bacteriorhodopsin: hydration-induced flexibility studied by quasielastic incoherent neutron scattering using oriented purple membranes. Proc. Natl. Acad. Sci. U. S. A. 93:7600-7605.
14. Wood, K., C. Caronna, P. Fouquet, W. Haussler, F. Natali, J. Ollivier, A. Orecchini, M. Plazanet and G. Zaccai. 2008. A benchmark for protein dynamics: ribonuclease A measured by neutron scattering in a large wavevector-energy transfer range. Chem. Phys. 345:305-314.
15. Barth, A. 2007. Infrared spectroscopy of proteins. Biochim. Biophys. Acta. 1767:1073-1101.