Changes in dynamics of α-chymotrypsin due to covalent inhibitors investigated by elastic incoherent neutron scattering

C. D. Andersson, Nicolas Martinez, Dominik Zeller, Stina Holmgren Rondahl, Michael M. Koza, Bernhard Frick, Fredrik Ekström, Judith Peters, Anna Linusson

Changes in dynamics of α-chymotrypsin due to covalent inhibitors investigated by elastic incoherent neutron scattering

C. D. Andersson[a], N. Martinez[bc], D. Zeller[bc], S. H. Rondahl[d], M. M. Koza[b], B. Frick[b], F. Ekström[d], J. Peters[bc] and A. Linusson[a]

An essential role of enzymes is to catalyze various chemical reactions in the human body and inhibition of the enzymatic activity by small molecules is the mechanism of action of many drugs or tool compounds used to study biological processes. Here, we investigate the effect on the dynamics of the serine protease a-chymotrypsin when in complex with two different covalently bound inhibitors using elastic incoherent neutron scattering. The results show that the inhibited enzyme displays enhanced dynamics compared to the free form. The difference was prominent at higher temperatures (240–310 K) and the type of motions that differ include both small amplitude motions, such as hydrogen atom rotations around a methyl group, and large amplitude motions, such as amino acid side chain movements. The measurements were analyzed with multivariate methods in addition to the standard univariate methods, allowing for a more in-depth analysis of the types of motions that differ between the two forms. The binding strength of an inhibitor is linked to the changes in dynamics occurring during the inhibitor-enzyme binding event and thus these results may aid in the deconvolution of this fundamental event and in the design of new inhibitors.

1. Introduction

Enzymes are dynamical by nature and catalyze various chemical reactions in biological systems, *in vivo* or in biochemical applications *in vitro*. The motions of enzymes are functionally important and especially in the events leading up to catalysis, where conformational changes in the enzymes may facilitate or impede this process.[1–3] Inhibition of the catalytic ability by reversible (non-covalent) or irreversible (covalent) inhibitors is an invaluable method to control enzymatic activity and to study enzymatic function. Understanding these mechanisms is of fundamental importance in many research fields, including drug discovery,[4] enzyme engineering,[5] and enzyme function studies, since it may aid in the design of more potent inhibitors and functionally modified enzymes. The effects of inhibitor binding on local and overall enzyme dynamics have been investigated with various experimental techniques including nuclear magnetic resonance (NMR)[6] and incoherent neutron scattering.[7] It has become evident that these effects are not easily predicted and that enzyme inhibition can cause increased,[8–10] decreased,[11,12] or no change in dynamics.[13,14] For instance, we have recently shown with neutron scattering experiments that the covalently bound nerve agent Soman causes a stiffening of the enzyme acetylcholinesterase,[15] and does not affect the dynamics of butyrylcholinesterase at comparable temperature ranges.[16] Here, we have investigated the dynamical effect that covalent inhibition causes in the serine protease α-chymotrypsin.

Classic serine proteases such as chymotrypsin, subtilisin, carboxypeptidase Y, and Clp protease, are involved in biological processes including digestion, reproduction, immune response, apoptosis and hemostasis,[17] where they degrade proteins and peptides by cleaving peptide bonds. Serine proteases bearing the chymotrypsin fold (*e.g.* chymotrypsin, trypsin, thrombin, and elastase) belong to the clan PA, and the cleavage reaction these enzymes perform is catalyzed by a clan-specific catalytic triad consisting of amino acids His–Asp–Ser present in the enzymes' active site.[18] This clan of enzymes has been extensively studied and includes targets for drug research programs,[19] for example, lowering of blood pressure (thrombin and coagulation factor Xa), and pancreatitis (trypsin-like). These enzymes are also of interest in biotechnological applications making use of their catalytic ability in, for example, prodrug design,[20–22] enzyme engineering,[23,24] and functional nanomaterials.[25] The structure of α-chymotrypsin was revealed in 1967[26] and the enzyme consists of two longer and one short 13 amino acid chains (Fig. 1). The catalytic site is

[a] *Department of Chemistry, Umeå University, SE-90187 Umeå, Sweden. E-mail: anna.linusson@umu.se*
[b] *Institut Laue Langevin, F-38042 Grenoble Cedex 9, France. E-mail: peters@ill.fr*
[c] *Univ. Grenoble Alpes, IBS and LiPhy, F-38044 Grenoble, France*
[d] *CBRN Defence and Security, Swedish Defence Research Agency, SE-90621 Umeå, Sweden*

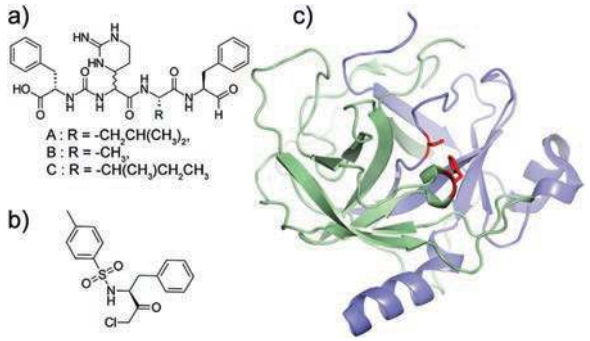

Fig. 1 Molecular structure of (a) chymostatin, a mixture of molecules A, B and C with different side chains indicated by R, (b) TPCK, and (c) α-chymotrypsin (PDB code 1YPH)[33] where the three protein chains are colored in white, green and blue, respectively, and the amino acids that participate in catalysis and that the inhibitors bind to are indicated in red.

positioned between the two long chains in a shallow gorge, with a deeper part where inhibitors or parts of inhibitors may bind.[27]

Numerous investigations have dealt with the binding mechanism of inhibitors of α-chymotrypsin[28,29] and the inhibitor-enzyme complex that are formed (see for instance ref. 27 and 30–32) but not much is known about the effect of inhibitor-binding on the enzyme dynamics at sub-nanosecond timescales. In this study, the motions of α-chymotrypsin were investigated in its free form and when inhibited with two different covalent inhibitors, 1-chloro-3-tosylamido-4-phenyl-2-butanone (TPCK) and chymostatin (Fig. 1a and b). These two inhibitors bind in the enzyme catalytic site; TPCK to His57[28] and chymostatin to Ser[195] (Fig. 1c).[29] The two inhibitors were selected based on that they bind to different amino acids and differ in size and, thus, they supposedly perturb the structure of α-chymotrypsin differently.

The flexibility of α-chymotrypsin was investigated with neutron spectroscopy, which is a powerful method to probe molecular motions on the atomic level. The incoherent neutron scattering cross section of the hydrogen atom is much higher than that of all other types of atoms present in biological systems,[34] and thus incoherent neutron scattering experiments reveal mainly the motions of hydrogen atoms and of the molecular groups to which they are bound. Enzymatic activities occur on micro- to millisecond time scale, but a recent study supported a correlation between activity and molecular dynamics at a nanosecond time scale within the family of human cholinesterases.[35] To evaluate the results of the neutron scattering measurements we used the classical univariate analysis methods such as summed intensities and atomic mean-square displacements (MSD).[36] In addition, we used multivariate methods including principal component analysis (PCA)[37,38] and orthogonal partial least squares discriminant analysis (OPLS-DA),[39,40] methods previously not used in this context, to elucidate more detailed effects of the inhibitors on the dynamics of α-chymotrypsin on the pico- to nanosecond time scale. In this way, we explored the data using two substantially different methods, the first one averaging over limited range of scattering angles, the second one exploiting all accessible angles or scattering vectors in a statistical meaningful way. This is an alternative to the self-distribution function procedure,[41] which permits to separate different molecular motions in various time windows from elastic incoherent neutron scattering data. Another successful method was suggested by L. Hong *et al.*[42] to identify global and internal protein motions from spin-echo and simulation data through wave vector dependent diffusion coefficients. Here we wanted to investigate how far an analysis distinguishing all wave vectors could permit to shed light on different motions and contributions within a rather limited space-time window defined through the resolutions of the three spectrometers used.

2. Experimental section

2.1 Sample preparation

Chymotrypsin is readily available as powder and very stable as a solid and in solution (pH 2–9) at temperatures up to 54.3 °C,[43] making the enzyme suitable for neutron scattering experimental conditions. 500 mg of α-chymotrypsin from Bovine pancreas and treated with 1-chloro-3-tosylamido-7-amino-2-heptanone to inhibit residual trypsin activity (Worthington Biochemical Corporation) was dissolved in 24 mL of a 25 mM ammonium acetate/deuterated water (D_2O) buffer. After 10 min of stirring on ice, the sample flash frozen in liquid nitrogen and freeze dried for 24 h. The dry protein powder was stored in −20 °C. Four samples were prepared from this dry protein powder: two equivalent samples including the inhibitor-free α-chymotrypsin, hereafter called CT1 and CT2, one sample with chymotrypsin inhibited by 1-chloro-3-tosylamido-4-phenyl-2-butanone (TPCK, CAS 402-71-1) and one with chymotrypsin inhibited by chymostatin (CAS 9076-44-2), giving samples CT/TP and CT/CS, respectively (Fig. 1). The average molecular weight (MW) of the three compounds in chymostatin (A, MW 607.7, 82%; B, MW 593.7, 11%; and C, MW 607.7, 7%) was estimated to 606.16 g mol^{-1}. Inhibitor stock solutions were prepared as follows. 6.06 mg (0.01 mmol) chymostatin was dissolved in 300 μL dimethylsulfoxide (DMSO) and then dissolved in ammonium acetate/D_2O buffer. 3.52 mg (0.01 mmol) TPCK was dissolved in 267 μL DMSO/EtOH 1 : 1 and then suspended in 12 mL ammonium acetate/D_2O buffer (resulting in a turbid solution). The inhibited chymotrypsin samples were prepared by addition of two (mole) equivalents of inhibitor to chymotrypsin in solution; dry chymotrypsin was dissolved in the inhibitor solution whereby all solutions became clear. Samples were freeze dried according to the protocol above. Protein samples were transferred to flat neutron scattering sample holders (3×4 cm^2) of 1 mm thickness and put under vacuum until completely dry (monitoring sample weight) followed by rehydration with D_2O to a final amount of D_2O/sample ratio (g g^{-1}) of 0.40, corresponding to at least one hydration layer around the protein. Final sample weights in the sample holder were 126.5 mg (CT1), 163.4 mg (CT2), 185.2 mg (CT/CS), and 147.9 mg (CT/TP).

2.2 Elastic incoherent neutron scattering experiments

We measured our enzyme samples with elastic incoherent neutron scattering (EINS) on three instruments, IN6,[44] IN13,[45] and IN16B[46] at the Institut Laue Langevin (ILL, France). The three

experiments were done consecutively using the same samples. The highest temperature reached was 310 K, corresponding to the human body temperature; therefore, no enzyme denaturation was expected. Both inhibited α-chymotrypsin samples (CT/TP and CT/CS) and one "free" α-chymotrypsin sample (CT2) were measured on all three spectrometers. The additional α-chymotrypsin sample (CT1) was included in the IN6 experiment to analyze inter-sample variation. The three spectrometers were chosen for their distinctly different time and spatial scale coverage. All three instruments are sensitive to local motions, but they give different information regarding the time window in which motions of different amplitudes occur. IN6 is a cold neutron time-of-flight spectrometer permitting to determine the incident wavelength and, thus, the instrumental energy resolution by a monochromator. In the present experiment, we opted for an incident wavelength of 5.1 Å corresponding to an energy resolution- and time window of 90 μeV and 7.5 ps (calculated according to Magazù *et al.*,[47]), respectively. Motions at this time scale correspond to very fast local motions of atoms. The elastic scattering was obtained on this time-of-flight spectrometer by integrating the peak in the range from −0.12 to 0.12 meV. A figure presenting IN6's resolution function and the integration limits is shown in the ESI† (Fig. S1). Each instrument gives access to a different Q range, Q being the momentum transfer between the incident and the scattered neutron in units of $\hbar$. Q is inversely related to space dimensions or in our case to the amplitudes of motion, which can be calculated through Bragg's law $L \propto 2\pi/Q$.[48] IN6 permits to measure intensities at Q values of the detector in between 0.4 and 2.0 $Å^{-1}$, corresponding to length scales between 16 and 3 Å. IN13 is a thermal backscattering spectrometer using the wavelength of 2.23 Å. The time window is about 100 ps and it covers a spatial scale of 1–30 Å. Such a time range typically covers local elastic vibrations and rotations of amino acid side chains at the surface.[49] IN16B has the highest energy resolution of about 1 μeV full width at half maximum and gives access to motions up to 1 ns, with length scales comparable to IN6 (3–16 Å). These motions can be associated to torsional vibrations of buried molecular sub-groups.[49] Following the evolution of the dynamics over time permits the detection of any correlations between different time scales, and measurements at the different Q values inform about the extension of the motions. However, as motions corresponding to the same Q value are observed on different energy resolutions and thus time scales, they are nonetheless not directly comparable and have to be considered and commented individually.

The samples were hydrated with D_2O at 0.4 g g^{-1} to highlight the motions of the hydrogens in the sample and not in the surrounding water. This prevented whole enzyme rotation and translation in the samples, so that we could focus on internal protein dynamics. The EINS data collected on the three instruments were analyzed individually, however following the same procedure. It is possible to show that the summed intensities are, up to a Q-range dependent multiplier, inversely proportional to the root of the MSD, but are less affected by statistical errors than the MSD.[15] They were calculated first by summing over the whole Q-range of each instrument. Furthermore within the so called Gaussian approximation,[50] which assumes that the distribution of the atoms around their average position follows a Gaussian distribution, the scattered elastic incoherent intensity (S_{el}) is given by the dynamic structure factor at zero energy exchange (S_0)

$$S_{el}(Q, \omega = 0 \pm \Delta E) \approx S_0 \exp\left(-\frac{1}{3}\langle u^2\rangle Q^2\right) \quad (1)$$

where ω and Q are the exchanged energy and momentum in units of $\hbar$, respectively, $\langle u^2\rangle$ is the MSD, and ΔE is the half width at half maximum of the instrumental energy resolution. The approximation is strictly valid for $Q \to 0$ and it holds up to $\langle u^2\rangle\, Q^2 \approx 1$. We determined the Gaussian approximation to be valid for the Q-ranges of 0.49–1.24 $Å^{-1}$ on IN6, 0.5–1.67 $Å^{-1}$ on IN13 and 0.7–1.48 $Å^{-1}$ on IN16B. The MSD can be obtained for each temperature by the slope of the semi-logarithmic plot of the incoherent scattering function through

$$\langle u^2\rangle \approx -3\frac{\mathrm{d}\ln S_{el}(Q, \omega = 0 \pm \Delta E)}{\mathrm{d}Q^2} \quad (2)$$

The MSD can be a measure for the flexibility of the biological system at a given temperature.[51] To obtain the intensities scattered by the sample only, scattering from the empty sample holder was subtracted, and the data were normalized to the lowest measured temperature data (IN6: 80 K, IN13: 20 K, IN16B: 40 K). At such low temperature, all motions are frozen and the neutron intensities thus reflect the relative detector efficiency and the instrumental resolution. The lowest temperature was different on the various spectrometers for technical reasons, but the differences in the data are negligible within the range of cryo-temperatures. The acquisition was continuous on IN6 and IN16B with ramps of 1 K min^{-1} below 80 K and 0.5 K min^{-1} above 80 K and on discrete points on IN13 (with steps of 10 or 20 K) with 1 h counting time below 280 K and 2 h above 280 K. Absorption correction was based on the correction formula of Paalman–Pings.[52] The complete data reduction was carried out using the LAMP software available at ILL.[53]

2.3 Multivariate data analysis

PCA[37,38] is a projection method that visualizes the main variation in multivariate data by transforming a multidimensional set of correlated variables to a smaller set of uncorrelated new variables. These are called principal components (PCs) and the PCs are eigenvectors of the original data. Here, the data matrix ($\mathbf{X}$) consists of intensities and the matrix row elements are the measurements in sequential order at the different temperatures (T). Column elements (variables) are the Q-values at the detector (Q:s), yielding $\mathbf{X}(Q_i, T_j)$, where $i = 1,\ldots, M$ dimensions of Q space and $j = 1,\ldots, N$ number of measurements. The intensities were normalized to the intensity at the lowest measured temperature and were mean-centered prior to modeling. The first PC (PC1) explain (per definition) the main variation in the data. After calculation of PC1, each measured data point received a new value, a score value s to build up a score vector s_1, which is the eigenvector that has the largest eigenvalue. The contribution of the original variables, Q:s, to

the new score vector (the position of the samples in the score vector) is described by the values in the loading vector **l**. The second PC (PC2) is placed orthogonally to PC1, so that it captures the remaining main variation (second largest eigenvalue), the third PC (PC3) is placed orthogonally to PC2, and so on. These PCs give rise to a new decomposition matrix (**SL**′) according to

$$\mathbf{X} = \bar{\mathbf{X}} + \mathbf{SL}' + \mathbf{E} \quad (3)$$

where $\bar{\mathbf{X}}$ is the **X** matrix average, **S** is the score matrix, **L**′ is the loading matrix, and **E** is the residual. The *l*- and *s*-values are visualized in two separate plots; the loading plot shows the *Q*:*s* and their weights on the PCs and the score plot shows how the EINS measurements differ in terms of their intensities in the *Q*-range.

OPLS-DA[39,40] makes use of a predefined matrix (**Y**) containing class assignment based on the samples. This information is used to decomposed the **X** matrix containing the measured *Q* intensities in such way to that it describes large variation in **X** and correlate with **Y**. **X** is decomposed into three parts according to

$$\mathbf{X} = \mathbf{S_P L_P}' + \mathbf{S_O L_O}' + \mathbf{E} \quad (4)$$

where $\mathbf{S_P}$ is the predictive score matrix, $\mathbf{L_P}'$ the predictive loading matrix, $\mathbf{S_O}$ the corresponding **Y**-orthogonal score matrix, $\mathbf{L_O}'$ the loading matrix of the **Y**-orthogonal components, and **E** is the residual matrix of **X**. Thus, OPLS-DA discriminates between the inter-class variation (class differences) shown by the predictive scores ($\boldsymbol{s}_{p,1}$) and intra-class variation (differences within a class) shown by the orthogonal scores ($\boldsymbol{s}_{o,1}$). The predictive loadings ($\boldsymbol{l}_{P,1}$), or in this case the weights ($\boldsymbol{w}^*$), can be analyzed to identify the *Q*-values (*i.e.*, the movement length scale) that gives rise to the classes. The absolute value of individual weight indicates the importance of that specific variable to the predefined class assignment, the larger value, the higher importance. The sign of the weight shows if the related variable is positively or negatively correlated to that particular class.

The quality and statistical significance of the multivariate models was determined from the proportion of the variation in the original data that was explained by the model, *i.e.*, the cumulative sum of squares of the entries ($R^2\mathbf{X}$(cum)) and eigenvalues (PCA), and the cumulative sum-of-squares of all the **y**-variables explained by the extracted OPLS-DA components ($R^2\mathbf{Y}$(cum)) (OPLS-DA). The models were further tested using leave-*n*-out cross-validation[54] (*n* equal to 1/7 of the data set) giving rise to a q^2(cum)-value comprising all model PCs or OPLS-DA components. Multivariate modeling was performed in the SIMCA software[55] and a more detailed description of PCA, OPLS-DA, and cross validation is presented in the ESI.†

3. Results and discussion

The dynamics of free (samples CT1 and CT2) and inhibited α-chymotrypsin (samples CT/CS and CT/TP) was measured on hydrated powders on the three instruments: IN6, IN13, and IN16B. The resulting EINS data were analyzed by extracting the intensities summed over all scattering angles (the entire *Q*-range) and MSD based on a shorter *Q*-range. Inspection of the three figures showing the summed intensities (Fig. 2) revealed the following: (i) negligible difference in the dynamics of the two samples of the free enzyme (CT1 and CT2) was visible on IN6. (ii) Covalent inhibitors affected the enzyme so that it became more flexible and this was true for both inhibited enzyme samples (CT/TP and CT/CS) and on all instruments. (iii) The differences in dynamics between the inhibited samples were within the error bars, but the tendency was that the effect of TPCK was larger than that of chymostatin. (iv) The results from IN6 and IN13 showed an effect of hydration; the CT2 sample had an increase in intensity on these instruments between 220 K and 270 K, which can be associated with water Bragg peaks. The hydration effect is seen in *Q*-ranges not included in the MSD calculations.

PCA was used to visualize general trends and groupings in the EINS data using the whole *Q*-range to visualize differences in the dynamics of the α-chymotrypsin samples. The PCA models were generally of high statistical quality where ∼99% of the original variation in the data was explained with a cross-validation q^2 of 0.98–0.99. Most of the variation was described by PC1 (97–99%) and a few percent by PC2 and PC3. Model details including statistics and number of samples and measurements are presented in the ESI,† Tables S1–S3. The main difference in enzyme dynamics between the different measurements, shown in PC1 (*e.g.*, for IN6 in Fig. 3a), was related to the different temperatures at which a sample was measured, but also the differences between free and inhibited α-chymotrypsin. This was true for all samples measured on all three instruments (Fig. 3 and 4).

The enzyme dynamics increased with the temperature; the intensities were negatively correlated with the temperature, as shown previously for proteins.[56] This relation was manifested in the PCA by the fact that the high *Q*:*s* were generally found at the right most in the loading plot, connected to the low temperatures to the right in the score plot. The different PCAs based on data of the three instruments, also showed groups consisting of the free (CT2) and inhibited α-chymotrypsin (CT/CS and CT/TP) and these groups were prominent mainly at higher temperatures (Fig. 3 and 4).

3.1 Measurement reproducibility

We analyzed the results of the two free α-chymotrypsin samples CT1 and CT2 on IN6 to determine the variation in measurement results on two similar samples prepared in parallel. Except for the Bragg-peak effects (see separate analysis below) the two samples appear equal within error bars from the summed intensities (Fig. 2a). The MSD and PCA showed that CT2 was slightly more dynamic in general (Fig. 2b and 3a) and more similar to the inhibited α-chymotrypsin samples in the MSD- and score plots, mainly at low temperatures. The difference between CT1 and CT2 may stem from differences in sample preparation, weights, and hydration. Nevertheless, we can conclude that the difference is

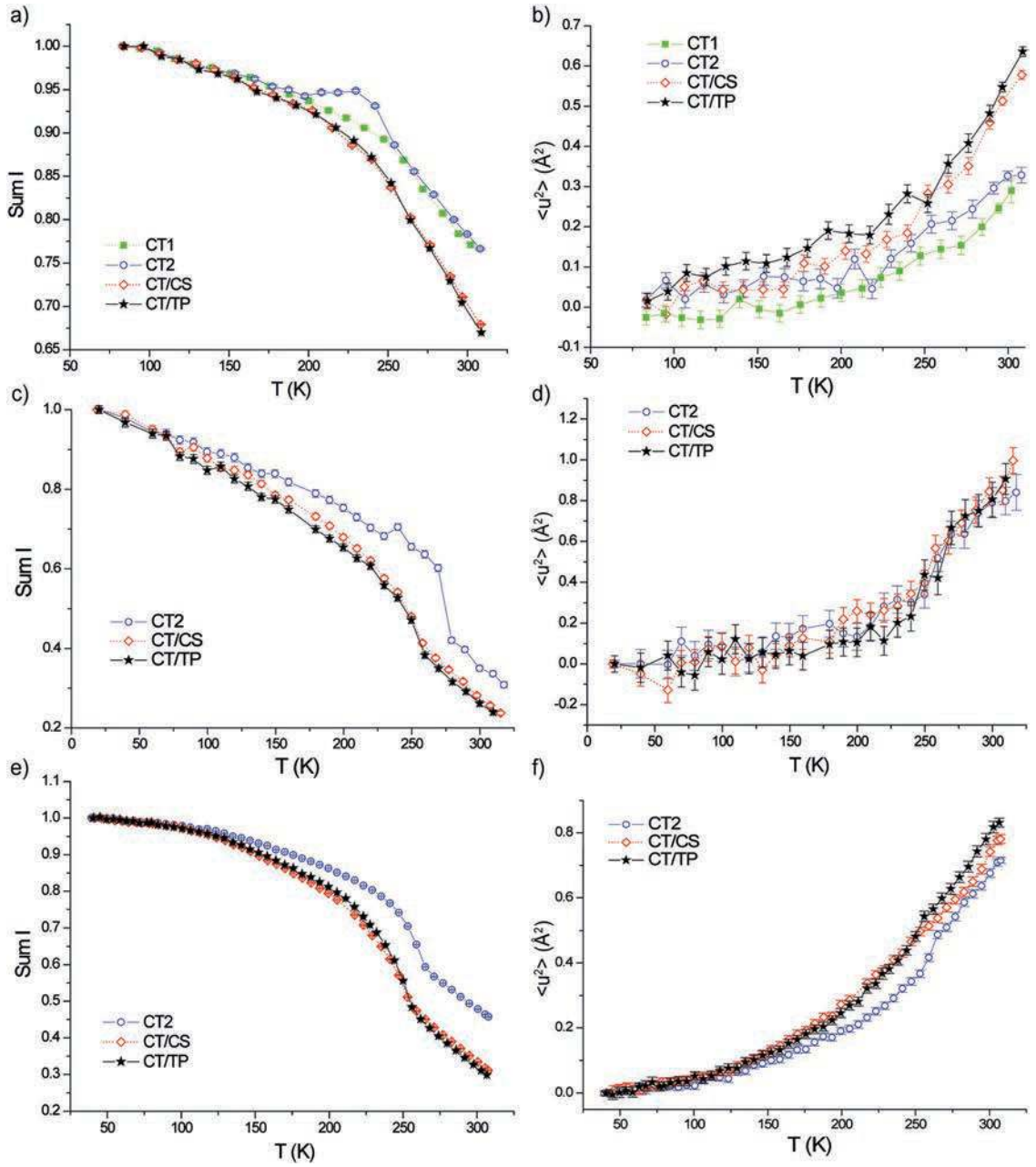

Fig. 2 Neutron intensities in arbitrary units summed over the available scattering angles (left column) and the extracted MSD (right column), both as function of temperature. Data are from IN6 (a and b), IN13 (c and d), and IN16B (e and f). The error bars were based on Poisson statistics applied to the raw data. The error bars of the MSD, extracted according to eqn (2), were obtained by linear regression when fitting the data.

small, and most importantly, considerably smaller than the difference between free and inhibited enzyme. Notably, the difference between CT1 and CT2 diminishes with increasing temperature, which is seen in both univariate (Fig. 2) and the multivariate analysis (Fig. 3a).

3.2 Effects of water phase transitions

On IN13 and IN16B, particularly for the CT2 sample, there was a sharp decrease in the summed intensities around 250–275 K (Fig. 2c and e) due to the melting of water. An effect of a water phase transition that is different from melting was seen for the CT2 sample measured on IN6 around 240 K (Fig. 2a and 3b). In fact, water may undergo several phase transitions but not all phases are thermodynamically stable. At a very low temperature, ice has an amorphous structure and at 150 K, it transforms into a quasi-cubic structure. Beyond 200 K the quasi-cubic structure passes into the hexagonal phase, which is very fast above 240 K, also reported by Koza *et al.*[57] Here, only the cubic to hexagonal phase transition makes fluctuations that contribute in a significant way to the dynamics probed by the neutrons. This effect was also weakly present in the other samples, which is only apparent by the multivariate analysis in Fig. 3b, where all samples show higher

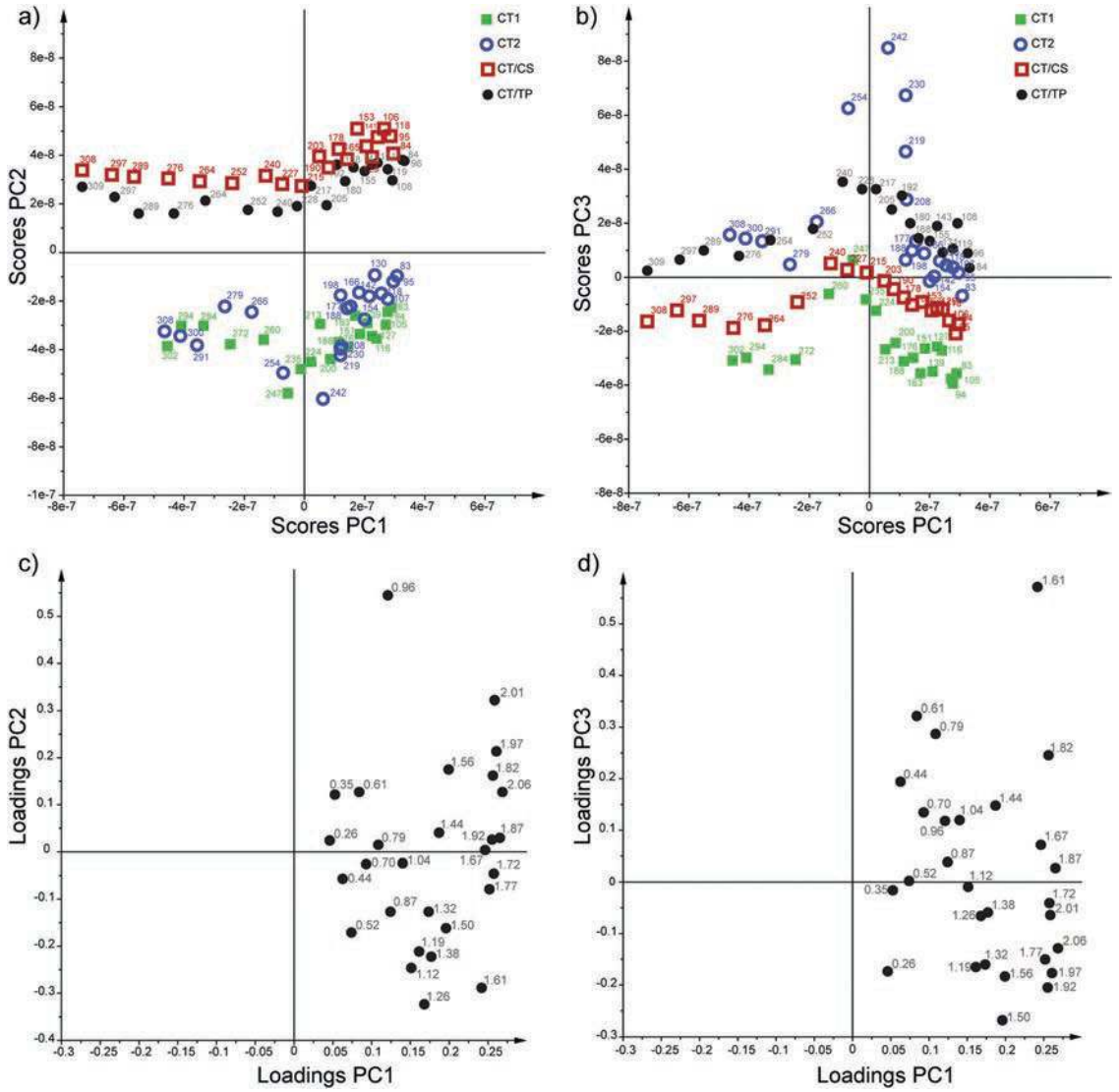

Fig. 3 PCA analysis of the measurements from IN6 showing in (a) and (b) score plots of the samples displaying the main differences between the samples (indicated with sample symbol and measurement temperature) based on neutron intensities, and in (c) and (d) loading plots with loading values for each variable Q indicated with the Q value in $Å^{-1}$. PC; principal component.

score values in principal component three (PC3) in that temperature range. The sharp increase in the intensities were mainly manifested in Q equal to 1.61 $Å^{-1}$, which matches with values for ice Bragg peaks that have been reported.[57] Such an effect is, however, negligible for the MSD extracted from the restricted Q-range, since the ice Bragg peaks appear in a Q-range not included in the MSD analysis. We concluded that the difference in dynamics between the two inhibitor-free samples CT1 and CT2 mainly occurred at mid-range temperatures due to Bragg peaks, and does not stem from a difference between the samples with any biological mechanistic relevance, but rather from slight sample dissimilarities.

3.3 Differences in dynamics between free and inhibited α-chymotrypsin

The summed intensities over the whole Q-range and the PCA of the neutron scattering data showed that there was a clear difference in the dynamics between the inhibited and free form of α-chymotrypsin; both covalent inhibitors affected the enzyme so that it became more flexible. The difference in enzyme motion is pronounced in the summed intensities measured on all instruments at mid- to high temperatures, but not at the lowest temperatures (Fig. 2a, c and e).

The MSD analysis of the IN6 data also showed clear differences between the inhibited and the free samples (Fig. 2b), where the MSD increases by $\sim$0.3 $Å^2$ at the highest temperature. In the IN13 data, the MSD had larger standard deviations, which did not allow a significant separation between inhibited and free α-chymotrypsin (Fig. 2d). On IN16B, the difference in the total elastic intensity was the largest between 200 K and 260 K. A significant difference in the MSD was also apparent at the highest temperatures (Fig. 2f), corresponding to a difference in MSD of $\sim$0.065 $Å^2$ between the CT/CS and CT2 samples and of $\sim$0.05 $Å^2$ between CT/TP and CT2 samples. In the PCA, notably

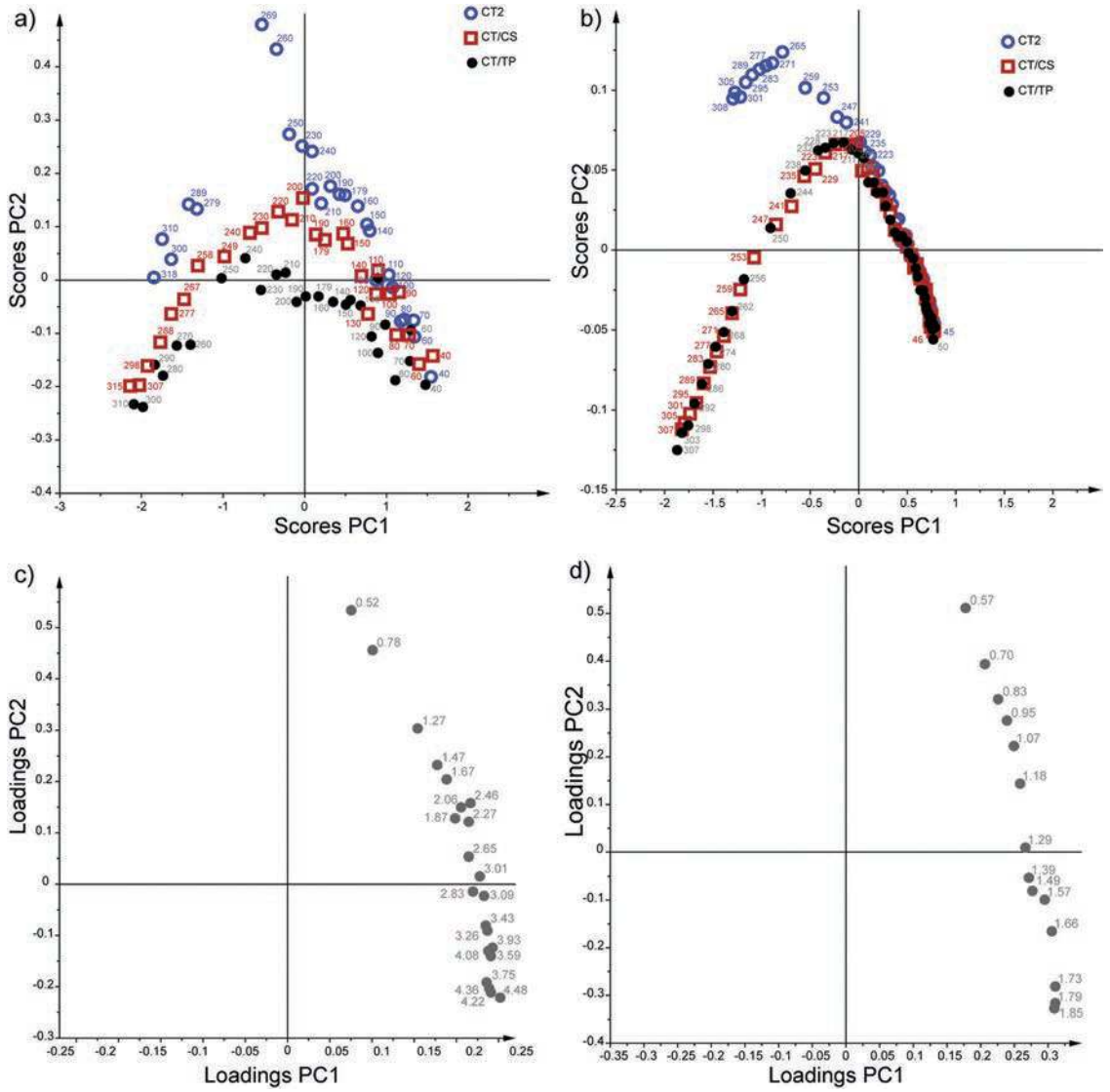

Fig. 4 PCA analysis of the measurements with score plots of the samples displaying the main differences between the samples (indicated with sample symbol and measurement temperature) based on neutron intensities from (a) IN13 and (b) IN16B, and loading plots with loading values for each variable Q indicated with the Q value in $Å^{-1}$ from (c) IN13 and (d) IN16b. PC; principal component.

including intensities of the whole Q-range, the difference between the free and inhibited enzyme was pronounced across all temperatures in the IN6 measurements (Fig. 3a). From the IN13 and IN16B measurements, mainly at mid- to high temperatures (240–310 K), there are two groups of measurements including free (CT2) and inhibited (CT/CS and CT/TP) α-chymotrypsin samples (Fig. 4a and b). On IN13, the difference is less pronounced but clearly there, especially at higher temperatures (~280–315 K). To explore the cause of the higher mobility of the inhibited α-chymotrypsin further, and to more clearly pinpoint the kind of atomic motions that were involved, the samples were classified into two classes corresponding to the free and inhibited α-chymotrypsin sample(s). The analysis was focused on the data at the biologically relevant temperature interval 270–310 K.

The variation in Q intensities that could explain class separations (the $R^2\mathbf{X}$ of the OPLS-DA predictive component) amounted to 60%, 47%, and 76% of the total variation collected at IN6, IN13, and IN16B, respectively, showing that a substantial amount of the information in the EINS data was related to the class differences. Model statistics are presented in the ESI,† Tables S1–S3. All of the classification models showed, as expected, a clear separation between inhibited and free α-chymotrypsin, with the CT2 (or CT1 and CT2) samples to the left and CT/CS and CT/TP samples grouped to the right in the score plot (Fig. 5a, c and e). All the weights in right panels (Fig. 5b, d and f) had negative values and this relates to the position of the inhibitor-free samples to the left and negative side in the score plot. This meant that the free enzyme samples (at 270–310 K) generally had higher intensities in the whole Q-range and were thus less mobile compared to the inhibited enzymes. Note that the same spatial scale or Q-value seen on different instruments corresponds to different time scales

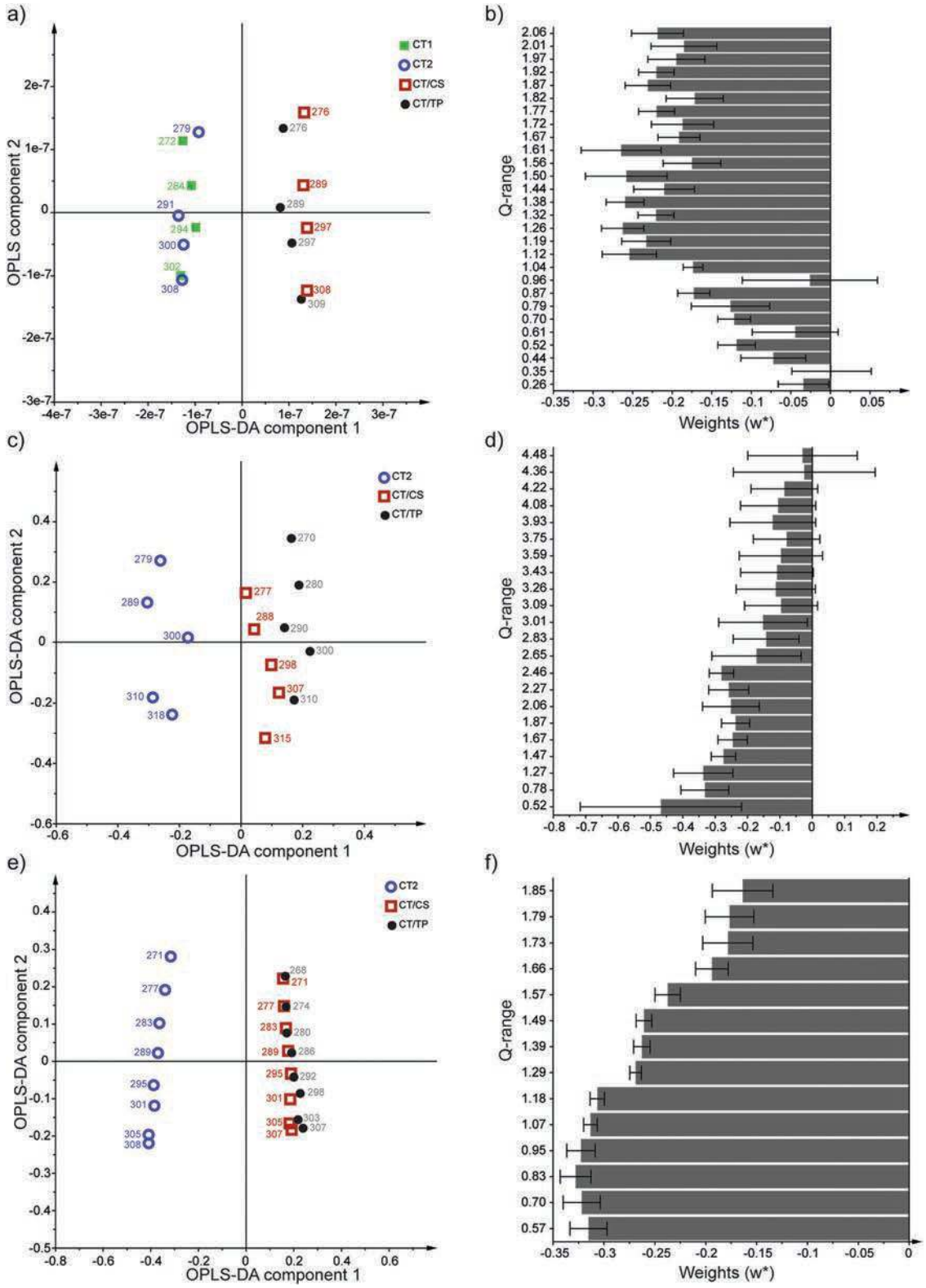

Fig. 5 Classification of inhibitor-free *versus* inhibited α-chymotrypsin in the temperature interval 270–310 K with scores, showing class separations, and weight for OPLS-DA component 1, showing scattering angle (*Q*-range) contributions to class separation, from measurements on IN6 (a and b, respectively), IN13 (c and d, respectively) and IN16B (e and f, respectively). Error bars are standard errors with a confidence level of 95% calculated by Jack-knifing.

according to the instrumental resolution. Furthermore, from the OPLS-DA weight plots (Fig. 5), it was clear that the high-range *Q*:*s* are most important for the motions seen on IN6. On IN13, the *Q*:*s* contributed similarly within the *Q*-ranges accessible on IN6 or

IN16B but higher-range Q:s decreased in weight, whereas a decrease with Q was clear on IN16B. In the 7.5 ps time window (IN6) the inhibited enzymes were more dynamic than the free form due to more small-amplitude high velocity motions (higher intensities at high scattering angles; Fig. 5a and b), such as the rotations of hydrogens bound to, for example, carbon in methyl groups. Water translational motions is monitored at a Q-range below $\sim$1.5 Å^{-1} but the short time window on IN6 cannot capture this motion in total, as shown by small and non-significant weights in the OPLS-DA weights plot (Fig. 5b). The difference between the inhibited and free samples seen in the PCA analysis for the whole temperature range (Fig. 3a), could be corroborated by an OPLS-DA model including the whole temperature range that was statistically significant and showed a separation between the free and inhibited α-chymotrypsin (details are presented in the ESI,† Fig. S4).

Even smaller amplitude motions measured with the broader Q-range monitor (Q of 2.27–4.48 Å^{-1}) at the IN13 instrument (100 ps time window) were not affected upon covalent inhibition, as these OPLS-DA weights were non-significant in the classification model (Fig. 5b and c) and are thus not important in the differentiation between free and inhibited α-chymotrypsin.

Neutron scattering intensities measured during the longest time window, 1 ns, at IN16B, showed that the increased dynamics of α-chymotrypsin upon inhibition was not only due to small amplitude motions but due to larger, slower motions increased as well. This can be concluded because all OPLS-DA weights in this Q-range (0.57–1.85 Å^{-1}) were significant (Fig. 5f).

The classification method OPLS-DA[39,40] was particularly suitable in our case, as this method filtered the data from scattering intensity differences between measurements that were not related to class separation (*e.g.*, the temperature dependences). The fact that there was a marked difference between the free and the inhibited enzyme with respect to the high Q-values on IN6 (Fig. 4b and Fig. S4, ESI†) might be interpreted as follows: Compared to the free enzyme, the inhibited underwent the dynamical transition at lower temperature (*i.e.*, around 200 K)[58] and in a more cooperative manner, meaning more particles undergoing the transition simultaneously. We therefore speculate that the bound inhibitors induced small collective motions in the enzyme, possibly also disturbed the water network, so that less energy was required to undergo the dynamical transition. The motions captured by IN16B concerned scales above 3.3 Å monitored during 1 ns and corresponding to amino acid side chains- and water molecule movements, and the inhibited enzyme was more dynamic when considering these kinds of movements. Importantly, such distinctions were not possible to identify from the MSD solely, which is based on the slope of the intensities with respect to Q^2.

3.4 Effects of different inhibitors

There was a weak tendency that α-chymotrypsin became more flexible when inhibited by TPCK compared to chymostatin, as evident by the difference between the two samples in the summed intensities at the highest temperatures from the IN6 and IN16B measurements (Fig. 2b and f). A statistically

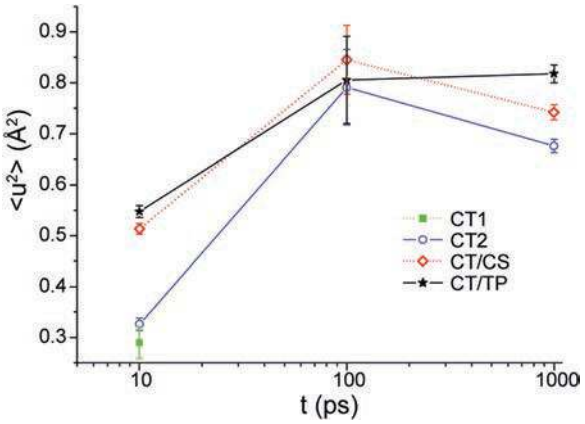

Fig. 6 MSD at 300 K for the different samples as function of the instrumental time window where 10, 100, and 1000 ps corresponds to the time window on IN6, IN13 and IN16B, respectively.

significant difference in MSD could not be seen between CT/TP and CT/CS at IN13 and at the highest temperature. These results are summarized in Fig. 6, showing the MSD of the samples at 300 K measured on the three instruments. No significant overall increase in $\langle u^2 \rangle$ is noted in between the IN13 and IN16B results, and only judging from the MSD results, this would indicate that no additional motions were recorded between 100 ps and 1 ns.

Nevertheless, the OPLS-DA (Fig. 5) showed that the Q-ranges contributing the most are not the same on the three instruments and thus only on average does the flexibility appear to be about the same on IN13 and IN16B. The more detailed multivariate analysis permits to shed light on this point. These findings might explain why different diffusion coefficients are often extracted from quasi-elastic neutron scattering data taken with different instrumental resolutions although the sample is identical.[59] The CT/TP sample showed a higher MSD compared to the CT/CS sample (on IN6 and IN16B) indicating that the former induced more motion in the α-chymotrypsin than the latter. Here, we note that TPCK is markedly smaller than chymostatin (Fig. 1), but still induce more motion. Nevertheless, the lack of structural data regarding the inhibitor-chymotrypsin complex and our data treatment does not allow for more precise conclusions, but we refer to the discussion in Peters *et al.*[15] where arguments are forwarded why inhibitors may act differently on the water network and how this could influence the global dynamics.

4. Conclusions

The dynamics of the serine protease α-chymotrypsin in its free form and when inhibited by covalently bound inhibitors TPCK and chymostatin was measured with EINS on the spectrometers IN6,[44] IN13,[45] and IN16B.[46] As expected, the dynamics for all samples increased with temperature, but the inhibited enzyme was shown to be more dynamical compared to the free, as was evident from both the summed intensities and the multivariate analysis including PCA and OPLS-DA. Other factors such as

sample preparation, hydration level, and sample weights could be ruled out as the cause for this difference. The increase in dynamics of the inhibited α-chymotrypsin was seen on all measured time-scales capturing movements such as methyl group rotations, water movements, and amino acid side chain movements. Contrary to the univariate methods, which average over effects and specific space dimensions the multivariate analysis could reveal subtle effects, such as the presence of Bragg peaks in all measurements and which *Q*-ranges and thus which type of motion constituted the difference in the enzyme dynamics. The OPLS-DA models pinpointed scattering angles that contained information on the difference in dynamics between inhibited and free enzyme, and gave statistical significance to this finding. At high temperatures on a short time-scale (IN6), mainly small amplitude motions, such as hydrogens rotating about carbon in methyl groups, contributed to the higher mobility of the inhibited enzyme. On IN16B, covering the longest time-scales, more motions that correspond to the movements of amino acid side chains could be seen in the inhibited enzyme. Thus, we conclude that the inhibited enzymes studied here underwent the dynamical transition at lower temperatures and in a more cooperative way leading to bigger amplitudes of motions. Here, we propose two simultaneous scenarios regarding the changes of the potential energy landscape of the enzymes. The inhibitors' interactions with the enzyme permitted more collective motions and more particles to overcome potential energy barriers in the conformational landscape simultaneously during the transition. Furthermore, the inhibitors influenced the water network around the enzymes in a way that permitted more degrees of motional freedom leading to a lowering of the potential energy barrier heights. Molecular dynamics simulations could eventually help to disentangle the different scenarios.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was funded by the Swedish Research Council (Dnr: 2014-4675). We are gratefully acknowledging the ILL for allocation of beam time.

References

1 V. Daggett, S. Schröder and P. Kollman, *J. Am. Chem. Soc.*, 1991, **113**, 8926–8935.
2 M. Kokkinidis, N. M. Glykos and V. E. Fadouloglou, *Adv. Protein Chem.*, 2012, **87**, 181–218.
3 A. Kohen, *Acc. Chem. Res.*, 2015, **48**, 466–473.
4 R. A. Copeland, M. R. Harpel and P. J. Tummino, *Expert Opin. Ther. Targets*, 2007, **11**, 967–978.
5 M. McAuley and D. J. Timson, *Appl. Biochem. Biotechnol.*, 2017, **181**, 83–90.
6 D. D. Boehr, H. J. Dyson and P. E. Wright, *Chem. Rev.*, 2006, **106**, 3055–3079.
7 R. M. Daniel, R. V. Dunn, J. L. Finney and J. C. Smith, *Annu. Rev. Biophys. Biomol. Struct.*, 2003, **32**, 69–92.
8 L. P. Yu, C. X. Zhu, Y. C. TseDinh and S. W. Fesik, *Biochemistry*, 1996, **35**, 9661–9666.
9 S. G. Yun, D. S. Jang, D. H. Kim, K. Y. Choi and H. C. Lee, *Biochemistry*, 2001, **40**, 3967–3973.
10 E. Balog, T. Becker, M. Oettl, R. Lechner, R. Daniel and J. Finney, *et al.*, *Phys. Rev. Lett.*, 2004, **93**, 028103.
11 A. Eletsky, A. Kienhofer, D. Hilvert and K. Pervushin, *Biochemistry*, 2005, **44**, 6788–6799.
12 Y. L. Miao, Z. Yi, C. Cantrell, D. C. Glass, J. Baudry and N. Jain, *et al.*, *Biophys. J.*, 2012, **103**, 2167–2176.
13 J. H. Davis and D. A. Agard, *Biochemistry*, 1998, **37**, 7696–7707.
14 M. Trapp, M. Trovaslet, F. Nachon, M. M. Koza, L. van Eijck and F. Hill, *et al.*, *J. Phys. Chem. B*, 2012, **116**, 14744–14753.
15 J. Peters, N. Martinez, M. Trovaslet, K. Scannapieco, M. M. Koza and P. Masson, *et al.*, *Phys. Chem. Chem. Phys.*, 2016, **18**, 12992–13001.
16 F. Gabel, M. Weik, P. Masson, F. Renault, D. Fournier and L. Brochier, *et al.*, *Biophys. J.*, 2005, **89**, 3303–3311.
17 L. Hedstrom, *Chem. Rev.*, 2002, **102**, 4501–4523.
18 N. D. Rawlings, F. R. Morton, C. Y. Kok, J. Kong and A. J. Barrett, *Nucleic Acids Res.*, 2007, **36**, D320–325.
19 B. Turk, *Nat. Rev. Drug Discovery*, 2006, **5**, 785–799.
20 G. B. Jones, M. O. Mitchell, J. S. Weinberg, A. V. D'Amico and G. J. Bubley, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 1987–1989.
21 Y. Y. Ge, X. H. Wu, D. Z. Zhang and L. Q. Hu, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 941–944.
22 J. Zhang, Y. F. Liu, L. Bo and C. A. Qiao, *J. Appl. Polym. Sci.*, 2011, **121**, 1992–1998.
23 C. S. Cummings, H. Murata, K. Matyjaszewski and A. J. Russell, *ACS Macro Lett.*, 2016, **5**, 493–497.
24 M. A. M. Latif, B. A. Tejo, R. Abedikargiban, M. B. A. Rahman and N. M. Micaelo, *J. Biomol. Struct. Dyn.*, 2014, **32**, 1263–1273.
25 R. Wang, Y. F. Zhang, D. N. Lu, J. Ge, Z. Liu and R. N. Zare, *Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol.*, 2013, **5**, 320–328.
26 B. W. Matthews, P. B. Sigler, R. Henderson and D. M. Blow, *Nature*, 1967, **214**, 652–656.
27 J. J. Birktoft and D. M. Blow, *J. Mol. Biol.*, 1972, **68**, 187–240.
28 G. Schoellmann and E. Shaw, *Biochemistry*, 1963, **2**, 252–255.
29 N. P. Tomkinson, I. J. Galpin and R. J. Beynon, *Biochem. J.*, 1992, **286**, 475–480.
30 J. J. Birktoft, D. M. Blow, R. Henderson and T. A. Steitz, *Philos. Trans. R. Soc., B*, 1970, **257**, 67–76.
31 A. Tulinsky and R. A. Blevins, *J. Biol. Chem.*, 1987, **262**, 7737–7743.
32 K. C. H. Chua, M. Pietsch, X. Z. Zhang, S. Hautmann, H. Y. Chan and J. B. Bruning, *et al.*, *Angew. Chem., Int. Ed.*, 2014, **53**, 7828–7831.
33 A. Razeto, B. Galunsky, V. Kasche, K. S. Wilson and V. S. Lamzin, *High resolution structure of bovine alpha-chymotrypsin*, 2006, DOI: 10.2210/pdb1yph/pdb.

34 V. F. Sears, *Neutron News*, 1992, **3**, 26–37.
35 J. Peters, M. Trovaslet, M. Trapp, F. Nachon, F. Hill and E. Royer, *et al.*, *Phys. Chem. Chem. Phys.*, 2012, **14**, 6764–6770.
36 S. W. Lovesey, *Theory of neutron scattering from condensed matter*, Oxford Science Publications, Oxford, UK, 1984, vol. I.
37 S. Wold, K. Esbensen and P. Geladi, *Chemom. Intell. Lab. Syst.*, 1987, **2**, 37–52.
38 J. E. Jackson, *A user's guide to principal components*, John Wiley & sons, Inc., New York, USA, 1991.
39 J. Trygg and S. Wold, *J. Chemom.*, 2002, **16**, 119–128.
40 M. Bylesjö, M. Rantalainen, O. Cloarec, J. K. Nicholson, E. Holmes and J. Trygg, *J. Chemom.*, 2006, **20**, 341–351.
41 S. Magazù, G. Maisano, F. Migliardo, G. Galli, A. Benedetto and D. Morineau, *et al.*, *J. Chem. Phys.*, 2008, **129**, 155103.
42 L. Hong, N. Smolin and J. C. Smith, *Phys. Rev. Lett.*, 2014, **112**, 158102.
43 A. Kumar and P. Venkatesu, *Chem. Rev.*, 2012, **112**, 4283–4307.
44 http://www.ill.eu/in6/home/, U.
45 F. Natali, J. Peters, D. Russo and F. Sonvico, *Neutron News*, 2008, **19**, 14–18.
46 B. Frick, E. Mamontov, L. van Eijck and T. Seydel, *Z. Phys. Chem.*, 2010, **224**, 33–60.
47 S. Magazù, F. Migliardo and A. Benedetto, *Rev. Sci. Instrum.*, 2011, **82**, 105115.
48 G. L. Squires, *Introduction to the theory of thermal neutron scattering*, Dover publications, Mineola, New York, USA, 1996.
49 J. A. McCammon and S. C. Harvey, *Dynamics of proteins and nuclear acids*, Cambridge University Press, Cambridge, UK, 1987.
50 A. Rahman, K. S. Singwi and A. Sjölander, *Phys. Rev.*, 1962, **126**, 986–996.
51 G. Zaccai, *Science*, 2000, **288**, 1604–1607.
52 H. H. Paalman and C. J. Pings, *J. Appl. Phys.*, 1962, **33**, 2635–2639.
53 LAMP, the Large Array Manipulation Program. http://www.ill.eu/data_treat/lamp/the-lamp-book/.
54 H. T. Eastment and W. J. Krzanowski, *Technometrics*, 1982, **24**, 73–77.
55 SIMCA 14, Umetrics AB. Box 7960, Umeå, Sweden.
56 F. Gabel, D. Bicout, U. Lehnert, M. Tehei, M. Weik and G. Zaccai, *Q. Rev. Biophys.*, 2002, **35**, 327–367.
57 M. M. Koza, B. Geil, H. Schober and F. Natali, *Phys. Chem. Chem. Phys.*, 2005, **7**, 1423–1431.
58 D. J. Bicout and G. Zaccai, *Biophys. J.*, 2001, **80**, 1115–1123.
59 M. Trapp, M. Tehei, M. Trovaslet, F. Nachon, N. Martinez and M. M. Koza, *et al.*, *J. R. Soc., Interface*, 2014, **11**, 20140372.