Abstract

The analysis of conformational transitions provides important information on the mode of action of proteins. Amide hydrogen exchange mass spectrometry (HX-MS) allows to detect conformational changes and to resolve the kinetics if the transition occurs in the accessible time domain. However, manual data analysis of HX-MS experiments is tedious, especially for large proteins, and therefore, analysis is usually restricted to a subset of the information available from the mass spectra. In response to this problem, several software tools, including HeXicon from our group, have recently been presented, aiming to automate data analysis of HX-MS experiments and to extract all accessible information. Bimodal isotope distributions as arising from EX1 exchange mechanisms or different coexisting conformations in a protein are of specific interest, because they report on the kinetics of conformational changes, but also provide an even greater challenge for automated data analysis. In this study, we tuned HeXicon in order to search specifically for bimodal isotope distributions in large data sets. We applied the modified program to a dataset from the *E. coli* Hsp90 homologue HtpG and compared the results with manual data analysis. All seven manually found bimodal cases were detected as bimodal by HeXicon. In addition, HeXicon also located nine previously unknown bimodal distributions, illustrating the benefit of automated data processing.

Keywords:
amide hydrogen exchange; bimodal isotope distribution; automated data analysis; coexisting conformations;

1. Introduction

Most proteins are intrinsically dynamic under physiological conditions continuously opening and closing many hydrogen bonds of their secondary structure elements. This flexibility of proteins is considered to be essential for their biological function such as enzymatic activity and allosteric regulation [1]. Hydrogen exchange followed by mass spectrometry (HX-MS) is an established, powerful method that enables the analysis of structural dynamics of proteins [2-4]. Since amide protons involved in hydrogen bonds or buried in the interior do not exchange, HX-MS reports on the folding-unfolding dynamics of secondary structure elements as well as entire domains.

To localize fast and slow exchanging regions within a protein, samples are digested by a protease after the exchange reaction under low pH quench conditions. Pepsin, which is generally used for proteolytic digestion due to its activity at low pH, has low sequence specificity and generates many overlapping peptides. For larger proteins the number of generated fragments that are then analyzed by liquid chromatography mass spectrometry (LC-MS) is immense and manual data analysis becomes very time consuming. In general, a minimum number of peptides that covers the protein sequence sufficiently is analyzed by determination of the centroid. Therefore, in many cases a lot of information contained in the spectra remains unused. As a consequence, a number of different semi-automatic software tools or programs for processing HX-MS data have been devised including HX-Express [5], The Deuterator [6, 7] and Hydra [8].

A major benefit of HX-MS approaches is their ability to detect coexisting populations of a single protein, which exhibit different exchange behaviors. Such different populations can arise from slow local unfolding due to intrinsic protein instability or from induced conformational changes that are slow compared to the exchange reaction kinetics. In HX-MS experiments, two coexisting populations with different exchange kinetics are characterized by a bimodal peak distribution within the isotope cluster.

However, the current general approach for automated exchange rate determination is based on the detection and tracing of the centroid of a protein's peptides over all incubation times for which measurements are available. Centroid-based deuteration level determination reduces the deuteration distribution for each incubation time to a single point measurement. Consequently, all information contained in the deuteration distribution is lost. To recover the information on bimodal isotope distributions, width of the isotope distribution was proposed as an estimate of the half-life of the conformational change and the number of hydrogens involved [9].

We recently proposed HeXicon [10] as a tool for automated deuteration distribution estimation at increased sequence coverage. HeXicon is based on an LC-MS extension of the NITPICK algorithm [11]. It extracts quantitative LC-MS features for each experiment, finds correspondences across different D_2O incubation times, estimates the associated deuteration distributions and derives deuteron incorporation rates. The estimation of the deuteration distribution as compared to the estimation of the centroid has the advantage of preserving the information on bimodality of isotope distributions.

In this study, we exploit HeXicon's feature extraction abilities to specifically search a large HX-MS dataset for peptides with a bimodal isotope distribution. As sample data we used the LC-MS measurements of pulse-labeling HX-MS experiments with *E. coli* HtpG, a protein that belongs to the family of 90 kDa heat shock proteins [12]. Upon addition of ATP HtpG undergoes a two-step conformational change into a highly protected state. The first step with a half-life of 3 ms occurs too fast for HX-MS experiments as

performed in this study [compare 13], but the second step with a half-life of approximately 120 s is accessible to HX analysis [12]. Manual data analysis was restricted to peptides previously identified by an MS/MS experiment and MASCOT search and yielded 7 peptides with a bimodal distribution. The aim of this study was to screen the entire dataset for additional peptides with a bimodal isotope distribution and to specifically identify these peptides by subsequent MS/MS analysis.

2. Experimental Approach

The data analyzed in this study differs significantly from HX experiments in the EX1 regime. Fig. 1a shows the experimental scheme. ATP was added to HtpG and pre-incubated for 0 to 300 s at 30°C before the reaction was diluted into D_2O and incubated for exactly 10 s for pulse-labeling of the protein. Subsequently, the reaction was quenched by the addition of a low-pH-phosphate buffer and analyzed by LC-MS as described previously [14, 15]. As binding of ATP induces the transition of HtpG from the relaxed state with high flexibility and rapid deuteron incorporation to the tensed state with low flexibility and slow deuteron incorporation, isotope peak distributions migrate to lower m/z values with increasing pre-incubation time in the presence of ATP (Fig. 1b).

HeXicon automates the estimate of the deuteration distribution for the complete HX-MS dataset with strong sensitivity and balanced specificity. However, it is subject to a few limitations that motivated us to introduce further improvement on its workflow. Firstly, HeXicon does not have a built-in LC-MS alignment module and often fails to establish the correct inter-sample correspondences when the raw data exhibit severe retention time offsets. We used Parametric Time Warping (PTW)[16], a popular algorithm for chromatogram alignment, to align the total ion chromatograms of all samples to the reference ATP-free sample. Publicly available R package ptw was used on TICs, extracted from the raw data by the Analyst software suite. The retention time values of the raw data in mzXML format were updated before being processed by HeXicon. In principle, any algorithm for sample alignment can be used, and in future we hope to be able to provide an alignment module directly in HeXicon [17].

Secondly, HeXicon was originally developed for continuous-labeling time course experiments (individual samples differ by D_2O incubation time) and accordingly assumes monotonic increase of deuteron incorporation. However, this assumption is violated in our experiment setup since some peptides exhibit decreasing deuteron incorporation with increasing ATP incubation times. To overcome this limitation, one could process the samples with HeXicon in a pairwise manner such that samples incubated in the presence of ATP are compared only with the reference sample. The resulting deuteration distributions for all time points could be combined for each peptide identified by MS/MS or found by the coverage extension procedure. An alternative approach is to introduce the samples into HeXicon with the deuteration time parameter modified in such a way, that the time labels sequence is reversed while keeping the time interval constant. In this approach the deuteron incorporation behavior complies with the assumption of HeXicon. We tried both approaches and found that although the results are the same for very abundant peptides, the second approach is more robust for peptides with other high intensity peptides in the vicinity, where peaks found in other samples could help correct the peptide assignment.

Thirdly, we adapted the early stopping criterion used in the NITPICK algorithm to prevent under-fitting in case when the observed signal consists of numerous overlapping and correlated basis functions (separate peptide signals). In the previous version of HeXicon, the run-time was decreased by making the algorithm opt for the early break from the LARS regression iterations, when the correlation between the next basis function to add and the functions already in the list was larger than a given threshold. While this did not present problems in the regular NITPICK use case, HeXicon results sometimes suffered from under-fitting, as not all the necessary basis functions were considered. This limitation became especially clear during the analysis of bimodal peptide distributions, as their spectra present an even more complex mixture of possibly correlated basis functions. We introduced an additional check on the minimum of the Bayesian Information Criterion and the possibility for the algorithm flow to go back to the LARS iterations if the minimum was not deep enough. This modification is beneficial for all HeXicon use cases and is now part of the HeXicon package available on our website.

The last significant modification was concerned with the peptide quality score estimation. The final processing step of HeXicon ranks the peptides found in the previous steps with the help of a random forest classifier. However, some of the features used for classification, such as, for example, the variance of the deuteration distribution, are not applicable for the bimodal distribution search. Moreover, since in this limited case we are only interested in the peptides whose distributions follow a certain pattern, a ranking of all found peptides is not needed and can be replaced by a post-processing filtering procedure, tuned to the bimodal pattern of interest. The procedure checks at the deuteration distributions of peptides across all samples. It selects the peptides, for which at least for some samples a “2 (local) maxima/2-3 minima” or “1 maximum/2 minima” pattern is observed and several such samples share a maximum position, while at least one other has a different maximum. Sensitivity of the procedure is controlled by the minimal number of samples with the same maximum, after which the distribution is considered bimodal. HeXicon now provides the necessary option to export also the “low-quality” results, and the Matlab scripts, which we used for the subsequent filtering and plotting, are available on our website.

The data processing consisted of two parts: a large scale screening procedure, where all the samples were processed by HeXicon at the same time, and targeted peptide quantification, where HeXicon was applied on a restricted m/z and retention time range for selected peptides of interest. All data used in this study was collected on a QSTAR Pulsar mass spectrometer. The analyst .wiff files were converted to mzXML format by mzWiff tool [18], and no further pre-processing such as smoothing was applied.

3. Results and Discussion

3.1 Proof of principle

To test the performance of the modified HeXicon algorithm we first restricted the search to m/z values known to represent peptides with bimodal isotope distributions. All seven previously known peptides were identified. Four of these show deuteration distributions that fit well to two Gaussian peaks with the maximum of the earlier time points (red colors) at higher deuteron numbers than the maximum of later time points (blue colors) as expected (Fig. 1c, suppl. Fig. 1b and data not shown). Thereby, in some cases the distribution of the hydrogen exchange in the absence of ATP (0 s time point) is

shifted to higher deuteron numbers as compared to the distribution of the 10 s time point, indicating the fast phase of the conformational change that cannot be analyzed with our HX-MS setup (Fig. 2b and suppl. Fig. 1b). For three peptides the deuteration distribution is not Gaussian shaped but a separation between earlier and later time points is visible (Fig. 2, suppl. Fig. 2 and data not shown). Tuning HeXicon parameters did not lead to more regular shaped deuteration distributions for these peptides, but the separation between fast and slow exchange behaviors was present for all HeXicon parameter settings.
Close inspection of the spectra for which HeXicon did not yield clearly separate Gaussian distributions reveals that these spectra are either of low intensity with poor signal to noise ratios (suppl. Fig. 2a) or contain an equally spaced isotope cluster originating from a different peptide which disturbed the determination of the deuteration distribution (Fig. 2a). These isotope clusters were clearly separated for the undeuterated peptides but merged after deuteration.

3.2 Analysis of a full data set

After verifying that the adapted HeXicon was able to detect bimodal isotope distributions we applied the program to the entire dataset with very sensitive post-processing filter settings to avoid missing interesting peptides. The filter produced a plot for a peptide, if in more than two samples the deuteration distributions had the same maximum and at least one more had a different maximum. It only considered distributions following the pattern of "2 (local) maxima/2-3 minima" or "1 maximum/2 minima". Such a sensitive setting led to a certain amount of false positive results, which were discarded by manual inspection of the produced distribution plots. Examples of the inspected plots can be seen in the figures of this contribution (Fig.1c and similar). HeXicon identified 27 additional, previously unidentified peptides with deuteration distribution that suggests two populations with distinct exchange kinetics. Nine of these peptides were verified to exhibit ATP-pre-incubation-time-dependent bimodal isotope distributions. For eight of these peptides the identity was determined by a follow-up MS/MS analysis with their monoisotopic masses in the inclusion list (Fig. 3). Three of the eight peptides were instances of known species at previously unobserved charge states. As expected, the deuteration distributions determined by HeXicon for the same peptide sequence at different charge states are highly similar (compare Fig. 1 and suppl. Fig. 3, and suppl. Fig. 4 and suppl. Fig. 5). The remaining five peptides overlap with the previously identified peptides (Fig. 3). In particular, one of them represented a rather short sub-sequence (residues 13-18) of a known longer peptide (residues 1-19) (Fig. 4). This peptide shows one amide hydrogen that becomes completely protected in the dead-time of the experiments (see differences between 0 s ATP and 10 s ATP in Fig. 4a and b) and a second amide hydrogen that is protected in about half of the peptide population after 10 s and becomes protected in the rest of the population during the following 300 s pre-incubation time in ATP. The last exchanging amide hydrogen of this peptide also becomes increasingly protected during the pre-incubation in ATP and the population of peptides, which has not exchanged any amide hydrogens, increases with time to about 30% of the total population. The transition between 10 and 300s is close to the smallest change in deuteron incorporation possible and was detected nicely by HeXicon software. Identifying the peptide residues 13-18 allowed us also to better understand the mechanism of N-terminal dimerization of HtpG. This peptide encompasses a region in HtpG that forms a small α-helix in the homologous yeast Hsp82, which was crystallized

in complex with AMPPNP and the cochaperone Sba1 (suppl. Fig. 6). We previously assumed that this α-helix becomes stabilized in the Hsp82 structure upon docking of the two N-terminal domains and by interaction with the cochaperone Sba1. The data for the HtpG peptide identified by HeXicon clearly show that most of the stabilization occurs in the dead time of the experiment, which is considerably before N-terminal docking as determined previously by fluorescence measurements [12]. The stabilization of this α-helix may even be a prerequisite for the docking. This will be explored in a future mutagenesis study.

To extract the half-life of the conformational change the normalized relative abundance of the part of the deuteration distribution corresponding to the low exchanging population (indicated by the gray horizontal bar in Figs. 1c, 2b, 4b, and suppl. Figs. 1b to 5b) was summed and plotted against the pre-incubation time in ATP (Figs. 1d, 2c, 4c and suppl. Figs). Single exponential functions were fitted to the data. Thereby the first data point (deuteration in the absence of ATP) was omitted for peptides that show a strong change in deuteration within the dead-time of the experiment due to the fast first step of the transition.

For the cases with two clearly separated Gaussian peaks (Fig. 1c and suppl. Figs. 1, 3, 4, and 5) the range to be summed up for the determination of the kinetics should obviously encompass one of the two peaks, which are indicated by the cross-over point of the deuteration distributions. For cases with no clear separation of the deuteration distribution into Gaussian peaks it is less clear, which range should be summed up for the analysis. Therefore, we tested how the summation range affects the calculated rate constants in the analysis (suppl. Fig. 7 and data not shown). Interestingly, the summation range did not severely affect the determined rate constant except for the smallest range.

This approach is not only suitable for pulse-labeling experiments, for which bimodal isotope distributions indicate the coexistance of alternative conformations and allows the determination of the interconversion rate between the alternative conformations. It could also be used for continuous-labeling experiments to distinguish EX1 and EX2 exchange mechanisms. The results for EX1 regimes would look reverse as compared to our results, since the bimodal distribution of the isotope peaks in EX1 exchange is characterized by the appearance of high exchanged species, while in our pulse-labeling experiments low exchanged species are observed after increasing pre-incubation time in the presence of ATP. The cause for EX1 behavior in continuous-labeling exchange experiments is either reversible local or global unfolding with refolding rates much smaller than the intrinsic chemical exchange rates or irreversible local or global unfolding [14, 19-21]. More difficult would be the analysis of a mixed EX1-EX2 exchange behavior. In such an exchange regime the distribution of the isotope peaks would be bimodal with changing ratios between low and high exchanged species and concomitant shift of one or both maxima of the bimodal distribution to higher m/z values. While HeXicon should be able to find the correct deuteration distribution, the subsequent analysis of the kinetics would require an adaptation of the summation range to the shifting deuteration distribution. This should not be a problem when the deuteration distribution exhibits two clear Gaussian peaks since the summation range could be oriented relative to the peak maxima. In other cases one could sum a fixed range of the deuteration scale starting with the lowest detected number of incorporated deuterons.

In addition to our HeXicon analysis, we analyzed the width of the isotope peak distribution as proposed by Weis and coauthors using the HX-Express software tool [5, 9] (Figs. 1e, 2d, 4d and suppl. Figs). For most peptides the half-lives determined by

fitting the HeXicon data coincide reasonably well with those estimated by the width of the isotope distributions.

3.3 Analysis of false positive candidates.

Some of the peptides suggested by HeXicon do not really exhibit a bimodal exchange behavior (see Fig. 5). The reason for these false positive results is in most cases a second peptide with a close monoisotopic mass co-eluting with the peptide of interest. For example, as shown in Fig. 5a, the isotope profiles of peptides at m/z 526.30 and 528.34 overlap for D_2O and ATP incubated samples, misleading HeXicon to incorrectly assign all observed peaks to the same peptide. The seemingly bimodal distribution is then just a sum of the deuteration distributions of two individual peptides. Another example of this behavior is demonstrated in Fig. 5c-5d, where a co-eluting peptide of a lower mass interferes with the isotope distribution of the peptide at m/z 960.00, and again causes HeXicon to find a false positive bimodal isotope distribution. The underlying cause for such incorrect assignments is the fairly large m/z error tolerance of 0.1 Da used by HeXicon when establishing correspondences between the peptides in the reference sample and their possible deuterated forms in other samples. Based on our experience with other, non-HX experiments, we anticipate that this will be considerably less of a problem with experiments that are performed on a higher resolution instrument. As the measured signal peaks become narrower, the isotope distributions of neighboring peptides will no longer overlap, unless it's a perfect overlap of the same peptide in different deuteration states. We could then lower the tolerance for m/z deviations between the supposed signals from the same peptide across different samples (currently, a difference of up to 0.1Da is considered acceptable) and thus improve the accuracy of the peptide matching across the samples.

Other examples for false-positive assignments of bimodal exchange behaviors could result from carry-over contaminations. Such samples would exhibit bimodal isotope distribution with seemingly unexchanged species appearing. Since in these cases the bimodal distributions are real, the problem cannot be solved by mass spectrometer or analysis software but only by improving the HPLC conditions, for example by introducing a run without injection in between two analysis runs. HeXicon could help to detect such problems since the amplitude of the appearing unexchanged fraction should not depend on the incubation times or conditions.

4. Conclusions

In summary it can be stated that HeXicon is very suitable for detecting bimodal isotope distributions. False positives generally originate from co-eluting peptides with similar m/z causing merging isotope distributions after deuteration. A poor signal to noise ratio may also contribute to erroneous estimation of bimodal deuteration distributions. While the first problem will be reduced in instruments with higher resolution, the second requires improved sensitivity or modified sample handling. The ability of HeXicon to automatically extract peptides with a bimodal isotope distribution may lead to a reversal of the workflow. Instead of first establishing the sequence coverage by MS/MS experiments, one could first perform an HX experiment, screen the data with HeXicon searching for specific features, and subsequently identify only the peptides of interest to the researcher by a targeted inclusion list MS/MS experiment. Such an approach would be particularly advantageous for very large proteins or a complex of several proteins.

Acknowledgements

We like to thank Thomas Ruppert for his help in the Facility for Mass Spectrometry and Proteomics of the ZMBH. This work was funded by the Deutsche Forschungsgemeinschaft (SFB638 to M.P.M., HA-4364/2 to A.K., F.A.H.) and the Alexander von Humboldt Foundation (3.1-DEU/1134241 to M.K.), the Heinz Goetze Memorial Fellowship (to X.L.), the Helmholtz Alliance on Systems Biology (to X.L., F.A.H.) and Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (to A.K.).

Figure Legends

Fig. 1: Comparison of HeXicon results and manual analysis of a previously found peptide with a bimodal isotope distribution. (**a**) Experimental design of the pulse labeling reaction. The HtpG protein was equilibrated at pH 7.5 and 30°C. ATP was added and the reaction incubated at 30°C for 0 to 300 s. Then the reaction was diluted into D_2O and incubated for 10 s at 30°C. The reaction was subsequently quenched by addition of the low pH phosphate buffer and analyzed by LC-ESI-MS as described previously [15]. (**b**) Selected mass spectra for the peptide residues 20 to 31 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**c**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 490-510 in m/z, unlimited in retention time dimension). (**d**) Analysis of the kinetics shown in (c). The abundance of deuteron incorporations indicated by the gray bar in (c) relative to the total distribution, which corresponds to the fraction of HtpG molecules in the tensed, slow exchanging state (% tensed), is plotted versus the pre-incubation time. The solid line is a single exponential fit to the data. (**e**) Analysis of the width of the isotope distribution according to a previously published procedure [9]. The width of the distribution at 20% and 50% maximal peak intensity is shown.

Fig. 2: Analysis of a result with non-Gaussian deuteration distribution modes. (**a**) Selected mass spectra for the peptide residues 89 to 98 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 520-540 in m/z, unlimited in retention time dimension). (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. The solid line is a single exponential fit to the data. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [9]. The width of the distribution at 20% and 50% maximal peak intensity is shown. A color version of the figure is found in supplemental material.

Fig. 3: Sequence coverage of peptic peptides of HtpG. The amino acid sequence of *E. coli* HtpG is shown with the secondary structure elements as found in the crystal structure of nucleotide-free HtpG (PDB entry code 2IOQ; [22]) above the sequence (dashed lines are not resolved in the crystal structure). The dashed and dotted lines below the sequence indicate the peptic peptides that were analyzed previously [12]. Dotted lines indicate previously known peptides with bimodal distribution. Black solid lines indicate the additional peptides found in this study by HeXicon and verified manually to be correct. All peptides were identified by tandem mass spectrometry.

Fig. 4: Comparison of HeXicon results and manual analysis of a previously unknown peptide with bimodal isotope distribution. (**a**) Selected mass spectra for the peptide residues 13 to18 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**)

Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 420-440 in m/z, unlimited in retention time dimension). (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. The solid line is a single exponential fit to the data minus the first data point. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [9]. The width of the distribution at 20% and 50% maximal peak intensity is shown. A color version of the figure is found in supplemental material.

Fig. 5: Analysis of false positive candidates. (**a**) Selected mass spectra for the peptide with mass 526.30^{2+} in H_2O, 10 s D_2O in the absence of ATP (0) and 120 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times (0 to 300s as indicated) as evaluated by HeXicon (As these peptides were found in the screening of the whole dataset, HeXicon was run without limits in either m/z or retention time). (**c**) and (**d**) same as (a) and (b) for the peptide with mass 960.00^{2+}. A color version of the figure is found in supplemental material.

References

[1] K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins, Nature 450 (2007) 964-972.

[2] M.M. Krishna, L. Hoang, Y. Lin, S.W. Englander, Hydrogen exchange methods to study protein folding, Methods 34 (2004) 51-64.

[3] J.R. Engen, Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS, Anal. Chem. 81 (2009) 7870-7875.

[4] A.N. Hoofnagle, K.A. Resing, N.G. Ahn, Protein analysis by hydrogen exchange mass spectrometry, Annu. Rev. Biophys. Biomol. Struct. 32 (2003) 1-25.

[5] D.D. Weis, J.R. Engen, I.J. Kass, Semi-automated data processing of hydrogen exchange mass spectra using HX-Express, J. Am. Soc. Mass. Spectrom. 17 (2006) 1700-1703.

[6] B.D. Pascal, M.J. Chalmers, S.A. Busby, C.C. Mader, M.R. Southern, N.F. Tsinoremas, P.R. Griffin, The Deuterator: software for the determination of backbone amide deuterium levels from H/D exchange MS data, BMC Bioinformatics 8 (2007) 156.

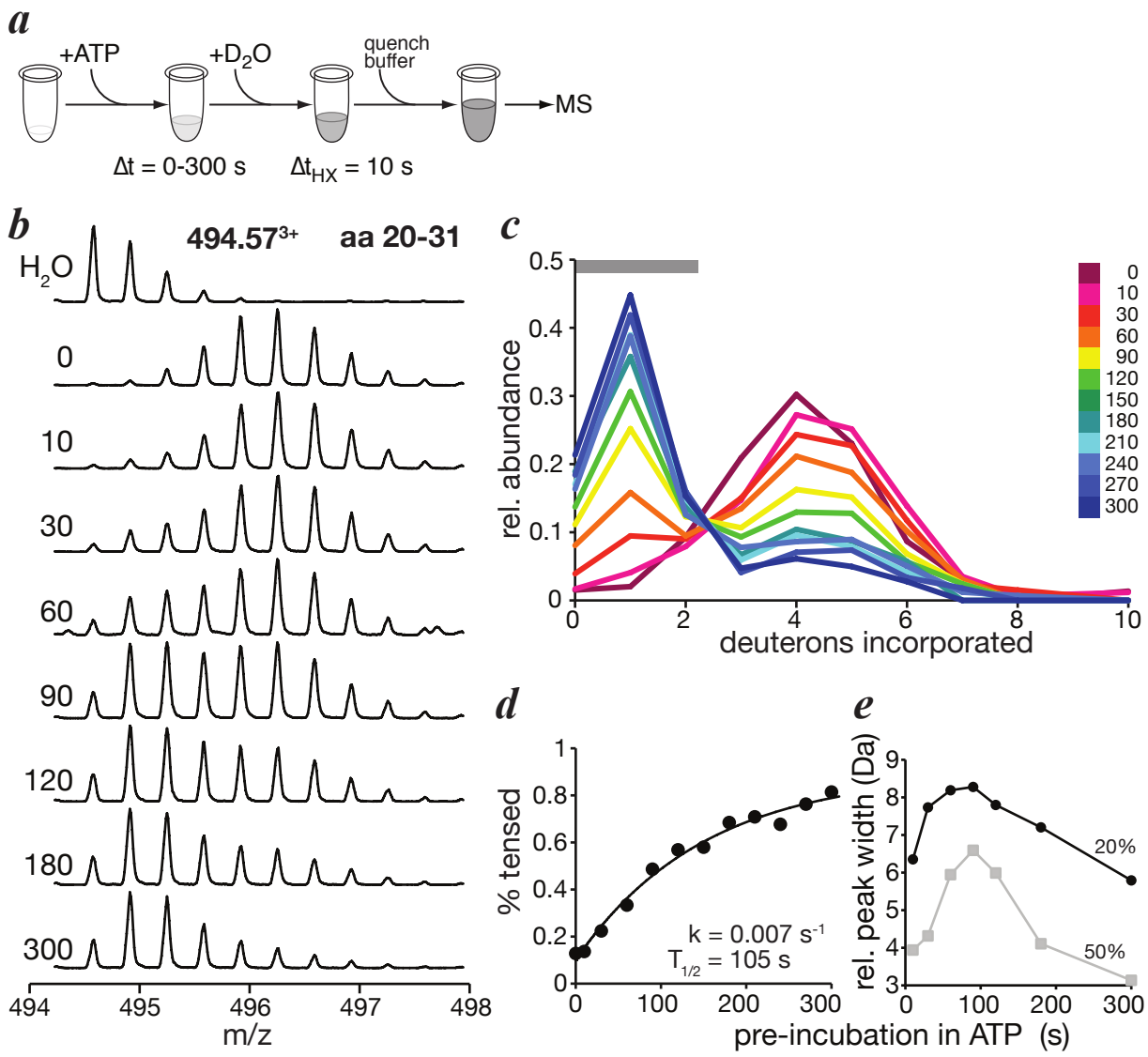
[7] B.D. Pascal, M.J. Chalmers, S.A. Busby, P.R. Griffin, HD desktop: an integrated platform for the analysis and visualization of H/D exchange data, J. Am. Soc. Mass. Spectrom. 20 (2009) 601-610.

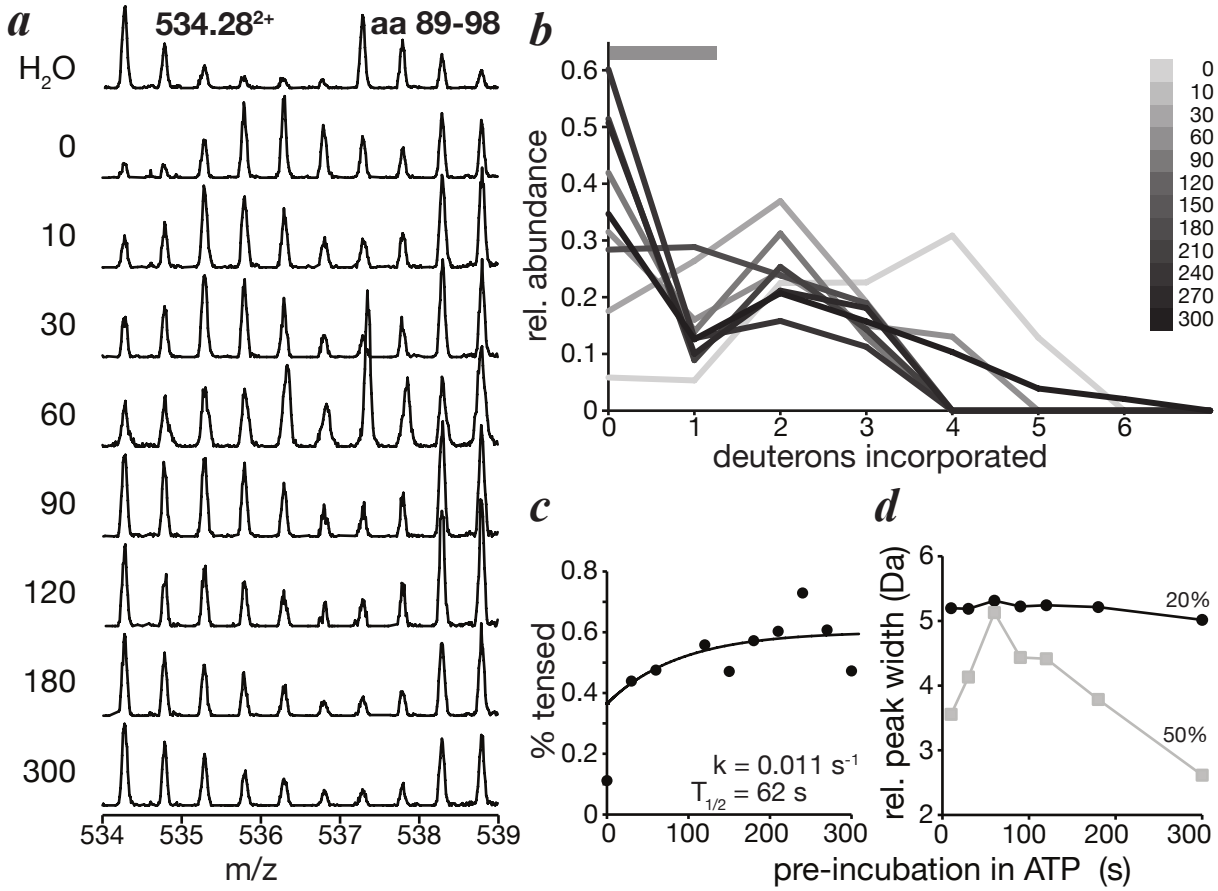
[8] G.W. Slysz, C.A. Baker, B.M. Bozsa, A. Dang, A.J. Percy, M. Bennett, D.C. Schriemer, Hydra: software for tailored processing of H/D exchange data from MS or tandem MS analyses, BMC Bioinformatics 10 (2009) 162.

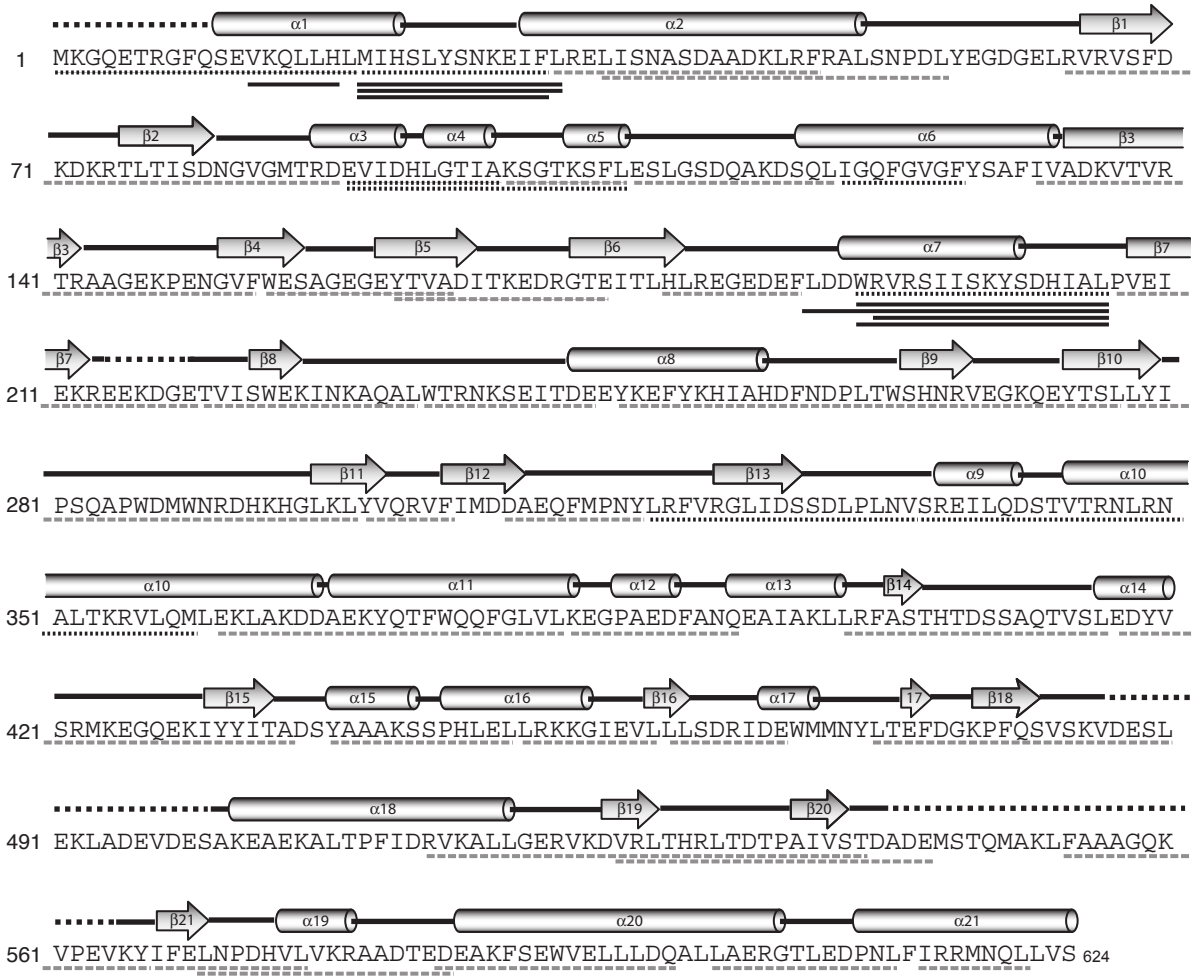
[9] D.D. Weis, T.E. Wales, J.R. Engen, M. Hotchko, L.F. Ten Eyck, Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis, J. Am. Soc. Mass. Spectrom. 17 (2006) 1498-1509.

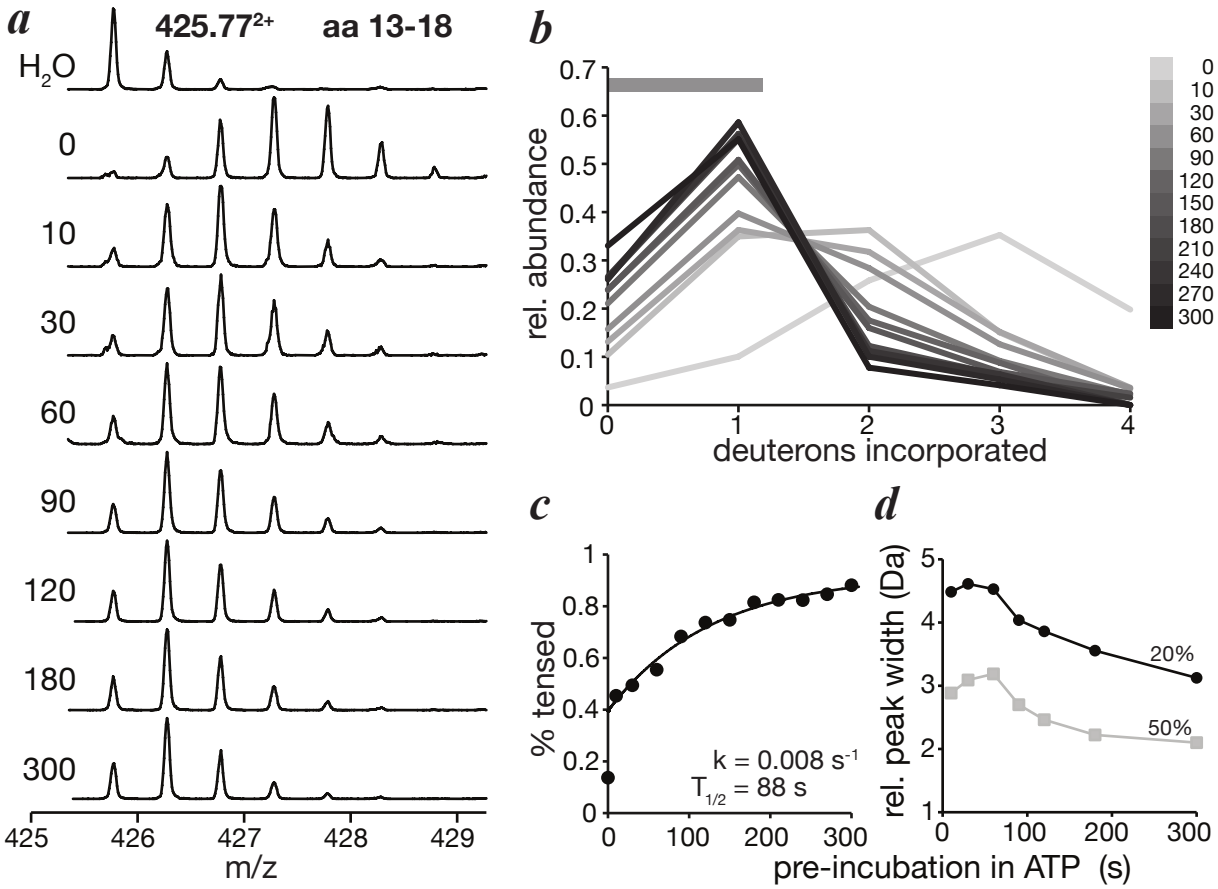
[10] X. Lou, M. Kirchner, B.Y. Renard, U. Kothe, S. Boppel, C. Graf, C.T. Lee, J.A. Steen, H. Steen, M.P. Mayer, F.A. Hamprecht, Deuteration Distribution Estimation with Improved Sequence Coverage for HX/MS Experiments, Bioinformatics (2010).

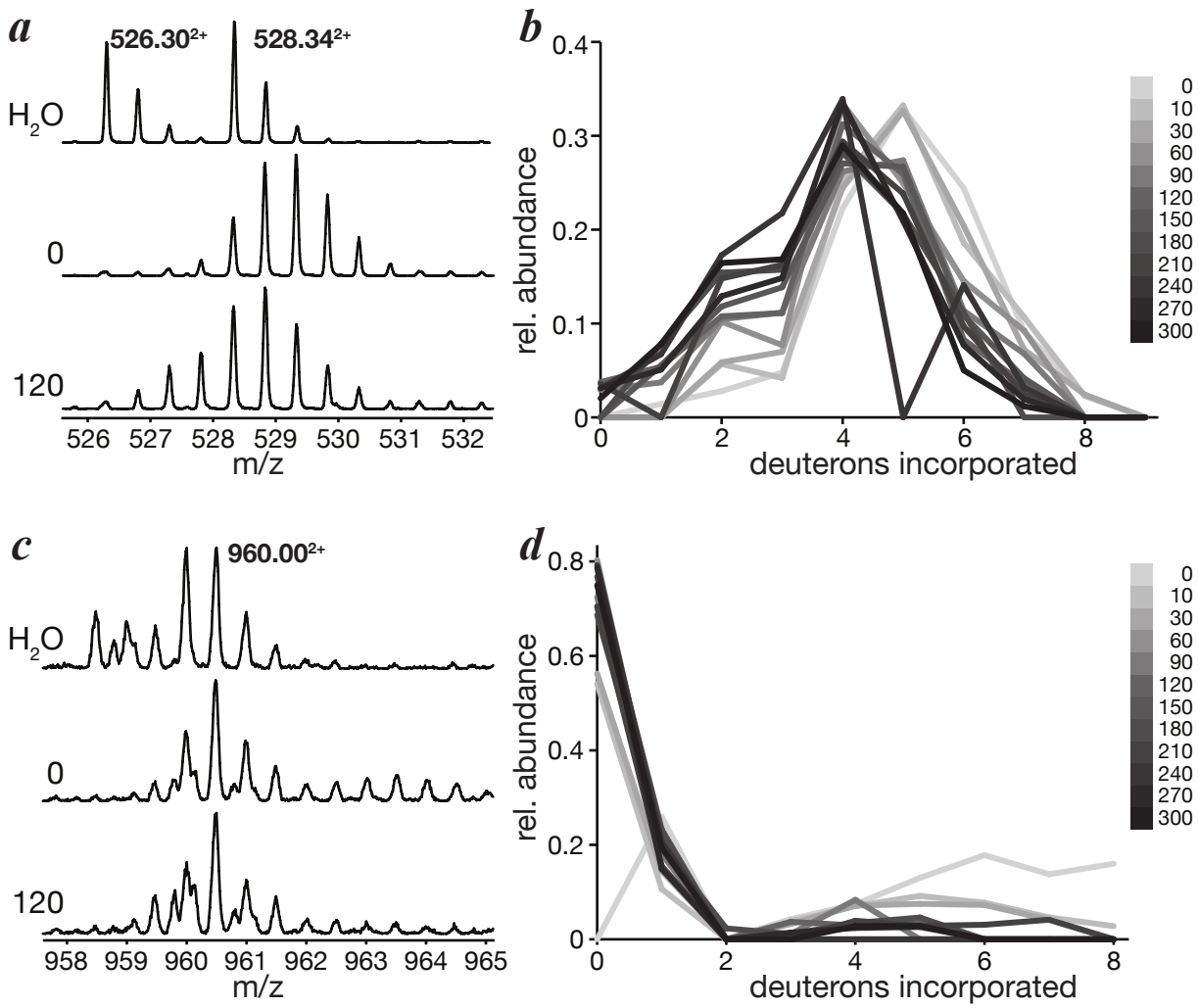
[11] B.Y. Renard, M. Kirchner, H. Steen, J.A. Steen, F.A. Hamprecht, NITPICK: peak identification for mass spectrometry data, BMC Bioinformatics 9 (2008) 355.
[12] C. Graf, M. Stankiewicz, G. Kramer, M.P. Mayer, Spatially and kinetically resolved changes in the conformational dynamics of the Hsp90 chaperone machine, EMBO J. 28 (2009) 602-613.
[13] H. Roder, G.A. Elove, S.W. Englander, Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR, Nature 335 (1988) 700-704.
[14] W. Rist, T.J.D. Jørgensen, P. Roepstorff, B. Bukau, M.P. Mayer, Mapping temperature-induced conformational changes in the Escherichia coli heat shock transcription factor sigma 32 by amide hydrogen exchange, J. Biol. Chem. 278 (2003) 51415-51421.
[15] W. Rist, C. Graf, B. Bukau, M.P. Mayer, Amide hydrogen exchange reveals conformational changes in hsp70 chaperones important for allosteric regulation, J. Biol. Chem. 281 (2006) 16493-16501.
[16] T.G. Bloemberg, J. Gerretzen, H.J.P. Wouters, J. Gloerich, M. van Dael, H.J.C.T. Wessels, L.P. van den Heuvel, P.H.C. Eilers, L.M.C. Buydens, R. Wehrens, Improved parametric time warping for proteomics, Chemometrics and Intelligent Laboratory Systems (2010) in press.
[17] B.M. Voss, B.Y. Renard, A. Kreshuk, M. Hanselmann, U. Koethe, H. Steen, J.A.J. steen, M. Kirchner, F.A. Hamprecht, Simultaneous Multiple Alignment for LC/MS Peak Lists., Conference of the American Society for Mass Spectrometry (ASMS) 2009 [A journal article and the corresponding C++ package are currently in preparation] (2009).
[18] A. Keller, J. Eng, N. Zhang, X.J. Li, R. Aebersold, A uniform proteomics MS/MS analysis platform utilizing open XML file formats, Mol. Syst. Biol. 1 (2005) 2005 0017.
[19] T.E. Wales, J.R. Engen, Hydrogen exchange mass spectrometry for the analysis of protein dynamics, Mass. Spectrom. Rev. 25 (2006) 158-170.
[20] J.R. Engen, T.E. Smithgall, W.H. Gmeiner, D.L. Smith, Identification and localization of slow, natural, cooperative unfolding in the hematopoietic cell kinase SH3 domain by amide hydrogen exchange and mass spectrometry, Biochemistry 36 (1997) 14384-14391.
[21] T.E. Wales, J.R. Engen, Partial unfolding of diverse SH3 domains on a wide timescale, J. Mol. Biol. 357 (2006) 1592-1604.
[22] A.K. Shiau, S.F. Harris, D.R. Southworth, D.A. Agard, Structural Analysis of E. coli hsp90 reveals dramatic nucleotide-dependent conformational rearrangements, Cell 127 (2006) 329-340.

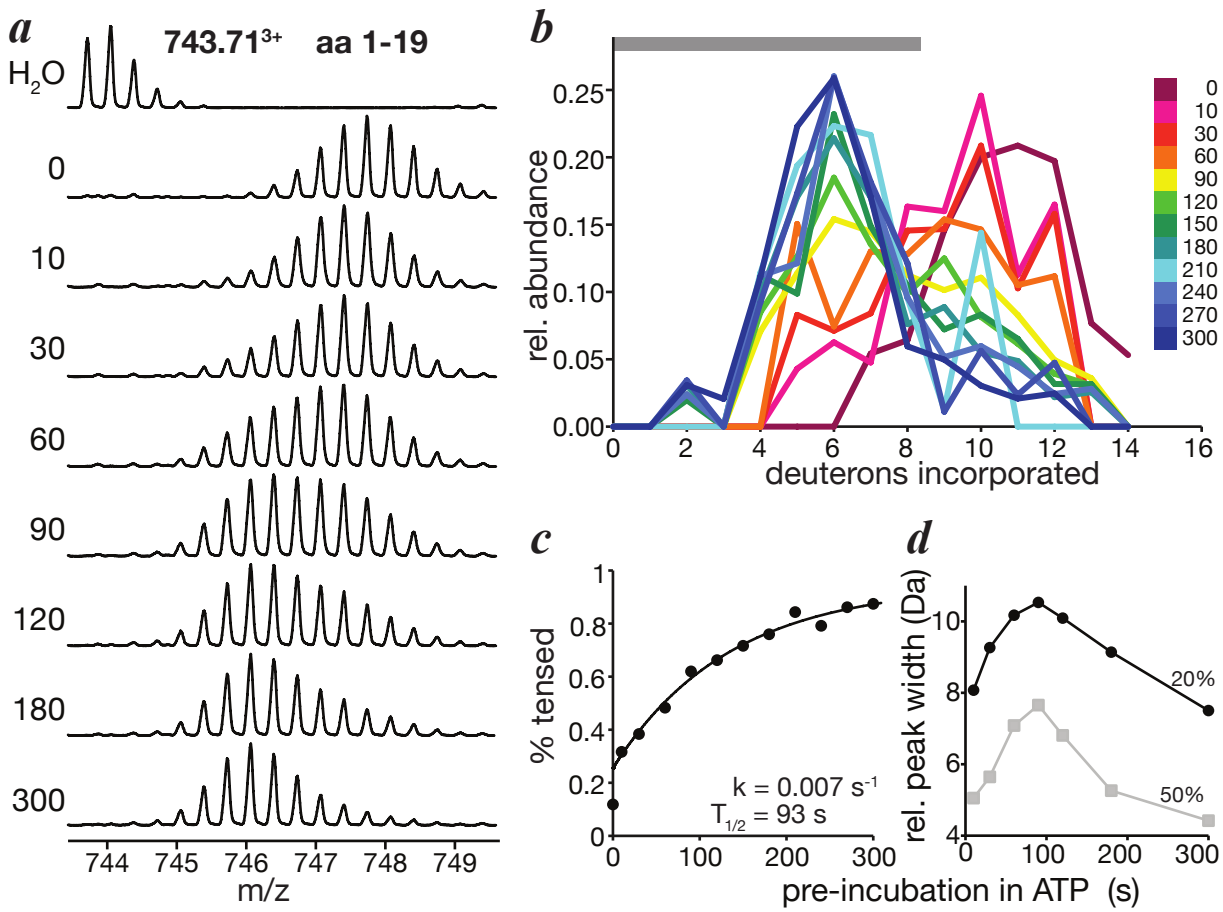

a
+ATP
+D2O
quench buffer
MS
Δt = 0-300 s
ΔtHX = 10 s
b
494.57 3+
aa 20-31
H2O
0
10
30
60
90
120
180
300
494
495
496
497
498
m/z
c
rel. abundance
deuterons incorporated
0
10
30
60
90
120
150
180
210
240
270
300
d
% tensed
k = 0.007 s-1
T1/2 = 105 s
e
rel. peak width (Da)
20%
50%
pre-incubation in ATP (s)


a
534.28²⁺
aa 89-98
H2O
0
10
30
60
90
120
180
300
534
535
536
537
538
539
m/z
b
rel. abundance
deuterons incorporated
0
10
30
60
90
120
150
180
210
240
270
300
c
% tensed
k = 0.011 s-1
T1/2 = 62 s
d
rel. peak width (Da)
20%
50%
pre-incubation in ATP (s)

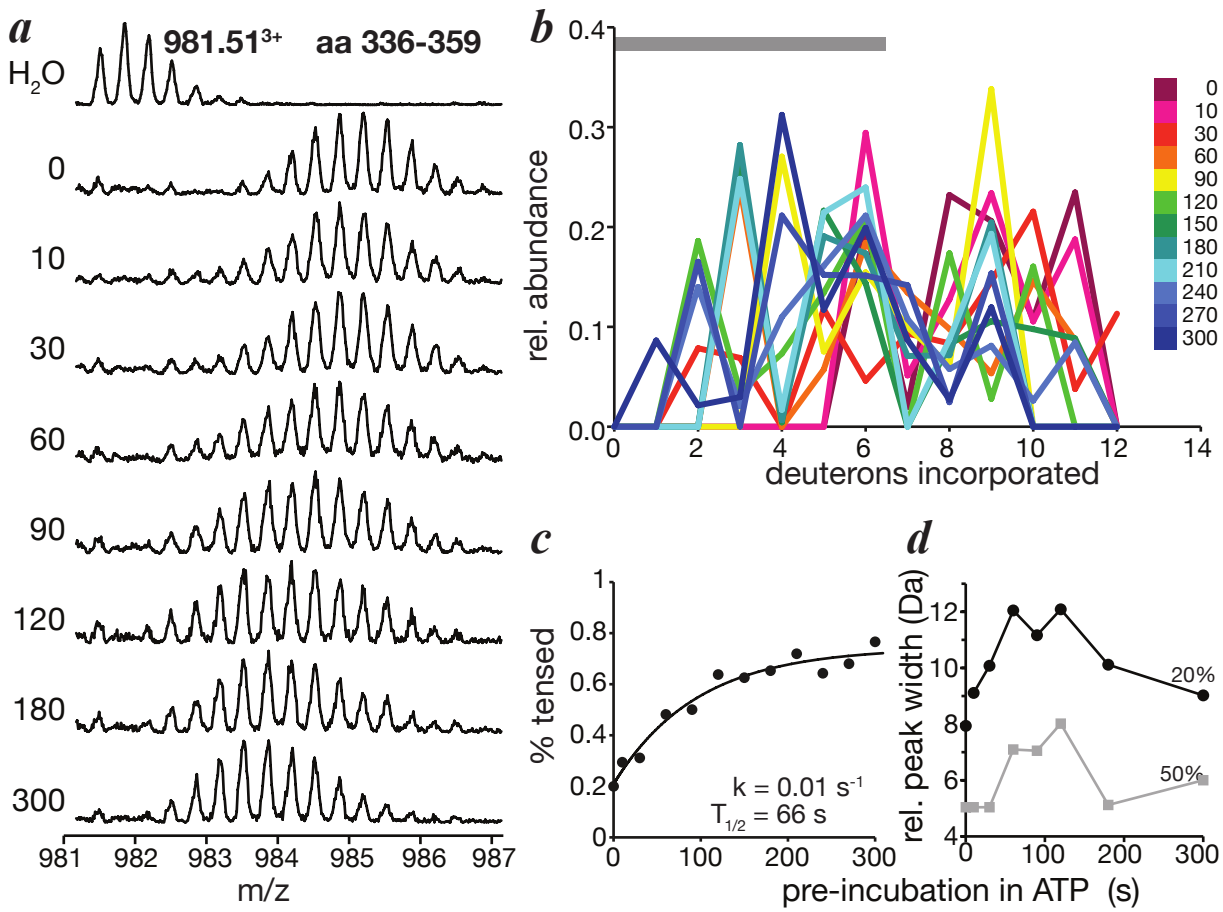

α1 α2 β1
1 MKGQETRGFQSEVKQLLHLMIHSLYSNKEIFLRELISNASDAADKLRFRALSNPDLYEGDGELRVRVSFD
β2 α3 α4 α5 α6 β3
71 KDKRTLTISDNGVGMTRDEVIDHLGTIAKSGTKSFLESLGSDQAKDSQLIGQFGVGFYSAFIVADKVTVR
β3 β4 β5 β6 α7 β7
141 TRAAGEKPENGVFWESAGEGEYTVADITKEDRGTEITLHLREGEDEFLDDWRVRSIISKYSDHIALPVEI
β7 β8 α8 β9 β10
211 EKREEKDGETVISWEKINKAQALWTRNKSEITDEEYKEFYKHIAHDFNDPLTWSHNRVEGKQEYTSLLYI
β11 β12 β13 α9 α10
281 PSQAPWDMWNRDHKHGLKLYVQRVFIMDDAEQFMPNYLRFVRGLIDSSDLPLNVSREILQDSTVTRNLRN
α10 α11 α12 α13 β14 α14
351 ALTKRVLQMLEKLAKDDAEKYQTFWQQFGLVLKEGPAEDFANQEAIAKLLRFASTHTDSSAQTVSLEDYV
β15 α15 α16 β16 α17 17 β18
421 SRMKEGQEKIYYITADSYAAAKSSPHLELLRKKGIEVLLLSDRIDEWMMNYLTEFDGKPFQSVSKVDESL
α18 β19 β20
491 EKLADEVDESAKEAEKALTPFIDRVKALLGERVKDVRLTHRLTDTPAIVSTDADEMSTQMAKLFAAAGQK
β21 α19 α20 α21
561 VPEVKYIFELNPDHVLVKRAADTEDEAKFSEWVELLLDQALLAERGTLEDPNLFIRRMNQLLVS 624


a
425.77²⁺ aa 13-18
H2O
0
10
30
60
90
120
180
300
425
426
427
428
429
m/z
b
rel. abundance
deuterons incorporated
0
10
30
60
90
120
150
180
210
240
270
300
c
% tensed
k = 0.008 s⁻¹
T1/2 = 88 s
d
rel. peak width (Da)
20%
50%
pre-incubation in ATP (s)

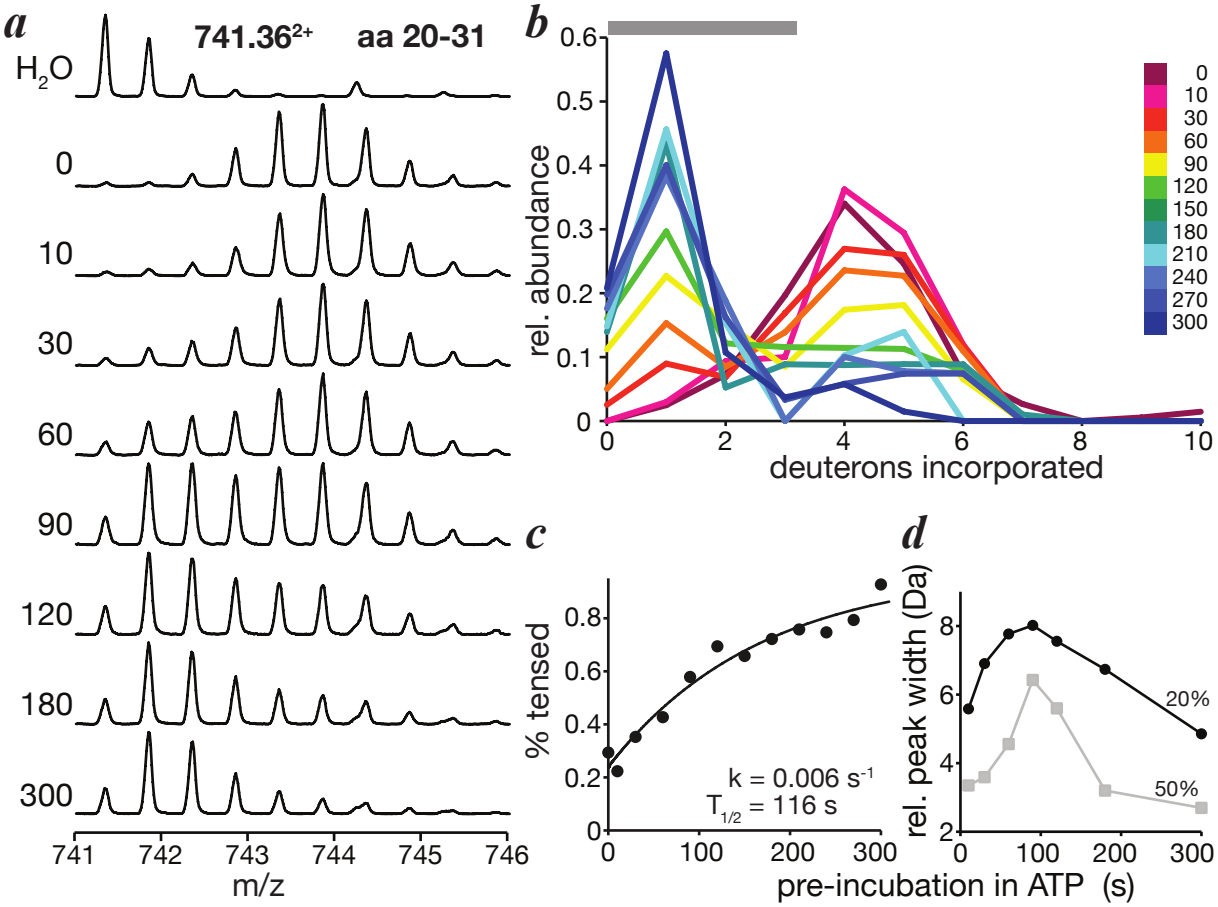

a
526.30²⁺
528.34²⁺
H2O
0
120
526 527 528 529 530 531 532
m/z
b
rel. abundance
0.4
0.3
0.2
0.1
0
0 2 4 6 8
deuterons incorporated
0 10 30 60 90 120 150 180 210 240 270 300
c
960.00²⁺
H2O
0
120
958 959 960 961 962 963 964 965
m/z
d
rel. abundance
0.8
0.6
0.4
0.2
0
0 2 4 6 8
deuterons incorporated
0 10 30 60 90 120 150 180 210 240 270 300



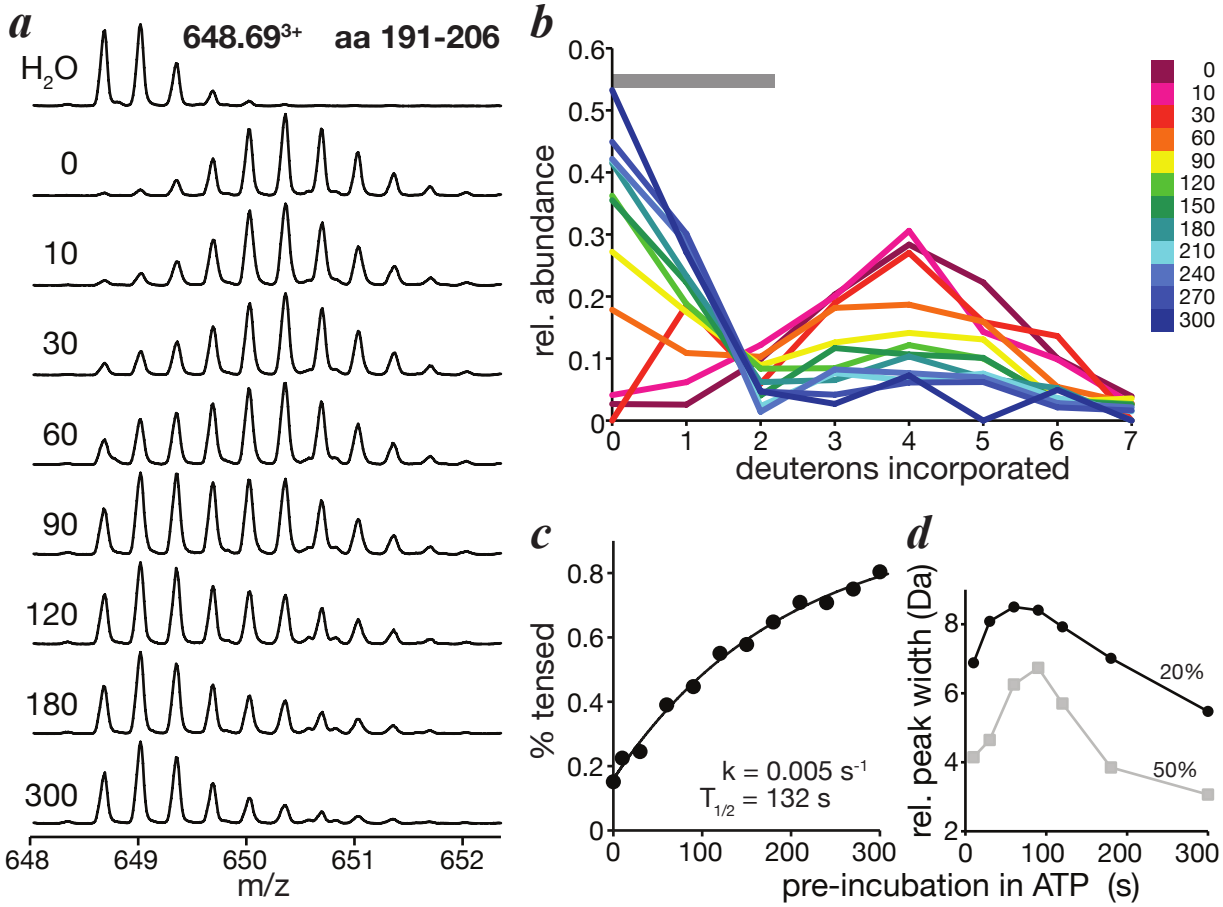
Suppl. Fig. 1: Comparison of HeXicon results and manual analysis of a previously known peptide with bimodal isotope distribution. (**a**) Selected mass spectra for the peptide residues 1 to 19 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon. [Spectrum limits for HeXicon were set to 740-760 in m/z, retention time dimension 700-840; retention time limits had to be introduced to avoid an overlap with a very abundant peptide at 741.36, suppl. Fig. 3. It is not always possible to set the retention time limits in such a way that on one hand the peptide is isolated and on the other hand HeXicon has enough information to compute global spectrum parameters, such as the width of the peak shape function. In most cases retention time limits are not needed, as HeXicon can do the segmentation into 1 peptide regions quite well, but such strong peptides as 741.36 can shadow the rest]. (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. The solid line is a single exponential fit to the data minus the first data point. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [11]. The width of the distribution at 20% and 50% maximal peak intensity is shown.



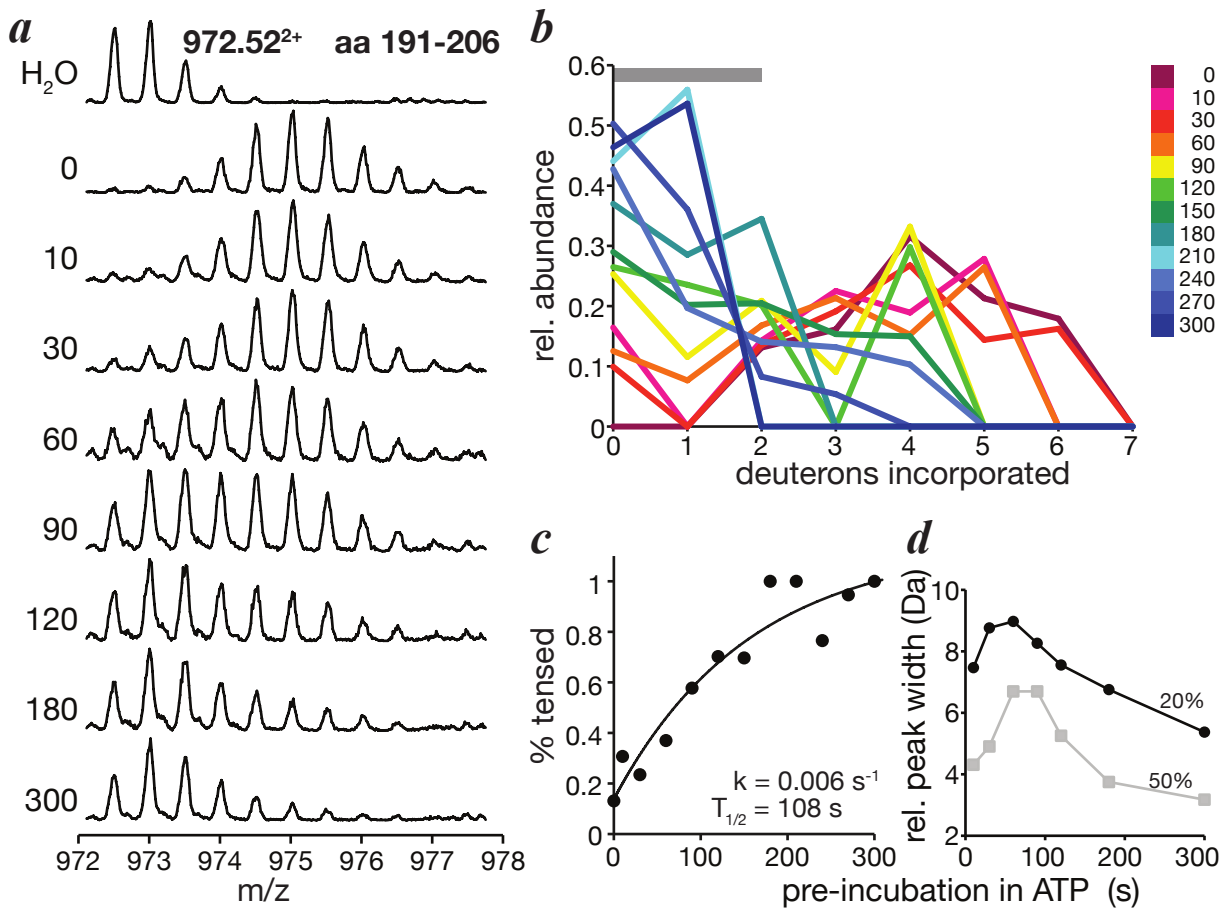
Suppl. Fig. 2: Comparison of HeXicon results and manual analysis of a previously known peptide with bimodal isotope distribution. (**a**) Selected mass spectra for the peptide residues 336 to 359 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 975-995 in m/z, unlimited in retention time dimension). (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. The solid line is a single exponential fit to the data. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [11]. The width of the distribution at 20% and 50% maximal peak intensity is shown.



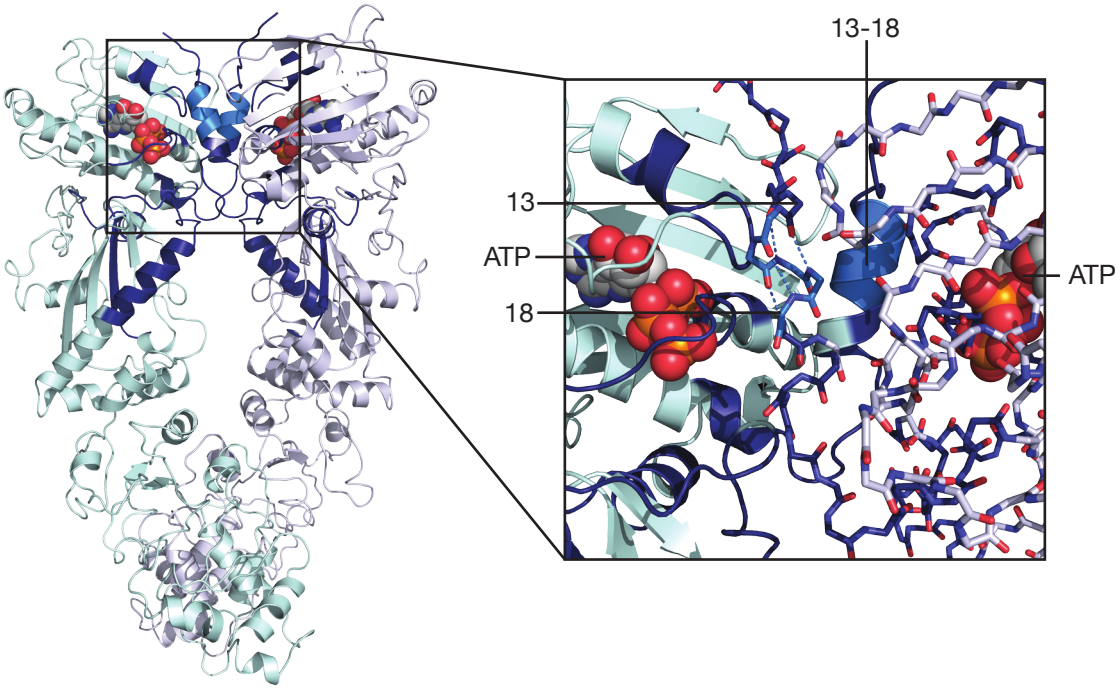
Suppl. Fig. 3: Comparison of HeXicon results and manual analysis of a previously unknown peptide with bimodal isotope distribution. (**a**) Selected mass spectra for the peptide residues 20 to 31 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 740-760 in m/z, unlimited in retention time dimension). (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. The solid line is a single exponential fit to the data. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [11]. The width of the distribution at 20% and 50% maximal peak intensity is shown.



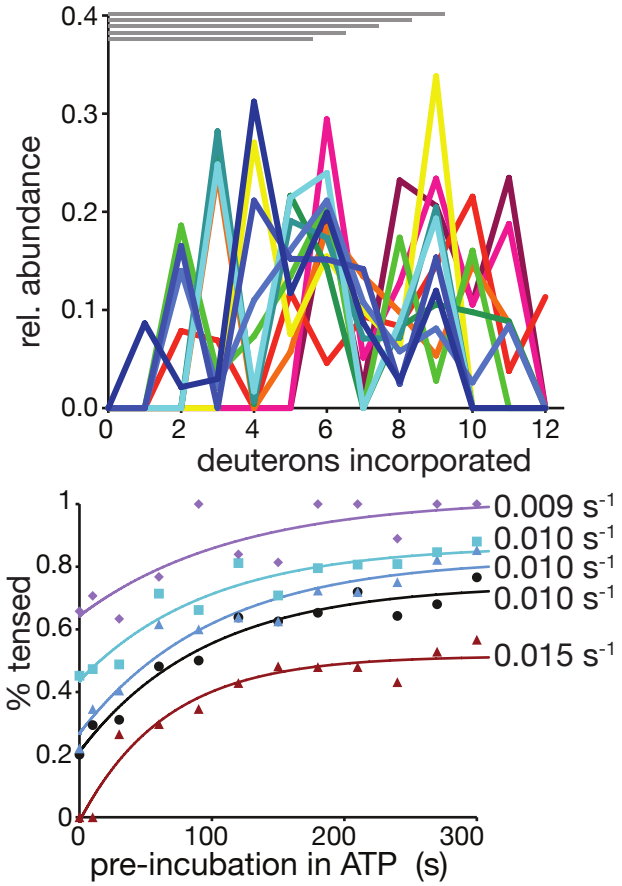
Suppl. Fig. 4: Comparison of HeXicon results and manual analysis of a previously known peptide with bimodal isotope distribution. (**a**) Selected mass spectra for the peptide residues 120 to 127 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 640-660 in m/z, unlimited in retention time dimension). (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. Since the deuteration distributions for early and late time points did not overlap the kinetics could not be fitted. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [11]. The width of the distribution at 20% and 50% maximal peak intensity is shown.



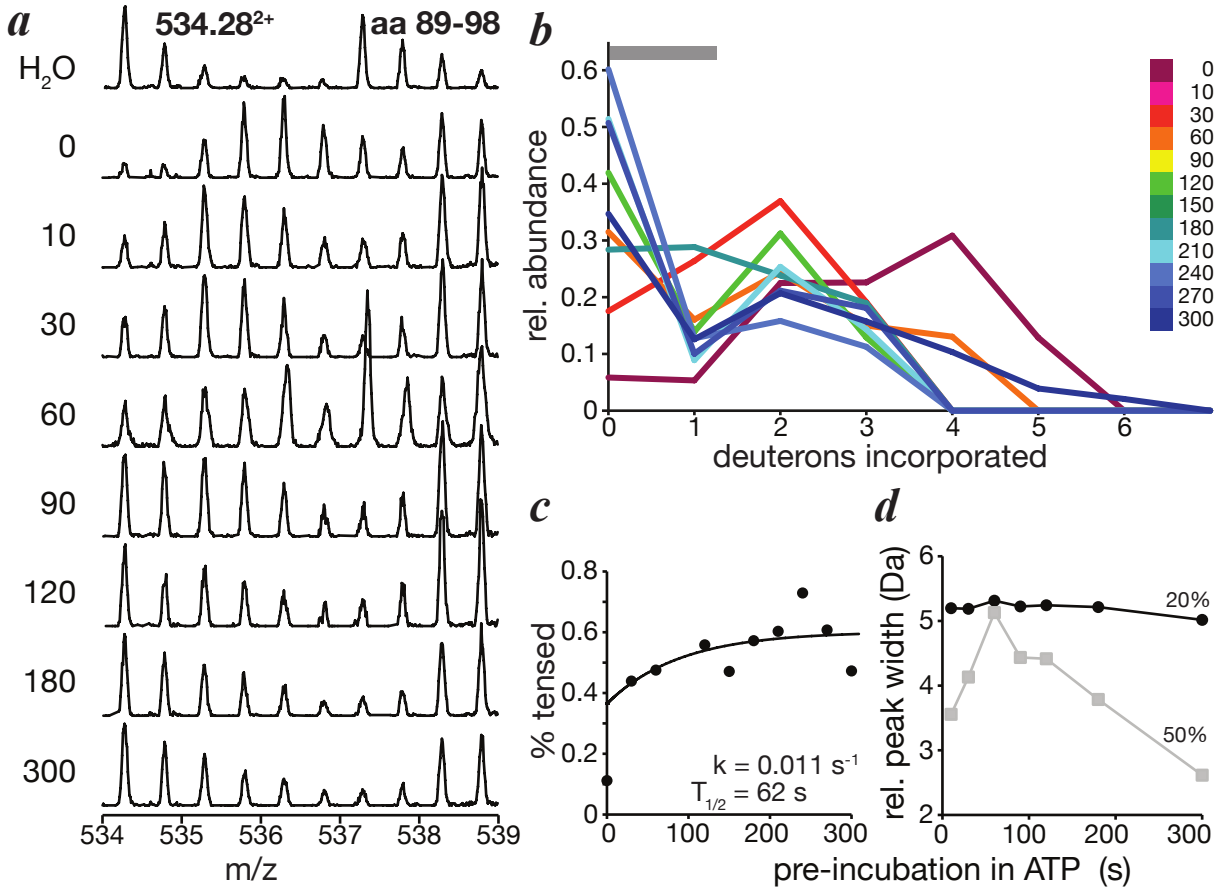
Suppl. Fig. 5: Comparison of HeXicon results and manual analysis of a previously unknown peptide with bimodal isotope distribution. (**a**) Selected mass spectra for the peptide residues 191 to 206 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 965-985 in m/z, unlimited in retention time dimension). (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. Since the deuteration distributions for early and late time points did not overlap the kinetics could not be fitted. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [11]. The width of the distribution at 20% and 50% maximal peak intensity is shown.



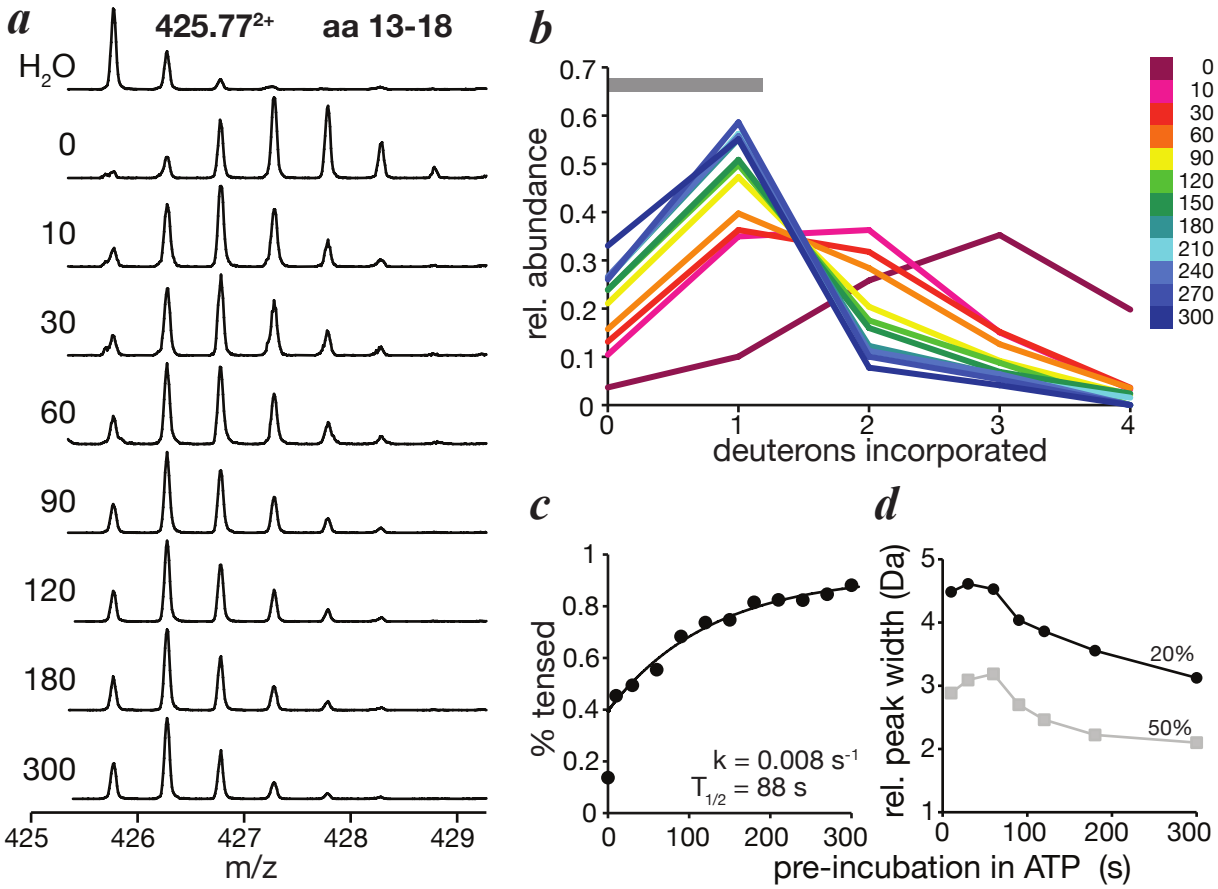
Suppl. Fig. 6: Structural interpretation of the data. Left panel, secondary structure representation of an homology model of *E. coli* HtpG onto the crystal structure of yeast Hsp82 (PDB entry code 2CG9). All peptides exhibiting a bimodal isotope distribution in the pulse-labeling HX experiments are highlighted in blue. The short peptide, residues 13-18, found by HeXicon is shown in light blue. Right panel, zoom into the indicated area of the left panel. One of the two protomers is shown as stick representation of the peptide backbone with hydrogen bonds within the α-helix of region 13-18 indicated by blue dashed lines.



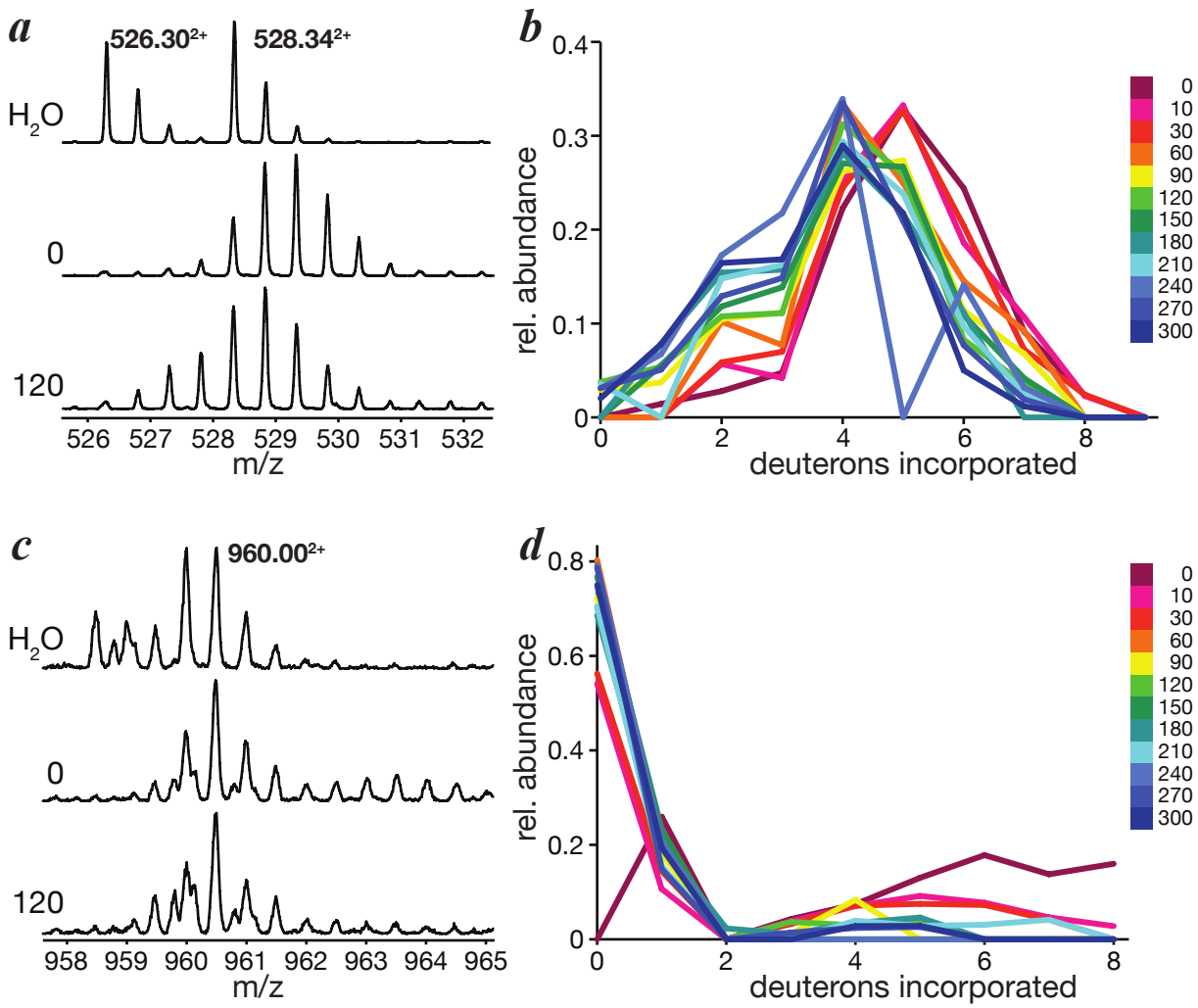
Suppl. Fig. 7: Influence of integration range on the determined transition rate. Upper panel, Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon; values from supplemental figure 2b. Lower panel, analysis of the kinetics shown in the upper panel using different integration ranges integrating the values for 0-5 (bottom) to 0-9 (top) deuterons as indicated by the gray bars in the upper panel. The line represent a single exponential fit to the data and the rate constants are given.



Color version of Fig. 2: Analysis of a result with non-Gaussian deuteration distribution modes. (**a**) Selected mass spectra for the peptide residues 89 to 98 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 520-540 in m/z, unlimited in retention time dimension). (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. The solid line is a single exponential fit to the data. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [9]. The width of the distribution at 20% and 50% maximal peak intensity is shown.



Color version of Fig. 4: Comparison of HeXicon results and manual analysis of a previously unknown peptide with bimodal isotope distribution. (**a**) Selected mass spectra for the peptide residues 13 to18 in H_2O, 10 s D_2O in the absence of ATP (0) and 10 to 300 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times as evaluated by HeXicon (Spectrum limits for HeXicon were set to 420-440 in m/z, unlimited in retention time dimension). (**c**) Analysis of the kinetics shown in (b). The abundance of deuteron incorporations indicated by the gray bar in (b) relative to the total distribution is plotted versus the pre-incubation time. The solid line is a single exponential fit to the data minus the first data point. (**d**) Analysis of the width of the isotope distribution according to a previously published procedure [9]. The width of the distribution at 20% and 50% maximal peak intensity is shown.



Color version of Fig. 5: Analysis of false positive candidates. (**a**) Selected mass spectra for the peptide with mass 526.302+ in H_2O, 10 s D_2O in the absence of ATP (0) and 120 s pre-incubated in the presence of ATP before pulse-labeling in D_2O for 10 s. (**b**) Normalized deuteration distribution for the different pre-incubation times (0 to 300s as indicated) as evaluated by HeXicon (As these peptides were found in the screening of the whole dataset, HeXicon was run without limits in either m/z or retention time). (**c**) and (**d**) same as (a) and (b) for the peptide with mass 960.002+.