

# CalFitter: a web server for analysis of protein thermal denaturation data

Stanislav Mazurenko<sup>1,†</sup>, Jan Stourac<sup>1,2,†</sup>, Antonin Kunka<sup>1,2,†</sup>, Sava Nedeljković<sup>1,3</sup>, David Bednar<sup>1,2</sup>, Zbynek Prokop<sup>1,2,\*</sup> and Jiri Damborsky<sup>1,2,\*</sup>

<sup>1</sup>Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment (RECETOX), Faculty of Science, Masaryk University, Brno, Czech Republic, <sup>2</sup>International Centre for Clinical Research, St. Anne's University Hospital Brno, Brno, Czech Republic and <sup>3</sup>Department of Information Systems, Faculty of Information Technology, Brno University of Technology, Brno, Czech Republic

Received February 07, 2018; Revised April 11, 2018; Editorial Decision April 23, 2018; Accepted April 24, 2018

## ABSTRACT

Despite significant advances in the understanding of protein structure-function relationships, revealing protein folding pathways still poses a challenge due to a limited number of relevant experimental tools. Widely-used experimental techniques, such as calorimetry or spectroscopy, critically depend on a proper data analysis. Currently, there are only separate data analysis tools available for each type of experiment with a limited model selection. To address this problem, we have developed the CalFitter web server to be a unified platform for comprehensive data fitting and analysis of protein thermal denaturation data. The server allows simultaneous global data fitting using any combination of input data types and offers 12 protein unfolding pathway models for selection, including irreversible transitions often missing from other tools. The data fitting produces optimal parameter values, their confidence intervals, and statistical information to define unfolding pathways. The server provides an interactive and easy-to-use interface that allows users to directly analyse input datasets and simulate modelled output based on the model parameters. CalFitter web server is available free at <https://loschmidt.chemi.muni.cz/calfitter/>.

## INTRODUCTION

Proteins are the main building blocks of living organisms and are widely used in numerous biomedical and biotechnological applications. Since they are made up of only 20 amino acids, the enormous variety of their functions mainly stems from their unique structures. The interest in

protein spatial organization is usually twofold: the exact position of active sites and connected residues can shed light on protein function such as enzymatic activity, intracellular transport or molecular signalling (1,2), and exact knowledge of structural elements provides methods of locating the possible sources of protein (in)stability and designing more stable variants using protein engineering (3,4). Moreover, protein misfolding and aggregation have also been reported as primary causes of several neurodegenerative diseases (5).

Streamlined protein denaturation experimental techniques to study protein (un)folding, misfolding, and aggregation include differential scanning calorimetry (DSC), fluorescence/absorbance spectroscopy, light scattering, and circular dichroism (CD) (6–10). They allow the recording of corresponding signals when a protein undergoes denaturation, e.g. due to an increase in temperature. Since those techniques produce an aggregated output from a highly complex underlying process of unfolding, they necessitate the development of software for data modelling and analysis (11,12). Such analysis usually involves the selection of an appropriate unfolding model that best fits the observed data and allows quantification of unfolding pathways in terms of the number of intermediate states, Gibbs free energy barriers separating those states, and corresponding melting temperatures (7,13,14). The importance of such information can hardly be overestimated: intermediates are often the culprits of aggregation, and energy barriers directly define protein half-lives. Hence, both provide attractive targets for protein engineering (15). Moreover, as far as molecular dynamic simulations are concerned, the number of unfolding intermediates can be used as input to cluster analysis in Markov state models, while experimental half-lives can provide guidance for the necessary length of simulations and their conditions (16,17).

Apart from general purpose but programming-intensive tools for data analysis, such as Matlab, Origin or Igor Pro,

\*To whom correspondence should be addressed. Tel: +420 549 4930 41; Fax: +420 549 4962 03; Email: [jiri@chemi.muni.cz](mailto:jiri@chemi.muni.cz), Website address: <https://loschmidt.chemi.muni.cz/calfitter/> (type of web server: data analysis and visualization).

Correspondence may also be addressed to Zbynek Prokop. Email: [zbynek@chemi.muni.cz](mailto:zbynek@chemi.muni.cz)

†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

there are only a few software packages that handle different types of thermal denaturation experiments. These tools were designed to study kinetics or chemical equilibrium — KinTek Explorer, DynaFit — or are linked to a specific type of measurement — MicroCal DSC Origin and CDpal (18–20). Although the former are based on equations that can be adapted for unfolding, data and differential equations with a gradual temperature change are not supported, and the set of parameters offered requires additional manipulations for translation into those usually used to describe unfolding, e.g. Gibbs free energy differences and enthalpy changes. The latter tools offer a limited set of models for fitting, e.g. only reversible denaturation despite increasing cases of proteins unfolding irreversibly. Moreover, such tools are unable to fit global data from different sources, e.g. equilibrium and kinetic data, which may sometimes lead to simplified conclusions.

There are several advantages of global fitting over analyses of separate data sets. Fitting curves to multiple-step models is error-prone because the signal measured can be insensitive to some of the intermediates on the unfolding pathway. Moreover, consecutive steps sometimes overlap significantly and produce apparent single transition which cannot be resolved by fitting into just one data type, and different types of experimental signals must be analyzed simultaneously to overcome this problem (21). Modelling and data fitting of single experimental types individually may also fail to separate parameters enforcing reparametrization of a model for available combinations of constants, e.g. the equilibrium constant  $K_{eq}$  instead of the rate ratios  $k_{fwd}/k_{rev}$ . Simultaneous fitting into a combination of different data types eventually leads to a single set of the original parameter values without any need for reparametrization (22,23). Finally, many independent variables that have to be introduced in separate data fitting contributes to increased uncertainty and can be avoided in the global data fitting.

There is a need for a single, universal platform that handles thermal denaturation data analysis with multiple test models, a user-friendly interface, and the option to join different types of experimental data in one data fitting session. We have developed the web server CalFitter for an interactive data analysis of the commonly used thermal denaturation techniques, such as DSC, CD, and kinetic temperature-jump ( $T$ -jump) experiments. This first-of-a-kind web server offers flexible visualization of the data, quick data pre-treatment for removal of irrelevant and poor-quality data, data simulation, and fitting based on a wide range of fully reversible, irreversible, and partially-reversible unfolding models, as well as statistical data analysis of the goodness of fit. Its data fitting functionality was validated using denaturation data for six wild-type proteins from different structural families, and seven mutant proteins.

## WORKFLOW

The basic workflow of CalFitter is outlined in Figure 1. There are three main phases in the process. First, the user uploads experimental datasets, plots them, and treats the data using built-in data pre-treatment options. Then the user selects a model for the data fitting, supplies initial pa-

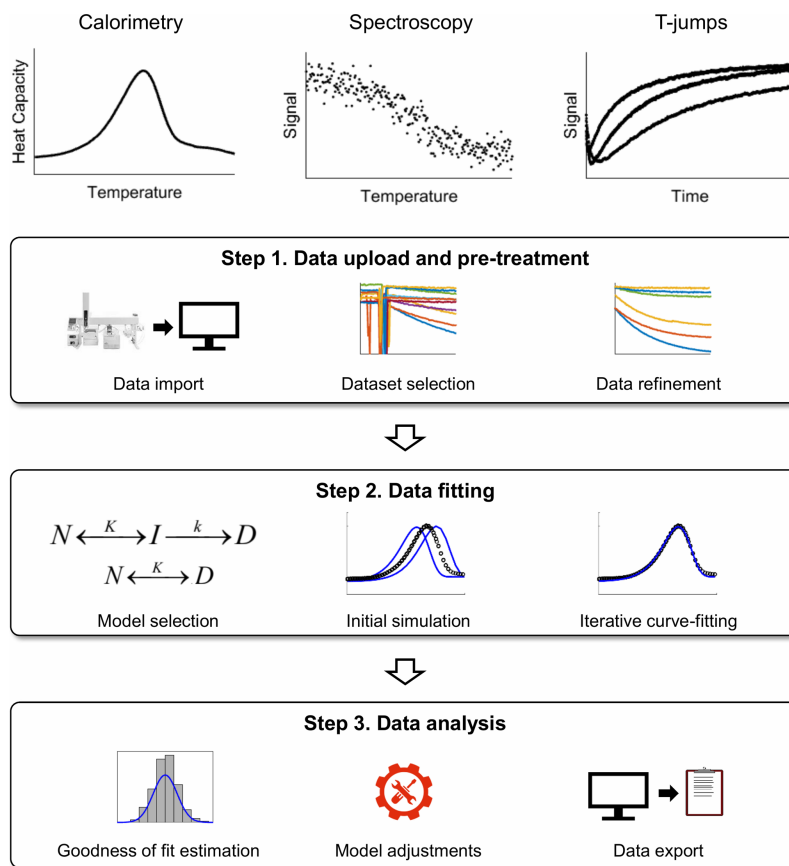
rameter estimates, simulates the modelled dynamics, and starts the fitting process. The server performs numerical data fitting by minimizing the normalized sums of squared residuals. Once the data fitting is complete, the server returns optimal parameter values, confidence intervals calculated from asymptotic normal distribution, and statistical information about the goodness of fit. This information can be further used to estimate the outcome of the fitting and help identify the necessary adjustments to the selected model required for refitting. Different fitting sessions can be conducted in independent tabs after data uploading to compare the outcomes merely by switching the tabs. Once the fitting is complete, the results of the analysis can be exported to a single excel file for further use. Moreover, each session is given a unique ID and thus can be repeatedly accessed at the server.

### Step 1: Data upload and pre-treatment

The input to the server consists of experimental datasets of three types: (i) temperature-dependent heat capacity data from DSC, (ii) temperature-dependent spectroscopic signals (ellipticity, fluorescence, or absorbance) from spectroscopic scanning measurements, and (iii) time-dependent spectroscopic signals from folding/unfolding  $T$ -jumps. In the first step, the user interactively uploads datasets obtained from experiments and specifies the corresponding data columns, units, and experimental setup parameters such as scan rates for DSC and spectroscopy measurements or temperatures for  $T$ -jumps. In order to eliminate systematic errors during global fitting, all datasets must be collected under the same experimental conditions, e.g. pH, ionic strength, and buffer composition. The concentration dependence of protein unfolding must be verified to avoid aggregation or other association/dissociation effects.

The data files can be uploaded in either CalFitter native format (24), plain comma-separated values format (CSV), or several formats exported directly from the build-in software that comes with instruments, e.g. Chirscan or Bio-Kine. The user can select the units from the most widely used ones for temperature ( $^{\circ}\text{C}$  or  $\text{K}$ ), heat capacity difference ( $\text{J}$ ,  $\text{kJ}$ ,  $\text{cal}$  or  $\text{kcal/mol K}$ ), and time ( $\text{ns}$ ,  $\mu\text{s}$ ,  $\text{ms}$ ,  $\text{s}$ ,  $\text{min}$ ,  $\text{h}$ ). There is no upper limit on the number of points in the datasets, although larger datasets take longer to calculate and can bias the fitting statistically. The web server provides a detailed Help page with guidelines about data formatting and uploading.

The user can then plot the uploaded datasets and select those that will be used for data analysis. Any combination of the dataset types can be used for global fitting. Moreover, the user can exclude some parts of the datasets such as temperature ranges with a poor experimental signal. Finally, visual data normalisation can be carried out at this step, which is of great importance in global fits, because collected experimental values usually have different units. In particular, DSC data for each scan rate can be superimposed using vertical shifts, spectroscopy data can be normalized by subtracting signal means and dividing by signal standard deviations for each dataset, and  $T$ -jump traces can be shifted vertically to the same starting point. This has only a visual effect since all the datasets are normalized automat-



**Figure 1.** CalFitter workflow. The software provides an integrated analysis of data using three different types of experimental techniques: (I) calorimetry, (II) spectroscopy and (III) T-jumps. The procedure consists of three steps: (1) data upload and pre-treatment, (2) data fitting and (3) data analysis. A detailed description of the individual steps is provided in the text.

ically during the fitting as described in the section Global fitting of the Supplementary Data.

## Step 2: Data fitting

Once the user is satisfied with the datasets selected and their quality, data fitting can be carried out. The procedure is similar to existing fitting software such as Origin or KinTek Explorer with a special focus on the determination of thermodynamic and kinetic parameters. First, a potential model is selected based on the desired number of steps on the unfolding pathway and their reversibility (Table 1). CalFitter currently offers models that are based on discrete macrostates on the unfolding pathways, i.e. native, intermediate, denatured, etc. Analysis based on the statistical free energy surface models of microstates (25) is beyond the scope of the web server. Second, initial parameter estimates are specified (Figure 1). The server produces initial values, however, the user needs to check and modify those values as there are currently no algorithms providing reliable initial parameter estimation for thermal denaturation models. The web server provides the option to simulate the output datasets based on the input initial parameters to assist the user at this stage. This allows the display of modelled and input data together on one graph. Finally, the user specifies whether some parameters should remain fixed during the fitting with an ad-

justable number of iterations. More details on mathematical modelling and data fitting can be found in the Supplementary Data and the relevant literature (11,12,24,26–28).

## Step 3: Analysis of the results

Once the fitting is complete, the web server updates the parameter values, their confidence intervals and provides statistical information from the fitting such as Akaike (AIC) and Bayesian (BIC) information criteria and residual plots. At this point, the user can either accept the model or change it and carry out refitting. The most common strategy is to start from the simple model with a few steps and then add additional steps while checking the goodness of fit visually or using the AIC and BIC values (29). The common sign of over-parameterization is drifting parameter values and large confidence intervals. In this case, either the model should be simplified by removing some steps, or some fitting parameters should be held constant. Finally, the user can also check the sensitivity of the output to input parameters or undo the last fit before exporting the results in standard formats.

The output consists of the unfolding pathway, the updated parameter values that best describe the data included, their confidence intervals, and other statistical information from the fitting such as goodness of fit. There are four data

**Table 1.** The description of the models and the corresponding parameters implemented in CalFitter

Model	Description	Model parameters <sup>a</sup>	Data sets
<b>1 step</b>			
N → D	A fully irreversible transition	$T_f, E_a, \Delta H^\ddagger, \Delta C_p$	All
N = D	A fully reversible transition with equilibrium	$T_m, \Delta H, \Delta C_p$	Calorimetry & spectroscopy
N = D (Van't Hoff's)	A fully reversible transition with equilibrium and van't Hoff's enthalpy	$T_m, \Delta H, \Delta H_{vh}, \Delta C_p$	Calorimetry & spectroscopy
N < = > D	A general transition with forward and reverse components	$T_{fwd}, E_{fwd}, T_{rev}, E_{rev}, \Delta C_p$	All
<b>2 steps</b>			
N → I → D	A fully irreversible transition	Step 1: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$ Step 2: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	All
N = I → D	A transition with a reversible step in equilibrium and an irreversible step	Step 1: $T_m, \Delta H, \Delta C_p$ Step 2: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	Calorimetry & spectroscopy
N < = > I → D	A general Lumry-Eyring model	Step 1: $T_{fwd}, E_{fwd}, T_{rev}, E_{rev}, \Delta C_p$ Step 2: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	All
N = I = D	A fully reversible transition	Step 1: $T_m, \Delta H, \Delta C_p$ Step 2: $T_m, \Delta H, \Delta C_p$	Calorimetry & spectroscopy
<b>3 steps</b>			
N → I <sub>1</sub> → I <sub>2</sub> → D	A fully irreversible transition	Step 1: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$ Step 2: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$ Step 3: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	All
N = I <sub>1</sub> → I <sub>2</sub> → D	A transition with the reversible first step in equilibrium and the irreversible second and third steps	Step 1: $T_m, \Delta H, \Delta C_p$ Step 2: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$ Step 3: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	Calorimetry & spectroscopy
N < = > I <sub>1</sub> → I <sub>2</sub> → D	A general Lumry-Eyring model with two intermediates	Step 1: $T_{fwd}, E_{fwd}, T_{rev}, E_{rev}, \Delta C_p$ Step 2: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$ Step 3: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	All
N = I <sub>1</sub> = I <sub>2</sub> → D	A transition with two reversible steps in equilibrium and an irreversible step	Step 1: $T_m, \Delta H, \Delta C_p$ Step 2: $T_m, \Delta H, \Delta C_p$ Step 3: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	Calorimetry & spectroscopy
<b>4 steps</b>			
N → I <sub>1</sub> → D	A two-branch irreversible unfolding pathway	Step 1: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	All
N → I <sub>2</sub> → D		Step 2: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	
		Step 3: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	
		Step 4: $T_f, E_a, \Delta H^\ddagger, \Delta C_p$	

<sup>a</sup> $T_m$  – the melting temperature,  $T_f$  – the reference temperature of an irreversible step at which the corresponding rate is 1 (fwd. – forward rates; rev. – reverse rates),  $\Delta H$  – the enthalpy change (at  $T_m$  if  $\Delta C_p$  is nonzero; vh – van't Hoff's);  $\Delta H^\ddagger$  – the activation enthalpy change (at  $T_f$  or  $T_m$  for irreversible and general steps, respectively, if  $\Delta C_p$  is nonzero);  $E_a$  – the activation energy;  $\Delta C_p$  – the heat capacity change. Since  $T$ -jumps are based on the relaxation kinetics, they cannot be simulated by the models with reversible steps assumed in equilibrium.

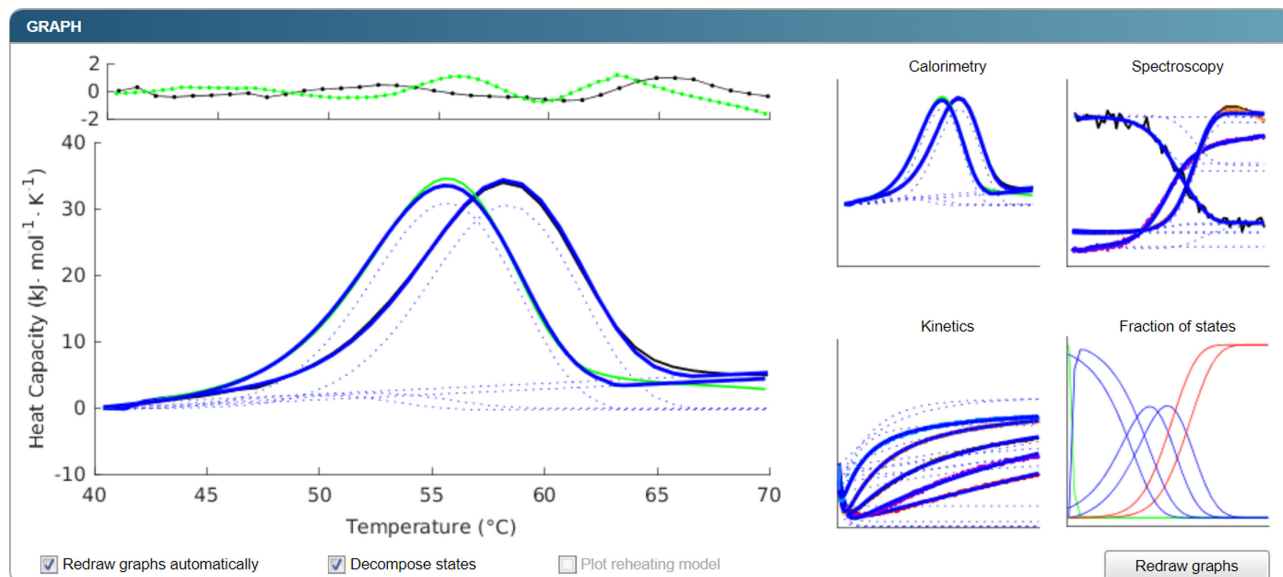
types available for visualising the output (Figure 2). The first type is raw experimental and pre-treated data, each dataset being separate or combined with other datasets of the same type. The second type uses modelled signals matching the selected datasets given the current parameter values. A state decomposition may be carried out with this type whereby signal contributions from each step of the unfolding pathway to the overall modelled signal are shown. The third type uses the modelled protein state fractions as functions of temperature or time. Finally, the fourth type uses residuals from the fit, which might shed further light on the quality of the fit and main discrepancies between the modelled and experimental data. The user can then export the graphs and the corresponding datasets as an archive with figures and settings files or as an Excel file.

## EXPERIMENTAL VALIDATION

The CalFitter performance was thoroughly validated with all the three data types. DSC data analysis was tested on previously published datasets: the thermograms of wild-type hen egg lysozyme, wild-type haloalkane dehalogenases LinB, DbjA, DhaA, fibroblast growth factor FGF2 and variants of DhaA and FGF2 engineered for higher ther-

mostability (24,30). Calorimetry curves were curve-fitted and compared with the output from MicroCal DSC Origin and the standalone Matlab-based CalFitter 1 (24). Spectroscopy data analysis was tested on four variants of DhaA and compared with the output from CD-pal. Finally,  $T$ -jump data analysis was validated using four different global datasets consisting of several traces of DhaA wild-type and compared against the values obtained using KinTek Explorer. In all the cases, the output values produced by the CalFitter web differed from the previously published by <0.1% on average for temperature related parameters, and by <6.8% on average for energy and heat-related parameters (Table 2). Maximal discrepancies for the latter were mainly due to smaller parameter values and/or wide confidence intervals. Moreover, in those cases, the CalFitter web simulation with the parameters from the other tools produced visually the same quality of the fit suggesting that the differences stem from the numerical procedures used for fitting rather than from a different model behaviour.

The global fitting provided by the web server was also used to analyse experimental data for stable variants of FGF2 protein designed recently using computer-assisted protein engineering (30). This analysis revealed new biophysical insights, namely the presence of an unfolding inter-



**Figure 2.** An example of the graphical output. Initial datasets with modelled signal in blue are depicted as icons on the right-hand side and can be zoomed-in and displayed on the left-hand side. The zoomed-in version also depicts the residuals at the top so that a user can estimate the quality of the fit and the presence of any systematic errors or unexplained data variation. In the presented case, waves are apparent in the residual plot that are indicative of an approximately 5% misfit at high temperatures. When the option 'Decompose states' is selected, the contribution of each step to the overall signal is plotted in dotted lines. Apart from the plots of data and corresponding modelled signals, modelled fractions of states are presented in one of the graphs.

**Table 2.** Experimental validation of CalFitter web server. Discrepancies are given in terms of absolute % difference for parameters obtained for energies and temperatures

Data type	Software used for comparison	Number of datasets	Temperature variables ( $T_m$ , $T_f$ )		Energy variables ( $E_a$ , $\Delta H$ , $\Delta H_{vh}$ )	
			Average discrepancy	Maximal discrepancy	Average discrepancy	Maximal discrepancy
DSC <sup>a</sup>	MicroCal DSC	67	0.06%	0.42%	3.03%	15.06%
	Origin					
CD <sup>b</sup>	Matlab-based	44	0.00%	0.01%	0.00%	0.07%
	CalFitter 1					
T-jumps <sup>c</sup>	KinTek Explorer	35	0.09%	0.21%	6.74%	10.87%

<sup>a</sup>based on  $\Delta H$ ,  $\Delta H_{vh}$ , and  $T_m$  from a non-two state model with  $\Delta H_{vh}$ .

<sup>b</sup>based on  $T_m$  and  $\Delta H$  for a one-step fully reversible model.

<sup>c</sup>data from global fitting based on  $E_a$  and  $T_a$  for a two-step fully irreversible model.

mediate, and demonstrated a good agreement between the *in silico* predicted Gibbs free energy differences and the differences in the transition barriers for the first unfolding step estimated from experiments. Another case study of thermal denaturation of haloalkane dehalogenase DhaA112 engineered for stability is described in the Supplementary Data. This new case reveals an unfolding intermediate and provides quantitative estimates of unfolding rates.

## CONCLUSIONS AND OUTLOOK

CalFitter is a web server that offers users a one-stop-shop for data analysis from commonly used temperature denaturation experiments. Not only does it offer a wider range of models for each separate data type when compared to most of the existing analogues, but it also enables the combination of different dataset types, such as equilibrium and T-jump data, in a single global data analysis. This feature has never been implemented for thermal unfolding stud-

ies before, to the best of the authors' knowledge. The fitting procedures used were optimised and validated using several dozen datasets from different sources, including recently published data as well as cross-validation using the existing software for each data type analysis.

The server is complemented by an easy-to-use graphical interface that allows users to interactively pre-treat the data by excluding irrelevant parts or artefacts, selecting the desired subset for analysis and fitting, and simulating the behaviour of the models when parameters change. The hidden mathematical calculations and fitting makes the process of data analysis accessible to users without any prior expertise of mathematical modelling. The web server graphical output consists of four different plot types to provide the user with a full image of the modelled pathway and its correspondence to the supplied data.

In the future, we will implement an 'advanced mode' with a model editor that the users can manually input any model of their choice using a simple text entry with an in-

tuitive syntax similar to KinTek Explorer or DynaFit. We are also working on a module for rapid initial parameter estimations. This module will estimate starting values based on the shape of the curves, rendering the web server even more user-friendly for researchers with limited experience in modelling data from thermal unfolding studies. Finally, we plan to add additional modelling capabilities to the existing modules, e.g., double T-jumps and singular value decomposition of CD spectra, as well as to develop modules for analysis of other types of experimental data, e.g., protein chemical denaturation and hydrogen-deuterium exchange mass spectrometry.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## FUNDING

Ministry of Education of the Czech Republic [LO1214, LQ1605, LM2015051, LM2015047 and LM2015055]; Czech Grant Agency [GA16-07965S]; European Union [720776 and 722610]; A.K. is a Brno Ph.D. Talent Scholarship Holder and funded by the Brno City Municipality. Funding for open access charge: Ministry of Education of the Czech Republic.

*Conflict of interest statement.* None declared.

## REFERENCES

- Dill, K.A. and MacCallum, J.L. (2012) The protein-folding problem, 50 years on. *Science*, **338**, 1042–1046.
- Orengo, C.A., Pearl, F.M.G., Bray, J.E., Todd, A.E., Martin, A., Lo Conte, L. and Thornton, J.M. (1999) The CATH database provides insights into protein structure/function relationships. *Nucleic Acids Res.*, **27**, 275–279.
- Fersht, A.R., Matouschek, A. and Serrano, L. (1992) The folding of an enzyme: I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.*, **224**, 771–782.
- Yang, H., Liu, L., Li, J., Chen, J. and Du, G. (2015) Rational design to improve protein thermostability: recent advances and prospects. *ChemBioEng Rev.*, **2**, 87–94.
- Knowles, T.P., Vendruscolo, M. and Dobson, C.M. (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.*, **15**, 384–396.
- Sancho, J. (2013) The stability of 2-state, 3-state and more-state proteins from simple spectroscopic techniques... plus the structure of the equilibrium intermediates at the same time. *Arch. Biochem. Biophys.*, **531**, 4–13.
- Temel, D.B., Landsman, P. and Brader, M.L. (2016) Orthogonal methods for characterizing the unfolding of therapeutic monoclonal antibodies: differential scanning calorimetry, isothermal chemical denaturation, and intrinsic fluorescence with concomitant static light scattering. *Methods Enzymol.*, **567**, 359–389.
- Gelman, H. and Gruebele, M. (2014) Fast protein folding kinetics. *Q. Rev. Biophys.*, **47**, 95–142.
- Goyal, M., Chaudhuri, T.K. and Kuwajima, K. (2014) Irreversible denaturation of maltodextrin glucosidase studied by differential scanning calorimetry, circular dichroism, and turbidity measurements. *PLoS One*, **9**, e115877.
- Dimitriadis, G., Drysdale, A., Myers, J.K., Arora, P., Radford, S.E., Oas, T.G. and Smith, D.A. (2004) Microsecond folding dynamics of the F13W G29A mutant of the B domain of staphylococcal protein A by laser-induced temperature jump. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 3809–3814.
- Lepock, J.R., Ritchie, K.P., Kolios, M.C., Rodahl, A.M., Heinz, K.A. and Kruuv, J. (1992) Influence of transition rates and scan rate on kinetic simulations of differential scanning calorimetry profiles of reversible and irreversible protein denaturation. *Biochemistry*, **31**, 12706–12712.
- Sanchez-Ruiz, J.M. (1992) Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophys. J.*, **61**, 921–935.
- Privalov, P.L. and Dragan, A.I. (2007) Microcalorimetry of biological macromolecules. *Biophys. Chem.*, **126**, 16–24.
- Tsytlonok, M. and Itzhaki, L.S. (2013) The how's and why's of protein folding intermediates. *Arch. Biochem. Biophys.*, **531**, 14–23.
- Neudecker, P., Robustelli, P., Cavalli, A., Walsh, P., Lundstrom, P., Zarrine-Afsar, A., Sharpe, S., Vendruscolo, M. and Kay, L.E. (2012) Structure of an intermediate state in protein folding and aggregation. *Science*, **336**, 362–366.
- Bowman, G.R., Beauchamp, K.A., Boxer, G. and Pande, V.S. (2009) Progress and challenges in the automated construction of Markov state models for full protein systems. *J. Chem. Phys.*, **131**, 124101.
- Wei, G., Xi, W., Nussinov, R. and Ma, B. (2016) Protein ensembles: how does nature harness thermodynamic fluctuations for life? the diverse functional roles of conformational ensembles in the cell. *Chem. Rev.*, **116**, 6516–6551.
- Johnson, K.A. (2009) Fitting enzyme kinetic data with KinTek global kinetic explorer. *Methods Enzymol.*, **467**, 601–626.
- Kuzmič, P. (2009) DynaFit—a software package for enzymology. *Methods Enzymol.*, **467**, 247–280.
- Niklasson, M., Andresen, C., Helander, S., Roth, M.G., Zimdahl Kahlin, A., Lindqvist Appell, M., Mårtensson, L. and Lundström, P. (2015) Robust and convenient analysis of protein thermal and chemical stability. *Prot. Sci.*, **24**, 2055–2062.
- Harder, M.E., Deinzer, M.L., Leid, M.E. and Schimerlik, M.I. (2004) Global analysis of threestate protein unfolding data. *Prot. Sci.*, **13**, 2207–2222.
- Li, A., Ziehr, J.L. and Johnson, K.A. (2017) A new general method for simultaneous fitting of temperature and concentration dependence of reaction rates yields kinetic and thermodynamic parameters for HIV reverse transcriptase specificity. *J. Biol. Chem.*, **292**, 6695–6702.
- Yi, Q., Scalley, M.L., Simons, K.T., Gladwin, S.T. and Baker, D. (1997) Characterization of the free energy spectrum of peptostreptococcal protein L. *Fold. Des.*, **2**, 271–280.
- Mazurenko, S., Kunka, A., Beerens, K., Johnson, C.M., Damborsky, J. and Prokop, Z. (2017) Exploration of protein unfolding by modelling calorimetry data from reheating. *Sci. Rep.*, **7**, 16321.
- Ibarra-Molero, B., Naganathan, A.N., Sanchez-Ruiz, J.M. and Muñoz, V. (2016) Modern analysis of protein folding by differential scanning calorimetry. *Methods Enzymol.*, **567**, 281–318.
- Rodríguez-Larrea, D., Ibarra-Molero, B., de Maria, L., Borchert, T.V. and Sanchez-Ruiz, J.M. (2008) Beyond Lumry-Eyring: an unexpected pattern of operational reversibility/irreversibility in protein denaturation. *Prot. Struct. Funct. Bioinf.*, **70**, 19–24.
- Lyubarev, A.E. and Kurganov, B.I. (2000) Analysis of DSC data relating to proteins undergoing irreversible thermal denaturation. *J. Therm. Anal. Cal.*, **62**, 51–62.
- Milardi, D., La Rosa, C. and Grasso, D. (1994) Extended theoretical analysis of irreversible protein thermal unfolding. *Biophys. Chem.*, **52**, 183–189.
- Kirk, P., Thorne, T. and Stumpf, M.P. (2013) Model selection in systems and synthetic biology. *Curr. Opin. Biotechnol.*, **24**, 767–774.
- Dvorak, P., Bednar, D., Vanacek, P., Balek, L., Eiselleova, L., Stepankova, V., Sebestova, E., Kunova Bosakova, M., Konecna, Z., Mazurenko, S. et al. (2018) Computer-assisted engineering of hyperstable fibroblast growth factor 2. *Biotechnol. Bioeng.*, **115**, 850–862.