

ProThermDB: thermodynamic database for proteins and mutants revisited after 15 years

Rahul Nikam[†], A. Kulandaisamy[†], K. Harini, Divya Sharma and M. Michael Gromiha^{ID*}

Department of Biotechnology, Bhupat and Jyoti Mehta School of BioSciences, Indian Institute of Technology Madras, Chennai 600 036, Tamilnadu, India

Received September 14, 2020; Revised October 14, 2020; Editorial Decision October 16, 2020; Accepted October 26, 2020

ABSTRACT

ProThermDB is an updated version of the thermodynamic database for proteins and mutants (ProTherm), which has ~31 500 data on protein stability, an increase of 84% from the previous version. It contains several thermodynamic parameters such as melting temperature, free energy obtained with thermal and denaturant denaturation, enthalpy change and heat capacity change along with experimental methods and conditions, sequence, structure and literature information. Besides, the current version of the database includes about 120 000 thermodynamic data obtained for different organisms and cell lines, which are determined by recent high throughput proteomics techniques using whole-cell approaches. In addition, we provided a graphical interface for visualization of mutations at sequence and structure levels. ProThermDB is cross-linked with other relevant databases, PDB, UniProt, PubMed etc. It is freely available at <https://web.iitm.ac.in/bioinfo2/prothermdbContext.html> without any login requirements. It is implemented in Python, HTML and JavaScript, and supports the latest versions of major browsers, such as Firefox, Chrome and Safari.

INTRODUCTION

Thermodynamic data for proteins are important for understanding the mechanism of protein folding, delineating the factors influencing the stability of proteins and mutants, development of computational tools, relating mutational effects on structure, stability, function and diseases as well as to design new stable mutants for different applications (1–5). With the advancements in high throughput biophysical techniques, the availability of thermodynamic data of proteins is increasing rapidly. The efficient compilation of these thermodynamic data with sequence and structure information could serve as a valuable resource for protein researchers.

We have developed the thermodynamic database for proteins and mutants, ProTherm, and updated continuously till 2006 (6–10). It has been effectively used to relate physico-chemical properties with protein stability (11–13), development of computational tools for predicting the change in melting temperature (14–17) and free energy change upon mutation (18–28) and understanding the role of protein stability in disease causing mutations (29–31). Recently, Kulandaisamy *et al.* (2020) developed a thermodynamic database, MPTherm, specifically for membrane proteins (32).

In this work, we have updated and reconstructed the database, ProThermDB, which contains the experimental thermodynamic parameters for proteins and mutants along with sequence and structure information, experimental methods and conditions, literature information and visualization of proteins and mutants. We have included >14 500 new data (84% increase) and among them 2200 are published in the last three years. In addition, we have included organism-based data and cell line data in the current version of the database. Furthermore, we have provided a user-friendly web interface to search the database and download the required data. It is cross-linked with UniProt (33) and Protein Data Bank (PDB) (34) as well as to the PubMed so that users can get the source of each entry directly. Besides, we have included a separate page for individual entry, where users can get protein information, experimental conditions, literature information, and visualization of three-dimensional structure of the protein with mutation information. We have also rectified the known errors reported in the literature (35). The ProThermDB is freely available at <https://web.iitm.ac.in/bioinfo2/prothermdbContext.html>.

CONTENTS OF THE DATABASE

ProThermDB provides two types of thermodynamic data for wild-type and mutants proteins based on experimental methods: (i) experimental thermodynamic data obtained with purified proteins using traditional methods such as circular dichroism (CD), differential scanning calorimetry (DSC) and fluorescence spectroscopy, and (ii) data pro-

*To whom correspondence should be addressed. Tel: +91 44 2257 4138; Fax: +91 44 2257 4102; Email: gromiha@iitm.ac.in

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

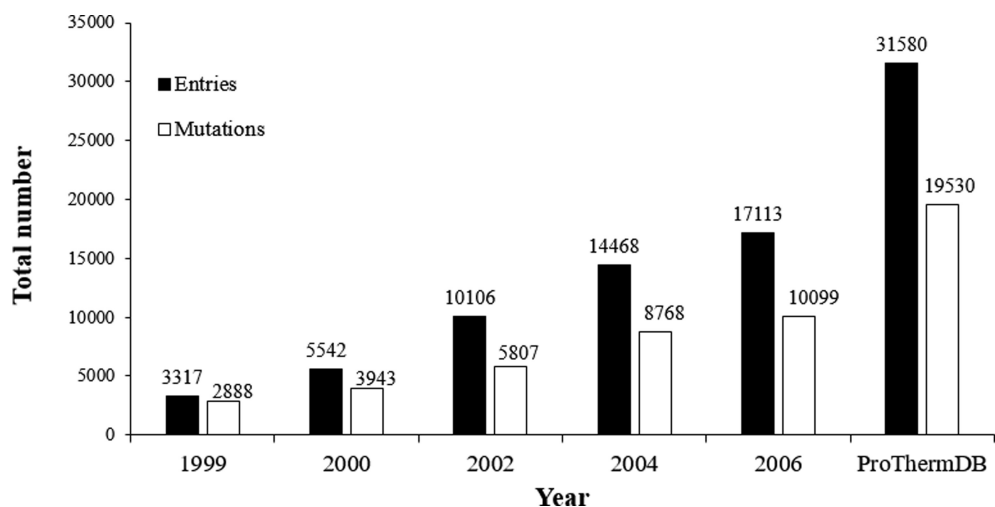


Figure 1. Growth of the ProThermDB.

Table 1. Comparison of data in ProThermDB with ProTherm 2006

| Data type | ProTherm 2006 | ProThermDB | % increase in ProThermDB |
|---|---------------|---------------|--------------------------|
| Total | 17113 | 31580 | 84.54 |
| Data type or Type of mutation | | | |
| Wild-type | 7077 (41.35) | 12050 (38.16) | 70.27 |
| Single | 7969 (46.57) | 16028 (50.75) | 101.13 |
| Double | 1139 (6.66) | 2046 (6.48) | 79.63 |
| Multiple (≥3) | 928 (5.42) | 1456 (4.61) | 56.90 |
| Mutations in secondary structures ^a | | | |
| Helix | 2470 (34.16) | 5540 (37.72) | 124.29 |
| Sheet | 1410 (19.50) | 3095 (21.07) | 119.50 |
| Turn | 1100 (15.21) | 2154 (14.67) | 95.82 |
| Coil | 2251 (31.13) | 3898 (26.54) | 73.17 |
| Mutations based on solvent accessibility ^a | | | |
| Buried | 2969 (41.06) | 6313 (42.98) | 112.63 |
| Partially buried | 2172 (30.04) | 4254 (28.96) | 95.86 |
| Exposed | 2090 (28.90) | 4120 (28.05) | 97.13 |

Percentage of data for each category is shown in parenthesis.
^aSecondary structure and solvent accessibility are computed using DSSP for proteins with known three-dimensional structures (40).

duced by mass spectrometry (MS)-based high-throughput techniques such as thermal proteome profiling (TPP) (36) and limited proteolysis (LiP) (37) using a whole-cell approach without protein purification. Further, data from high-throughput techniques are grouped together based on the source organism and cell lines.

Each entry in ProThermDB is identified with a unique accession number and it includes the following information:

- a) **Protein information:** protein name, source, UniProt ID, length of the protein sequence, enzyme commission number, molecular weight, mutation based on UniProt and/or PDB, and type of mutation such as single, double and multiple.
- b) **Structural information:** PDB codes for both wild-type and mutant structures (if available), number of chains, chain name, mutation details (wild-type and mutant residues along with mutant position), solvent accessibility and secondary structure of wild-type residue and 3D

- visualization of the mutation using JSmol interface (38). We utilized the SIFTS database for mapping residue positions between UniProt and PDB (39). Further, we observed that 95% of data have the same protein constructs used in thermodynamic experiments and reported in PDB.
- c) **Experimental conditions:** temperature (T), pH, buffer name, additives, measurement, method and remarks.
 - d) **Literature information:** PubMed identifier, name of the author(s), journal name, year of publication, location of data and keywords.
 - e) **Experimental thermodynamic data:** free energy of unfolding in water (ΔG^{H_2O}), free energy obtained with thermal denaturation (ΔG), melting temperature (T_m), the slope (m) and midpoint of denaturation (C_m). In order to account the effect of mutations on protein stability, we computed the change in free energy of unfolding in water ($\Delta\Delta G^{H_2O}$), change in free energy upon thermal denaturation ($\Delta\Delta G$) and change in melting temperature (ΔT_m). The positive and negative values of $\Delta\Delta G^{H_2O}$, $\Delta\Delta G$ and ΔT_m indicate stabilizing and destabilizing mutations, respectively.

RECENT DEVELOPMENTS

In the latest version of the database, we have developed a new user-friendly web interface with several search and display options. It contains 31 580 entries, including 84% of new data. For each entry, a separate page is added to provide the information on protein sequence and structure details, experimental conditions, thermodynamic parameters, literature information, and visualization of three-dimensional structures. We have unified protein details based on UniProt and PDB databases and corrected all the known errors (35) in 10% of the data such as (i) mutation position in a sequence based on UniProt and PDB databases, (ii) unified the units of free energy values (in kcal/mol), temperatures (°C) and other thermodynamic parameters, (iii) rectified errors in stability data, (iv) protein names, (v) UniProt ID and (vi) references. We have also provided an option for ‘data upload’ to receive the PubMed ID or DOI (Digital

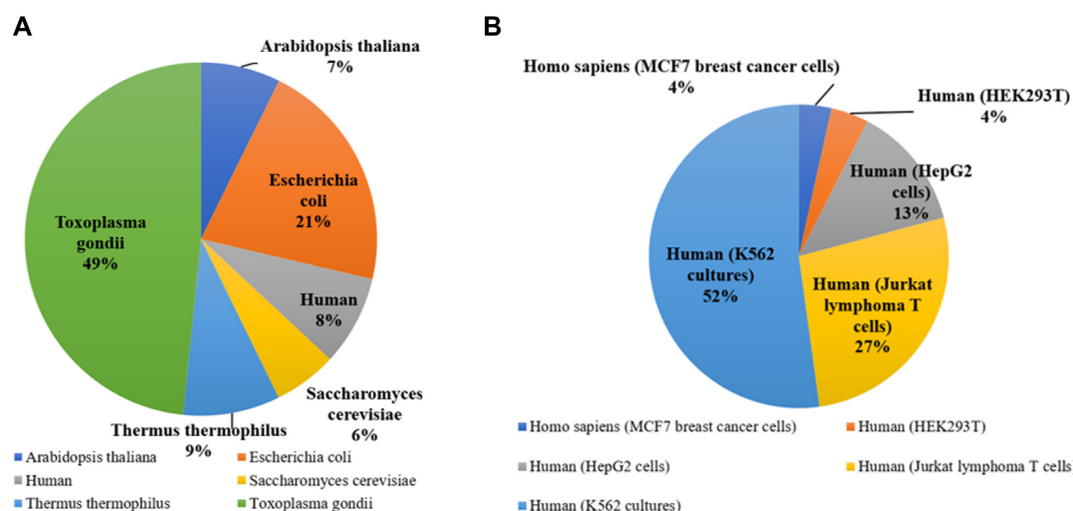


Figure 2. Distribution of thermodynamic data obtained with whole-cell approaches based on (A) organisms and (B) cell lines.

Object Identifier) from the users and include the data in the database. In addition, we included the thermodynamic data obtained from MS-based high-throughput techniques based on source organisms and cell lines.

In the 'statistics page', we furnished a brief report on the contents of the ProThermDB database (<https://web.iitm.ac.in/bioinfo2/prothermdb/Statistics.html>) using pie charts. The figures illustrate the percentage for each type of mutation (wild-type, single, double and multiple mutations), location of single mutations with respect to solvent accessibility and secondary structure and average stability of stabilizing and destabilizing point mutations. Further, we provide a 20×20 amino acid substitution matrix for the frequency of single mutations, and 20×20 substitution matrices for average stabilizing and destabilizing values for each type of single mutation with respect to $\Delta\Delta G^{\text{H}_2\text{O}}$, $\Delta\Delta G$ and ΔT_m .

DATABASE STATISTICS

Figure 1 illustrates the overall growth of ProThermDB, and the recent version includes 84% of new data. We also showed the contents of the database based on mutation type, solvent accessibility, and secondary structure in Table 1. The latest version contains 38%, 51% and 11% of wild type, single and multiple mutations, respectively, which is similar (<4%) to the previous version (Table 1). On the other hand, ProThermDB has an increase of 70%, 101% and 69% data in wild type, single and multiple mutations, respectively. According to solvent accessibility, 43%, 29% and 28% of single mutations are located in buried, partially buried and exposed regions of protein structures, respectively (Table 1). We observed an increase of 113%, 96% and 97% on the number of data in these locations compared to the previous version. Based on secondary structure 38% of single mutations adopt α -helical conformation followed by 27% in coils and 21% in β -strands (Table 1). Although the percentage of mutations is similar in previous and latest versions, the recent version has an increase of 124%, 73% and 119% of data in α -helices, coils and β -strands, respectively.

Further, the current version of ProThermDB contains 12 150 and 1 05 444 thermodynamic data from six organisms and five human cell lines, respectively. Figure 2 shows the distribution of thermodynamic data depends on source organisms and different cell lines that are derived from high throughput biophysical techniques. The *Toxoplasma gondii* has 5878 (49%) thermodynamic data followed by 2591 (21%) in *Escherichia coli*. K562 cell line of human has data for the maximum of 52% (Figure 2).

We have compared the number and percentage of data in ProThermDB with MPTherm database based on the types of mutations and the results are presented in Supplementary Table S1. ProThermDB is dominated with single mutations followed by wild-type data whereas an opposite trend was observed in MPTherm. The percentage of double mutations is similar in both the databases. On the other hand, number of multiple mutations is less than double mutations in ProThermDB whereas MPTherm contains more number of multiple mutations than double mutations. Considering all the data together, number of data in ProThermDB is 4.4 times more than MPThermDB. Consequently, data for wild type, single, double and multiple mutations in ProThermDB are 3.2, 6.9, 5.7 and 2.1 times more than MPTherm (32).

Data retrieval from ProThermDB

The detailed information about the search and display options and an example for data retrieval from ProThermDB is illustrated in Figure 3. In this example, we build a query using a combination of multiple search options such as the protein name with 'carbonic anhydrase 2', secondary structure 'helix' (Figure 3A). Also, we selected the desired columns in the display options (Figure 3B). After submitting the query, the results are displayed in a table format (Figure 3C). We also provided an option to download the search results. In the result page, each entry accession number has a hyperlink for their respective external page, which contains the complete information (Supplementary

A SEARCH OPTIONS

Entry: 27543

UniProt: P00918

PDB Code: 1ZCA

Protein: Carbonic anhydrase 2 **Query 1**

Source: Homo sapiens (Human)

Mutation type: Any **Query 2**

Sec Str: ☒ Helix ☐ Sheet ☐ Turn ☐ Coil

Accessibility: ☐ Any ☐ Buried ☐ Partially Buried ☐ Exposed

ASA: 0 To 254

pH (0-13): 0 To 13

T: -16 To 134

Measure: ☐ Absorbance ☐ CD ☐ DSC ☐ Fluorescence ☐ Others

Method: ☐ Thermal ☐ GdnHCl ☐ Urea ☐ Others

Tm: -52 To 220

ΔT_m : -286 To 72

$\Delta \Delta G$: -23 To 43

$\Delta \Delta G^{H2O}$: -12 To 23

ΔH : -75 To 1905

ΔC_p : -2 To 86

ΔG^{H2O} : -20 To 85

State: ☐ 2 ☐ >3

Reversibility: Select

PubMed ID: 31778702

Keyword: Carbonic anhydrase defi

B DISPLAY OPTIONS

Protein information

☒ Entry ☒ Protein ☒ UniProt ☒ PDB

☒ Source ☒ Mutation ☒ Sec Str ☒ ASA

☐ EC Number

Experimental conditions

☒ pH ☒ T ☐ Measure ☐ Method

☐ Buffer_Name ☐ Buffer_conc ☐ Ion_Name ☐ Ion_conc

Thermodynamic parameters

☐ T_m ☐ ΔT_m ☐ ΔH ☐ ΔC_p

☐ ΔH_{VH} ☐ ΔG ☒ $\Delta \Delta G$ ☐ Δm

☐ C_m ☐ ΔG^{H2O} ☒ $\Delta \Delta G^{H2O}$ ☒ State ☒ Reversibility

Literature

☒ PubMed Id ☒ Key Words ☒ Reference ☒ Author

☐ Remarks

Select All

☐ ALL ☐ Reset

C Result

| NO | PROTEIN | UniProt_ID | PDB_wild | SOURCE | MUTATION | SEC_STR | ASA | pH | T (C) | $\Delta \Delta G$ (kcal/mol) | $\Delta \Delta G^{H2O}$ (kcal/mol) | STATE | REVERS |
|-------|----------------------|------------|----------|----------------------|--------------------------|---------|------|-----|-------|------------------------------|------------------------------------|-------|--------|
| 15059 | Carbonic anhydrase 2 | P00918 | 1ZCA | Homo sapiens (Human) | H94Y (Based on UniProt) | Helix | 28.0 | 7.5 | - | - | - | - | no |
| 15060 | Carbonic anhydrase 2 | P00918 | 1ZCA | Homo sapiens (Human) | G144R (Based on UniProt) | Helix | 0.0 | 7.5 | - | - | - | - | no |
| 15062 | Carbonic anhydrase 2 | P00918 | 1ZCA | Homo sapiens (Human) | H34Y (Based on UniProt) | Helix | 28.0 | 7.5 | - | - | - | - | no |
| 15063 | Carbonic anhydrase 2 | P00918 | 1ZCA | Homo sapiens (Human) | G144R (Based on UniProt) | Helix | 0.0 | 7.5 | - | - | - | - | no |
| 27542 | Carbonic anhydrase 2 | P00918 | 1ZCA | Homo sapiens (Human) | H94Y (Based on UniProt) | Helix | 28.0 | 7.5 | 25 | - | -1.17 | 3 | yes |
| 27543 | Carbonic anhydrase 2 | P00918 | 1ZCA | Homo sapiens (Human) | G144R (Based on UniProt) | Helix | 0.0 | 7.5 | 25 | - | -0.26 | 3 | yes |
| 27545 | Carbonic anhydrase 2 | P00918 | 1ZCA | Homo sapiens (Human) | H94Y (Based on UniProt) | Helix | 28.0 | 7.5 | 25 | - | 0.02 | 3 | yes |
| 27546 | Carbonic anhydrase 2 | P00918 | 1ZCA | Homo sapiens (Human) | G144R (Based on UniProt) | Helix | 0.0 | 7.5 | 25 | - | -0.88 | 3 | yes |

Download results

Download Now! For each thermodynamic parameter, the standard deviation values are given in parentheses.

Figure 3. An example of data retrieval from the ProThermDB database using different search and display options.

Table S2). The structural visualization is available in the external page of each entry.

DATA AVAILABILITY

The ProThermDB database is developed using Python, HTML and JavaScript programming languages, and it supports the latest version of major browsers such as Firefox, Chrome, and Safari. The web interface is available at <https://web.iitm.ac.in/bioinfo2/prothermdb/index.html>. The database will be maintained and updated regularly. The updated information will be reflected on the homepage of the database. Any constructive comments and suggestions are welcome and should be sent to gromiha@iitm.ac.in or pbl.prothermdb2020@gmail.com.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We acknowledge the reviewers for their constructive comments. We thank the Indian Institute of Technology Madras and the High-Performance Computing Environment (HPCE) for computational facilities. Also, we thank

all the members from Protein Bioinformatics Lab for discussions and valuable suggestions.

FUNDING

Indian Institute of Technology Madras.

Conflict of interest statement. None declared.

REFERENCES

- Pfeil, W. (2001) In: *Protein Stability and Folding, Supplement 1: A Collection of Thermodynamic Data*. Springer, NY.
- Lorch, M., Mason, J.M., Clarke, A.R. and Parker, M.J. (1999) Effects of core mutations on the folding of a β -sheet protein: implications for backbone organization in the I-state. *Biochemistry*, **38**, 1377–1385.
- Lorch, M., Mason, J.M., Sessions, R.B. and Clarke, A.R. (2000) Effects of mutations on the thermodynamics of a protein folding reaction: implications for the mechanism of formation of the intermediate and transition states. *Biochemistry*, **39**, 3480–3485.
- Yamada, Y., Banno, Y., Yoshida, H., Kikuchi, R., Akao, Y., Murate, T. and Nozawa, Y. (2006) Catalytic inactivation of human phospholipase D2 by a naturally occurring Gly901Asp mutation. *Arch. Med. Res.*, **37**, 696–699.
- Singh, S.M., Bandi, S., Shah, D.D., Armstrong, G. and Mallela, K.M. (2014) Missense mutation Lys18Asn in dystrophin that triggers X-linked dilated cardiomyopathy decreases protein stability, increases protein unfolding, and perturbs protein structure, but does not affect protein function. *PLoS One*, **9**, e110439.

6. Gromiha, M.M., An, J., Kono, H., Oobatake, M., Uedaira, H. and Sarai, A. (1999) ProTherm: thermo-dynamic database for proteins and mutants. *Nucleic Acids Res.*, **27**, 286–288.
7. Gromiha, M.M., Uedaira, H., An, J., Selvaraj, S., Prabakaran, P. and Sarai, A. (2002) ProTherm, thermodynamic database for proteins and mutants: developments in version 3.0. *Nucleic Acids Res.*, **30**, 301–302.
8. Sarai, A., Gromiha, M.M., An, J., Prabakaran, P., Selvaraj, S., Kono, H., Oobatake, M. and Uedaira, H. (2002) Thermodynamic databases for proteins and protein–nucleic acid interactions. *Biopolymers*, **61**, 121–126.
9. Bava, K.A., Gromiha, M.M., Uedaira, H., Kitajima, K. and Sarai, A. (2004) ProTherm, version 4.0: thermodynamic database for proteins and mutants. *Nucleic Acids Res.*, **32**, 120–121.
10. Kumar, M.S., Bava, K.A., Gromiha, M.M., Prabakaran, P., Kitajima, K., Uedaira, H. and Sarai, A. (2006) ProTherm and ProNIT: thermodynamic databases for proteins and protein–nucleic acid interactions. *Nucleic Acids Res.*, **34**, D204–D206.
11. Ponnuswamy, P.K. and Gromiha, M.M. (1994) On the conformational stability of folded proteins. *J. Theor. Biol.*, **166**, 63–74.
12. Gromiha, M.M., Oobatake, M., Kono, H., Uedaira, H. and Sarai, A. (1999) Relationship between amino acid properties and protein stability: buried mutations. *J. Protein Chem.*, **18**, 565–578.
13. Shen, B., Bai, J. and Vihinen, M. (2008) Physicochemical feature-based classification of amino acid mutations. *Protein Eng. Des. Sel.*, **21**, 37–44.
14. Saraboji, K., Gromiha, M.M. and Ponnuswamy, M.N. (2006) Average assignment method for predicting the stability of protein mutants. *Biopolymers*, **82**, 80–92.
15. Masso, M. and Vaisman, I.I. (2014) AUTO-MUTE 2.0: a portable framework with enhanced capabilities for predicting protein functional consequences upon mutation. *Adv Bioinformatics*, **2014**, 278385.
16. Pucci, F., Bourgeois, R. and Rooman, M. (2016) Predicting protein thermal stability changes upon point mutations using statistical potentials: introducing HoTMuSiC. *Sci. Rep.*, **6**, 23257.
17. Kulandaisamy, A., Zauha, J., Frishman, D. and Michael Gromiha, M. (2020) MPTherm-pred: analysis and prediction of thermal stability changes upon mutations in transmembrane proteins. *J. Mol. Biol.*, doi:10.1016/j.jmb.2020.09.005.
18. Capriotti, E., Fariselli, P. and Casadio, R. (2005) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.*, **33**, W306–W310.
19. Parthiban, V., Gromiha, M.M. and Schomburg, D. (2006) CUPSAT: prediction of protein stability upon point mutations. *Nucleic Acids Res.*, **34**, W239–W242.
20. Cheng, J., Randall, A. and Baldi, P. (2006) Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins*, **62**, 1125–1132.
21. Yin, S., Ding, F. and Dokholyan, N.V. (2007) Modeling backbone flexibility improves protein stability estimation. *Structure*, **15**, 1567–1576.
22. Dehouck, Y., Kwasigroch, J.M., Gilis, D. and Rooman, M. (2011) PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics*, **12**, 151.
23. Pires, D.E., Ascher, D.B. and Blundell, T.L. (2014) mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics*, **30**, 335–342.
24. Laimer, J., Hofer, H., Fritz, M., Wegenkittl, S. and Lackner, P. (2015) MAESTRO-multi agent stability prediction upon point mutations. *BMC Bioinformatics*, **16**, 116.
25. Folkman, L., Stantic, B., Sattar, A. and Zhou, Y. (2016) EASE-MM: sequence-based prediction of mutation-induced stability changes with feature-based multiple models. *J. Mol. Biol.*, **428**, 1394–1405.
26. Steinbrecher, T., Zhu, C., Wang, L., Abel, R., Negron, C., Pearlman, D., Feyfant, E., Duan, J. and Sherman, W. (2017) Predicting the effect of amino acid single-point mutations on protein stability—large-scale validation of MD-based relative free energy calculations. *J. Mol. Biol.*, **429**, 948–963.
27. Rodrigues, C.H., Pires, D.E. and Ascher, D.B. (2018) DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.*, **46**, W350–W355.
28. Montanucci, L., Capriotti, E., Frank, Y., Ben-Tal, N. and Fariselli, P. (2019) DDGun: an untrained method for the prediction of protein stability changes upon single and multiple point variations. *BMC Bioinformatics*, **20**, 335.
29. Casadio, R., Vassura, M., Tiwari, S., Fariselli, P. and Luigi Martelli, P. (2011) Correlating disease-related mutations to their effect on protein stability: A large-scale analysis of the human proteome. *Hum. Mutat.*, **32**, 1161–1170.
30. Steff, S., Nishi, H., Petukh, M., Panchenko, A.R. and Alexov, E. (2013) Molecular mechanisms of disease-causing missense mutations. *J. Mol. Biol.*, **425**, 3919–3936.
31. Martelli, P.L., Fariselli, P., Savojardo, C., Babbi, G., Aggazio, F. and Casadio, R. (2016) Large scale analysis of protein stability in OMIM disease related human protein variants. *BMC Genomics*, **17**, 239–247.
32. Kulandaisamy, A., Sakthivel, R. and Gromiha, M.M. (2020) MPTherm: database for membrane protein thermodynamics for understanding folding and stability. *Brief. Bioinform.*, bbaa064.
33. UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.*, **47**, D506–D515.
34. Burley, S.K., Berman, H.M., Kleywegt, G.J., Markley, J.L., Nakamura, H. and Velankar, S. (2017) Protein Data Bank (PDB): the single global macromolecular structure archive. *Methods Mol. Biol.*, **1607**, 627–641.
35. Yang, Y., Urolagin, S., Niroula, A., Ding, X., Shen, B. and Vihinen, M. (2018) Pon-tstab: protein variant stability predictor. Importance of training data quality. *Int. J. Mol. Sci.*, **19**, 1009.
36. Savitski, M.M., Reinhard, F.B., Franken, H., Werner, T., Savitski, M.F., Eberhard, D., Molina, D.M., Jafari, R., Dovega, R.B., Klaeger, S. et al. (2014) Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science*, **346**, 1255784.
37. Liu, F. and Fitzgerald, M.C. (2016) Large-scale analysis of breast cancer-related conformational changes in proteins using limited proteolysis. *J. Proteome Res.*, **15**, 4666–4674.
38. Hanson, R.M., Prilusky, J., Renjian, Z., Nakane, T. and Sussman, J.L. (2013) JSmol and the next-generation web-based representation of 3D molecular structure as applied to Proteopedia. *Israel J. Chem.*, **53**, 207–216.
39. Dana, J.M., Gutmanas, A., Tyagi, N., Qi, G., O'Donovan, C., Martin, M. and Velankar, S. (2019) SIFTS: updated structure integration with function, taxonomy and sequences resource allows 40-fold increase in coverage of structure-based annotations for proteins. *Nucleic Acids Res.*, **47**, D482–D489.
40. Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–637.