

# ProNAB: database for binding affinities of protein–nucleic acid complexes and their mutants

Kannan Harini, Ambuj Srivastava, Arulsamy Kulandaisamy and M. Michael Gromiha <sup>\*</sup>

Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India

Received August 12, 2021; Revised September 08, 2021; Editorial Decision September 10, 2021; Accepted September 10, 2021

## ABSTRACT

**Protein–nucleic acid interactions are involved in various biological processes such as gene expression, replication, transcription, translation and packaging. The binding affinities of protein–DNA and protein–RNA complexes are important for elucidating the mechanism of protein–nucleic acid recognition. Although experimental data on binding affinity are reported abundantly in the literature, no well-curated database is currently available for protein–nucleic acid binding affinity. We have developed a database, ProNAB, which contains more than 20 000 experimental data for the binding affinities of protein–DNA and protein–RNA complexes. Each entry provides comprehensive information on sequence and structural features of a protein, nucleic acid and its complex, experimental conditions, thermodynamic parameters such as dissociation constant ( $K_d$ ), binding free energy ( $\Delta G$ ) and change in binding free energy upon mutation ( $\Delta\Delta G$ ), and literature information. ProNAB is cross-linked with GenBank, UniProt, PDB, ProThermDB, PROSITE, DisProt and Pubmed. It provides a user-friendly web interface with options for search, display, sorting, visualization, download and upload the data. ProNAB is freely available at <https://web.iitm.ac.in/bioinfo2/pronab/> and it has potential applications such as understanding the factors influencing the affinity, development of prediction tools, binding affinity change upon mutation and design complexes with the desired affinity.**

## INTRODUCTION

Protein–nucleic acid interactions play essential roles in fundamental cellular processes such as regulation of gene expression, replication, translation, DNA repair and packaging. The functions of protein–nucleic acid complexes are mainly dictated by their binding affinities (1). Experimentally, the strength of the protein–nucleic acid interactions

is determined using isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), electrophoretic mobility shift assay (EMSA) and fluorescence. The binding affinity data such as dissociation constant ( $K_d$ ) and binding free energy ( $\Delta G$ ) are successfully used to understand the recognition mechanism of protein–DNA and protein–RNA complexes (2).

Further, amino acid and/or nucleotide mutations in protein–nucleic acid complexes alter their binding affinities and some of them lead to diseases including cancer and neurodegenerative disorders (3,4). Hence, binding affinities of protein–nucleic acid complexes are essential for understanding the disease-causing mechanisms and developing disease-specific drug design strategies. For example, designing aptamers with high affinity is reported to be a promising therapy for different diseases (5–7). With advancements in high-throughput experimental methods, vast amount of thermodynamic data on protein–DNA and protein–RNA complexes are reported in the literature. An effective compilation of these data and constructing a public repository would aid researchers in gaining insights for understanding the relationship among binding affinity, structure, function and diseases.

The binding affinities of protein–nucleic acid complexes are accumulated in few databases such as ProNIT (8), dbAMEPNI (9) and PDBbind (10), and these databases have several limitations on scope, update and availability. ProNIT has not been updated after 2006 and dbAMEPNI has a limited number of data on alanine mutations alone. PDBbind is not specific for protein–nucleic acid complexes and also it has only the binding affinity data for complexes with known structural information. On the other hand, experimental binding affinities are necessary to develop machine learning methods for predicting the binding affinity (11–13) and change in binding affinity upon mutation (14).

In this study, we developed a comprehensive database for protein–nucleic acid binding affinity, ProNAB, which contains binding affinity data such as dissociation ( $K_d$ ) and association ( $K_a$ ) constants, enthalpy, binding free energy ( $\Delta G$ ), and change in free energy upon mutation ( $\Delta\Delta G$ ) along with sequence and structural information of proteins

<sup>\*</sup>To whom correspondence should be addressed. Tel: +91 44225 74138; Fax: +91 44225 74102; Email: [gromiha@iitm.ac.in](mailto:gromiha@iitm.ac.in)

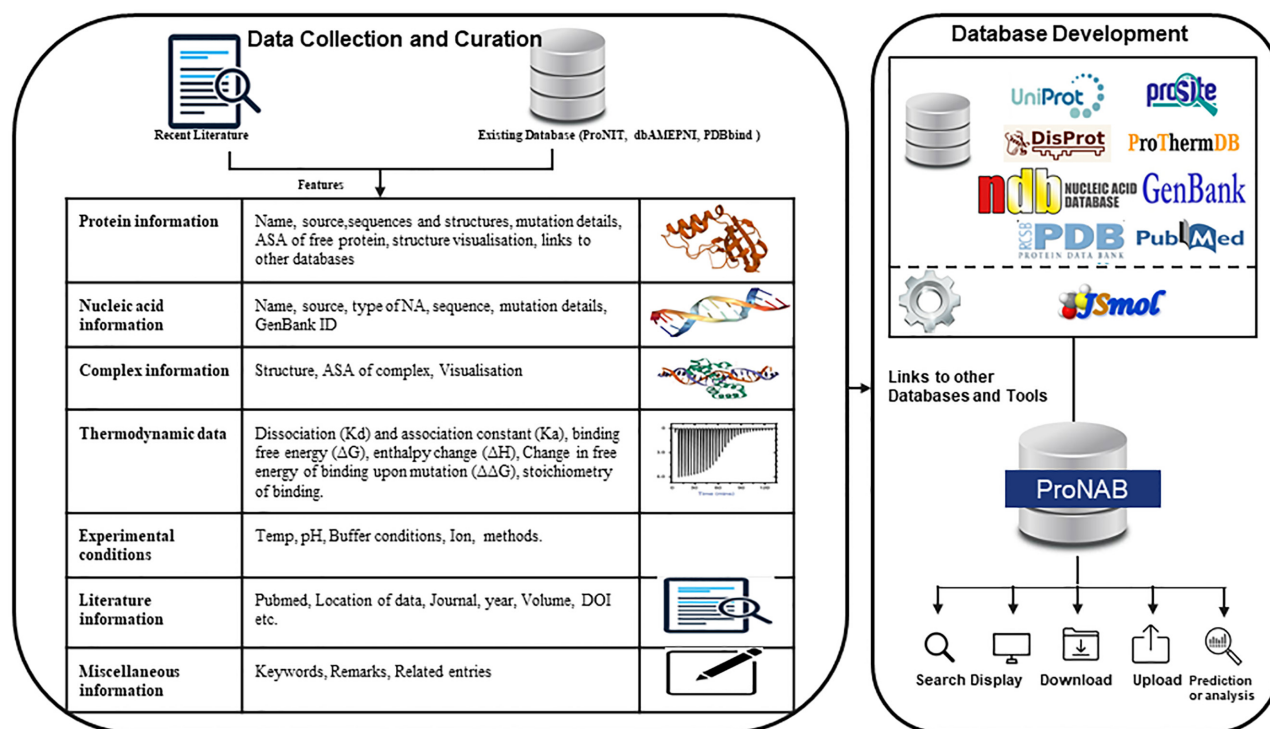


Figure 1. Overall workflow of ProNAB database.

and nucleic acids, experimental conditions and literature information. Each entry in ProNAB provides comprehensive features from both protein and nucleic acid sequence and structure. Further, it is cross-linked with various sequence and structure databases of proteins and nucleic acids. The web interface provides flexible options for researchers to search based on different parameters along with options for sorting and visualization. ProNAB is available at <https://web.iitm.ac.in/bioinfo2/pronab/>.

## CONTENTS OF THE DATABASE

ProNAB provides experimentally determined binding affinity data for wild-type and mutant protein–nucleic acid complexes. These data are obtained from a detailed survey of literature and existing/ obsolete databases.

We have retrieved research articles and reviews related to binding affinity of protein–nucleic acid complexes using keyword searches with AND/OR operations (e.g. binding affinity, protein–DNA, protein–RNA, dissociation constant,  $K_d$ , free energy of binding, ITC, SPR etc.) from PubMed. Further, we obtained the papers listed in Protein Data Bank (PDB) for known three-dimensional structures of protein–nucleic acid complexes and their binding affinities. In addition, we checked the ‘Table of contents’ of specific journals (e.g. Biochemistry, Nucleic Acids Research, Journal of Molecular Biology, Journal of Biological Chemistry, PNAS etc.), which publish articles related to experimental data on binding affinity of protein–nucleic acid complexes. From each article, we manually curated the information about the name of the protein, nucleic acid, complex, experimental conditions, measurement, method, thermody-

amic data, literature information and location of the data in the research article.

The detailed workflow of the ProNAB is provided in Figure 1. For each entry, ProNAB contains the following seven different levels of information. Each entry in ProNAB is identified with a unique entry number, which contains the information provided in Table 1.

- Protein information:** Protein name, name of the organism (source), sequence, structure, accession numbers of UniProt (15), PROSITE (16), ProThermDB (17), DisProt (18), enzyme commission number (19) and PDB (20), secondary structure, solvent accessibility and mutation information (wild-type, single, double and multiple mutations along with mutant positions). We utilized the SIFTS database (21) for mapping residue positions between UniProt and PDB.
- Nucleic acid information:** Nucleic acid name, source, type of nucleic acid such as DNA or RNA, GenBank ID, sequence, mutation and type of mutation such as single, double and multiple along with mutation position.
- Complex information:** PDB (20) and NDB (22) codes for both wild-type and mutant structures of protein–nucleic acid complexes (if available), 3D visualization of the complex using JSmol interface (23), secondary structure and solvent accessibility of the mutant in a complex calculated using DSSP (24) for the proteins with known three-dimensional structures.
- Experimental conditions:** Temperature, pH, buffer name, additives, ions and method.
- Thermodynamic data:** The thermodynamic parameters of binding affinity are represented as dissociation ( $K_d$ )

**Table 1.** Description of data in ProNAB with an example entry showing the binding affinity data of ‘Cysteine-tRNA ligase’ protein–RNA complex

Description	Example
Entry id	12197
Protein Name	Cysteine–tRNA ligase
Synonyms	CysteinyI–tRNA synthetase; CysRS
EC number	6.1.1.16
Protein Source	<i>Escherichia coli</i> (strain K12)
Sequence	MLKIFNTLTRQKEEFKPIHAGEVGMVYVCGITVYDLCHIGHGRTFVAFDVVARYLRF LGYKLLKYVRNITDIDDKIIKRANENGESFVAMVDRMIAEMHKDFDALNLRPDMEP RATHHIAEIIELTEQLIAKGHAYVADNGDVMFDV...
Length	461
Mass (Da)	52,202
UniProt ID	P21888
PROSITE ID	-
DisProt ID	-
PDB of Free Protein	1LI5
ASA of Free protein (Å <sup>2</sup> )	29
ProTherm Id	-
Mutation in protein	N351A
Nucleic acid Name	TRNA-cys
Nucleic acid Source	Synthetic
Type of Nuclei acid	RNA
Sequence	GGCGCGUUAACAAAGCGGUUAUGUAGCGGAUUGCAAUCCGUCUAGUCCGGU UCGACUCCGGAAC....
Mutation in Nucleic acid	G48C
Genbank ID	56966181
PDB Complex	1U0B
NDB Complex	PR0135
ASA of Complex (Å <sup>2</sup> )	40
Sec str	Coil
pH	7.5
Temperature (K)	298
Buffer	20 mM Tris–Hcl
Ion name	50 mM NaCl
Method	Fluorescence
$K_d$ wild (M)	$2.7 \times 10^{-7}$
$K_d$ mutant (M)	$8.16 \times 10^{-6}$
$K_a$ wild (M <sup>-1</sup> )	$4 \times 10^6$
$K_a$ mutant (M <sup>-1</sup> )	$1 \times 10^5$
$\Delta G$ wild (kcal/mol)	–8.96
$\Delta G$ mutant (kcal/mol)	–6.94
$\Delta\Delta G$ (kcal/mol)	2.02
$\Delta H$ wild (kcal/mol)	-
$\Delta H$ mutant (kcal/mol)	-
Stoichiometry	-
Reference	Nat Struct Mol Biol. 2004 Nov;11(11):1134–41.
Title	Shape-selective RNA recognition by cysteinyI-tRNA synthetase.
Authors	Hauenstein S, Zhang CM, Hou YM, Perona JJ
Keywords	CysRS; tRNA aminoacylation; elongation factor;
PubMed	15489861
DOI	<a href="http://dx.doi.org/10.1038/nsmb849">http://dx.doi.org/10.1038/nsmb849</a>
Location of data	Table 2; Page No.: 1138
Remarks	-
Related Entries	12194; 12195; 12196

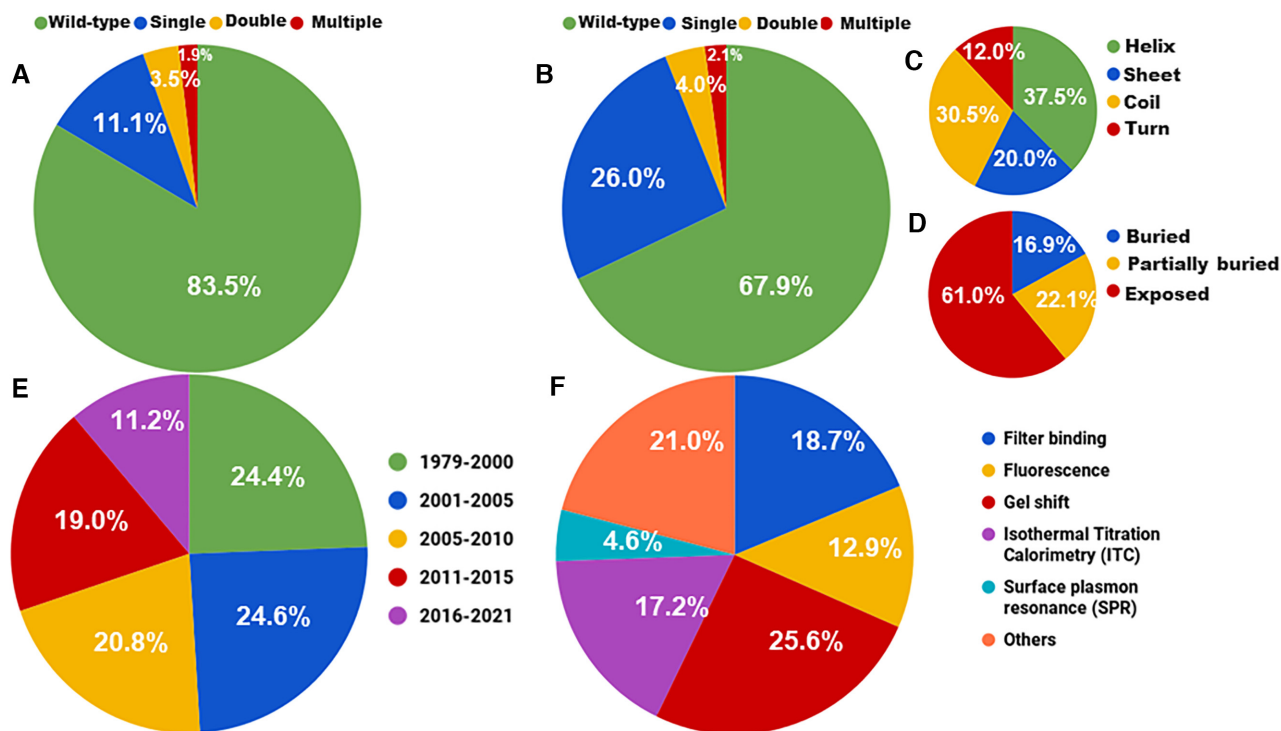
and association constants ( $K_a$ ), enthalpy ( $\Delta H$ ), the free energy of binding ( $\Delta G$ ) and the change in free energy of binding ( $\Delta\Delta G$ ) for the mutants.

- f) **Literature information:** PubMed identifier, name of the author(s), journal name, year of publication, location of the data, keywords and Digital Object Identifier (DOI).
- g) **Miscellaneous Information:** Remarks and entry numbers related to the same protein in ProNAB.

## DATABASE STATISTICS

ProNAB contains 20 090 entries, which include 14 606 and 5323 entries for protein–DNA and protein–RNA binding

affinities, respectively along with 161 entries for hybrid complexes (protein–DNA–RNA). It has binding affinity information for 1027 unique nucleic acid binding proteins from 1250 literature sources published during 1979–2021. A total of 798 unique protein–DNA and 340 protein–RNA complex structures are present in ProNAB. Based on wild-type, the current version contains 13 642, 12 318 and 4304 data from proteins, DNA and RNA, respectively. Further, 6448, 2288 and 5323 mutation data are available for proteins, DNA and RNA, respectively. Among them, 76.4%, 15.4% and 8.15% are single, double and multiple mutations, respectively. Detailed statistics on wild-type and mutant data for proteins and nucleic acids are presented in Figure 2A



**Figure 2.** Statistics of ProNAB database based on the distribution of (A) wild-type and mutant data in Proteins, (B) wild-type and mutant data in nucleic acids, (C) secondary structure of mutants, (D) solvent accessibility of mutants, (E) publication years and (F) methods

and B. Figure 2C and D shows the representation of mutants based on secondary structure and solvent accessibility, respectively. Figure 2E and F provides the information about the distribution of data based on the year of publication and experimental methods used to determine the binding affinity, respectively.

## LINKS TO OTHER DATABASES

ProNAB is cross-linked with different sequence, structure, stability, and other relevant databases. Entries are linked with (i) UniProt, which provides both sequence and functional information of the protein, (ii) PDB, three-dimensional structural information, (iii) ProThermDB to understand the stability of the proteins and mutants, (iv) PROSITE, which provides details about the motif present in the proteins, (v) DisProt to show the disorderness of the protein, (vi) GenBank to obtain the information on nucleic acid sequences, (vii) PDB and NDB for the protein–nucleic acid complexes and (viii) PUBMED for the literature information.

## DATA RETRIEVAL

The detailed information about the search and display options and an example for data retrieval from ProNAB are illustrated in Figure 3. In this example, we build a query using a combination of multiple search options as ‘Get the change in binding free energy upon mutation ( $\Delta\Delta G$ ) within the range of  $-5$  to  $-1$  kcal/mol in proteins with ‘single mutations’ measured at the temperature and pH in the range of ‘293–298 K

and ‘5–7’ respectively’ (Figure 3A). In addition, we selected the desired columns in the display options as  $\Delta\Delta G$ , Temperature, pH and other default options (Figure 3B). The data is sorted based on  $\Delta\Delta G$  values (Figure 3C). After submitting the query, the results are displayed in a table format (Figure 3D). We also provided an option to download the search results in CSV format and users can easily parse the same for further analysis. On the result page, each entry accession number has a hyperlink for their respective external page (Figure 3E), which contains the complete information. The structural visualization is available for each entry.

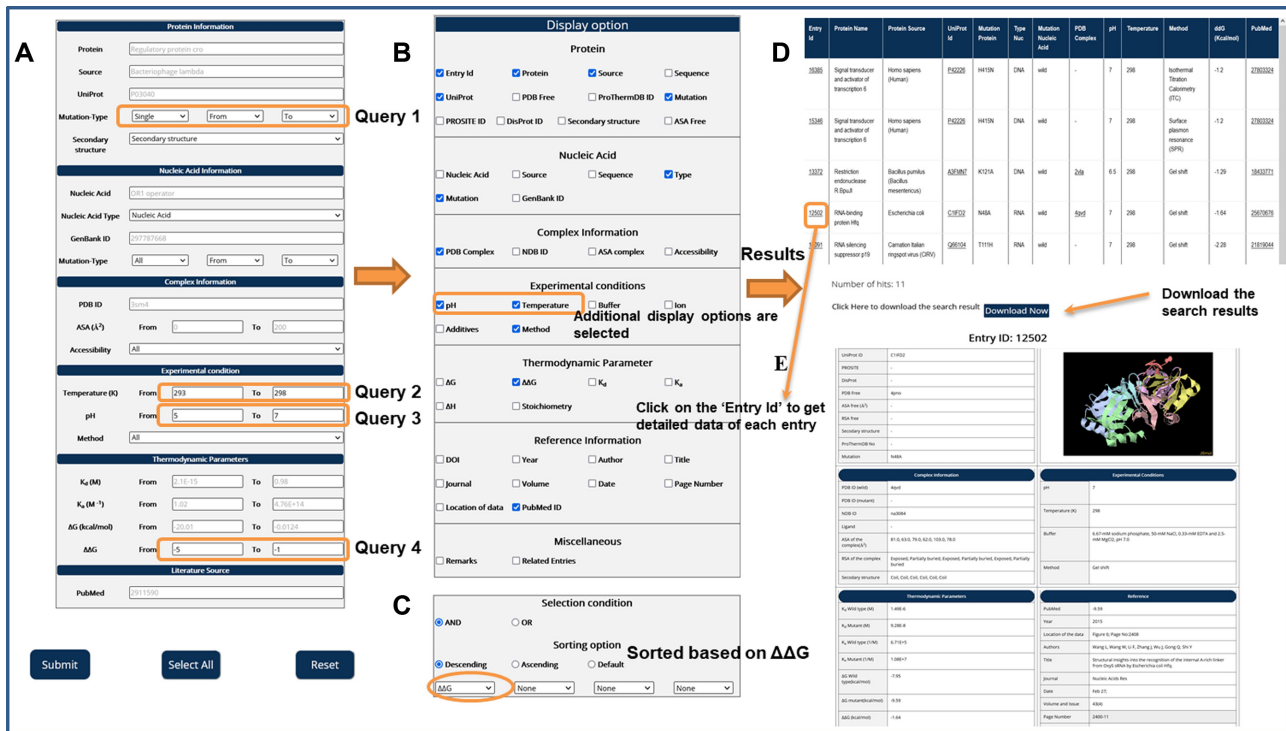
## DATA DOWNLOAD AND UPLOAD

Users can upload their new experimentally determined binding affinity data of protein–nucleic acid complexes into the ProNAB database. For uploading the data, depositors are requested to use the ‘Data Upload’ option in the web page and supply the following information: Protein–nucleic acid complex name, UniProt/PDB code, PubMed or Digital Object Identifier (DOI) number. We have also provided an option to download the entire ProNAB data by submitting a request to the corresponding author through the ‘Data Download’ option in the web page.

## COMPARISON WITH EXISTING DATABASES

A detailed comparison of ProNAB with other existing databases is given in Table 2. ProNIT is currently not accessible on the web and ProNAB has an increase of 66% data compared to the previous version of ProNIT with a





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

**Table 2.** Comparison of ProNAB with other existing databases

Features	ProNIT	dbAMEPNI	PDBbind	ProNAB
Availability	No	Yes	Yes	Yes
Number of Entries	12 174	578	973	20 090
Number of unique PDB structures	124 (108 protein–DNA, 16 protein–RNA)	152 (101 protein–DNA, 51 protein–RNA)	973(670 protein–DNA, 293 protein–RNA)	1138 (798 protein–DNA, 340 protein–RNA)
Type of mutation	All	Only for alanine mutation	All	All
Change in binding affinity upon mutation ( $\Delta\Delta G$ )	No	Yes	No	Yes
Availability of exact location of data	No	Yes	No	Yes
Structure visualization	No	No	No	Yes
Protein information and buffer conditions	Yes	No	No	Yes
Option for upload, to help maintenance	No	No	No	Yes
Literature Year	1983–2013	1983–2017	1993–2019	1979–2021

considerable increase in entries with known structural information. dbAMEPNI has only 578 binding affinity data specific for alanine mutations, whereas ProNAB has 2397 entries for alanine mutations and also has binding affinity data for all types of mutations along with the wild-type data. The other database, PDBbind, has binding affinity data for biomolecular complexes (protein–protein, protein–ligand, protein–nucleic acids) in PDB, i.e. only for complexes with known structural information. On the other hand, ProNAB is a comprehensive database for the binding affinity of protein–nucleic acid complexes with sequence as well as structural information. Further, ProNAB also provides other features such as change in binding affinity upon mutations and the exact location of the data in the literature.

Also, our database has options to upload new data and links for the structural visualization of complexes. The ProNAB database is linked to several databases such as ProThermDB, DisProt, PROSITE and GenBank.

### APPLICATIONS

ProNAB has several potential applications and some of them are listed below:

- (i) Explore the relationship between binding affinity and structural/sequence-based features of proteins and nucleic acids to understand the molecular mechanism of protein–nucleic acid interactions (25,26). ProNAB

provides a wealth of data for the binding affinities of protein–nucleic acid complexes, which can be used to elucidate the important features based on structure and function, which governs the affinities.

- (ii) Study the effect of mutation on the binding affinity of the complex (27,28). ProNAB serves as a potential resource for providing experimentally determined binding affinities of protein–nucleic acid complexes and their mutants. These data are useful for both large-scale analysis as well as in-depth analysis of a specific complex.
- (iii) Develop computational tools and reliable machine learning models for predicting the binding affinity of protein–nucleic acid complexes and binding affinity change upon mutations (11–14).
- (iv) Design DNA/RNA aptamers with the desired affinity (29,30). Using the binding affinity data of experimentally determined DNA/RNA aptamers available in ProNAB, computational and experimental methods could be developed to design aptamers with desired affinities.
- (v) Investigate the relationship between binding affinity change and disease causing mutations as reported for protein–protein complexes (31).

## DATA AVAILABILITY

ProNAB is available at <https://web.iitm.ac.in/bioinfo2/pronab/>. The database is developed using HTML, CSS, PHP, MySQL and JavaScript and it supports the latest version of major browsers such as Firefox, Chrome and Opera. The database will be maintained and updated regularly. Each update 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](mailto:gromiha@iitm.ac.in).

## ACKNOWLEDGEMENTS

We acknowledge the Indian Institute of Technology Madras and the High-Performance Computing Environment (HPCE) for computational facilities. We thank all the members from Protein Bioinformatics Lab for providing valuable suggestions and Divya Sharma for maintaining the database. The work is partially supported by the Science and Engineering Research Board (SERB), Ministry of Science and Technology, Government of India to MMG (No. CRG/2020/000314)

## FUNDING

Indian Institute of Technology, Madras; Science and Engineering Research Board (SERB), Ministry of Science and Technology, Government of India [CRG/2020/000314 to M.M.G.]. Funding for open access charge: IIT Madras and Department of Science and Technology, Government of India.

*Conflict of interest statement.* None declared.

## REFERENCES

1. Crocker, J., Noon, E.P. and Stern, D.L. (2016). The soft touch: low-affinity transcription factor binding sites in development and evolution. *Curr. Top. Dev. Biol.*, **117**, 455–469.
2. Duss, O., Michel, E., Diarra dit Konté, N., Schubert, M. and Allain, F.H. (2014). Molecular basis for the wide range of affinity found in Csr/Rsm protein–RNA recognition. *Nucleic Acids Res.*, **42**, 5332–5346.
3. Cordeiro, Y., Macedo, B., Silva, J.L. and Gomes, M.P.B. (2014). Pathological implications of nucleic acid interactions with proteins associated with neurodegenerative diseases. *Biophys. Rev.*, **6**, 97–110.
4. Qin, H., Ni, H., Liu, Y., Yuan, Y., Xi, T., Li, X. and Zheng, L. (2020). RNA-binding proteins in tumor progression. *J. Hematol. Oncol.*, **13**, 90.
5. Wongphatcharachai, M., Wang, P., Enomoto, S., Webby, R.J., Gramer, M.R., Amonsin, A. and Sreevatsan, S. (2013). Neutralizing DNA aptamers against swine influenza H3N2 viruses. *J. Clin. Microbiol.*, **51**, 46–54.
6. Xu, J., Zhang, X., Zhou, S., Shen, J., Yang, D., Wu, J., Li, X., Li, M., Huang, X., Sealy, J.E. et al. (2017). A DNA aptamer efficiently inhibits the infectivity of Bovine herpesvirus 1 by blocking viral entry. *Sci. Rep.*, **7**, 11796.
7. Liu, X., Wang, Y.L., Wu, J., Qi, J., Zeng, Z., Wan, Q., Chen, Z., Manandhar, P., Cavener, V.S., Boyle, N.R. et al. (2021). Neutralizing aptamers block S/RBD-ACE2 interactions and prevent host cell infection. *Angew. Chem. Weinheim. Bergstr. Ger.*, **133**, 10361–10366.
8. Kumar, M.D., 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.
9. Liu, L., Xiong, Y., Gao, H., Wei, D.Q., Mitchell, J.C. and Zhu, X. (2018). dbAMEPNI: a database of alanine mutagenic effects for protein–nucleic acid interactions. *Database*, **2018**, bay034.
10. Liu, Z., Li, Y., Han, L., Li, J., Liu, J., Zhao, Z., Nie, W., Liu, Y. and Wang, R. (2015). PDB-wide collection of binding data: current status of the PDBbind database. *Bioinformatics*, **31**, 405–412.
11. Deng, L., Yang, W. and Liu, H. (2019). PredPRBA: prediction of protein–RNA binding affinity using gradient boosted regression trees. *Front. Genet.*, **10**, 637.
12. Yang, W. and Deng, L. (2020). PreDBA: A heterogeneous ensemble approach for predicting protein–DNA binding affinity. *Sci. Rep.*, **10**, 1278.
13. Nithin, C., Mukherjee, S. and Bahadur, R.P. (2019). A structure-based model for the prediction of protein–RNA binding affinity. *RNA*, **25**, 1628–1645.
14. Jiang, Y., Liu, H.F. and Liu, R. (2021). Systematic comparison and prediction of the effects of missense mutations on protein–DNA and protein–RNA interactions. *PLoS Comput. Biol.*, **17**, e1008951.
15. UniProt Consortium (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.
16. Sigrist, C.J., Cerutti, L., de Castro, E., Langendijk-Genevaux, P.S., Bulliard, V., Bairoch, A. and Hulo, N. (2010). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res.*, **38**, D161–D166.
17. Nikam, R., Kulandaisamy, A., Harini, K., Sharma, D. and Gromiha, M.M. (2021) ProThermDB: thermodynamic database for proteins and mutants revisited after 15 years. *Nucleic Acids Res.*, **49**, D420–D424.
18. Hatos, A., Hajdu-Soltesz, B., Monzon, A.M., Palopoli, N., Alvarez, L., Aykac-Fas, B., Bassot, C., Benitez, G.I., Bevilacqua, M., Chasapi, A. et al. (2020) DisProt: intrinsic protein disorder annotation in 2020. *Nucleic Acids Res.*, **48**, D269–D276.
19. Morgat, A., Lombardot, T., Coudert, E., Axelsen, K., Neto, T.B., Gehant, S., Bansal, P., Bolleman, J., Gasteiger, E., de Castro, E. et al. (2020) Enzyme annotation in UniProtKB using Rhea. *Bioinformatics*, **36**, 1896–1901.
20. Rose, Y., Duarte, J.M., Lowe, R., Segura, J., Bi, C., Bhikadiya, C., Chen, L., Rose, A.S., Bittrich, S., Burley, S.K. et al. (2021) RCSB Protein Data Bank: architectural advances towards integrated searching and efficient access to macromolecular structure data from the PDB archive. *J. Mol. Biol.*, **433**, 166704.
21. 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.
22. Narayanan, B.C., Westbrook, J., Ghosh, S., Petrov, A.I., Sweeney, B., Zirbel, C.L., Leontis, N.B. and Berman, H.M. (2014) The Nucleic Acid

- Database: new features and capabilities. *Nucleic Acids Res.*, **42**, D114–D122.
23. 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. *Isr. J. Chem.*, **53**, 207–216.
24. Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure - pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
25. Jakubec, D., Laskowski, R.A. and Vondrasek, J. (2016) Sequence-specific recognition of DNA by proteins: binding motifs discovered using a novel statistical/computational analysis. *PLoS One*, **11**, e0158704.
26. Zeiske, T., Baburajendran, N., Kaczynska, A., Brasch, J., Palmer, A.G. 3rd, Shapiro, L., Honig, B. and Mann, R.S. (2018) Intrinsic DNA shape accounts for affinity differences between Hox-cofactor binding sites. *Cell Rep.*, **24**, 2221–2230.
27. Remali, J., Aizat, W.M., Ng, C.L., Lim, Y.C., Mohamed-Hussein, Z.A. and Fazry, S. (2020) In silico analysis on the functional and structural impact of Rad50 mutations involved in DNA strand break repair. *PeerJ.*, **8**, e9197.
28. Pejaver, V., Urresti, J., Lugo-Martinez, J., Pagel, K.A., Lin, G.N., Nam, H.J., Mort, M., Cooper, D.N., Sebat, J., Iakoucheva, L.M. *et al.* (2020) Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nat Commun.*, **11**, 5918.
29. Byun, J. (2021) Recent progress and opportunities for nucleic acid aptamers. *Life (Basel)*, **11**, 193.
30. Zhou, J. and Rossi, J. (2017) Aptamers as targeted therapeutics: current potential and challenges. *Nat. Rev. Drug Discov.*, **16**, 181–202.
31. Jemimah, S. and Gromiha, M.M. (2020) Insights into changes in binding affinity caused by disease mutations in protein-protein complexes. *Comput. Biol. Med.*, **123**, 103829.