Thermodynamic Database for Protein-Nucleic Acid Interactions

P. Prabakaran [1]
praba@rtc.riken.go.jp

Jianghong An [1]
ajh@rtc.riken.go.jp

M. Michael Gromiha [1]
gromiha@rtc.riken.go.jp

S. Selvaraj [1]
selvaraj@rtc.riken.go.jp

Hatsuho Uedaira [1]
uedaira@rtc.riken.go.jp

Hidetoshi Kono [2]
hkono@rtc.riken.go.jp

Akinori Sarai [1]
sarai@rtc.riken.go.jp

[1] Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

[2] Department of Chemistry, University of Pennsylvania, 231 South 34 street, Philadelphia PA, 19104-6323, USA

1 Introduction

Protein-nucleic acid interaction plays an essential role in the regulation of gene expression. Structural and binding information provides insight into the mechanism of protein-nucleic acid recognition. Although there are structural databases for protein-nucleic acid complexes, there exists no database for protein-nucleic acid binding. Thus, we have decided to develop the Thermodynamic Database for Protein-Nucleic Acid Interactions. It contains several important thermodynamic parameters, experimental conditions, structural information about proteins, nucleic acids and of the complex, and literature information. This database will be helpful for understanding the mechanism of protein-nucleic acid recognition. A WWW interface will help users to search the data based on various conditions with different sorting options for outputs.

2 Contents

Each entry of the database is referred by a serial number and contains the following information: (i) *Protein information*: name, source, fragment and sequence of the protein, enzyme code (EC), Protein Data Bank (PDB) code for wild and mutant structures, information about monomeric, oligomeric states, details of mutation with mutant residue name, number and secondary structure having mutant sites, accessibility, Thermodynamic Database for Proteins and Mutants (ProTherm) database number and Protein Mutant Database (PMD) number. (ii) *Nucleic acid information*: name, source and sequence of the nucleic acid, information on mutation and sequence of the mutant nucleic acid, GeneBank number and Nucleic Acid Database (NDB) number. (iii) *Complex information*: codes for PDB and NDB, Protein-Nucleic Acid Complex Database number, details of ligand molecules, accessibility of relevant mutant residue in the complex and conformational changes of protein as well as nucleic acid upon binding. (iv) *Experimental conditions*: temperature (T), the pH value, details about buffer, ions, additives and experimental method. (v) *Binding data*: dissociation constant (K_{d}), association constant (K_{a}), Gibbs free energy change (ΔG), enthalpy change (ΔH) and heat capacity change (ΔC_{p}) for wild and mutant entities, stoichiometry, activity (k_{m} and k_{cat}). (vi) *Literature information*: reference, authors, keywords and remarks.

3 Search and Display options

A WWW interface can be used to search data for various conditions with different sorting options for output according to users' purpose and convenience. The following are the searching options available in the database: (i) retrieving data for a particular protein by its name, source, sequence or PDB code, (ii) specifying the type of mutation as single, double, multiple or wild type. Further, specifying any wild type residue and/or mutant residue is possible, (iii) specifying the secondary structures possessing mutations as helix, strand, turn and coil, (iv) searching data based on a range of solvant accessible surface area (ASA; in % or $Å^2$) values of relevant mutant residue, (v) retrieving the complex information by providing ligand name, conformational changes and PDB code, (vi) extracting data based on various experimental conditions, namely, method, T, pH, buffer, ion and additives, and a particular range of values of these parameters can be specified, (vii) selecting data for a particular range of binding parameters K_{d}, K_{a}, ΔG, ΔH and ΔC_{p}, and (viii) searching data with authors, keywords and year of publication. Finally, the output format can be specified by selecting various display options and by sorting with ASA_free, ASA_complex, T, pH, Ion_conc, K_{d}_wild and year of publication.

4 Links

Thermodynamic Database for Protein-Nucleic Acid Interactions is integrated with other databases we have developed, ProTherm Database [1], 3DinSight [2], and Protein-Nucleic Acid Recognition Database. This database is also cross-linked with NCBI PUBMED literature database, EC, PMD, PDB and NDB.

5 Accessibility

Thermodynamic Database for Protein-Nucleic Acid Interactions is freely accessible at http://www.rtc.riken.go.jp/binding.html. Suggestions and comments are appreciated and should be sent to Akinori Sarai (sarai@rtc.riken.go.jp).

References

[1] Gromiha, M.M., An, J., Kono, H., Oobatake, M., Uedaira, H., and Sarai, A., ProTherm: Thermodynamic database for proteins and mutants, *Nucleic Acids Res.*, 27:286–288, 1999.

[2] An, J., Nakama, T., Kubota, Y., and Sarai, A., 3DinSight: an integrated relational database and search tool for structure, function and property of biomolecules", *Bioinformatics*, 14:188–195, 1998.