International Standard Serial Number (ISSN): 2249-6807

International Journal of Institutional Pharmacy and Life Sciences 5(3): May-June 2015

INTERNATIONAL JOURNAL OF INSTITUTIONAL PHARMACY AND LIFE SCIENCES

Life Sciences **Research Article……!!!**

Received: 28-04-2015; Revised: 14-05-2015; Accepted: 15-05-2015

ADSBET: AUTOMATED DETERMINATION OF SALT-BRIDGE ENERGY TERMS

Arnab Nayek, Parth Sarthi Sen Gupta, Shyamashree Banerjee, Sunit Das, Vishma Pratap Sur, Pratyay Seth, Rifat Nawaz Ul Islam and Amal K Bandyopadhyay*

Department of Biotechnology, The University of Burdwan, Burdwan 713101, West Bengal, India

Keywords:

Salt-Bridge; Energy-Terms; Poison-Boltzmann-Equation; Bridge energy; Protein energy; Desolvation energy

For Correspondence:

Amal K Bandyopadhyay

Department of Biotechnology, The University of Burdwan, Burdwan 713101, West Bengal, India

E-mail:

akbanerjee@biotech.buruniv.ac.in

ABSTRACT

Salt-bridges are specific electrostatic interaction between groups of opposite charges. Net-interaction-energy ($\Delta\Delta G_{net}$) of a salt-bridge is partitioned into bridge ($\Delta\Delta G_{brd}$), desolvation ($\Delta\Delta G_{dsolv}$) and protein ($\Delta\Delta G_{prot}$) energy-terms of which estimation of the later two are possible only by computational means. Computation of salt-bridge energy-terms using Poison-Boltzmann-Equation (PBE) solver method employs highly involved process in that prior computation of energy-terms, one needs to determine protein-specific salt-bridges, input-parameters, hydrophobic-isosteres-mutated charge-radius files. At this juncture, an efficient fully automated all-in-one-procedure that could analyze large dataset in a single run would be useful. To the best of our knowledge, such procedure is truly lacking in public domain. At this end, our fully automated procedure: ADSBET uses "APBS" method to compute component as well as net energy-terms for any numbers of salt-bridges in any numbers of structure files and redirect output in excel format. The output is not only useful for post run application of these energy terms but also helps to extract net energetic contribution of networked salt bridges. It works in UNIX like environment including CYGWIN. Overall, our ADSBET provides detailed energetic of salt-bridges from multiple crystal structures and find application in the field of computational structural biology. ADSBET is freely available at (http://sourceforge.net/projects/adsbet/).

INTRODUCTION

Salt bridge interactions that occur at **≤4Å** distance between side chain charges of acidic and basic residues[1-2] are maximally stable at near neutral pH conditions[3-4]. There is software for extraction of details of atomic[1] and residue specific[2] salt bridges from crystal structure. The net interaction energy of a given salt bridge is the sum of three component energy terms such as bridge ($\Delta\Delta G_{brd}$), desolvation ($\Delta\Delta G_{dsolv}$) and background terms ($\Delta\Delta G_{prot}$), They could either be stabilizing ($\Delta\Delta G_{net}$ negative) or destabilizing ($\Delta\Delta G_{net}$ positive)[5-6] and hence contributes both on stability and packing of functional state of proteins[2]. Of these component energy terms, $\Delta\Delta G_{brd}$ is direct term and is always contributing. On the other hand, $\Delta\Delta G_{dsolv}$ is due to desolvation of charges and thus always costly. The $\Delta\Delta G_{prot}$ is due to interaction of salt bridge partners with other charges in proteins in its vicinity[6-7] could either be stabilizing or costly. The latter two indirect terms also act as sensor for salt bridge's micro environment[6]. Thus stabilizing and destabilizing effects largely depends on location (core or surface of protein), microenvironment (favorable or unfavorable charges near salt bridges) and geometry of salt bridges[6-7].

How to measure these component as well as net energy terms? Double mutation cycle and pKa approaches are popularly used for experimental measurement of salt bridges. Unfortunately, none of these methods could determine indirect energy terms (i.e. $\Delta\Delta G_{dsolv}$ and $\Delta\Delta G_{prot}$) and hence the $\Delta\Delta G_{net}$ one[8]. Thus computational method is the only choice for the purpose. It can be done at any given pH[8] using general purpose Poison Boltzmann Equation (PBE) solvers such as APBS[9] and Delphi[10].The former is advantageous over the later due to its availability in public domain and also it allows measurement at any given ionic strength[9]. Energetic analyses of large database of salt bridges from different domains of lives and SCOP classes would reveal insight into protein adaptation, engineering and stability[7-8]. At this point mention may be made of the fact that application of these general purpose solvers is very few in the context of salt-bridge analyses[6-8]. The major bottleneck might be the great deal of manual involvement for the initial preparatory and final analytical steps. For each cycle of computation (i.e. per salt bridge), one needs i] to locate residue specific salt bridges in structure files ii] to prepare hydrophobic isosteres mediated mutation files (5 per salt bridge), iii] to generate structure specific input parameters such as grid points, center of mass and v] to prepare a combined input file using these and other user and default parameters. Further, upon completion of the run, integration of the product of atomic potential (ϕ) and their respective charges (**q**) are to be performed for each of these energy

terms on per salt bridge basis. These facts might answer as to why only limited numbers of structure files are addressed to date.

In this work we present a fully automated procedure for computation of component as well as net energy terms using the mostly accepted model[5-7] and Poison Boltzmann Equation Solver i.e. APBS. The details on computation of energy terms, extraction of residue specific salt bridges from crystal structures, preparation of input files using default, program-generated and user parameters are also been reported. We also present details and application of output. Overall, our automation which can analyze any number of structure file for any number of salt bridges in them has potential application in the field of computation structural biology.

METHODS

Model for computation of energy terms

Let us consider a salt bridge $[\boldsymbol{P^i_{ms}}]$-$[\boldsymbol{N^j_{ms}}]$ where $\boldsymbol{P^i_{ms}}$ and $\boldsymbol{N^j_{ms}}$ are positive and negative charge residues at position $\boldsymbol{i}$ and $\boldsymbol{j}$ respectively. Here, $\boldsymbol{m}$ and $\boldsymbol{s}$ indicate main and side chain atoms respectively.

Bridge energy model

$\boldsymbol{\Delta\Delta G^{P^iN^j}_{brd}}$ is resulted from interaction of side chain atoms of positive ($\boldsymbol{P^s}$) and negative charge ($\boldsymbol{N^s}$) residues in the folded state of protein. According to the working model[5], in the native protein-charge-radius (PQR) file, charges of either $\boldsymbol{P^s}$ or $\boldsymbol{N^s}$ is retained[6-7] and others are mutated by hydrophobic isosteres. The generated PQR file is then subjected for manually configured multigrid Poison Boltzmann calculation under single Debay-Huckel boundary condition (mPBsDH) by APBS method[9]. The atomic potential file thus obtained is used for computation of bridge energy term ($\boldsymbol{\Delta\Delta G^{P^iN^j}_{brd}}$) using following formula:

$$\boldsymbol{\Delta\Delta G^{P^iN^j}_{brd}} = [\textstyle\sum_s(\boldsymbol{q_{N^s}} * \boldsymbol{\phi_{N^s}})]$$ [When atomic charges of $\boldsymbol{P^s}$ is used for mPBsDH]

Equation 1

Or

$$\boldsymbol{\Delta\Delta G^{P^iN^j}_{brd}} = [\textstyle\sum_s(\boldsymbol{q_{P^s}} * \boldsymbol{\phi_{P^s}})]$$ [When atomic charges of $\boldsymbol{N^s}$ is used for mPBsDH]

Equation 2

Background energy model

$\boldsymbol{\Delta\Delta G^{P^iN^j}_{prot}}$ is the interaction energy of $\boldsymbol{P^s}$ and $\boldsymbol{N^s}$ with $\boldsymbol{O^s}$ (other side chain atoms in protein i.e. all except $\boldsymbol{P^s}$ and $\boldsymbol{N^s}$) in the folded state of protein[5]. In this case mPBsDH is solved on a generated PQR file that retains atomic charges only for $\boldsymbol{P^s}$ and $\boldsymbol{N^s}$. The potential file thus obtained was used for calculation of $\boldsymbol{\Delta\Delta G^{P^iN^j}_{prot}}$ using the following formula:

$\Delta\Delta G_{prot}^{P^iN^j} = [\sum_s (q_{O^s} * \phi_{O^s})]$ [When atomic charges of P^s and N^s are used for mPBsDH] Equation 3

Desolvation energy model

It is calculated separately on P^s and N^s for both folded and unfolded state of protein[5] using the following formula:

$$\Delta\Delta G_{dsolv}^{P^iN^j} = \left[\left(G_{n_{sol}}^{p^s} - G_{n_{vac}}^{P^s}\right) + \left(G_{n_{sol}}^{N^s} - G_{n_{vac}}^{N^s}\right)\right] - \left[\left(G_{u_{sol}}^{p^s} - G_{u_{vac}}^{P^s}\right) + \left(G_{u_{sol}}^{N^s} - G_{u_{vac}}^{N^s}\right)\right]$$

Equation 4

$G_{n_{sol}}^{p^s}$ is the reaction field energy of side chain atoms of P^i in native solvated condition

$G_{n_{vac}}^{p^s}$ is the reaction field energy of side chain atoms of P^i in native vacuum condition

$G_{n_{sol}}^{N^s}$ is the reaction field energy of side chain atoms of N^j in native solvated condition

$G_{n_{vac}}^{N^s}$ is the reaction field energy of side chain atoms of N^j in native vacuum condition

In unfolded state of protein (right part of right hand side equation) terms has similar meaning. In this calculation, PQR file possess ***m*** of (***i-1***), both ***m*** and ***s*** of ***i*** and ***m*** of (***i+1***) with atomic charges only on ***s***. In this case (***i-1***) and (***i+1)*** acts as local backbone for the residue (P^i). Similar consideration is given on other salt bridge partner (N^j) [5]. Calculations for both folded and unfolded states were performed at identical center of grid points.

Net salt bridge energy model

Net salt bridge energy was obtained by summing component energy terms[7] i.e.

$$\Delta\Delta G_{net} = \Delta\Delta G_{brd} + \Delta\Delta G_{prot} + \Delta\Delta G_{dsolv}$$ Equation 5

Residue specific salt bridges determination

Side chain of acidic (acceptor: D & E) and basic (donor: R, K, H) residues in folded protein form salt bridges. ADSBET adapts features of SBION[1] and SBION2[2] to obtain residue specific 1:1 salt bridges for computation of energy terms. Networked and multivalent salt bridge contribution was resolved based on their architecture during post run analysis.

ADSBET is fully automated single step procedure that incorporates above methods and models for computation of component energy terms. A detailed flowchart for the functioning of the procedure is presented in figure 1.

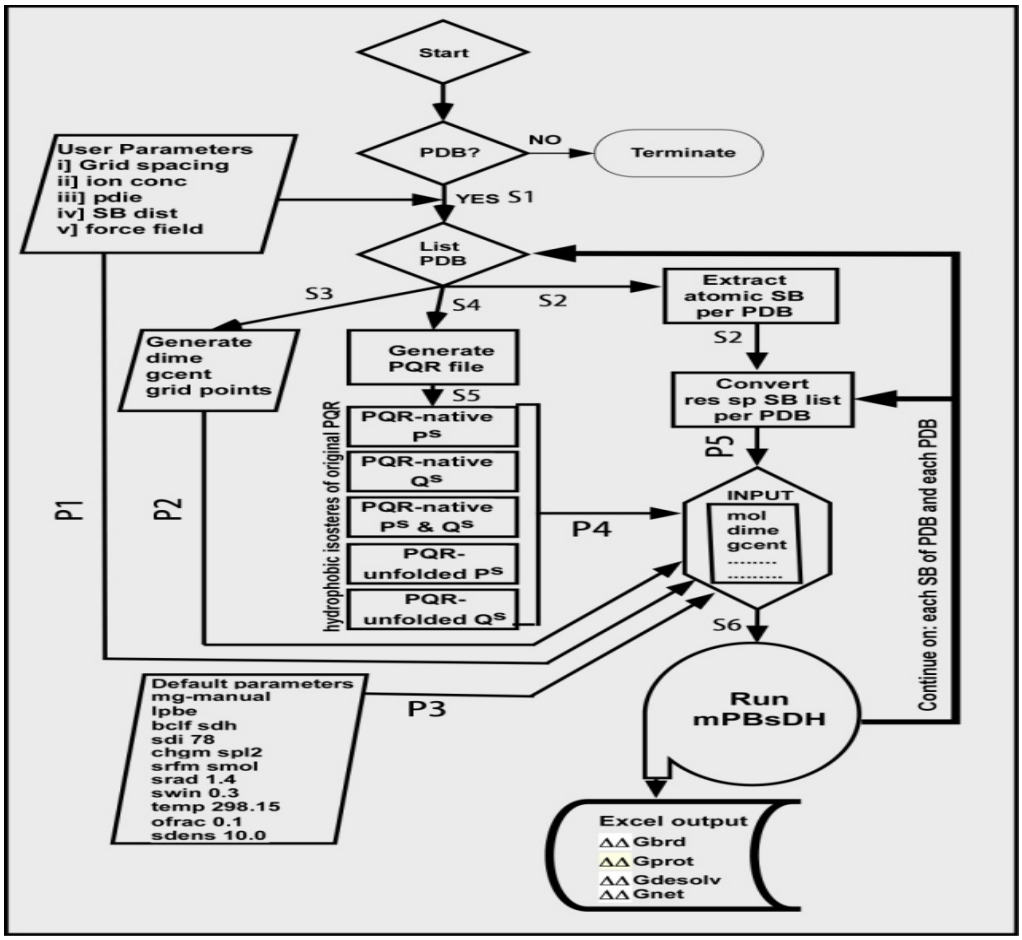

Figure 1 Flowchart shows details for manually configured multigrid Poison Boltzmann calculation (APBS) under single Debay-Huckel boundary condition (*mPBsDH*). *pdie* protein dielectric constant; *gcent* grid placement center. *SB dist* salt bridge distance.

Basically the program performs following sequential steps:

i. Makes list of all PDB files that are present in the working directory (Figure 1; S1),
ii. checks and corrects each PDB file for field and sequence gap errors,
iii. extracts atomic salt bridges and then convert these into residue specific list ones (Figure 1; S2),
iv. generates PDB structure specific APBS-input parameters such as dime, gcent etc (Figure 1; S3),
v. generates PQR from corrected PDB file using user supplied force field (Figure1; S4),
vi. uses the original PQR to generate 5 different hydrophobic isosteres mutated PQR files (Figure1; S5),
vii. prepares combined APBS-input per salt bridge that includes parameters P1, P2, P3, P4 and P5 (Figure 1),

viii. Runs APBS using the input (Figure 1; S6) and redirect energy terms in excel file. Parameters (P1, P2 and P3) and methods (equation 1 through 5) are also redirected for post run use,

ix. Repeats steps (iii) to (viii) for completion of calculation on all salt-bridges in a PDB file and then start with another PDB from the list (step (i)). The loop is continued until all salt bridges per structure and for all structure files are exhausted.

SYSTEM IMPLEMENTATION

The program is written using AWK[11] programming language along with the use of few standard UNIX commands. It runs smooth and fast in CYGWIN and other C-SHELL and B-SHELL UNIX environments. The working environment of ADSBET seems justified compare to a typical web-design as Poison Boltzmann Equation Solvers (e.g. APBS) is a time requiring process as it has to perform lengthy computation on structure even with a moderate grid point resolution[9].

Input

Poison Boltzmann Equation could be solved using various alternative parameters settings[8] of which the program adopt most accepted one[5-7]. ADSBET uses APBS[9] input parameters that could be classified into three different categories: i] default, ii] user setting and iii] auto-generated parameters. Default parameters are those (Figure1; P3) that are procured from earlier works[5-7]. Users setting ones (Figure2) are user parameters obtained from command prompt during the run of the program. These parameters are very sensitive for salt bridge energetic. Auto-generated parameters includes PDB list, PDB dimension related parameters (Figure1; S3), residue specific salt bridge list (Figure1; S2) and hydrophobic isosteres mutated PQR files (Figure1; S5).

```
user parameters

grid Spacing:
        lower value eg .33 =>more resolution but more run time
        higher the value eg 0.955 =>slightly lower resolution but much faster run
        adjustment to be done based on size of first unique chain of structure file
        0.8 a reccommended one: .9
              moderately high resolution... ok
mobile ion Conc in Molar [ eg .2]: .2
protein dielectric constant ie pdie [ eg 4]: 4
ion-pair distance <=4A for salt bridge: 4
force field for PDB2PQR [charmm/parse/tyl06/peoepb/swanson/amber]
        recommended -- parse or charmm : charmm
```

Figure 2 Grid spacing, Mobile ion concentration, protein dielectric constant and ion-pair distance are user parameters that are up taken during the run of ADSBET.

The force field is needed for the generation of PQR file using PDB2PQR[12]. Although in principle any of the 6 force fields could be used, in the context of salt bridge energetic either "parse" or "charmm" were used[5-7].

Output

The purpose of ADSBET is to obtain results on component and net energy terms of all salt bridges in a structure file and for all structure file. Table 1 show a part of the output (ITEM-III) of ADSBET run on 1CZP.pdb and 1DOI.pdb that were present in the working directory. The output is compact that chronologically arrange structure files and salt bridges along with their energy terms in Kcal Mol^{-1} unit. Always positive $\Delta\Delta G_{dsolv}$ term appears first followed by always negative $\Delta\Delta G_{brd}$ term**.**

Table 1 shows ADSBET generated output (ITEM-III) from crystal structures (1CZP.pdb, 1DOI.pdb) that were present in the working directory. Component and net energy terms per salt bridge per structure file are presented in compact chronological manner.

ITEM-III		Results:			
	ION-PAIR	**$\Delta\Delta G_{dsolv}$**	**$\Delta\Delta G_{brd}$**	**$\Delta\Delta G_{prot}$**	**$\Delta\Delta G_{net}$**
1CZP	**LYS93_GLU10**	4.75	-7.44	-1.75	-4.45
	ARG42_ASP28	7.01	-3.92	-8.73	-5.64
	ARG42_GLU31	9.44	-10.88	-2.03	-3.48
	LYS93_ASP96	5.57	-3.93	-9.89	-8.25
1DOI	**ARG126_GLU92**	10.87	-12.42	-5.36	-6.91
	LYS112_GLU4	5.46	-5.96	-5.84	-6.34
	ARG99_ASP79	5.13	-6.5	-0.24	-1.61
	ARG64_GLU53	9.94	-11.35	-0.5	-1.92

The term $\Delta\Delta G_{prot}$ which could either be positive or negative (in this case all are negative) appears after bridge term. The last term is the net salt bridge interaction term ($\Delta\Delta G_{brd}$) is obtained by the sum of these three terms. Residue specific salt bridge partners are showing their type followed by sequence position in protein chain. Shaded salt bridges in 1CZP.pdb forming a networked salt bridge (where ARG42 forming connection with both ASP28 and GLU31) with net interaction energy of -9.12 Kcal Mol^{-1}.

The output file also includes ITEM-I and ITEM-II information (https://sourceforge.net/projects/adsbet/files/all_sb_eng_table.xls/download). ITEM-I provide details on model of computation of energy terms and ITEM-II provides list of APBS parameters used for the current run. This part would be necessary for post run application of the results (Table1).

FUTURE DEVELOPMENT

Although ADSBET in its current form generate compact, user-friendly excel table for component as well as net energy terms for any number of salt bridge in any number of structure files, following developments of the program are intended: (1) When number of structures and salt bridges forming large dataset, automated sorting of networked salt bridges would be necessary. (2) Partitioning of $\mathbf{\Delta\Delta G_{prot}}$ term into different classes (such as acidic, basic etc) might reveals class specific microenvironments around salt bridges for their stabilizing and destabilizing effects. (3) Based on the inclusion of atoms of side chain, presence of permanent dipoles and isolated charges different model for calculation could be developed for computation of salt bridge energy. Presently we are engaged in developing web interface to integrate all the software tools[1, 2 and 13] of our laboratory such that their availability could reach to academic users within a unique web service.

CONCLUSION

We developed a fully automated procedure that takes PDB files as input and redirect compact, user friendly excel output for component as well as net energy terms of salt-bridges. It adopts most accepted models for computation of component energy terms. The program is efficient in that it works on any number of structure files with any number of salt bridges in them. Further, automated generation of residue specific salt bridges, structure specific parameters, hydrophobic isosteres mutated PQR files, collection of user parameters from command prompt, and assembling all these into a single input file prior to the run has been other attribute of the program. The output file thus obtained is useful for post run application. Moreover, identification of networked salt bridges and their net contribution could be worked out from the output. Overall, ADSBET is capable to extract detailed energetic of salt bridges employing large database whose outcome has great implication in computational structure biology.

AUTHORS' CONTRIBUTION

AKB conceive the project and prepare manuscript. AKB, AN, SB and PSSG developed the tools. AN, PSSG, SB, SD, PS, VPS and AKB participate in extensive testing of ADSBET for its system compatibility, data analysis and output design. All the authors have read and approved the final manuscript.

ACKNOWLEDGEMENTS

Authors are grateful for computational facility laboratory in Department of Biotechnology, The University of Burdwan. PSSG, SB thankfully acknowledge State funded JRF (Channeled through Higher Education Department, Govt. of West Bengal) by University of Burdwan.

COMPETING INTERESTS

The authors declare that no competing interests exist.

REFERENCES

1. Sen Gupta PS, Mondal S, Mondal B, Islam RNU, Banerjee S, Bandyopadhyay AK. SBION: A Program for Analyses of Salt-Bridges from Multiple Structure Files. Bioinformation 2014;10:164-166
2. Sen Gupta PS, Nayek A, Banerjee S, Seth P, Das S, Sur VP, Bandyopadhyay AK. SBION2: Analyses of Salt Bridges from Multiple Structure Files, Version 2. Bioinformation 2015;11:39-42
3. Bandyopadhyay AK, Krishnamurthy G, Sonawat HM. Structural stabilization of [2Fe-2S] ferredoxin from Halobacterium salinarum. Biochemistry 2001; 405:1284–1292.
4. Bandyopadhyay AK, Krishnamurthy G, Padhy LC, Sonawat HM. Kinetics of salt-dependent unfolding of [2Fe-2S] ferredoxin of Halobacterium salinarum. Extremophiles 2007;4:615–625
5. Hendsch ZS, Tidor B. Do salt-bridges stabilize proteins? A continuum electrostatic analysis. Protein Science 1994;3:211–226.
6. Nayek A, Sen Gupta PS, Banerjee S, Mondal B, Bandyopadhyay AK. Salt-Bridge Energetics in Halophilic Proteins. PLoS ONE 2014;9(4):e93862
7. Kumar S, Nussinov R. Salt-bridge stability in monomeric proteins. J Mol Biol 1999;293:1241–1255
8. Bosshard RH, Marti ND, Jelesarov I. Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. Mol. Recognition 2004;17: 1–16
9. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. PNAS USA 2001;98:10037-10041
10. Petukh M, Zhirinovsky M, Li C, Li L, Wang L, Alexov E. Predicting nonspecific ion binding using Delphi. Biophysical J. 2012;102:12
11. Aho AV, Kernighan BW, Weinberger PJ. Awk - A Pattern Scanning and Processing Language. Awk — A Pattern Scanning and Processing Language USD 1977;19-73.
12. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Research 2004;32: W665-W667.
13. Sen Gupta PS, Banerjee S, Islam RNU, Mondal S, Mondal B, Bandyopadhyay AK. PHYSICO: An UNIX based Standalone Procedure for Computation of Individual and Group Properties of Protein Sequences. Bioinformation 2014;10:105-107.