W372–W375 *Nucleic Acids Research, 2005, Vol. 33, Web Server issue*
doi:10.1093/nar/gki365

PCE: web tools to compute protein continuum electrostatics

Maria A. Miteva*, Pierre Tufféry[1] and Bruno O. Villoutreix

INSERM U648, University Paris 5, Paris 75006, France and [1]INSERM U726, EBGM, University Paris 7, France

Received February 4, 2005; Revised and Accepted February 24, 2005

ABSTRACT

PCE (protein continuum electrostatics) is an online service for protein electrostatic computations presently based on the MEAD (macroscopic electrostatics with atomic detail) package initially developed by D. Bashford [(2004) *Front Biosci.*, 9, 1082–1099]. This computer method uses a macroscopic electrostatic model for the calculation of protein electrostatic properties, such as pK_a values of titratable groups and electrostatic potentials. The MEAD package generates electrostatic energies via finite difference solution to the Poisson–Boltzmann equation. Users submit a PDB file and PCE returns potentials and pK_a values as well as color (static or animated) figures displaying electrostatic potentials mapped on the molecular surface. This service is intended to facilitate electrostatics analyses of proteins and thereby broaden the accessibility to continuum electrostatics to the biological community. PCE can be accessed at http://bioserv.rpbs.jussieu.fr/PCE.

INTRODUCTION

Electrostatic interactions play a central role in protein structure and function (1–3). However, in spite of their importance, it is still difficult to treat electrostatic properties theoretically. One reason is that Coulomb's law is valid for charges immersed in an infinite medium of a uniform dielectric (4). Yet, many reactions take place at the interface between regions having different dielectric values. Analytical solution of the Poisson–Boltzmann equation can resolve this problem only in several particular cases. Thus, the numerical solution of the Poisson–Boltzmann equation, like the Finite Difference Poisson–Boltzmann method (FDPB) (5), appears to be a good model for describing macromolecules in water solutions (6,7). Some of these methods are implemented at: http://agave.wustl.edu/pdb2pqr/ (8), http://agave.wustl.edu/apbs/ (9) and http://honiglab.cpmc.columbia.edu/ (10).

The basic electrostatic property that can be calculated via PCE (protein continuum electrostatics) is the electrostatic potential, yet continuum electrostatic models have also proven to be useful for the calculation of pK_a in proteins. We have implemented on the PCE server (http://bioserv.rpbs.jussieu.fr/PCE) the MEAD (macroscopic electrostatics with atomic detail) package based on the FDPB method for the computation of electrostatic potential values and pH titration behavior of titratable groups in proteins (11). In this approach, the protein is modeled as a low dielectric material (dielectric usually between 2.0 and 4.0) with embedded partial charges surrounded by a high dielectric medium (the solvent is regarded as a continuum with a dielectric value around 80.0), the atomic details of the protein structure are used to define the boundary and charge placement and the FDPB method is used to solve the resulting partial differential equations for the electrostatic potential.

METHODS AND IMPLEMENTATION

The PCE web service is driven by a modular, Python-written collection of routines, which provides non-interactive, high-throughput usage of the package MEAD. MEAD is based on the FDPB method to compute electrostatic potential values and pH titration behavior of titratable groups in proteins (11).

In the MEAD package, the electrostatic interactions are calculated with a continuum dielectric model where the protein and the water phase are represented as homogenous materials with a low (by default $\varepsilon_{in} = 4$) and high (by default $\varepsilon_{out} = 80$) dielectric constant ε, respectively. The dielectric boundary between protein and solvent is the molecular surface, which depends on the atomic radii and the radius of a solvent-sized spherical probe. The Poisson–Boltzmann equation is solved by means of a finite difference algorithm (5,11–13). The user can use the default values for ε or, for example, use $\varepsilon_{in} = \varepsilon_{out}$ (in this case the Poisson–Boltzmann equation will be resolved in a homogenous dielectric medium) or evaluate the effect of changing ε_{in} from 2 to 10 or 20 to simulate protein flexibility (14).

*To whom correspondence should be addressed at INSERM U648, Faculté des Sciences Pharmaceutiques et Biologiques, 4 avenue de l'Observatoire, 75006 Paris, France. Tel: +33 1 53 73 95 73; Fax: +33 1 53 73 95 74; Email: maria.miteva@univ-paris5.fr

Electrostatic potentials calculations

In our implementation, the electrostatic potentials are computed by the program 'potential' of MEAD starting from the coordinate file. Our implementation adds hydrogen atoms (15) and assigns the Parse parameters (radii and charges) (16). In addition, it is possible to upload PQR files instead of PDB files, such format parameterized files can be obtained, e.g. at http://agave.wustl.edu/pdb2pqr/ (8) (at this site, users can also perform hydrogen-bond network optimization). The finite difference computation uses a cubic box with 1 Å lattice spacing with number of grid points depending on the protein size and centered on the protein.

Calculations of pK_as of titratable groups

The overall methodology for calculating pK_a, defined as the pH for which the average protonation of the group is 1/2, has been described elsewhere (12). This approach implies the calculation of the intrinsic pK_a values of all titratable groups (defined as the pK_a of one titratable group if all other titratable groups in the protein were held in their neutral form) as well as the pair-wise electrostatic interactions between the titratable groups:

(i) The intrinsic pK_a is calculated by $pK_{int} = pK_{mod} - (\Delta\Delta G_{Born} + \Delta\Delta G_{back})/2.3RT$, where the p$K_{mod}$ is the pK_a of model compound of amino acid in water determined experimentally. In the MEAD implementation, the pK_{mod} values are as follows: Asp, 4.0; Glu, 4.4; His, 6.6 for the Nε2 or 7.0 for Nδ1; Tyr, 9.6; Lys, 10.4; Arg, 12.0; C-terminus, 4.0; and N-terminus, 7.5. The correction $\Delta\Delta G_{Born}$ takes into account the Born self-energy owing to the transfer of the amino acid group from water (high dielectric medium) to the protein (low dielectric medium). The correction $\Delta\Delta G_{back}$ is the background interaction term owing to the interaction of the titratable group with non-titratable (permanent) charges in the protein. In our implementation, all permanent partial charges on all non-titratable groups are taken into consideration.

(ii) The pair-wise charge interactions between the titratable groups can be represented as a matrix, W_{ij}, where i and j label the interacting groups.

(iii) As a result, the pK_a of one titratable group i is represented by $pK_{a_i} = pK_{int,i} + \frac{1}{2}\sum q_i q_j W_{ij}$, where q_i and q_j are the pH-dependent charges of the titratable groups i and j, respectively. All electrostatic energies (W_{ij}, $\Delta\Delta G_{Born}$ and $\Delta\Delta G_{back}$) can be calculated by solving the Poisson–Boltzmann equation for the electrostatic potential generated by the protonated and deprotonated forms of each titratable group placed in protein and in water. For these computations, our implementation applies the MEAD program 'multiflex' to solve the electrostatic potentials by the FDPB method using two successive lattices with increasing grid resolutions, 1 and 0.25 Å, respectively.

(iv) As seen in (ii), the charges q_i and q_j of the titratable groups at any given pH have to be additionally computed. Proteins can contain several tens to hundreds of titratable groups, thus leading to the problem of multiple site titration for the computation of pK_a values. The exact calculation of a multiple site titration curve, given the intrinsic pK_as and the pair-wise charge interactions, could be carried out at any given pH by a Boltzmann-weighted average over all of the possible protonation states of the protein (12). However, such calculations grow exponentially with the number of titratable groups (2^N combinations for N titratable groups). In the MEAD package, this problem is overcome by introducing the Reduced Site approximation method (17). This approach assumes that a titratable group can be considered as non-titratable at a pH far away from its pK_{mod}. In this case, this group can be assumed to be non-titratable for this pH, and the number of possible combinations to compute is reduced to 2^{N-1}. Thus, our service proposes at present pK_a calculations for <50 titratable groups. The computed titration curves, in our implementation, give a pH value where the protonation state of the titratable groups is 0.5, which is assigned to be the p$K_{1/2}$ and can be interpreted as the pK_a of this group.

IMPLEMENTATION

Potential and pK_a calculations are presented as two distinct services. Both of them start from a coordinate file in the PDB Format (18). Users can either specify a PDB identifier or upload their own file (with or without hydrogens). The protein internal dielectric, solvent dielectric as well as ionic strength values can be adjusted interactively.

For the electrostatic potential calculations, it is in addition possible to specify coordinates at which the value of the potentials will be returned. Finally, the electrostatic potential values can then be used to color-code the molecular surface of the selected protein via DINO (Visualizing Structural Biology, http://www.dino3d.org). Some options allow to adjust the color gradients for positive and negative potential values or to specify a site in the protein that is of importance such that this residue becomes positioned at the center of the picture. Returned images are either static (four orthogonal views are proposed) or animated (rotation around the y-axis) in the GIF.

For the pK_a calculations, it is presently only possible to activate Cys and Tyr as titratable, Asp, Glu, Lys, Arg and His being always considered as titratable. This is likely to evolve in future versions of the interface but was presently retained so as to lower the risk of getting improper results.

RESULTS

Electrostatic potentials

The users can download as results, color figures (in different formats, different orientations and resolutions together with animated GIF files) of 3D electrostatic potentials distribution on the protein surface. The users can select the energy range for the color-coding as well as decide on which region to focus the figure. In addition, the PQR file can be obtained. If the users specify a set of x, y, z coordinates, the server can return electrostatic potential values at these points.

As an example, if the users select lysozyme (PDB code 7lyz) as input file, they can obtain figures as follows (Figure 1). Such computations can be used to compare electrostatic potentials for wild-type proteins and mutants, in such cases the users need to upload their mutant PDB files.

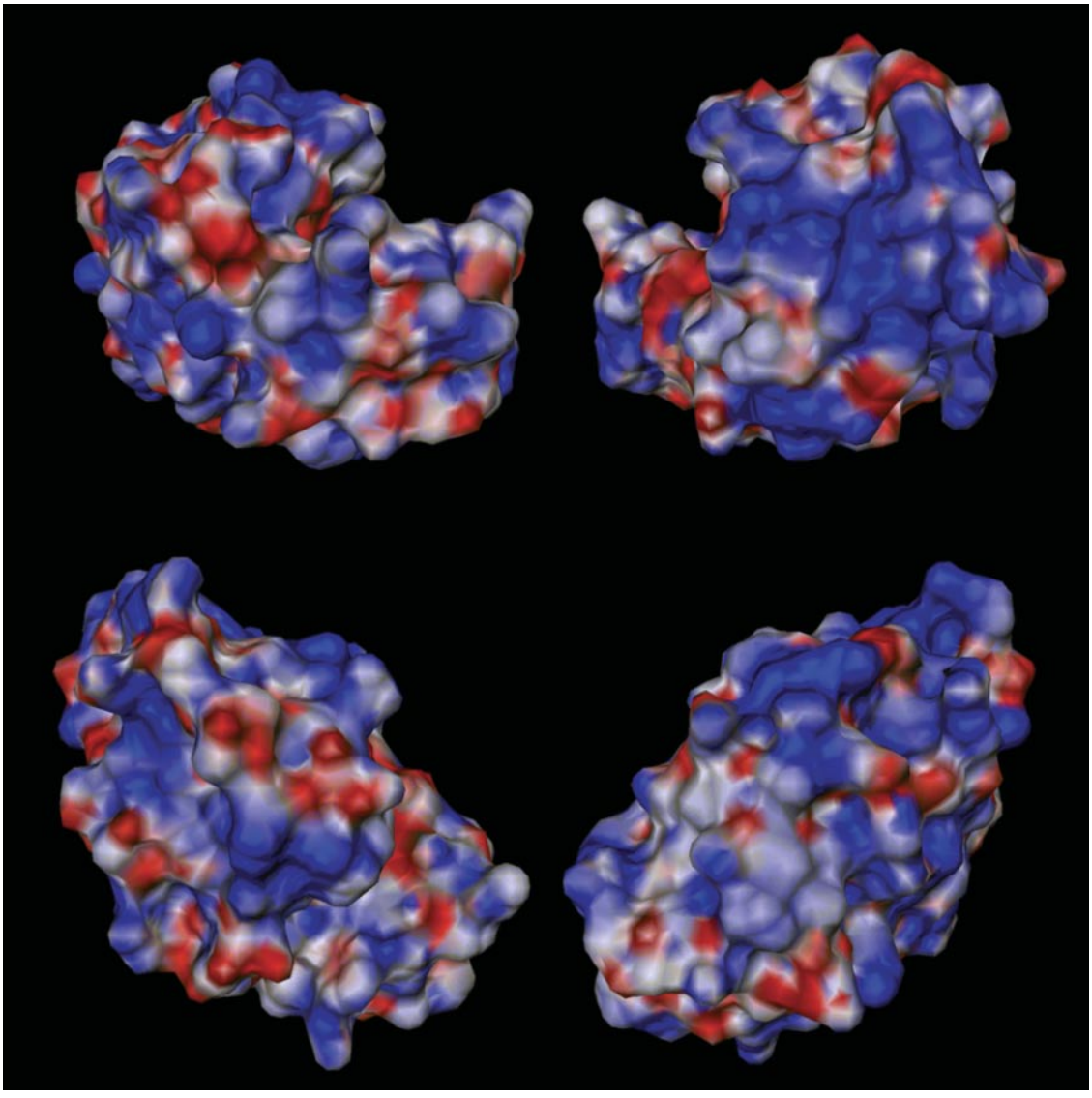

Figure 1. Electrostatic potential distributions on the lysozyme surface from −3.0 kcal/mol/e (red) to +3.0 kcal/mol/e (blue).

pK_as of titratable groups

The PCE server returns pK_{int}, p$K_{1/2}$, titration curves for all titratable groups and pair-wise charge interactions between titratable groups as well as the isoelectric point (pI). We have tested the implementation on several proteins taken from the PDB (18) and report the results for lysozyme (PDB entry 7lyz) and ribonuclease A (PDB entry 3rn3). Our computed pK_a values (ε_{in} = 4) are compared with pK_a values previously calculated or experimentally determined (Tables 1 and 2) (12,14). Small deviations between our data and the previously calculated p$K_{1/2}$ are owing to the differences in some parameters in the finite difference calculations. Overall, a good correspondence between our p$K_{1/2}$ and the experimental pK_a values can be seen, except for few titratable groups. This seems to be due, in part, to the fact that proteins should have considerable mobility for the solvent-exposed side chains (19). These fluctuations can be simulated by using a higher ε_{in} as proposed in (14) or by computing pK_a values on conformations selected from molecular dynamic trajectories (20). Other methods have been developed to account for protein flexibility during pK_a calculations (21–23).

CONCLUSION AND FUTURE DIRECTIONS

PCE is presently based on MEAD and allows users to compute protein electrostatic potentials via a FDPB solver starting with a PDB file as query input. The electrostatic potential values can then be used to color-code the molecular surface of the molecules presently generated within DINO. Static images in different formats and with different resolutions, centered on specific residues, as well as animated GIF files can be downloaded . The PARSE parameters are used for the computations, but we plan to implement other forcefield parameters in a near future. Currently, non-protein or nucleic acid atoms are excluded from the calculations, but work is in progress to circumvent this limitation. Additional optimizations are in progress as we plan to define automatically the charge assignments depending on the calculated pK_a values and

Table 1. pK_a calculation for lysozyme

Group	pK_{int}	p$K_{1/2}$	p$K_{1/2}$ (12)	p$K_{experimental}$ (12)
N-term	5.6	5.1	6.4	7.8–8.0
His-15	3.5	2.4	4.0	5.8
Glu-7	5.5	3.2	2.1	2.6
Glu-35	6.5	5.7	6.3	6.1
Asp-18	3.7	1.6	3.1	2.8–3.0
Asp-48	5.3	2.5	1.0	4.3
Asp-52	6.9	7.4	7.0	3.5–3.7
Asp-66	5.9	1.5	1.7	1.5–2.5
Asp-87	3.6	1.9	1.2	3.5–3.75
Asp-101	5.3	4.3	7.9	4.0–4.25
Asp-119	4.6	3.6	3.2	2.2–2.8
Tyr-20	12.5	12.7	14.0	10.3
Tyr-23	10.2	9.5	11.7	9.8
Tyr-53	12.9	>16	20.8	12.1
Lys-1	9.7	11.2	9.6	10.7–10.9
Lys-13	9.6	12.9	11.6	10.4–10.6
Lys-33	10.3	10.0	9.6	10.5–10.7
Lys-96	10.4	10.7	10.4	10.7–10.9
Lys-97	10.6	10.9	10.6	10.2–10.4
Lys-116	10.4	10.3	9.9	10.3–10.5
C-terminal	5.0	2.7	2.3	2.7–2.8

The p$K_{1/2}$ of a site is defined as the point in the calculated titration curve where the site is half-protonated. The pK_{int} is defined as the pK_a of one titratable group if all other titratable groups in the protein were held in their neutral form.

Table 2. pK_a calculation for ribonuclease

Group	pK_{int}	p$K_{1/2}$	p$K_{1/2}$ (14)	p$K_{experimental}$ (14)
N-terminal	7.5	6.0	6.7	7.6
His-12	6.8	3.0	−0.15	5.8/7.2
His-48	−2.0	<0.0	−4.6	6.3
His-105	5.3	6.0	4.3	6.6
His-119	6.7	7.7	9.1	6.1/7.6
Glu-2	7.5	0.6	2.0	2.8
Glu-9	5.1	4.7	4.8	4.0
Glu-49	5.0	5.7	6.6	4.7
Glu-86	4.6	3.0	4.9	4.1
Glu-111	4.6	3.8	4.5	3.5
Asp-14	5.0	3.3	7.5	<2
Asp-38	4.5	2.9	2.9	3.1
Asp-53	4.1	4.0	3.2	3.9
Asp-83	7.2	4.7	3.2	3.5
Asp-121	6.8	1.5	−0.65	3.1
C-terminal	4.2	1.7	1.9	2.4

The p$K_{1/2}$ of a site is defined as the point in the calculated titration curve where the site is half-protonated. The pK_{int} is defined as the pK_a of one titratable group if all other titratable groups in the protein were held in their neutral form.

then compute electrostatic potentials taking into account such information.

Users can also compute pK_{int}, p$K_{1/2}$, titration curves for all titratable groups and pair-wise charge interactions between titratable groups as well as isoelectric point of the protein. At present, no more than 50 titratable sites can be considered. However, we are working on the implementation of a Monte Carlo method to allow pK_a computations for larger proteins.

ACKNOWLEDGEMENTS

This work was supported by the Inserm Institute via a 'postevert' to M.M. and the Avenir grant to B.V. RPBS was supported by 'Action Conjointe CNRS-INSERM, Bioinformatique 2002'. Funding to pay the Open Access publication charges for this article was provided by INSERM.

Conflict of interest statement. None declared.

REFERENCES

1. Honig,B. and Nicholls,A. (1995) Classical electrostatics in biology and chemistry. *Science*, **268**, 1144–1149.
2. Warshel,A. and Aqvist,J. (1991) Electrostatic energy and macromolecular function. *Annu. Rev. Biophys. Biophys. Chem.*, **20**, 267–298.
3. Bashford,D. (2004) Macroscopic electrostatic models for protonation states in proteins. *Front Biosci.*, **9**, 1082–1099.
4. Sharp,K.A. and Honig,B. (1990) Electrostatic interactions in macromolecules: theory and applications. *Annu. Rev. Biophys. Biophys. Chem.*, **19**, 301–332.
5. Warwicker,J. and Watson,H.C. (1982) Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. *J. Mol. Biol.*, **157**, 671–679.
6. Koumanov,A., Ruterjans,H. and Karshikoff,A. (2002) Continuum electrostatic analysis of irregular ionization and proton allocation in proteins. *Proteins*, **46**, 85–96.
7. Rocchia,W., Sridharan,S., Nicholls,A., Alexov,E., Chiabrera,A. and Honig,B. (2002) Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. *J. Comput. Chem.*, **23**, 128–137.
8. Dolinsky,T.J., Nielsen,J.E., McCammon,J.A. and Baker,N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res.*, **32**, W665–W667.
9. Baker,N.A., Sept,D., Joseph,S., Holst,M.J. and McCammon,J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
10. Nayal,M., Hitz,B.C. and Honig,B. (1999) GRASS: a server for the graphical representation and analysis of structures. *Protein Sci.*, **8**, 676–679.
11. Bashford,D. and Gerwert,K. (1992) Electrostatic calculations of the pK_a values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.*, **224**, 473–486.
12. Bashford,D. and Karplus,M. (1990) pK_a's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry*, **29**, 10219–10225.
13. Gilson,M.K. and Honig,B. (1988) Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins*, **4**, 7–18.
14. Antosiewicz,J., McCammon,J.A. and Gilson,M.K. (1996) The determinants of pK_as in proteins. *Biochemistry*, **35**, 7819–7833.
15. Gautier,R., Camproux,A.C. and Tuffery,P. (2004) SCit: web tools for protein side chain conformation analysis. *Nucleic Acids Res.*, **32**, W508–W511.
16. Sitkoff,D., Sharp,K.A. and Honig,B. (1994) Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.*, **98**, 1978–1988.
17. Bashford,D. and Karplus,M. (1991) Multiple-site titration curves of proteins. *J. Phys. Chem.*, **95**, 9556–9561.
18. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The protein data bank. *Nucleic Acids Res.*, **28**, 235–242.
19. Smith,L.J., Sutcliffe,M.J., Redfield,C. and Dobson,C.M. (1991) Analysis of phi and chi 1 torsion angles for hen lysozyme in solution from ^{1}H NMR spin–spin coupling constants. *Biochemistry*, **30**, 986–996.
20. Miteva,M.A., Brugge,J.M., Rosing,J., Nicolaes,G.A. and Villoutreix,B.O. (2004) Theoretical and experimental study of the D2194G mutation in the C2 domain of coagulation factor V. *Biophys. J.*, **86**, 488–498.
21. You,T.J. and Bashford,D. (1995) Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. *Biophys. J.*, **69**, 1721–1733.
22. Georgescu,R.E., Alexov,E. and Gunner,M.R. (2002) Combining conformational flexibility and continuum electrostatics for calculating p*K*(a)s in proteins. *Biophys. J.*, **83**, 1731–1748.
23. Alexov,E. (2003) Role of the protein side-chain fluctuations on the strength of pair-wise electrostatic interactions: comparing experimental with computed p*K*(a)s. *Proteins*, **50**, 94–103.