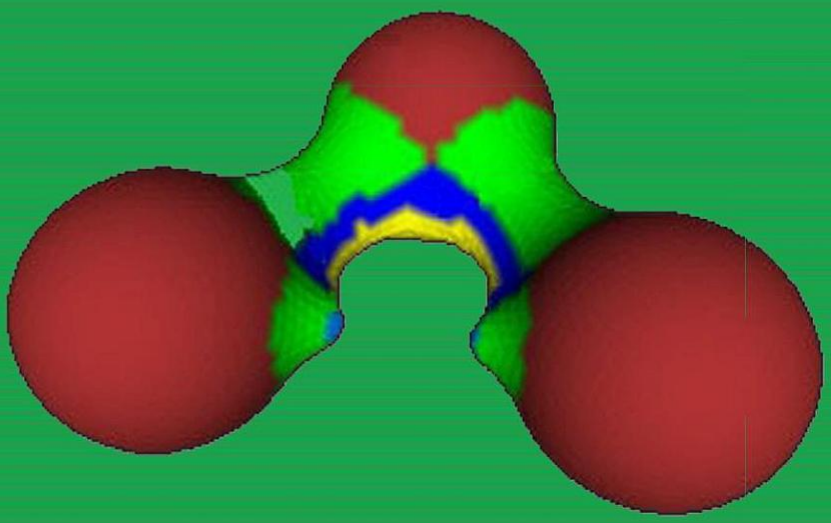
KUPERVASSER OLEG YURJEVICH
WANNER NATALIA EDUARDOVNA

Continuum Solution Models for Computer Aided Drug Design

The second edition, Colour

(The first edition: Aracne, Italy)



Attention!

This book can be sold here:

http://www.amazon.com/Continuum-Solution-Models-Computer-Design-ebook/dp/B014AMK7Y2/ref=sr_1_2?ie=UTF8&qid=1448289353&sr=8-2&keywords=kupervasser

If you buy the book, please, write customer review. To do it, please, push yellow button “Write a customer review” in the end of the Internet page.

Continuum Solution Models for Computer Aided Drug Design

The second edition, Colour (The first edition: Aracne, Italy)

KUPERVASSER OLEG YURJEVICH

WANNER NATALIA EDUARDOVNA

ISBN: 1515151700

ISBN-13: 978-1515151708

CONTENTS

ACKNOWLEDGMENTS

We would like to thank A.Vardy, A. N. Romanov, M.V. Basilevsky, F. V. Grigoriev, S. N. Zhabin, Ya. B. Martynov, K. M. Fedulov, I. V. Oferkin, A. V. Sulimov, and V. B. Sulimov for their fruitful job and ideas, which were very useful for creating the book.

1 CHAPTER INTRODUCTION

The main paradigm used in the modern development of new drugs consists in the following. Many diseases are associated with the functioning of certain proteins, so they must be blocked to cure diseases. For example, a target protein may belong to a virus, and its blockage disables the reproduction of a virus in an organism. Blockage is performed with the use of molecules that selectively bind to these proteins in an organism. Such molecules that serve as a basis for new drugs are called inhibitors. As a rule, inhibitors represent comparatively small organic molecules that bind to certain areas of target proteins. These areas are called binding sites or active sites. The search for such inhibitor molecules for a given target protein is the initial stage in the development of a new drug. The fast and efficient solution of this problem governs to a considerable degree the minimization of material expenditures and the duration of subsequent stages in the development of a new drug. With respect to time, the stage of developing new inhibitors takes nearly 50% of the total duration of the development of a new drug.

The time and material expenditures at the stage of searching for inhibitors can considerably be reduced with the use of computer-aided molecular modeling methods [1], among which docking is of first importance. Docking is the positioning of molecules that are candidates to inhibitors (they are often called ligands, from Latin *ligare* that means *to bind*) at the active site of a target protein and the estimation of their binding energy. The stronger a molecule

binds to a protein, the better is an inhibitor and the more efficient is the drug based on this inhibitor. Docking is performed with the special molecular modeling software [2] that is also used on supercomputers.

The precision of estimating the protein–inhibitor binding energy governs the efficiency of predicting the activity of an inhibitor: the higher is the binding energy, the more active is an inhibitor and the more efficient is the drug based on this inhibitor, as the required effect can be attained at a lower drug concentration. If the precision of calculating the protein–inhibitor binding energy is insufficiently high, the probability of predicting whether new synthesized compounds (ligands) will inhibit this protein is low, and a great deal of materials, which are used to synthesize ligands and measure their inhibiting activity, and corresponding time are wasted. A sufficiently high practical predictability is attained at an error in calculating the protein–inhibitor binding energy of lower than 1 kcal/mol. For this reason, the precision of calculating all the contributions to the protein–inhibitor binding energy must be maximally high in the molecular modeling of the interaction of ligands with target proteins.

The precision of estimating the protein–inhibitor binding energy in molecular modeling is governed by many factors, e.g., the quality of a force field used for the description of intra- and intermolecular interactions, the efficiency of searching for a global minimum in the course of positioning an inhibitor in the active site of a target protein, the estimated contribution of the entropy component to the free protein–inhibitor binding energy, etc. Since the binding of an inhibitor to a protein in experiments (*in vitro* and *in vivo*) occurs in an aqueous solution, the presence of a solvent (water) must be taken into account, when calculating the protein–inhibitor binding energy.

The effect of a solvent on the protein–inhibitor binding energy is predominantly determined via the desolvation energy, which represents the difference between the solvation energy of a protein–ligand complex and the solvation energies of individual protein and ligand. This contribution to the protein–ligand binding energy is caused by that a solvent (water) is forced out of the active site of a protein upon the binding of a ligand to a protein, and some atoms in a ligand and a protein's active site cease to interact with a solvent. Hence, to determine the desolvation energy, it is necessary to calculate the solvation energies of a protein, a ligand, and their complex.

To calculate the free solvation energy, it is necessary to construct a solvent model. This may be done explicitly, considering a solvent as a set of a great number of molecules. However, this method needs comparatively high expenditures of computational resources in modeling, as the calculation of the observed effects requires us to perform the averaging over the state of solvent

molecules, for example, by the molecular dynamics or Monte-Carlo methods. For this reason, implicit (often called continual) solvent models [3–14], in which a solvent is treated as a continuous (continual) medium with specified properties, including dielectric permittivity, are used more frequently.

The present work is devoted to the DISOLV software [9,14-16] that allows us to find the free solvation energy of molecules and its gradients over the displacements of atoms in molecules, as the DISOLV software will be used to optimize the structure of molecules in a solvent with both the force field methods and the quantum-chemical methods, in which the gradient-based local optimization algorithms are usually applied.

The free Gibbs energy ΔG_s for the process of solvation, i.e., the transition of a molecule fr om a vacuum into a solvent or, briefly, the solvation energy, is represented as the sum of the three components

$$\Delta G_s = \Delta G_{pol} + \Delta G_{np} + \Delta G_{cav},$$

where ΔG_{pol} is the polar component of the interaction of a dissolved substrate molecules with a solvent, ΔG_{np} is the non-polar component of the interaction of a dissolved molecule with a solvent due to van der Waals forces of intermolecular interaction, and ΔG_{cav} is the cavitation component of the free solvation energy due to the formation of a cavity comprising a dissolved molecule in the volume of a solvent.

In the DISOLV software, much attention is concentrated on the calculation of the polar component ΔG_{pol} of the interaction of a molecule with a solvent (or, briefly, the polarization energy) using several methods, and the other solvation energy components ΔG_{np} and ΔG_{cav} are taken into account by a simple and rather widely applied method (see below).

Within the framework of the used continual model, the solvation energy represents the energy of the electrostatic interaction of atomic charges in a molecule located in the cavity of a dielectric with surface charges induced by them on this surface.

By now, many existing implicit solvent models and their software implementations have been integrated into larger packages, e.g., the quantum-

chemical packages Gaussian [17], Gamess [18], MolPro [19], MOPAC [20], the molecular dynamics package Charmm [21], and the free programs for finding the polar component of the interaction with a solvent, e.g., DelPhi [22] or APBS [23], implement the numerical solution of the finite-difference approximation of the three-dimensional Poisson–Boltzmann equation.

The objective of the present book is to describe the original algorithms that are designed to calculate the polar component of the solvation energy of molecules and aimed at solving the corresponding equations on the two-dimensional surface of a solvent surrounding a molecule, their software implementation DISOLV written in C++, and its corresponding validation. At the end of the book, we perform the brief comparison of the characteristics of DISOLV with the corresponding characteristics of APBS [23] that implements the numerical solution of the Poisson–Boltzmann equation in a 3D space.

2 CHAPTER AN ALGORITHM FOR THE FORMATION OF A SMOOTH MOLECULAR SURFACE

2.1 Introduction

The objective of this chapter is to give the full and exhaustive description of an algorithm that allows us to form an optimally smooth molecular surface via primary and secondary rolling for its further use in:

(a) molecular editors for demonstration purposes

(b) the calculations of the solvation energy of a molecule (the difference between its free energies in solution and vacuum) and the analytical gradients of this energy.

We shall realize this smoothness via the primary rolling of a molecule with spheres, whose radius is equal to the size of a solvent molecule (for the outer surface of a molecule), and the secondary rolling of a molecule with a variable-radius sphere (for the inner surface of a molecule) with the elimination of all remaining irregularities.

We should emphasize the special importance of obtaining a smooth surface for these purposes. Really, in the case of a molecular editor, the smoothness of a surface is necessary for its triangulation and further comprehensive representation, which will not contain any physically meaningless irregularities dissipating our attention.

In the case of calculating the electrostatic component of the solvation energy and its analytical derivatives, the smoothness of a surface is a necessary algorithmic stability condition, as fictitious superficial irregularities accumulate fictitious charges, thus leading to instability in the operation of an algorithm and great numerical errors [24-26].

For this reason, the formation of a smooth surface is not only an interesting mathematical problem, but is also practically important. The primary and/or subsequent secondary rolling algorithm described in this chapter was used as the base for developing the following software: PQMS [27], MSMS [28], Totrov and Abagyan's program [29], SIMS [30], TAGSS [31–33] and its improved version implemented as a subroutine of DISOLV [9,15,16]. These software and algorithms served as the base for creating the successfully operating molecular editor [33] and licensing the program for the calculation of the solvation energy and its derivatives [9,15,16].

What is the originality of our work in comparison with the other studies [27-32,34-37] that analyze the primary and/or secondary rolling of a molecular surface? First, it consists in completeness: we have considered all the possible cases of irregularity and showed the possibility of smoothing for them. Second, we try to smooth them optimally. This means that we try not only to obtain smoothness, but also to avoid to an ultimately possible degree the appearance of surface domains that, although rather smooth, are nevertheless very similar to irregularities (i.e., to avoid the appearance of very narrow "channels" and surface domains with a very small curvature radius). For all we know, this problem has been solved fully and exhaustively only in our work.

However, we should note that the problem of constructing a smooth surface without any additional constraints may be solved trivially via the simple circumscription of a sphere around a molecule. What are these constraints?

First, the atomic radii are well-defined parameters, which may not be arbitrarily varied. When the given algorithm is used to image a molecule in a molecular editor, the radii of its atoms are determined by the sizes of electron shells (so-called van der Waals radii [38,39], different sets of which are available in the literature [38-40]). When they are used to calculate the solvation energy, the initial ("rough") radii are specified as for the first case, but then refined so

that the calculated solvation energies also correspond to their experimental values.

Second, if a molecule is submerged into a solution, the radius of solvent molecules is also a well-defined value, which may not be arbitrarily varied. This radius undoubtedly has some effect on the interface between a solution and a molecule itself and selected from the same above described reasoning as for the radius of atoms in a molecule.

Let us note that the above mentioned radii rather strongly restrict the possibility of varying the surface of a molecule. But they nevertheless hold enough opportunities for us to make it smooth.

Moreover, we wish to make it not simply smooth, but optimally smooth. In the given case, optimality is understood to mean that we (a) not only search for a smooth surface, but also try to reduce its curvature without losing surface features and, in addition, (b) strive to decrease the Cartesian distance between *non-adjacent surface domains*. These domains are close to each other spatially, but remote in the case of measuring the distance between them along the surface. To estimate the degree of smoothness, we specify *the maximum critical distance*, at which non-adjacent surface domains are allowed to approach each other in the algorithm. If this approach distance is smaller than the critical distance, there occurs a similar-to-irregularity situation, which requires secondary rolling.

There exist the two types of surfaces surrounding a molecule [41]:

(1) SAS, a solvent accessible surface, is formed by the centers of solvent molecules tangent to a substrate molecule. The number of solvent molecules tangent to the surface of a molecule is proportional to the SAS area.

(2) SES, a solvent excluded surface. The volume occupied by a solvent lies *outside* the volume enveloped by this surface. The substrate itself lies completely *inside* this volume.

SAS can be obtained by rolling a substrate molecule with a solvent molecule and marking the positions of its center. Rolling is the displacement of a solvent molecule along the surface of a substrate and its sequential contact with all the accessible points of this substrate (Fig. 2.1). For simplicity, a solvent molecule

may be replaced with a rolling sphere (a sphere circumscribed around a solvent molecule) [42-43].

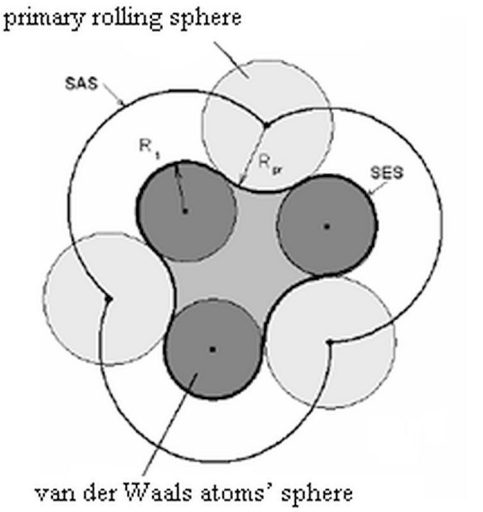


Fig.2.1 Primary rolling of a surface (see [9]).

A molecular SES may be described as follows [3,44]:

(1) Smoothedly, replacing it with simple structures like [45-52]

   (a) A sphere,
   (b) An ellipsoid, or
   (c) A cylinder;

(2) In details, reproducing all the inflections on the surface of a molecule

   (a) By coating atoms with van der Waals spheres;

(b) By coating chemical groups of atom with spheres;

(c) As described in the two previous methods, but filling the remaining empty space inside SES with fictitious spheres (as implemented in the GEPOL software) [34-37]; and

(d) With the molecular electron density level surface [53] determined via quantum mechanics or the other types of functions employed for the construction of level surfaces [54-58]. This method encounters serious difficulties: a level surface may also be non-smooth; its implicit definition complicates triangulation; it is difficult to fit functions giving us a surface that is close to real and determined by the van der Waals radii of atoms; and

(e) By bridging the spheres described in (a) and (b) with convex and concave surface elements [29,30].

The smoothest and most realistic surface can be obtained by method 2e considered in our work. This method allows us to obtain SES in the same manner as SAS by rolling the outer surface of a molecule with a sphere and taking:

(1) the positions of points of contact between a rolling sphere and atoms;

(2) the rolling sphere's geodesic arc segments that pass through two points of contact between a rolling sphere and atoms;

(3) the segments of the lower part of a rolling sphere between its geodesic arc segments that pass through two points of contact between a rolling sphere and atoms (*primary rolling*, Fig. 2.1).

Such a technique of determining SES was first proposed in [59]. The proposed method was further developed in [29, 60-64]. However, thus obtained SES may prove to be non-smooth [28]. For its further smoothing, the inner surface of a molecule may be rolled again (*secondary rolling*, Figs. 2.2 and 2.3), as was originally proposed in [30].

This may raise the question: why are we sure that secondary rolling will successfully smooth all irregularities and no tertiary, quaternary, and so forth rolling will be required? There is the following argument for this: we have no

possibility of varying the radii of atoms and primary rolling spheres. They are constant predetermined values. At the same time, the secondary rolling radius is an arbitrarily selected value and may be selected so as to be various even for the rolling of different domains of the same surface.

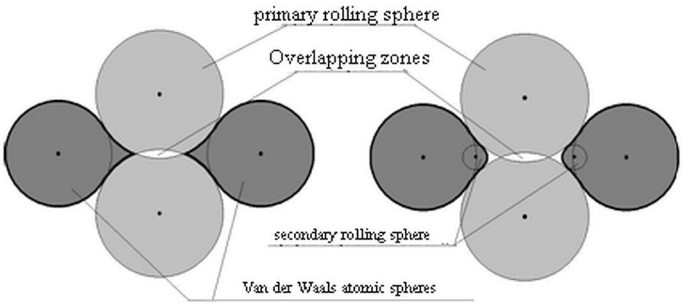


Fig. 2.2 Secondary rolling of a surface (see [9]).

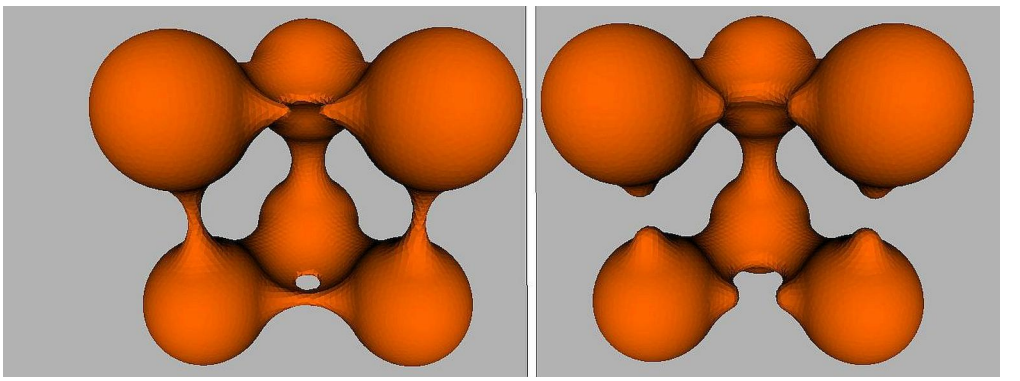


Fig. 2.3 Applying the method of secondary rolling to the geometric configuration of several atoms.

It is intuitively clear that we can "smooth" anything, selecting the secondary rolling radius as equal to an infinitely small value. Really, there are no problems for the smooth rolling of any irregularity within an infinitely small radius. However, we strive to obtain a surface that is not merely smooth, but optimally smooth. In other words, it must not contain any elements that, although rather

smooth, are nevertheless similar to irregularities. For this reason, we should try to perform secondary rolling with a maximum possible radius, taking into account the geometric hindrances of a surface. This makes the algorithm of secondary rolling nontrivial enough.

A rolling surface is composed by surface segments of the two types: spherical and toroidal. They are divided into fragments of the five types: spherical elements of van der Waals atomic spheres, concave spherical elements of primary rolling, toroidal elements of primary and secondary rolling, and convex spherical elements of secondary rolling (Fig. 2.4).

The primary and secondary rolling radii and the critical distance have a clear physical meaning. The primary rolling radius is equal to the radius of a sphere circumscribed around a solvent molecule. The secondary rolling radius and the above defined critical distance determine the minimum curvature of a molecular boundary. The lower limits for the secondary rolling radius and the above defined maximum critical distance are associated with the "smearing" of electron charge clouds, which can not give very acute angles and narrow necks and channels that may appear after primary rolling owing to the Heisenberg's uncertainty relation between the coordinate and impulse of an electron.

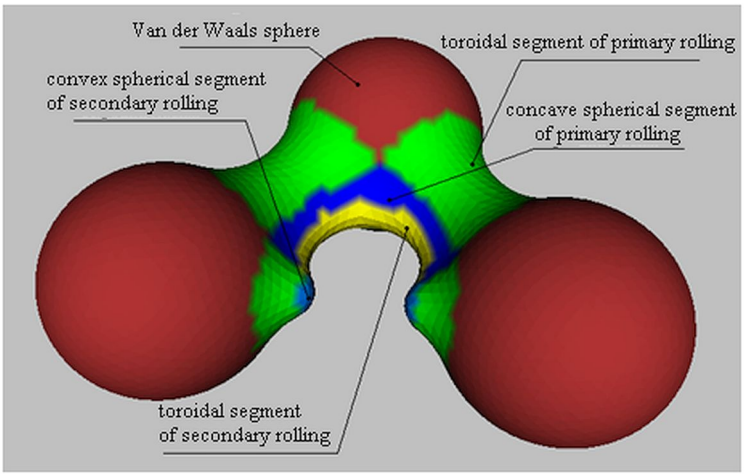


Fig. 2.4 Molecular surface composed of the five types of fragments: spherical elements of van der Waals atomic spheres, concave spherical elements of primary rolling, toroidal fragments of primary and secondary rolling, and convex spherical elements of secondary rolling.

10 CHAPTER CONCLUSIONS

In the present work, we considered one of the aspects of improving the precision of the computer-aided prediction of the inhibiting activity of molecules that are candidates to inhibitors for specified target proteins, namely, the allowance for the effect of a solvent on the protein–ligand binding energy. The physical foundations of the four implicit (continual) solvent models implemented in the DISOLV software, such as PCM, COSMO, SGB, and PCM with enlarged elements, were described in details, and the methods of solving the corresponding equations and the foundations of the algorithm used to construct the surfaces in these models and calculate the gradients of the energy of a molecule in a solvent were discussed. The latter gradients are needed in the DISOLV software for the local optimization of the energy of a molecule in a solution. The DISOLV validation results that showed not only the possibility of attaining a good precision of calculations (at arbitrary shifts of a triangulation grid) of better that several tenths of kilocalories per mole over reasonable time periods for such large macromolecules as proteins, but also a good agreement (root-mean-square deviation, 0.8 kcal/mol) of the calculated Gibbs energies of dissolution of a molecule in water, i.e., the energy of transfer of a molecule from a vacuum into water, with their experimental values for several hundreds of neutral molecules were represented. For molecular ions, the root-mean-square deviation between the calculated and experimental solvation energies is considerably higher (10 kcal/mol), but this value also lies within the limits of measurement error in most cases. On the whole, the validation results show that

the DISOLV software can be used in the post-processing regime to refine the protein–ligand binding energy estimate given by the docking program. Recently, the new method "Multicharge Approximation" [121] was suggested for free solvation energy calculation. The authors of [121] demonstrated that "Multicharge Approximation" has the same precision as DISOLV, but is much faster. However, there was considered the version of DISOLV with non-adaptive algorithm for the triangulation of a smooth molecular surface in [121]. Using the adaptive algorithm for the triangulation of a smooth molecular surface, considered in Chapter 2 of this book, allows overcoming this disadvantage.

REFERENCES

1) Sadovnichii, V.A. and Sulimov, V.B., Supercomputer Technologies in Medicine, in Supercomp'yuternye tekhnologii v nauke, obrazovanii, i promyshlennosti (Supercomputer Technologies in Science, Education, and Industry), Sadovnichii, V.A., Savin, G.I., and Voevodin, V.V., Eds., Moscow: Mosk. Gos. Univ., 2009, pp. 16–23.

2) Romanov, A.N., Kondakova, O.A., Grigoriev, F.V., Sulimov, A.V., Lushchekina, S.V., Martynov, Ya.B, and Sulimov V.B., Computer-Aided Design of Drugs: the SOL Docking Software, Vychisl. Metods & Progr., 2008, vol. 9, no. 2, pp. 64–84.

3) Tomasi, J. and Persico, M., Molecular Interactions in Solution: an Overview of Methods Based on Continuous Distributions of the Solvent, Chem. Rev., 1994, vol. 94, no. 7, pp. 2027–2094.

4) Cramer, C. and Truhlar, D., Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics, Chem. Rev., 1999, vol. 99, no. 8, pp. 2161–2200.

5) Onufriev, A., Continuum Electrostatics Solvent Modeling with the Generalized Born Model, in Modeling Solvent Environments. Applications to Simulations of Biomolecules, Weinheim: Wiley-VCH Verlag, 2010, pp. 127–166.

6) Basilevsky, M.V., Grigoriev, F.V., Kupervasser, O.Y. , Specific features of the dielectric continuum solvation model with a position-dependent permittivity function, The Journal of Physical Chemistry B, 2010, 114 (49), 16427-16435

7) Grigoriev, F.V., Kupervasser, O.Y. , Kikot', I. P. , An Iterative Method for Calculating the Polar Component of the Molecular Solvation Gibbs Energy Under a Smooth Change in the Dielectric Permittivity of a Solution, Russian Journal of Physical Chemistry B, 2011, Vol. 5, No. 6, pp. 1038-1044; Khimicheskaya Fizika, 2011, Vol. 30, No. 12, pp. 76-82

8) Kupervasser, O.Y. , Kikot', I. P. , Enlarged Surface Meshes and Normalization Conditions for Columns and Rows of Matrices in the COSMO Method, Russian Journal of Physical Chemistry B, 2011, Vol. 5, No. 5, pp. 873-879; Khimicheskaya Fizika, 2011, Vol. 30, No. 10, pp. 86-93

9) Kupervasser, O.Y., Zhabin, S.N., Martynov, Ya.B., Fedulov, K.M., Oferkin, I.F., Sulimov, A.V., Sulimov, V.B., A continuum solvent model: the DISOLV program - algorithms, implementation, and validation, Numerical Methods and Programming, 2011, V.12, pp.246-261

10) Kupervasser, O.Y., Wanner, N. E., Continual Model of Medium I: Algorithm for Formation of Smooth Molecular Surface, Science and business: development ways, N 11(17), 2012.

11) Kupervasser, O.Y. , Wanner, N. E., Continual Model of Medium II: Universal adaptive algorithm for triangulation of smooth molecular surface, Global Scientific Potential, N 11(20), 2012.

12) Kupervasser, O.Y., Wanner, N. E., Continual Model of Medium III: Calculation of analytical gradients of parameters of surface meshes on the molecular surfaces over atomic coordinates, Science Prospects, N 11(38), 2012.

13) Kupervasser, O.Y., Wanner, N. E., Continual Model of Medium IV: Calculation of analytical gradients for solvation energy over atomic coordinates, Science Prospects, N 12(39), 2012.

14) Kupervasser O.Yu., Gabin S.N., Sulimov V.B., Computer program DISOLV, The certificate on the state registration of the computer programs № 2010612994, it is registered in the register of the computer programs of Federal service on intellectual property, patents and trademarks, Russia, on May, 6th, 2010

15) Kupervasser, O.Yu., Zhabin, S.N., and Sulimov, V.B., State Registration Certificate for the DISOLV Software no. 2010612994, May 6, 2012.

16) Kupervaser, O.Yu., Zhabin, S.N., and Sulimov, V.B., Continual Models of Solvent. Calculating the Free Solvation Energy. Trudy II Vseross. konf. "Mnogomasshtabnoe modelirovanie protsessov i struktur v nanotekhnologiyakh" (Proc. II All-Russia Conf. "Multiscale Modeling of Processes and Structures in Nanotechnologies"), Moscow, 2009, p. 230.

17) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, K. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H.Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O.Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gompers, R. E. Sratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, A. D. Daniels, O. Farkas, A. D. Rabuck, K. Raghavachari, and J. V. Ortiz, Gaussian 03, Rev. B.03 (Gaussian Inc., Pittsburgh, PA, 2003).

18) General Atomic and Molecular Electronic Structure System GAMESS User's Guide, Ames, IA: Iowa State Univ., 2010 http://www.msg.chem.iastate.edu/gamess/documentation.html

19) Werner, H.J. and Knowles, P.J. MolPro Users Manual. Version 2010.1, Univ. College Cardiff Consult. Ltd, 2012 http://www.molpro.net/info/current/doc/manual/node1.html

20) Stewart, J.P., MOPAC Manual, 1993 nova.colombo58.unimi.it/manual/pdf/Mopac7.pdf

21) Im, W., Beglov, D., and Roux, B., Continuum Solvation Model: Computation of Electrostatic Forces from Numerical Solutions to the Poisson–Boltzmann Equation, Comp. Phys. Commun., 1998, vol. 111, nos. 1–3, pp. 59–75.

22) Rocchia, W., Sridharan, S., Nicholls, A., Alexov, E., Chiabrera, A., and Honig, B., 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. Comp. Chem., 2002, vol. 23, no. 1, pp. 128–137.

23) Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A., Electrostatics of Nanosystems: Application to Microtubules and the Ribosome, Proc. Natl. Acad. Sci. USA, 2001, vol. 98, no. 18, pp. 10037–10041.

24) Yu S., Geng W., Wei G.W. //J Chem Phys. 2007. V.126 P.244108

25) Yu S, Geng W, Wei G. W. //J Chem Phys. 2007. V.127 P.114106

26) Hou S. , Wang W. , Wang L. // Journal of Computational Physics. 2010. V.229. P.7162

27) Connolly M.L.//J. Mol. Graphics. 1993. V. 11. P.139

28) Sanner M. F., Olson A. J.,Spehner J. C.// Biopolymers. 1996 V.38, P.305-

29) Totrov M., Abagyan R. //Journal of structural Biology. 1996. V.116. P.138

30) Vorobjev Y. N.,Hermans J. //Biophysical Journal. 1997. V. 73 P.722

31) Zhabin S.N., Sulimov V.B. TAGSS: Svidetel'stvo o gosudarstvennoj registracii programm dlja JeVM № 2006613753// Zaregistrirovano v reestre programm dlja JeVM Federal'noj sluzhby po intellektual'noj sobstvennosti, patentam i tovarnym znakam 27 Oktjabrja 2006

32) Zhabin S.N., Sulimov V.B. Programma postroenija dostupnoj rastvoritelju poverhnosti dlja proizvol'nyh organicheskih molekul i interaktivnyj prosmotr polozhenij ligandov v aktivnyh centrah belkov// Sbornik materialov XIII rossijskogo nacional'nogo kongressa «Chelovek i lekarstvo», 3-5 aprelja 2006. C.15.

33) Zhabin S.N., Sulimov V.B. Programma dlja vizualizacii i redaktirovanija molekul «MOLRED»//Sbornik materialov II vserossijskoj konferencii «Mnogomoshtabnoe modelirovanie processov i struktur v nanotehnologijah» , 27-29 maja 2009. C. 166; S. Zhabin, V. Sulimov, "Construction of the adaptively triangulated smooth molecular surface: program TAGSS", (MEPhI) Electronic Journal "Scientific Visualization", V. 3, № 2, pp. 27-53,
http://sv-journal.org/2011-2/03.php?lang=en

34) Pascual-Ahuir J.L., Silla E.// J. Comput. Chem. 1990. V.11. P.1047

35) Silla E., Tunon I., Pascual-Ahuir J.L. //J. Comput. Chem. 1991. V.12. P.1077

36) Pascual-Ahuir J.L., Tunon I., Silla E. // J. Comput. Chem. 1994. V.15. P.1127

37) Cossi M., Mennuccci B., Gammi R. //Journal of Computational Chemistry. 1996. V.17. P.57

38) Bordner, A.J., Cavasotto, C.N., and Abagyan, R.A., Accurate Transferable Model for Water, n-Octanol, and n-Hexadecane Solvation Free Energies, J. Phys. Chem. B, 2002, vol. 106, no. 42, pp. 11009–11015.

39) Bondi A. // J. Phys. Chem. 1964. V.68 P.441

40) Cheng A., Best S.A., Merz Jr K.M., Reynolds C.H. // Journal of Molecular Graphics and Modelling. 2000. V.18. P.273

41) Scharlin, P., Battino, R., Silla, E., Tunon, I., and Pascual-Ahuir, J.L., Solubility of Gases in Water: Correlation Between Solubility and the Number of Water Molecules in the First Solvation Shell, Pure & App. Chem., 1998, vol. 70, no. 10, pp. 1895–1904.

42) Hermann R.B. //J. Phys. Chem. 1972. V.76. P.2754

43) Connolly M.L.// Science. 1983. V.221. P.709

44) Krone M.,Grottel S.,Ertl T.,"Parallel Contour-Buildup Algorithm for the Molecular Surface",Biological Data Visualization (BioVis), 2011 IEEE Symposium on, 17-22

45) Rinaldi D. , Rivail J.L.// Theor. Chim. Acta.1973. V.32 P.57

46) Rivail J.L., Rinaldi D. // Chem Phys. 1976. V.18. P.233

47) Harrison S. W., Nolte N.J., Beveridge D. // J. Phys. Chem. 1976 V.80. P.2580

48) Felder C.E. // J. Chem. Phys. 1981. V.75 P.4679

49) Rivail J.L., Terryn B. // Journal de Chimie Physique et de Physico-Chimie Biologique. 1982. V.79(1) P.1-6

50) Rivail J.L., Rinaldi D., V. Dillet V.// Journal de Chimie Physique et de Physico-Chimie Biologique. 1998,V.95(8) PP. 1818-1830

51) Rinaldi, D., Ruiz-Lopez M.F., Rivail J.L. // J. Chem. Phys. 1983. V.78. P.834

52) Kirkwood J.G.// J. Chem. Phys. 1934. V.2. P.351

53) Giard J., Macq B. // Int. J. Biomed. Imaging. 2010. V.2010 P.1

54) Varshney A., Brooks F.P. // IEEE Proceedings of the 4th conference on Visualization, 1993. P.300

55) X.-C. Tai, T. F. Chan. A survey on multiple level set methods with applications for identifying piecewise constant functions // Int. J. Numer. Anal. Model., 1(1):25–47, 2004

56) Grant J., Pickup B. //Journal of Phys Chem. 1995. V.99 P.3503

57) Can T., Chen C., Wang Y. //J. Mol. Graph. Model. 2006. V.24. P.442

58) Pan Q., Tai X.-C. // Commun. Comput. Phys. 2009. V.6 P.777

59) Richards E.M.// Ann. Rev. Biophys. Bioeng. 1977. V.6 P.151

60) Connolly M.L. //J. Appl. Crystallogr. 1983. V.16 P.548

61) Greer J., Bush B.L.// In Proceedings of the National Academy of Sciences USA. 1978 V. 75. P.303

62) Liang J., Edelsbrunner H., Fu P., Sudhakar P.V., Subramaniam S.// Proteins: Structure, Function, and Genetics. 1998. V.33. P.1

63) Ryu J., Park R., Cho C.-H., Kim D.-S. // KSIAM Meetings 2005, http://www.ksiam.org/conference/annual052/upfile/abstract_2.pdf

64) Bhat S., Purisima E.O. //Proteins: Structure, Function, and Bioinformatics, 2006. V.62. P.244

65) Edelsbrunner H. // Discrete and Computational Geometry. 1999. V.21. P. 87

66) Bates, P. W., Wei, G. W. and Zhao, S.//J. Comput. Chem. 2008. V.29. P.380

67) Cheng H.-L., X. Shi X. //In Proceedings of the conference on IEEE Visualization, 2004. P.481

68) Laug P., Borouchaki H.// Engineering with Computers. 2002. V. 18. P.199

69) Connolly M.L.// J. Appl. Cryst. 1985. V.18. P.499

70) Akkiraju N., Edelsbrunner H.// Journal Discrete Applied Mathematics - Special volume on computational molecular biology. 1996. V. 71. P.5

71) Akkiraju N.,Advisers: Edelsbrunner N. Doctoral Dissertation, Molecule surface triangulation from alpha shapes, USA: University of Illinois at Urbana-Champaign Champaign, 1996
http://portal.acm.org/citation.cfm?id=924371

72) Ryu J., Park R., Seo J., Kim C., Lee H.C., Kim D.-S.// ICCSA'07 Proceedings of the international conference on computational science and its applications, 2007. V.Part I. P.55-67

73) Cheng H.-L., Dey T.K., Edelsbrunner H., Sullivan J.//SODA '01 Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms: Society for Industrial and Applied Mathematics Philadelphia,2001. P.47

74) Ju T., Losasso F., Schaefer S., Warren J. //Proc. ACM Int.Conf. Computer Graphics and Interactive Techniques (SIGGRAPH "02), 2002. V.21. P.339

75) Lorensen W., Cline H.// Proc. ACM Int.Conf. Computer Graphics and Interactive Techniques (SIGGRAPH "87), 1987. V.21 P.163

76) Zhang Y., Xu G., Bajaj C. //Comput. Aided Geom. Des, 2003. Vol. 23. P.510

77) Frey P.J., Geotge P.-L. Mesh Generation. Application to Finite Elements. Wiley-ISTE , 2008. P.850

78) Lohner R. // Communications in Numerical Methods in Engineering 1996. V.12. P.683

79) Moller P., Hansbo P.// International Journal for Numerical Methods in Engineering 1995. V.38. P.3551

80) Cuiliere J. C.// Computer Aided Design 1998. V.30. P.139

81) Tristano J.R., Owen S.J., Canann S.A. // 7th International Meshing Roundtable: Sandia National Labs, 1998. P.429

82) Guan Z., Shan J., Zheng Y., Gu Y.// Int. J. Numer. Methods Eng. 2008. V.74. P.642

83) Hilton A., Stoddart A.J., Illingworth J., Windeatt T.// IEEE International Conference on Image Processing, 1996. P.381

ftp://ftp.ee.surrey.ac.uk/pub/vision/papers/hilton-icip96.ps.Z

84) Hartmann E.// Vis. Comput. 1998. V.14. P.95

85) Karkanis T., Stewart A.J.// IEEE Computer Graphics and Applications. 2001. v.21. P.60

86) Minxin Chen M., Lu B. // J. Chem. Theory Comput. 2011. V.7. P.203

87) Isenburg M., Lindstrom P., Gumhold S., Snoeyink J. // Proceedings of IEEE Visualization: Seattle, Washington, 2003 P.465

88) Blacker T., Stephenson M. // Int. J. Numerical Methods in Eng. 1991. V.32. P.811

89) Basilevsky M.V., Leontyev I.V., Luschekina S.V., Kondakova O.A., Sulimov V.B. // J. Comp. Chem. 2006. V. 27. P. 552

90) Klamt A., Schuurmann G. // J. Chem. Soc. Perkin Trans 2. 1993. P.799-805.

91) Truong T.N., V. Stefanovich E.V.// J. Chem. Phys. 1995. V.103. P.3709

92) Cammi R., J. Tomasi J.//J. Chem. Phys. 1994.V.100. P.7495

93) Cammi R., J. Tomasi J.// J. Chem. Phys. 1994. V.101. P.5900

94) Cammi R., J. Tomasi J.//Journal of Computational Chemistry 1995. V.16. P.1449

95) Totrov M., Abagyan R. // Biopolymers (Peptide Science). 2001. Vol. 60. P.124

96) Pomelli S. , Tomasi J.// Theor Chem Acc. 1997. V.96 P.39

97) Romanov A.N., Jabin S.N., Martynov Y.B., SulimovA.V., Grigoriev F.V., Sulimov V.B. // J. Phys. Chem. A. 2004. V.108. N. 43. P. 9323

98) Totrov M.// J. Comp. Chem. 2004. V.25 P.609

99) Halgren, T.A., Merck Molecular Force Field, J. Comput. Chem., 1996, vol. 17, nos. 5–6, pp. 490–519, 520–552, 553–586, 587–615, 616–641.

100) Onufriev, A., Case, D.A., and Bashford, D., Effective Born Radii in the Generalized Born Approximation: the Importance of Being Perfect, J. Comput. Chem., 2002, vol. 23, no. 14, pp. 1297–1304.

101) Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E., The Protein Data Bank, Nucl. Acids Res., 2000, vol. 28, no. 1, pp. 235–242 http://www.rcsb.org/pdb/home/home.do

102) Word, J.M., Lovell, S.C., Richardson, J.S., and Richardson, D.C., Asparagine and Glutamine: Using Hydrogen Atom Contacts in the Choice of Side-Chain Amide Orientation, J. Mol. Biol., 1999, vol. 285, no. 4, pp. 1735–1747.

103) Zhabin, S.N. and Sulimov, V.B., Implementation of Interactivity in the MOLRED Molecular Editor, Nauch. Visual., 2010, vol. 2, no. 1, pp. 59–81.

104) P. Tosco, N. Stiefl and G. Landrum, Bringing the MMFF force field to the RDKit: implementation and validation// Journal of Cheminformatics 2014, 6:37

105) Chudinov, G.E., Napolov, D.V., and Basilevsky, M.V., Quantum-Chemical Calculations of the Hydration Energies of Organic Cations and Anions in the Framework of a Continuum Solvent Approximation // Chem. Phys., 1992, vol. 160, no. 1, pp. 41–54.

106) Jorgensen, W.L., Ulmschneider, J.P., and Tirado-Rivers, J., Free Energies of Hydration from a Generalized Born Model and an All-Atom Force Field, J. Phys. Chem. B, 2004, vol. 108, no. 41, pp. 16264–16270.

107) J. R. Shewchuk, An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, Technical Report (School of Computer Sci., Carnegie Mellon University, 1994), p. 58.

108) D. V. Beklemishev, The Couse of Analytical Geometry and Linear Algebra (Nauka, Moscow, 1984), p. 319 [in Russian]. 108. N. 43. P. 9323.

109) M. V. Basilevsky, F. V. Grigoriev, E. A. Nikitina, and J. Leszczynski, J. Phys. Chem. B 114, 2457 (2010).

110) M. A. Vorotintsev and A. A. Kornyshev, Electrostatics of Media with the Spatial Dispersion (Nauka, Moscow, 1993) [in Russian].

111) Y. Kharkats, A. A. Kornyshev, and M. A. Vorotyntsev, J. Chem. Soc., Faraday Trans. 2 72, 361 (1976).

112) O. Kupervasser, The mysterious optimality of Naive Bayes: Estimation of the probability in the system of "classifiers", Pattern Recognition and Image Analysis, 24 (1), pp. 1-10 (2014),

http://arxiv.org/abs/cs/0202020v1 (2002)

113) E. D. Maslennikov, A. V. Sulimov, I. A. Savkin, M. A. Evdokimova, D. A. Zateyshchikov, V. V. Nosikov, V. B. Sulimov, An intuitive risk factors search algorithm: usage of the Bayesian network technique in personalized medicine, Journal of Applied Statistics,42 (1), pp. 71–87 (2015)

114) V. Ramensky, A. Sobol, N. Zaitseva, A. Rubinov, V. Zosimov, A novel approach to local similarity of protein binding sites substantially improves computational drug design results, Proteins: Structure, Function, and Bioinformatics, , 69(2), pp 349–357 (2007)

115) S. Nikitin1, N. Zaitseva, O. Demina, V. Solovieva, E. Mazin, S. Mikhalev, M. Smolov, A. Rubinov, P. Vlasov, D. Lepikhin, D. Khachko, V. Fokin, C. Queen,V. Zosimov, A very large diversity space of synthetically accessible compounds for use with drug design programs, Journal of Computer-Aided Molecular Design, 19, pp 47–63 (2005)

116) Raymer M. L., Doom T. E., Kuhn L. A., Punch W. F., "Knowledge discovery in medical and biological datasets using a hybrid bayes classifier/evolutionary algorithm," IEEE Transactions on Systems, Man, and Cybernetics, 33B, 802 (2003)

117) Domingos, P., and Pazzani, M., On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. Machine Learning 29, 103 (1997)

118) Zhang H., The Optimality of Naive Bayes, In FLAIRS Conference (2004)
http://www.cs.unb.ca/profs/hzhang/publications/FLAIRS04ZhangH.pdf

119) Kuncheva L.I., On the optimality of Naive Bayes with dependent binary features, Pattern Recognition Letters, 27,830 (2006)

120) Landau L.D., Lifshitz E.M., Statistical Physics, 5,Elsevier Science Technology, United Kingdom, (1996)

121) A. Yu. Mikhalev, I. V. Oferkin, I. V. Oseledets, et al., "Application of the Multicharge Approximation for Large Dense Matrices in the Framework of the Polarized Continuum Solvent Model," Vychisl. Metody Programm. 15, 9–21 (2014).

Continuum Solution Models for Computer Aided Drug Design

The main paradigm used in the modern development of new drugs consists in the association of a disease with the functioning of certain proteins, so they must be blocked to cure diseases. Blockage is performed with the use of molecules that selectively bind to these proteins in an organism (inhibitors). The time and material expenditures at the stage of searching for inhibitors can considerably be reduced with the use of computer-aided molecular modeling methods. The precision of estimating the protein–inhibitor binding energy governs the efficiency of predicting the activity of an inhibitor. The effect of a solvent on the protein–inhibitor binding energy is predominantly determined via the free solvation energy. To calculate the free solvation energy, it is necessary to construct a solvent model. Implicit (often called continual) solvent models, in which a solvent is treated as a continuum (continual, continuous) medium with specified properties, including dielectric permittivity, are used more frequently. We consider in the book these continuum solution models.

Kupervasser Oleg Yurjevich, physicist, was born on February 1st, 1966 in Moscow. He got his MSc at the Moscow Institute of Radioengineering, Electronics and Automation in 1989; the PhD, at Weizmann Institute of Science, Isral in 1999. He has a Certification in Bioinformatics (ATLAS College, 2002). He worked as scientist for Technion, in Haifa Israel, between 2005 and 2008; as Consultant of Rafael, Haifa, between 2005 and 2008; he worked at the Moscow State University between 2008 and 2010. He has been Algorithm Group Leader of Transas Group of Companies, in Moscow for the years 2011–2012 and he worked at LG Technology Center or Moscow in 2011. He was awarded the following prizes: Skolkovo Project Participator Association, 2012; included in the new 30th Pearl Anniversary Edition of "Marquis Who's Who in the World", 2013. His achievements are: resolution of the basic paradoxes of physics; the pole solution in formation of Saffman–Taylor fingers and the flame front propagation; patent for a continuum solvent model: DISOLV program–algorithm, implementation and validation; research in visual–navigation algorithm; design of a server for decomposition of a protein structure on a set of closed loops. His avocations are the human history and the futurology.

Wanner Natalia Eduardovna is Candidate of Biology Science. She started her scientific activities in 1983 by studying in Graduate Center of SSI All-Russian Research Institute of Veterinary, Sanitation, Hygiene and Ecology. Currently, she is a senior research scientist of Laboratory of Investigation of Aerosols in SSI All-Russian Research Institute of Veterinary, Sanitation, Hygiene and Ecology. She is also a faculty member of Department of Bioecology and Biosecurity in Moscow State University of food production. She has 7 patents and 37 publications.