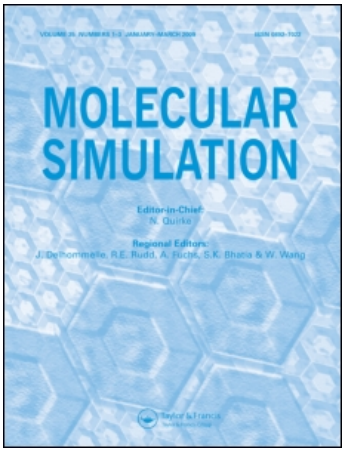


Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Cytochrome P450 52A3: Modelling of 3D Structure and Surface Mutations

Vladlen S. Skvortsov[a]; Natalya V. Belkina[a]; Alexis S. Ivanov[a]
[a] Institute of Biomedical Chemistry, Moscow, Russia

To cite this Article Skvortsov, Vladlen S. , Belkina, Natalya V. and Ivanov, Alexis S.(2000) 'Cytochrome P450 52A3: Modelling of 3D Structure and Surface Mutations', Molecular Simulation, 24: 4, 369 — 378
To link to this Article: DOI: 10.1080/08927020008022382
URL: http://dx.doi.org/10.1080/08927020008022382

Molecular Simulation, 2000, Vol. 24, pp. 369–378
Reprints available directly from the publisher
Photocopying permitted by license only

CYTOCHROME P450 52A3: MODELLING OF 3D STRUCTURE AND SURFACE MUTATIONS

VLADLEN S. SKVORTSOV*, NATALYA V. BELKINA and ALEXIS S. IVANOV

Institute of Biomedical Chemistry, Moscow, Russia

(Received April 1999; accepted May 1999)

Earlier W.-H. Schunck *et al.* [1] have prepared a water soluble enzymatically active fragment of cytochrome P450 52A3 (CYP52A3) which is lack of 66 amino acid residues, existed as a dimer in aqueous solution. Now we propose 3D structure of the fragment, which is based on multiple sequence alignment of the CYP52A3 with its homologues proteins of known 3D structure: CYP101, 102, 107A1 and 108. The structural model have been optimised and used as a prototype for computer simulation of point mutations. These mutations should bring some changes in the surface properties, interfering dimer formation. For this aim the point of 22 hydrophobic amino acid residues have been sequentially replaced with that of charged amino acids (GLU, ASP, ARG and LYS). The scoring of "mutants" was conducted based on the changes of protein surface hydrophobicity and protein-solvent interaction energy. An analysis of the surface hydrophobicity and protein-solvent interactions permit to select most sensitive three sites (171, 352 and particularly 164 amino acid residues). The dimerization of the following "mutant" fragments must be investigated experimentally.

Keywords: Cytochrome P450 52A3; surface mutations

INTRODUCTION

Earlier W.-H. Schunck and co-workers [1] have prepared an enzymatically active water-soluble fragment of cytochrome P450 52A3 (CYP52A3) which is lack of 66 N-terminal amino acid residues, which is dimerized in aqueous solution. The dimerization of the fragment defined by hydrophobic interactions. It is naturally to expect, that "mutations" can lead to a fall in

*Corresponding author.

surface hydrophobicity and thereby to suppress dimer formation without disturbing enzymatic activity.

Efficiency of such approach can be greatly increased by a preliminary analysis of amino acid replacings by computing technology. Unfortunately, X-ray structure of CYP52A3 is unknown, as well as of those of all another membrane-bounded cytochromes P450. Thus, computer modelling is an unique method to obtain 3D structure of the protein. X-ray data related to five bacterial cytochromes can be used for such a modelling based on the between mammalian and bacterial P450s homology.

The investigation presented includes:

1. Creation of homology based computer 3D model of water soluble fragment of cytochrome P450 52A3;
2. Revealing of the surface amino acid residues and computer simulation of their replacing leading to the decrease in dimerization potential;
3. Selection of the most effective replacement variants for the following testing by gene engineering.

METHODS

The first step was the creation 3D model of CYP52A3 water soluble fragment based on the multiple sequence alignment of CYP52A3 and P450's with known 3D structures (CYP101, 102, 107A1 and 108). Since the cytochrome P450's family is not of high homology a choice of the alignment procedure becomes a critical problem.

Comparison of several variants of pair alignment of CYP52A3 with CYP102 is represented in Figure 1 (here and hereinafter the 1st residues in a sequence of water-soluble fragment corresponds to 67th residues in that of CYP52A3). The result of multiple alignment of CYP52A3 with CYP102 and CYP101 based on the global alignment of cytochrome P450s family (more than 400 sequences) [2] was shown on the same figures. All these alignments are similar, but there are smaller numbers of gaps in alignment received by using program QUANTA than in others ones.

Preliminary modelling of three-dimensional structure was done by using program QUANTA [5]. As a main template for modelling structural conservative regions we used CYP102, the substrate specificity of which is just like to that of CYP52A3. The modelling of nonconservative parts was done with using of suitable fragments from PDB [6]. Resulted structural model has been optimised by using SYBYL software suit [7].

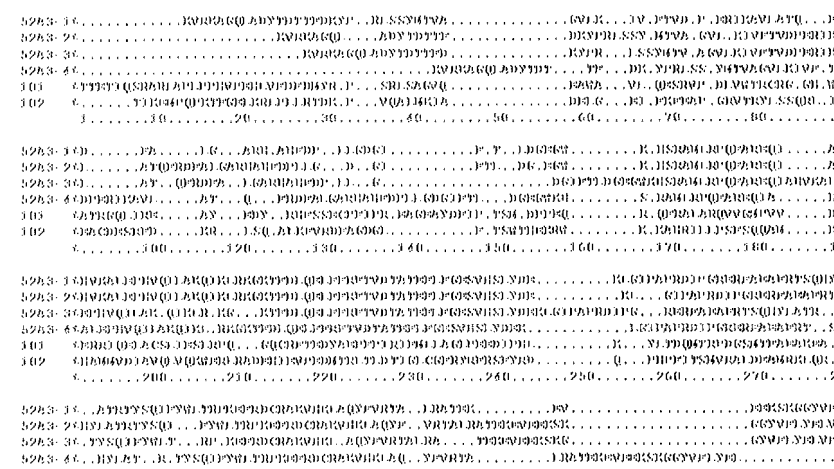


FIGURE 1 The comparison of four distinct alignments CYP52A3 (starting from 67 amino acid) with CYP102. 52A3-1 – pair alignment obtained by QUANTA program; 52A3-2 – alignment obtained by Dr. S. Graham-Lorence [3]; 52A3-3 – alignment obtained by Dr. B. Sobolev [4]; 52A3-4, 101, 102 – results of multiple alignment obtained by Dr. L. Koymans (CYP52A3, CYP101 and CYP101 respectively).

On the next stage we consecutively replace surface residues of LEU, ILE or VAL on GLU, ASP, ARG and LYS residues. Some scoring functions have been used for evaluation of the replacement consequences: energy transfer from an apolar to aqueous environments, areas of hydrophobic and hydrophilic surface regions and also the integral parameters hydrophily and hydrophoby of protein surface.

The transfer energy was calculated by FIESTA program, which was kindly afforded by doctor Sklenar [8]. The program allows to calculate rapidly the energy without explicit simulating the water environment. Other

parameters were calculated by using a self-made program based on a molecular lipophilicity potential (LP) [9]:

$$\mathrm{LP}_{HM} = \frac{\sum_{i=1}^{N} f_i \cdot g(d_i)}{\sum_{i=1}^{N} g(d_i)}, \quad \text{where } g(d_i) = \frac{e^{-\mathrm{C_1C_2}} + 1}{e^{-\mathrm{C_1}(d_i - \mathrm{C_2})} + 1}$$

d_i – the distance of a certain point from atom i;
$\mathrm{C_1}$ and $\mathrm{C_2}$ – the Fermi constants, which have been determined empirically as 1 and 4 respectively;
f_i – the Crippen atomic partial lipophilicities [10].

Integral parameters of hydrophily (L) and lipophilicity (U) was calculated as follows:

$$L = \sum_{\mathrm{LP}^i_{HM} < 0} \mathrm{LP}^i_{HM} \cdot ds_i \quad \text{and} \quad U = \sum_{\mathrm{LP}^i_{HM} \geq 0} \mathrm{LP}^i_{HM} \cdot ds_i,$$

where

LP^i_{HM} – the value of lipophilicity potential (LP_{HM}) on area element i,
ds_i – the value of area element i.

The value of hydrophily (S_L) and lipophilicity (S_U) areas was calculated as follows:

$$S_L = \sum_{\mathrm{LP}^i_{HM} < 0} ds_i \quad \text{and} \quad S_U = \sum_{\mathrm{LP}^i_{HM} \geq 0} ds_i.$$

RESULTS AND DISCUSSION

The 3D structure modelling of water-soluble CYP52A3 fragment was based on alignment procedure with bacterial cytochromes with the aid of QUANTA program [5]. The model resulted was optimised up to minimum of potential energy, thus indicating most stable state (see Fig. 2A). It can be seen the number of certain inexactnesses of the secondary structure. To overcome the troubles it was necessary to determine structure conservative regions with high reliability. For this aim four another models CYP52A3 were constructed. The models were based on two variants of alignments CYP52A3 with CYP101 and CYP102 (the former has been obtained by Dr. S. Graham-Lorence [3], the latter by Dr. L. Koymans [2]).

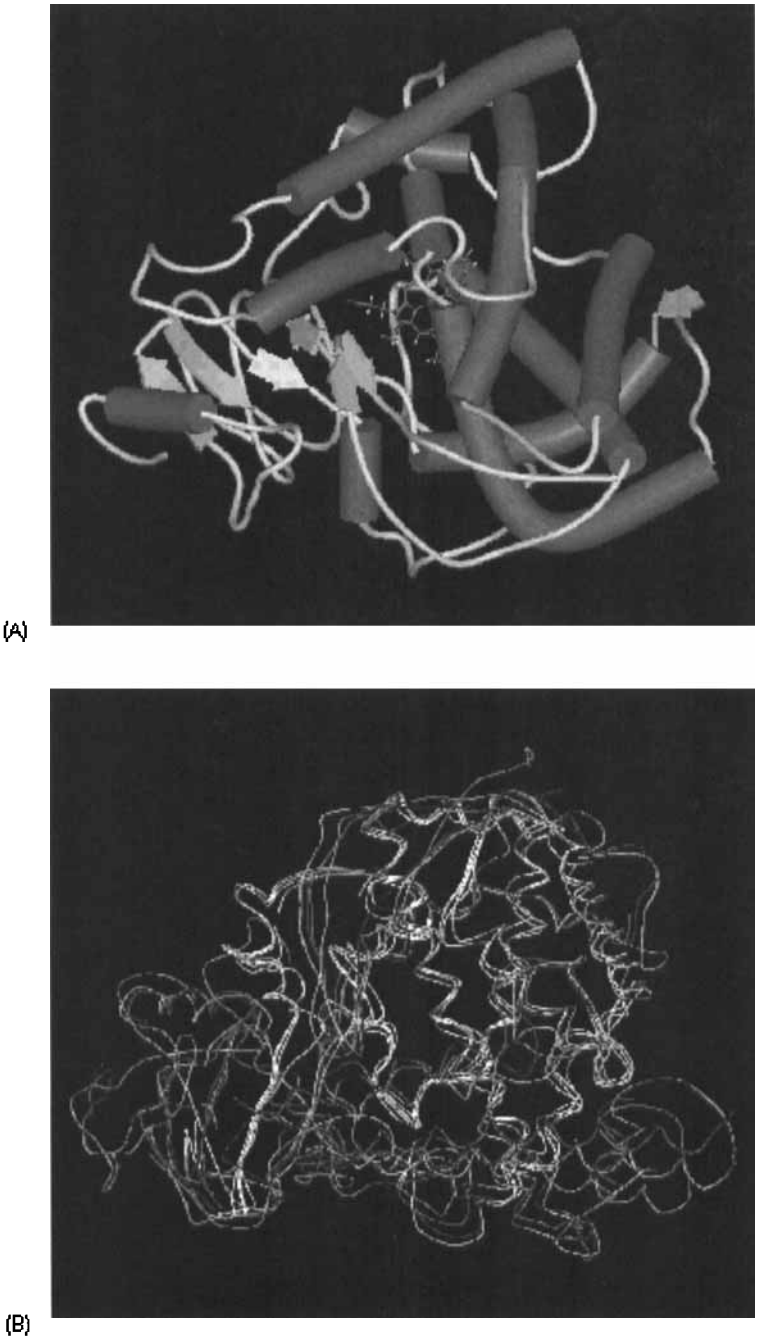


FIGURE 2 The results of CYP52A3 modelling: (A) The schematic model of CYP52A3. Secondary structure automatically determined by using the program WebLab Viewer: (B) The superposition of five models built by using different pair alignments with CYP101 or CYP102. structure-conservative regions of sequences are shown in white colour. (See Color Plate XII).

Structure conservative regions (SCRs) of model CYP52A3 were revealed on rms deviations of the backbone for each pair of structures. Superposition all five models CYP52A3 reveal SCRs of the protein and it was shown on Figure 2B.

Secondary structure elements and active site of CYP52A3 model were analysed on next stage. Distances between active heme oxygen and C-alpha atoms were measured for all amino acids of the protein. Then we determined a set of surface amino acid residues by calculating water accessible area for each amino acid. All these data used for replacement analysis are presented on Figure 3.

22 hydrophobic residues (LEU, ILE and VAL) were selected for replacing. The more than half of their surface is water-accessible. Each of these residues were consecutively replaced by ones charged (GLU, ASP, ARG and LYS), so we got 88 "mutant" proteins. All of obtained molecules were undergone local energy minimization near to the "point mutation". The characteristics of obtained mutants are presented on Figures 4 and 5. It should be pointed that introductions of different charged amino acid (positive or negative) residues may lead to differing effects dependent of its neighbourhood. The changes of area of hydrophilic and lipophilic surfaces are not present, because it was not significant.

We designated replacement, which results in the most advantageous transfer energy and the most in the maximal decrease of the total surface

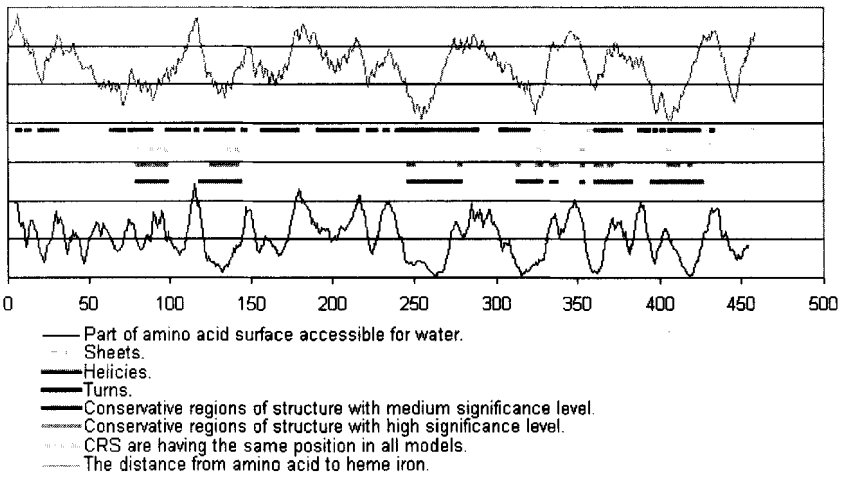


FIGURE 3 The parameters of the CYP52A3 model based on alignment received by QUANTA program.

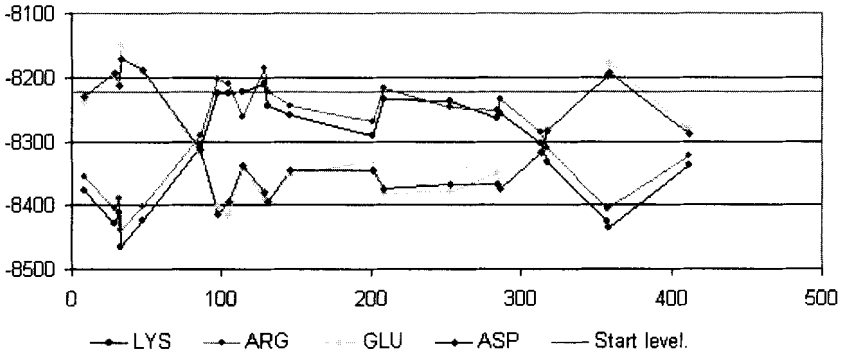


FIGURE 4 The energy transfer from an apolar to aqueous environment of CYP52A3 "mutants". All 88 replacements are present.

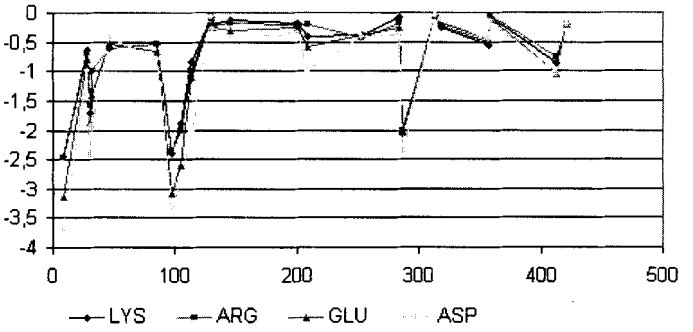
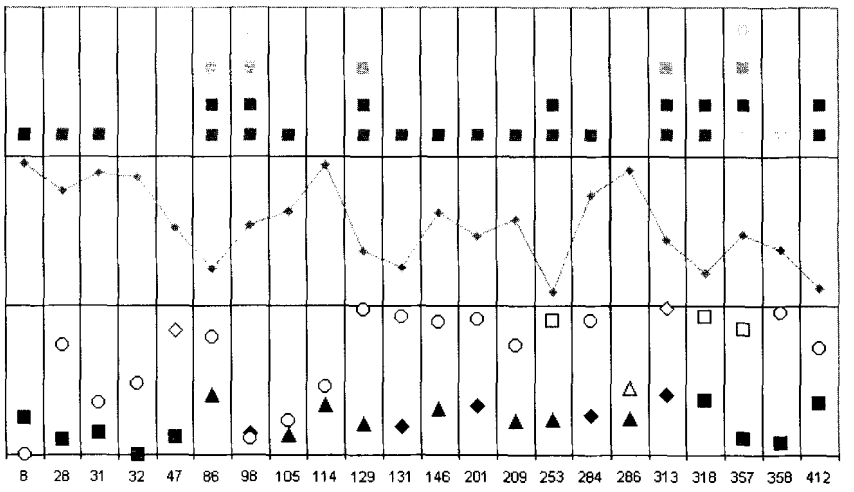


FIGURE 5 The changes of surface hydrophobicity of CYP52A3 "mutants" ($L+U$). All 88 replacements are present.

lipophilicity, as the most favourite replacement (see Fig. 6). We believe that the replacement of residues 8, 31, 98, 105 or 286 is a most effective pathway to decrease or prevent protein dimerization.

It is well known that structure of terminal regions of protein chain is changeable. Furthermore, the alignments indicated that namely N-terminal sequence is the most ambiguous. Hence replacements of 8 and 31 residues should be exclude from the list of possible "mutations". Other three amino acid residues (98, 105 and 268) are situated on protein surface nearly (see Fig. 7). The residues VAL98 and VAL105 are parts of stable alpha-helix; in addition VAL98 is a part of high conservative element of the protein structure. Thus VAL98 and VAL105 are the most perspective sites for replacements. VAL286 is less appropriate because it belongs to unstable surface loop, which was modelled not sufficiently reliably.



Sheets.
Helicies.
Turns.
Conservative regions of structure with medium significance level.
Conservative regions of structure with high significance level.
CRS are having the same position in all models.
The distance from amino acid to heme iron.
Energy of electrostatic interaction of molecule with water when new amino acid is LYS.
Energy of electrostatic interaction of molecule with water when new amino acid is ARG.
Energy of electrostatic interaction of molecule with water when new amino acid is GLU.
Energy of electrostatic interaction of molecule with water when new amino acid is ASP.
Reduction of surface's hydrophobicity of molecule when new amino acid is LYS.
Reduction of surface's hydrophobicity of molecule when new amino acid is ARG.
Reduction of surface's hydrophobicity of molecule when new amino acid is GLU.
Reduction of surface's hydrophobicity of molecule when new amino acid is ASP.

FIGURE 6 The best parameters of 22 possible replacements.

Thus we can recommend for experimental checking following amino acid replacements:

1. VAL 98 (164) → GLU or ASP;
2. VAL105 (171) → GLU or ASP;
3. VAL286 (352) → GLU or ASP.

Numbers of residues according to full sequence CYP52A3 are presented parenthetically. Order of the amino acid replacements coincides with conjectural probability of the success; *i.e.*, we suppose that VAL98 is the most perspective site to replacement.

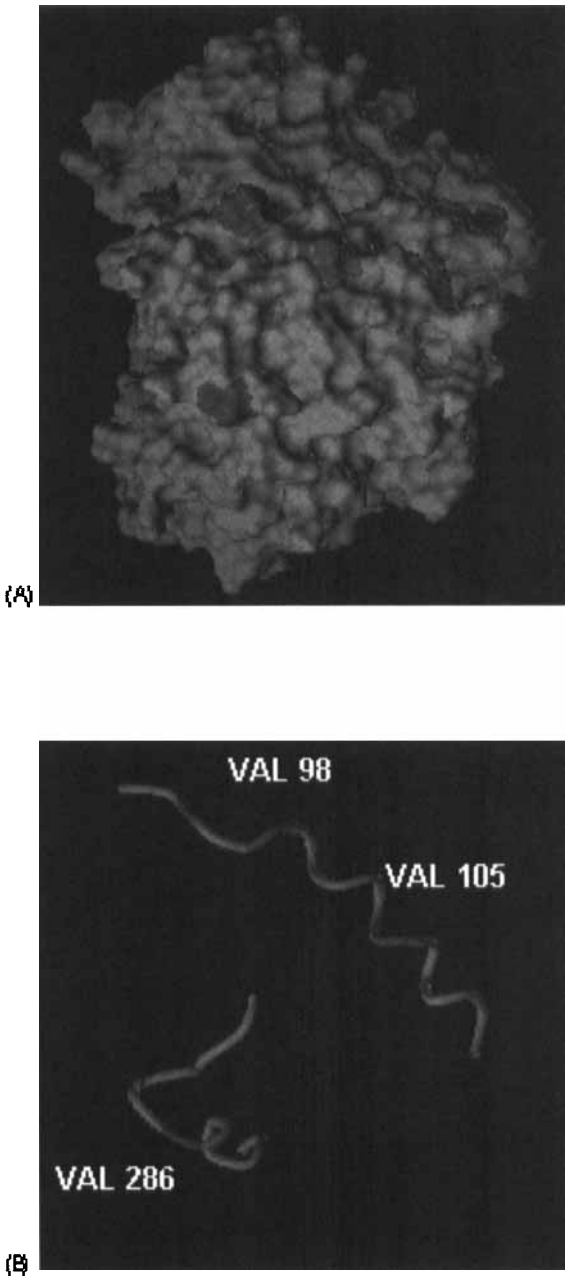


FIGURE 7 The most perspective sites for replacement: (A) Situating on the protein surface; (B) Situating in the elements of the secondary structure. (See Color Plate XIII).

Acknowledgments

We are grateful Dr. W.-H. Schunck and Dr. H. Sclenar for sufficient scientific and technical support. We also thank Dr. A. Shkrob for critical reading of the manuscript.

References

[1] Scheller, U., Juretzek, T. and Schunck, W.-H. (1996). *Methods Enzymology*, **272**, 65–67.
[2] Nelson, D. R., Koymans, L., Kamataki, T. *et al.* (1996). *Pharmacogenetics*. Feb. 6, **1**, 1–42.
[3] Graham-Lorence, S. E. and Peterson, J. A. (1996). *Methods Enzymol*, **272**, 315–326.
[4] Sobolev, B. N., Unpublished results.
[5] QUANTA – *Molecular Simulations Inc.*, 8 New England Executive Park, Burlington, MA 01803-5297, USA.
[6] Sussman, J. L., Lin, D., Jiang, J. *et al.* (1998). *Acta Cryst.*, **D54**, 1078–1084.
[7] Sybyl 6.5 – *Tripos Inc.*, 1699 South Hanley Road, St. Louis, Missouri, 63144. USA.
[8] Sklenar, H., Eisenhaber, F., Poncin, M. and Lavery, R. (1990). *Theoretical Biochemistry and Molecular Biophysics*. Eds.: Beveridge, D. L. and Lavery, R., Adenine Press, pp. 317–335.
[9] Heiden, W., Moeckel, G. and Brickmann, J. (1993). *J. Comp.-Aided Mol. Design*, **7**, 503–514.
[10] Viswanadhan, V. N., Ghose, A. K., Revankar, G. R. and Robins, R. K. (1989). *J. Chem. Inf. Comput. Sci.*, **29**, 163–172.