



(SI)Accurate calculations in drug discovery research: the biological roles of magnesium and calcium ions

Graham Richards

► To cite this version:

Graham Richards. (SI)Accurate calculations in drug discovery research: the biological roles of magnesium and calcium ions. *Molecular Physics*, 2009, 107 (08-12), pp.819-822. 10.1080/00268970802649617 . hal-00513242

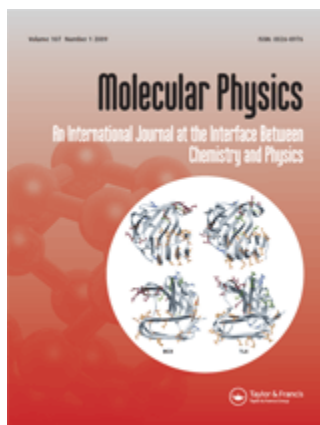
HAL Id: hal-00513242

<https://hal.science/hal-00513242>

Submitted on 1 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



(SI)Accurate calculations in drug discovery research: the biological roles of magnesium and calcium ions

Journal:	<i>Molecular Physics</i>
Manuscript ID:	TMPH-2008-0320.R1
Manuscript Type:	Special Issue Paper - Fritz Schaefer
Date Submitted by the Author:	20-Nov-2008
Complete List of Authors:	Richards, Graham; University of Oxford, Chemistry
Keywords:	



Accurate calculations in drug discovery research: the biological roles of magnesium and calcium ions

By W. GRAHAM RICHARDS

Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ

graham.richards@chem.ox.ac.uk

Abstract

The use of computational chemistry in drug discovery has in the past been largely restricted to the use of empirical potential functions or semi-empirical quantum mechanics. Computer power is now offering the opportunity to employ high quality *ab initio* methods of the sort pioneered by Fritz Schaefer. Here the general scene of computational drug discovery is set and a specific problem where accurate methods are essential is introduced. This problem is the contrasting roles of Mg^{2+} and Ca^{2+} in biology, the understanding of which requires very accurate calculation of binding free energies between the cations and organic anions as well as with proteins.

1 INTRODUCTION

Fritz Schaefer and I were both part of the earliest group of the young researchers who began to use *ab initio* molecular orbital calculations in the 1960s to answer chemical questions. We started from a very similar viewpoint: doing the most accurate calculations we could on inevitably small and isolated systems. His early papers include work on configuration interaction with Frank Harris [1], while I stuck to diatomics and in particular spin-orbit coupling phenomena.[2]

1
2
3 In the 1970s our research diverged. Fritz produced ever more accurate calculations on small
4
5 systems and I was pleased to be able to write in a 1979 article in *Nature* describing his work,
6
7 that he had ushered in the ‘Third Age of Quantum Chemistry’, where the calculations can be as
8
9 precise as experiment.[3] My own research followed the alternative path of going to bigger
10
11 and bigger molecules and systems with the inevitable loss of accuracy, but with the intention of
12
13 tackling problems in molecular biology and in particular in the field of drug discovery.
14
15

16
17 Now as we both reach mature years the two strands are once more converging and the power of
18
19 contemporary computational resources is permitting accurate calculations to be used to answer
20
21 questions in complex molecular systems which may incorporate macromolecules, solvent
22
23 molecules and ions.
24
25

26
27 The intention of this paper is to explore just how well accurate calculations are performing in
28
29 the area of drug discovery and by way of a specific example to show how there remain
30
31 problems which only development in theory can overcome.
32
33

34
35 Fritz’s contributions to computational chemistry have been enormous and it is a pleasure to
36
37 dedicate this article to him
38
39

40 41 42 43 2 COMPUTATIONAL DRUG DISCOVERY 44 45

46
47 With the essential completion of the Human Genome Project and the perhaps surprising
48
49 discovery that we only have a couple of tens of thousands of genes in our cells, attention and
50
51 funding has turned to the human proteome: the set of proteins for which our DNA provides the
52
53 code. Even if we consider all living systems and acknowledge the possibility of different
54
55 reading frames of the genome, there can exist only at most a few hundred thousand proteins.
56
57 All over the world synchrotron sources are being constructed with the prime intention of
58
59 determining the three-dimensional structures of these proteins which are to an overwhelming
60

degree the targets of drugs. In addition to attempts to solve structures of existing protein crystals, there are also continuing efforts to overcome the daunting technical hurdle of achieving suitable protein crystallization conditions for entire classes of proteins for which exceedingly few successful crystallizations have emerged, for example, membrane-spanning entities such as G-protein coupled receptors (GPCRs). There are thus a finite number of macromolecular targets for drugs and biologists are rapidly highlighting where blockage of a protein binding site might beneficially influence disease.

However, even if the number of drug targets is quite limited, the number of small drug-like molecules which might act as medicinal agents is almost limitless. Estimates have suggested that perhaps as many as 10^{40} different drug-like compounds could be included even in the molecular weight range between 150 and 800.

Computational chemistry can and is being employed to whittle down this vast number of compounds to a handful which can be synthesised and screened biologically. In our much publicised screensaver project we built a starting database of 3.5 billion candidate molecules, starting with the couple of million commercially available and then using published combinatorial chemistry schemes, before filtering the input information using Lipinski's rules.[4]

Computational drug discovery can then be considered as a funnel with increasingly restrictive and expensive filters.

Fig. 1

If we start with an enormous database and the crystal structure of a target protein with an inhibitor bound in its active site, the first filter is shape. Only molecules with a shape similar to the known inhibitor need be considered further. Recently we have developed some ultrafast shape similarity methods which are capable of handling billions of starting compounds in a

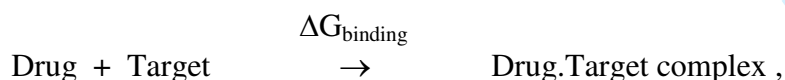
very short time.[5,6] Subsequent filtration may be performed using electrostatic similarity, pharmacophore matching and molecular docking and scoring.

With the number of interesting compounds reduced to a few hundreds, then it becomes possible to think in terms of accurate calculations which could bring us into the realm of Schaefer-level accuracy.

At the same time it must be remembered that finding a small molecule which will inhibit a protein is itself only a first stage which leads to a second funnel of myriad biological and safety assays in the drug discovery process (e.g. off-target pharmacology, distribution, metabolism, pharmacokinetics, etc). While accurate prediction of the binding energetics represents a great advance, it is only the first step forward toward becoming a drug.

3 FREE ENERGY CALCULATIONS

What (as theoreticians are always wont to say) the experimentalist would really like to be told would be the free energy of binding between the drug and the protein in aqueous solution, since the bigger the binding energy the better the drug with the smaller the dose and likelihood of side effects. Although 'in principle' (again as theoreticians often say) one should be able by molecular dynamics calculation to compute the free energy for



in practice this is such a slow (multisecond) process that the direct route is impossible.

At the accurate level, the almost as useful computation of relative free energies of binding of two drugs, say A and B, to the same target protein can be computed using the cycle illustrated in Figure 2.

Fig. 2

There are instances where such differences in binding energy have been so calculated to an accuracy of half a kilocalorie per mole, but in general despite its appeal this approach has not proved as fruitful as had been first anticipated. Possible problems include conformational changes in the protein target on binding the drug; difficulties in truly accounting for solvent waters and deficiencies in the intermolecular potentials used.

To investigate some of these difficulties we are here going to consider a case which should be interesting beyond being a model. It is a difference of fundamental importance in biology: the relative binding of Mg^{2+} and Ca^{2+} to proteins.

4 MAGNESIUM AND CALCIUM IN BIOLOGY

These two simple ions are major players in biology, but despite their similarity behave very differently, as summarized in Table 1.

Table 1

The only differences between the two ions are size and solvation and yet these can give rise to dramatically different behaviour. Table 2 gives some experimental binding constants between the pair of ions and some simple organic ions.[7]

Table 2

Note that sometimes Mg^{2+} binds better than Ca^{2+} and sometimes the reverse, with differences amounting to as much as five orders of magnitude. Here is a good puzzle for someone wanting to do accurate calculations.

5 COMPUTING BINDING CONSTANTS

Binding constants can be computed from the potential of mean force using either molecular dynamics or Monte Carlo methods. The latter have the advantage of giving exact NPT values.

We use the simple theory

$$\Delta G = RT \ln K$$

$$K = \int_{\infty}^{r_c} 4\pi r^2 \exp(-W_r / kT) dr$$

with W_r being the potential of mean force between the ions in solution. The molecular mechanics potentials can be checked by calculating such parameters as heats of hydration and volume of solution and it is a fact that most such potentials do not perform very satisfactorily, the reasons for which we will discuss later.

Figure 3 shows the results of such a calculation for Mg^{2+} interacting with acetate in water, and Figure 4 is a similar calculation for Ca^{2+} acetate.

Fig. 3

Fig. 4

Note that at long range the potential of mean force should match the dotted continuum value.

In this case the results in Table 3 are in very satisfactory agreement with experiment.

Table 3

When we come to try to do the same calculation for Mg^{2+} and Ca^{2+} binding to glycine, however, things are not as happy as shown in Table 4.

Table 4

Although it only represents a relatively small error in energy, the calculated binding energy is off by an order of magnitude.

If calculations of binding energy difference between two spherical ions of the same charge to simple partners do so badly, how can we hope to produce really useful binding energies of similar drug molecules to proteins? It is important to understand why the calculations perform so poorly.

6 SOURCES OF ERROR

One error which bedevils molecular simulation is the poor quality of the potentials used. They do not reproduce simple measurable quantities. It should be remembered that if we look at the simple Argon-Argon potential, we do not yet get the correct crystal structure of solid argon unless three-body forces are included and these are invariably ignored in molecular mechanics. In addition most calculations of molecular simulations are done in a box of a limited size so that there has to be a cut-off in the potential. This is particularly damaging when ions are involved since their interactions persist for long distances.

The greatest defect in the calculations described above is, however, the failure to incorporate the inevitable transfer of charge which will be found when a charged species is involved.

Empirical potentials are also defective in incorporating cation- π electron interactions. The way round those defects has to be to study the interactions using quantum mechanics which allows the negative electron charge to be mobile.

7 CONCLUSION

Fritz Schaefer has been the leading light in the use of molecular quantum mechanics to resolve chemical problems. In biology too, simplified methods are insufficient and we need quantum chemistry. In the past this has been problematic because of the computer time required to use *ab initio* methods. Now, however, innovations such as grid computing point the way to get

round the computational barrier. The techniques developed for small molecules can and should be deployed in areas such as drug discovery.

Acknowledgements

The potential of mean force calculations were performed by Dr James Bradley and the research was supported by the National Foundation for Cancer Research.

References

- [1] Schaefer, H.F., and Harris, F.E., 1967, *Chem. Phys. Lett.* **1**, 407-408.
- [2] Verhaegen, G., and Richards, W.G., 1966, *J. Chem. Phys.* **45**, 1828-1833.
- [3] Richards, W.G., 1979, *Nature*, **278**, 507.
- [4] Richards, W.G., 2002, *Nature Reviews Drug Discovery*, **1**, 551-555.
- [5] Ballester, P.J., and Richards, W.G., 2007, *Proc. R. Soc. A.*, **463**, 1307-1321.
- [6] Ballester, P.J., and Richards, W.G., 2007, *J. Comp. Chem.* **28**, 1711-1723.
- [7] Tam, S.C., and Williams, R.J.P., 1985, *Structure and Bonding*, **63**, 103-151.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Legends for Figures

Fig. 1 The drug discovery funnel

Fig. 2 Accurate binding energies. Simulation using free energy perturbation

Fig. 3 Mg^{2+} - acetate. Potential of mean force

Fig. 4 Ca^{2+} - acetate. Potential of mean force

Tables

Table 1 Mg^{2+} and Ca^{2+} contrasts

Table 2 Binding constants of Mg^{2+} and Ca^{2+}

Log_{10} binding const (bigger is stronger)

Table 3 Calculated binding constants for Mg^{2+} and Ca^{2+} to acetate

Table 4 Calculated binding constants for Mg^{2+} and Ca^{2+} to glycine zwitterion

Table 1

Mg²⁺ and Ca²⁺ contrasts

Mg ²⁺	Ca ²⁺
Slow ligand exchange	Fluxional complexes
Well defined octahedron	Varying coordination
Strong field ligands	Weak field ligands
Binds to anions in cells	Binds to lattices (bones & teeth)

Table 2

Binding constants of Mg^{2+} and Ca^{2+} Log_{10} binding const (bigger is stronger)

Ligand	Mg^{2+}	Ca^{2+}
Acetate	0.8	0.7
Glycine	3.4	1.4
Citrate	3.2	4.8
Imidodiacetate	2.9	2.6
Nitrilotriacetate	5.3	6.4
EDTA	8.9	10.7
EGTA	5.4	10.7
Malonate	2.8	2.5

Table 3

Calculated binding constants for Mg^{2+} and Ca^{2+} to acetate

	Log ₁₀ K	
	calc	expt
Mg^{2+}	0.8	0.8
Ca^{2+}	0.7	0.6

Table 4

Calculated binding constants for Mg^{2+} and Ca^{2+} to glycine zwitterion

	$\text{Log}_{10}K$	
	calc	expt
Mg^{2+}	2.5	3.4
Ca^{2+}	1.0	1.4

40 billion molecules

Molecular Physics

~100,000

Shape

~10,000

Electrostatics

~1,000

Pharmacophore

~100

Docking

10

Free energy of binding

URL: <http://mc.manuscriptcentral.com/tandf/tmph>

Target + Drug A $\xrightarrow{\Delta G_1}$ T.A complex

ΔG_3

ΔG_4

Target + Drug B $\xrightarrow{\Delta G_2}$ T.B complex

URL: <http://mc.manuscriptcentral.com/tandf/temph>

$$\Delta G_1 - \Delta G_2 = \Delta G_3 - \Delta G_4$$

