



HAL
open science

(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:	



1
2
3 **Accurate calculations in drug discovery research: the biological roles of magnesium and**
4 **calcium ions**
5
6

7
8 By W. GRAHAM RICHARDS
9

10
11 Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ
12

13
14 graham.richards@chem.ox.ac.uk
15
16

17
18
19
20 **Abstract**
21

22
23 The use of computational chemistry in drug discovery has in the past been largely restricted to
24 the use of empirical potential functions or semi-empirical quantum mechanics. Computer
25 power is now offering the opportunity to employ high quality *ab initio* methods of the sort
26 pioneered by Fritz Schaefer. Here the general scene of computational drug discovery is set and
27 a specific problem where accurate methods are essential is introduced. This problem is the
28 contrasting roles of Mg^{2+} and Ca^{2+} in biology, the understanding of which requires very
29 accurate calculation of binding free energies between the cations and organic anions as well as
30 with proteins.
31
32
33
34
35
36
37
38
39
40
41
42
43
44

45 1 INTRODUCTION
46

47
48 Fritz Schaefer and I were both part of the earliest group of the young researchers who began to
49 use *ab initio* molecular orbital calculations in the 1960s to answer chemical questions. We
50 started from a very similar viewpoint: doing the most accurate calculations we could on
51 inevitably small and isolated systems. His early papers include work on configuration
52 interaction with Frank Harris [1], while I stuck to diatomics and in particular spin-orbit
53 coupling phenomena.[2]
54
55
56
57
58
59
60

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

10
11 Now as we both reach mature years the two strands are once more converging and the power of
12 contemporary computational resources is permitting accurate calculations to be used to answer
13 questions in complex molecular systems which may incorporate macromolecules, solvent
14 molecules and ions.
15

16
17 The intention of this paper is to explore just how well accurate calculations are performing in
18 the area of drug discovery and by way of a specific example to show how there remain
19 problems which only development in theory can overcome.
20
21

22
23 Fritz’s contributions to computational chemistry have been enormous and it is a pleasure to
24 dedicate this article to him
25
26

27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 2 COMPUTATIONAL DRUG DISCOVERY 44 45

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

1
2
3 degree the targets of drugs. In addition to attempts to solve structures of existing protein
4
5 crystals, there are also continuing efforts to overcome the daunting technical hurdle of
6
7 achieving suitable protein crystallization conditions for entire classes of proteins for which
8
9 exceedingly few successful crystallizations have emerged, for example, membrane-spanning
10
11 entities such as G-protein coupled receptors (GPCRs). There are thus a finite number of
12
13 macromolecular targets for drugs and biologists are rapidly highlighting where blockage of a
14
15 protein binding site might beneficially influence disease.
16
17

18
19
20 However, even if the number of drug targets is quite limited, the number of small drug-like
21
22 molecules which might act as medicinal agents is almost limitless. Estimates have suggested
23
24 that perhaps as many as 10^{40} different drug-like compounds could be included even in the
25
26 molecular weight range between 150 and 800.
27
28

29
30 Computational chemistry can and is being employed to whittle down this vast number of
31
32 compounds to a handful which can be synthesised and screened biologically. In our much
33
34 publicised screensaver project we built a starting database of 3.5 billion candidate molecules,
35
36 starting with the couple of million commercially available and then using published
37
38 combinatorial chemistry schemes, before filtering the input information using Lipinski's
39
40 rules.[4]
41
42

43
44 Computational drug discovery can then be considered as a funnel with increasingly restrictive
45
46 and expensive filters.
47
48

49
50 Fig. 1
51

52
53 If we start with an enormous database and the crystal structure of a target protein with an
54
55 inhibitor bound in its active site, the first filter is shape. Only molecules with a shape similar
56
57 to the known inhibitor need be considered further. Recently we have developed some ultrafast
58
59 shape similarity methods which are capable of handling billions of starting compounds in a
60

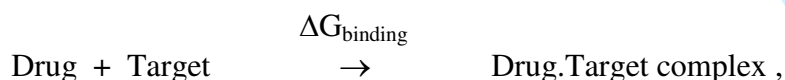
1
2
3 very short time.[5,6] Subsequent filtration may be performed using electrostatic similarity,
4
5 pharmacophore matching and molecular docking and scoring.
6
7

8 With the number of interesting compounds reduced to a few hundreds, then it becomes
9
10 possible to think in terms of accurate calculations which could bring us into the realm of
11
12 Schaefer-level accuracy.
13

14
15
16 At the same time it must be remembered that finding a small molecule which will inhibit a
17
18 protein is itself only a first stage which leads to a second funnel of myriad biological and safety
19
20 assays in the drug discovery process (e.g. off-target pharmacology, distribution, metabolism,
21
22 pharmacokinetics, etc). While accurate prediction of the binding energetics represents a great
23
24 advance, it is only the first step forward toward becoming a drug.
25
26
27
28
29
30

31 FREE ENERGY CALCULATIONS

32
33
34 What (as theoreticians are always wont to say) the experimentalist would really like to be told
35
36 would be the free energy of binding between the drug and the protein in aqueous solution,
37
38 since the bigger the binding energy the better the drug with the smaller the dose and likelihood
39
40 of side effects. Although 'in principle' (again as theoreticians often say) one should be able by
41
42 molecular dynamics calculation to compute the free energy for
43
44



45
46
47
48
49 in practice this is such a slow (multisecond) process that the direct route is impossible.
50

51
52
53 At the accurate level, the almost as useful computation of relative free energies of binding of
54
55 two drugs, say A and B, to the same target protein can be computed using the cycle illustrated
56
57 in Figure 2.
58
59
60

Fig. 2

1
2
3 There are instances where such differences in binding energy have been so calculated to an
4 accuracy of half a kilocalorie per mole, but in general despite its appeal this approach has not
5 proved as fruitful as had been first anticipated. Possible problems include conformational
6 changes in the protein target on binding the drug; difficulties in truly accounting for solvent
7 waters and deficiencies in the intermolecular potentials used.
8
9

10 To investigate some of these difficulties we are here going to consider a case which should be
11 interesting beyond being a model. It is a difference of fundamental importance in biology: the
12 relative binding of Mg^{2+} and Ca^{2+} to proteins.
13
14

25 4 MAGNESIUM AND CALCIUM IN BIOLOGY

26
27
28 These two simple ions are major players in biology, but despite their similarity behave very
29 differently, as summarized in Table 1.
30
31

32
33
34 Table 1

35
36 The only differences between the two ions are size and solvation and yet these can give rise to
37 dramatically different behaviour. Table 2 gives some experimental binding constants between
38 the pair of ions and some simple organic ions.[7]
39
40
41

42
43
44 Table 2

45
46 Note that sometimes Mg^{2+} binds better than Ca^{2+} and sometimes the reverse, with differences
47 amounting to as much as five orders of magnitude. Here is a good puzzle for someone wanting
48 to do accurate calculations.
49
50
51
52
53
54
55
56
57
58
59
60

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.

1
2
3 If calculations of binding energy difference between two spherical ions of the same charge to
4
5 simple partners do so badly, how can we hope to produce really useful binding energies of
6
7 similar drug molecules to proteins? It is important to understand why the calculations perform
8
9 so poorly.
10
11

12 13 14 15 16 6 SOURCES OF ERROR

17
18 One error which bedevils molecular simulation is the poor quality of the potentials used. They
19
20 do not reproduce simple measurable quantities. It should be remembered that if we look at the
21
22 simple Argon-Argon potential, we do not yet get the correct crystal structure of solid argon
23
24 unless three-body forces are included and these are invariably ignored in molecular mechanics.
25
26 In addition most calculations of molecular simulations are done in a box of a limited size so
27
28 that there has to be a cut-off in the potential. This is particularly damaging when ions are
29
30 involved since their interactions persist for long distances.
31
32

33
34 The greatest defect in the calculations described above is, however, the failure to incorporate
35
36 the inevitable transfer of charge which will be found when a charged species is involved.
37
38

39
40 Empirical potentials are also defective in incorporating cation- π electron interactions. The
41
42 way round those defects has to be to study the interactions using quantum mechanics which
43
44 allows the negative electron charge to be mobile.
45
46
47
48
49

50 51 7 CONCLUSION

52
53 Fritz Schaefer has been the leading light in the use of molecular quantum mechanics to resolve
54
55 chemical problems. In biology too, simplified methods are insufficient and we need quantum
56
57 chemistry. In the past this has been problematic because of the computer time required to use
58
59 *ab initio* methods. Now, however, innovations such as grid computing point the way to get
60

1
2
3 round the computational barrier. The techniques developed for small molecules can and should
4
5 be deployed in areas such as drug discovery.
6
7
8
9

10 **Acknowledgements**

11
12 The potential of mean force calculations were performed by Dr James Bradley and the research
13
14 was supported by the National Foundation for Cancer Research.
15
16
17
18
19
20
21

22 **References**

- 23
24
25 [1] Schaefer, H.F., and Harris, F.E., 1967, *Chem. Phys. Lett.* **1**, 407-408.
26
27
28 [2] Verhaegen, G., and Richards, W.G., 1966, *J. Chem. Phys.* **45**, 1828-1833.
29
30
31 [3] Richards, W.G., 1979, *Nature*, **278**, 507.
32
33
34 [4] Richards, W.G., 2002, *Nature Reviews Drug Discovery*, **1**, 551-555.
35
36
37 [5] Ballester, P.J., and Richards, W.G., 2007, *Proc. R. Soc. A.*, **463**, 1307-1321.
38
39
40 [6] Ballester, P.J., and Richards, W.G., 2007, *J. Comp. Chem.* **28**, 1711-1723.
41
42
43 [7] Tam, S.C., and Williams, R.J.P., 1985, *Structure and Bonding*, **63**, 103-151.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Legends for Figures
4
5
6
7

8
9 Fig. 1 The drug discovery funnel
10
11

12
13
14 Fig. 2 Accurate binding energies. Simulation using free energy perturbation
15
16

17
18
19
20 Fig. 3 Mg^{2+} - acetate. Potential of mean force
21
22

23
24
25
26 Fig. 4 Ca^{2+} - acetate. Potential of mean force
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

1
2
3 Tables
4
5
6
78
9 Table 1 Mg^{2+} and Ca^{2+} contrasts
10
1112
13
14 Table 2 Binding constants of Mg^{2+} and Ca^{2+}
15
1617 Log_{10} binding const (bigger is stronger)
18
1920
21
22
23 Table 3 Calculated binding constants for Mg^{2+} and Ca^{2+} to acetate
24
25
26
2728
29 Table 4 Calculated binding constants for Mg^{2+} and Ca^{2+} to glycine zwitterion
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

Table 1

 Mg^{2+} and Ca^{2+} 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+} Log₁₀ 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

	$\text{Log}_{10}K$	
	calc	expt
Mg^{2+}	0.8	0.8
Ca^{2+}	0.7	0.6

For Peer Review Only

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

For Peer Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

~100,000

Shape

~10,000

Electrostatics

~1,000

Pharmacophore

~100

Docking

10

Free energy of binding

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

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

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

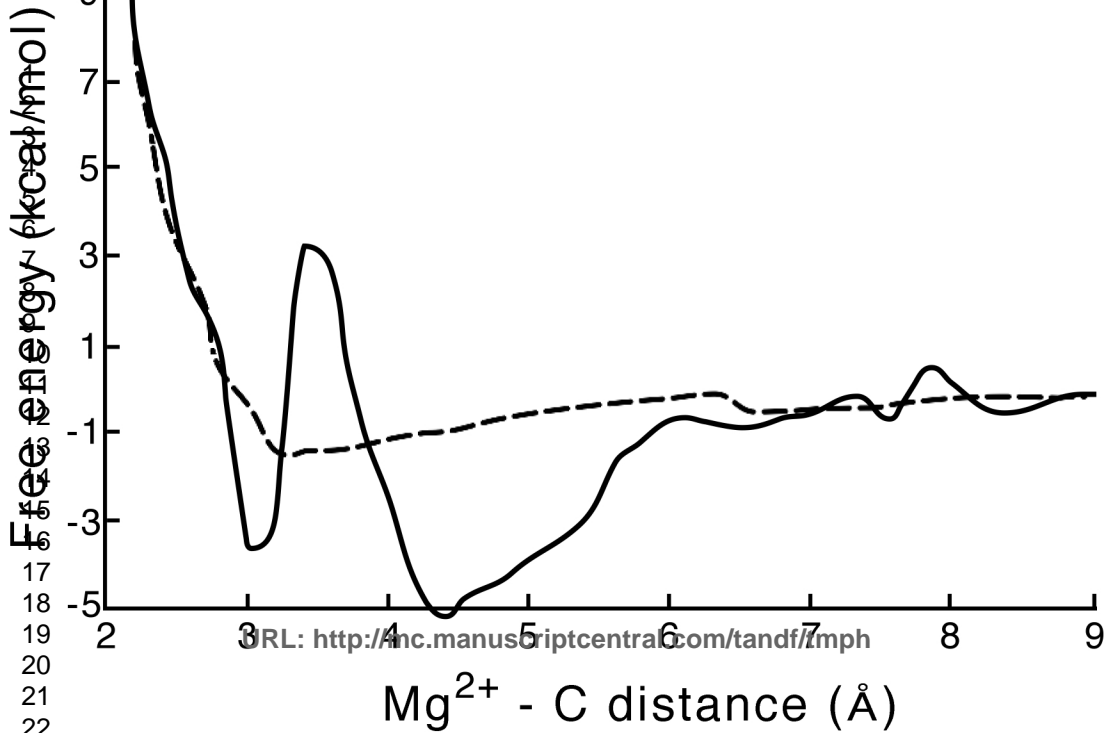
ΔG_3

ΔG_4

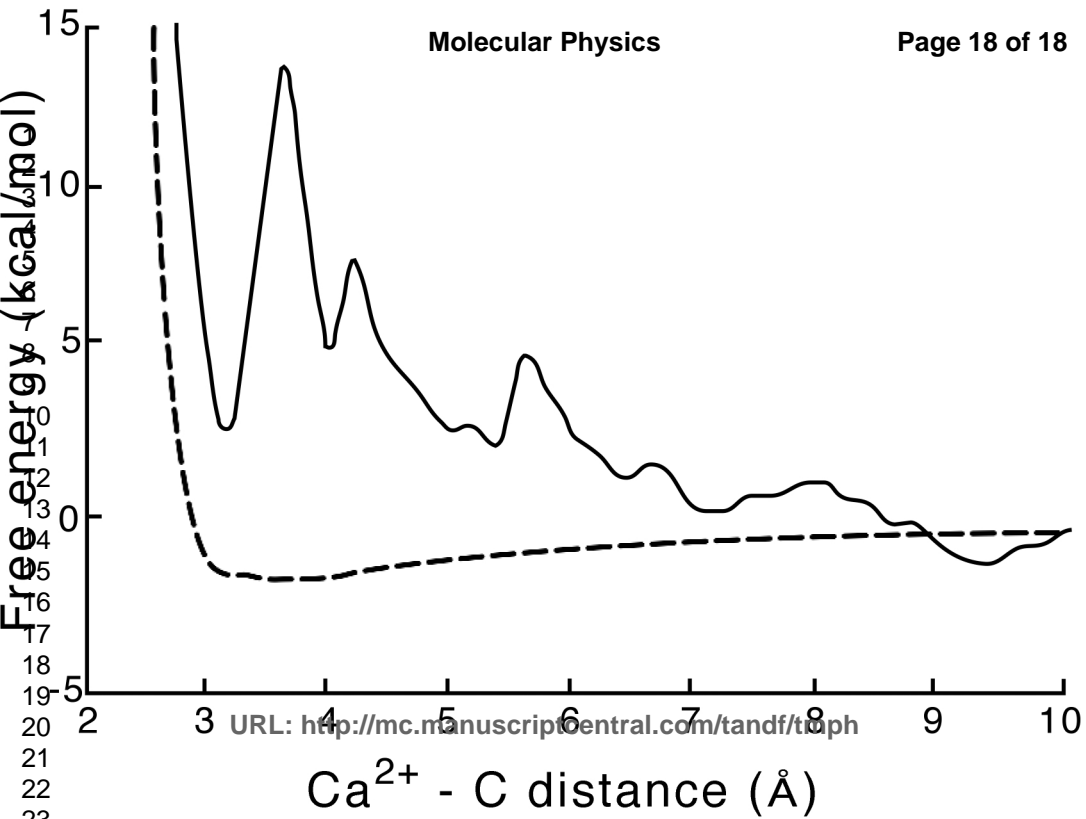
Target + Drug B $\xrightarrow{\Delta G_2}$ T.B complex

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

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



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

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