ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2011, *56*(4), 289-298

Dedicated to Professor Alexandru T. Balaban on the occasion of his 80th anniversary

REVIEW

QUANTUM CHEMISTRY APPROACHES TO LIGAND-RECEPTOR INTERACTIONS IN TIMIŞOARA

Liliana M. PĂCUREANU,[a*] Ludovic KURUNCZI[b] and Zeno SIMON[a]

[a] Computational Chemistry Department, Institute of Chemistry of Roumanian Academy, 24 Mihai Viteazul Ave., Timişoara, 300223, Timiş, Roumania, E-mail: zsimon@acad-icht.tm.edu.ro
[b] Chemical-Physics Department, Faculty of Pharmacy, University of Medicine and Pharmacy, 2 Eftimie Murgu, Timişoara, 300041, Timiş, Roumania, E-mail: kurunczi@umft.ro

Received September 23, 2010

An account of the most interesting results obtained at the Institute of Chemistry of Roumanian Academy in the last 5-10 years is given. They concern quantum chemical studies in water solution (saline bonds, Watson-Crick base pairs) and receptor-ligand docking studies. The focus is on comparison of calculated versus experimental results for different computational methods, on contribution of zero point energy and other corrections, on precision of computational results versus the required 1 kcal/mol for the, here pertinent, intermolecular interactions.

INTRODUCTION

Molecular recognition in receptor-ligand interactions is of primary importance for drug-design. X-ray crystallography studies yield the spatial structure of ever more receptor-ligand complexes and the interest to predict receptor-ligand affinities from chemical structure of ligand molecules is obvious.

The first steps for such predictions were the quantitative structure-activity relationships (QSAR) started in the sixties especially by the school of Hansch;[1] resolved structures of receptor-ligand complexes were extremely rare at that time. The 3D-QSAR methods, developed in the eighties-nineties, especially CoMFA, produce, implicitly an even if incomplete mapping of receptor binding sites;[2] despite interesting results, these are far from the required predictions of receptor-ligand affinities.

Quantum chemical calculations, including also methods to calculate hydration energies should be the answer for receptor-ligand interactions because such interactions take place, mostly, at water-protein interfaces. While quantum chemistry calculations, in vacuo, are successful in predicting unimolecular or bimolecular reaction rates for isolated molecule pairs,[3] the receptor-ligand systems seem to be too complex for presently available computing power. Even for interactions of relatively small molecules in water solution, existing results are not very encouraging, as predictions.

In Timişoara, at the Institute of Chemistry of Roumanian Academy, at Chemistry Departments of various universities, interest for quantum chemistry started in the early seventies. Quantum chemistry attempts to predict stabilities of some types of saline bonds implied in interactions between proteoglycans and attempts to predict

* Corresponding author: pacureanu@acad-icht.tm.edu.ro

stabilities, also in water solution, of Watson-Crick type nitrogen base-pairs, started in 2006. Results are somewhat surprising. The focus will be on comparison, whenever possible, with experimental data, comparison of results of various quantum-chemistry methods, on the effect of zero-point energy corrections, of the general precision of quantum-chemistry as compared to the precision of at least 1kcal/mol required by experimental results (here, rather in the intermolecular force domain).

COMPUTATIONAL METHODS

The PCM method[4] coupled with 6-31G* basis set *ab initio* or DFT quantum chemistry methods were used in most examples here described. This method has an iterative character: first charge distributions, geometry, etc., for the hydrated molecules are calculated *in vacuum*, followed by polarization of the surrounding water (considered as polarizable continuum with D=78.5); perturbation of the molecule's charge distribution by the polarized continuum (quantum chemistry calculations); recalculation of polarization of the surrounding water by the new charge distribution, etc.

For a dissociation process: AB→A+B, the free energy ΔG_d will be:

$$\Delta G_d = \Delta G(A) + \Delta G(B) - \Delta G(AB) \quad (1)$$

with:

$$\Delta G(A) = E_{tot}^{aq}(A) + ZPE(A) - T\Delta S(A) \quad (2)$$

and similarly for ΔG(B) and ΔG(AB). Here $E_{tot}^{aq}(A)$ represents the final energy of the molecule plus the polarizable continuum after the PCM - cycle, $ZPE(A)$ is the zero point energy correction of A (zero point vibrational energy) based on vibrational analysis of A, and $\Delta S(A)$ is the vibrational entropy contribution for A to which rotational and other components may be added. As to hydration free enthalpies, $\Delta G_h(A)$, these are calculated as differences between $E_{tot}^{aq}(A)$ in the polarized water and in vacuum $E_{tot}^{vac}(A)$ - for the minimal energy spatial configuration:

$$\Delta G_h(A) = E_{tot}^{aq}(A) - E_{tot}^{vac}(A) \quad (3)$$

Here, free electrons and nuclei are considered as zero for total energy $E_{tot}^{aq}(A)$ calculations. The method implies also some adjustable parameters, for example ionic and atomic van der Waals radii.

There are several packages based on different sets of basis formations, etc., for example the PC GAMESS-version 7.0[5] or Gaussian 03W.[6]

Other, more primitive, semiempirical methods here used:

Boundary element method (BEM) calculations were derived by a group of Canadian scientists.[7,8] The free energy of binding in water is calculated as the sum of molecular mechanics energy term ΔE_{bind}^{mm} (using, for example the AMBER 4.1 force field) and solvation free energy term ΔG_{bind}^{solv}:

$$\Delta G_{bind} = \Delta E_{bind}^{mm} + \Delta G_{bind}^{solv} = E_{int\,er}^{vdw} + E_{int\,er}^{coul} + \Delta E_{int\,ra}^{strain} + \Delta G_{bind}^{solv} \quad (4)$$

The solvation term ΔG_{bind}^{solv} is calculated (separately for the free molecules, ions, and for the complexes) using – at least for the highly charged species here considered – a continuum dielectric model with a solute interior dielectric constant D = 2 and for the surrounding water D = 78.5. The solving of the Poisson equation with the boundary element method (BEM) is required. Appropriate results depend strongly upon van der Waals (Born) radii selected for atoms and ions. One must remark that ΔG_{bind} is not calculated as a difference between (very large) G values for final and initial particles but as a semiempirical calculated "perturbation" between the interacting particles. Charge distributions for the molecules are calculated by quantum chemistry methods; these are not significantly changed by hydration, even in an iterative approach, as the molecules considered with D = 2 have a low degree of polarizability.

Combined PM3 plus BEM calculations was used in some of the studies described here. The semiempirical PM3 - method is considered most appropriate for complexes with hydrogen bonds;[9] the MNDO/PM3 method implemented in HyperChem 5.11 Pro[10] were here used. Solvation energies of particles or of the corresponding complexes were calculated by the BEM method of BEMcalc. Binding energies (ΔG_{bind} or ΔG_d) are differences between PM3 energies (in vacuum) for the particles, or the corresponding complexes, to which the differences between the BEM-hydration energies (free enthalpies, ΔG_h) are added.

SALINE BONDS MEDIATED BY CA^{2+} AND MG^{2+}

Dudev *et al.*[11] reviewed the biological roles of Ca and Mg ions considering the reaction of hydrated metal $[MW_6]^{2+}$ with biological ligands $HCOO^-$, $HCONH_2$, and imidazole. Binding constants of Mg^{2+} and Ca^{2+} with acetate, malonate, EDTA, etc. ions in aqueous solution were evaluated by Richards[12] by means of potential mean force. The lack of charge transfer incorporation and the inclusion of cut-off distances are the main drawbacks of this method. Meanwhile, the density functional theory and *ab initio* methods coupled with continuum dielectric method were able to give valuable information about physical basis of metal binding and selectivity of proteins by calculating the exchange free energy for metal-bound water for ligands of biological interest.[13-15]

Our interests were focused on saline bonds intercellular interactions mediated by peptidogycans that have a homospecific character. They imply saline bonds between Lewis base groups (-OH, etheric -O-, $-COO^-$, $-OSO_3^-$) and Ca^{2+} ions which cannot be substituted by Mg^{2+}.[16,17] Concerning experimental values for dissociation ΔG_d 's of such saline bonds, from dissociation - binding studies between oligomeric carbohydrates (including coated monolayers) in neutral water solution with 150mM NaCl and 10mM $CaCl_2$, Hernais *et al.*[18] inferred 6.5 to 8.5 kcal/mol per saline bond.

We performed ΔG_d - calculations for the following dissociation process:

$$[ML_2]^{2-2z} \rightarrow 2L^{z-} + M^{2+} \qquad (5)$$

with M^{2+}=Ca^{2+} and Mg^{2+}, L=CH_3COO^-; CH_3OSO_3; 1,2-ethanediol and 1,4-dioxane (with a BII-conformation) as model for groups forming saline bonds between peptidoglycan molecules. Our aim was to explain why saline bonds with Ca^{2+} are more stable than those with Mg^{2+} and to see if ΔG_d values are within a reasonable interval for such type of interactions (*i.e.* within ±10 kcal/mol).

As a molecular recognition problem, interactions between peptidoglycans (mediated by Ca^{2+} ions) have a marked homeospecific character.[16,17] This is also true for interactions between oligomeric carbohydrates.[18] The explanation may be quite simple: molecules of the same kind (shape) generally allow better contacts between themselves. Crystallization, also from mixed solutions, usually yields crystals containing a single kind of molecules.

In vacuo quantum chemistry results. It is certainly of interest to compare structural and electronic density results obtained by various *ab initio* and semiempirical quantum chemistry methods. Such studies, were performed for isolated molecules/complexes $ML_{2\ and}$ for isolated ligands by ZINDO, PM3 and *ab initio* with 3-21G and 6-31G* basis sets.[19,20] The HyperChem[10] and the GAMESS[5] packages were hereby used.

Structural data from X-ray crystallography and theoretical results of ZINDO, PM3, 3-21G and 6-31G* for bond lengths in the acetate ion[9] and in 1,4 dioxane[10] are listed in Table 1. As can be seen, the results of PM3, *ab initio* 3-21G, and 6-31G* are similar and close to experimental data, but those of ZINDO differ by about 0.5Å from experiment. For 1,4 dioxane experimental data correspond to the BII-form; data are corrected for zero point vibrational energy. Experimental geometry data from references 21 and 22.[21,22]

For electronic charges, results of ZINDO, *ab initio* 3-21G and 6-31G* calculations for the central atom of the ML_2 complexes are listed in Table 2.

Table 1

X-ray crystallography data (Å) and calculated bond lengths for acetate ion and 1,4 dioxane (adapted from ref. 19, 20)

Species/Bond	Exp	ZINDO	PM3	3-21G	6-31G*
Acetate					
C–O[a]	1.25[b]	1.31	1.26	1.25	1.23
C–C	1.511	1.48	1.54	1.58	1.55
C–H[a]	0.94[a]	1.10	1.10	1.09	1.09
1,4 - dioxane					
C–O[a]	1.44	1.39	1.42	1.44	1.40
C–C	1.522	1.49	1.53	1.52	1.52
C–H[a]	0.98	1.10	1.10	1.08	1.09

Superscript a indicates mean values; b indicates data corrected for zero point vibrations.

Table 2

Electronic charges for the central atom M in ML_2 complexes (data from ref. 19,20)

ML_2 complex	ZINDO	3-21G	6-31G*
$(CH_3COO)_2Ca$	1.01	1.48	1.43
$(CH_3COO)_2Mg$	0.39	1.15	1.06
$(CH_3OSO_3)_2Ca$	1.28	1.46	1.46
$(CH_3OSO_3)_2Mg$	0.65	1.10	1.04
$[ethanediol_2Ca]^{2+}$	1.47	1.62	1.63
$[ethanediol_2Mg]^{2+}$	0.45	1.35	1.46
$[1,4\text{-}dioxane_2Ca]^{2+}$	1.51	1.68	1.63
$[1,4\text{-}dioxane_2Mg]^{2+}$	0.44	1.35	1.33

The charges of the central atom (M) calculated by *ab initio* 3-21G and 6-31G* are similar, but results of the semiempirical ZINDO differ sensibly from *ab initio* results.

Dissociation energies (for $ML_2 \rightarrow M+2L$) *in vacuum*, calculated by ZINDO, *ab initio* 3-21G and 6-31G*(two variants) are listed in Table 3.

Table 3

Dissociation energies (kcal/mol) in vacuo for ML_2 complexes (data from ref.19,20)

ML_2 complex	ZINDO	3-21G	6-31G*[a]	6-31G*[b]
$(CH_3COO)_2Ca$	543	522	510	504
$(CH_3COO)_2Mg$	1012	646	613	602
$(CH_3OSO_3)_2Ca$	466	448	428	448
$(CH_3OSO_3)_2Mg$	870	588	529	538
$[(ethanediol)_2Ca]^{2+}$	152	226	185	347
$[(ethanediol)_2Mg]^{2+}$	614	342	271	485
$[(CH_2CH_2O)_2]_2Ca$	122	170	137	-
$[(CH_2CH_2O)_2]_2Mg$	742	282	218	-

ZINDO, 3-21G and 6-31G*[a] results are calculated with the HyperChem package,[10] and 6-31G*[b] results with the GAMESS package.[5]

As can be seen, the results of the three *ab initio* variants are similar, but those of the semiempirical ZINDO are significantly different. All ligands indicate larger dissociation energies for Mg^{2+} than for Ca^{2+} complexes and also for ML_2 complexes with anionic L-ligands than with neutral L-ligands. The dissociation energy interval for the three *ab initio* variants are in the range 150-600 kcal/mol, similar to the range 200-500 kcal/mol found by MM2-PM3 calculations for *in vacuo* dissociation of chitosan - Hg^{2+}, - Pb^{2+}, - Cd^{2+} complexes by Debbaudt, Ferreira and Gschaider.[23]

Hydration ΔG_h calculations. Semiempirical methods for ΔG_h calculations were reported by Florian and Warshell[24] and by Jorgensen *et al.*[25] in which both electrostatic (hydrogen-bonding) and hydrophobic interactions with the surrounding water are considered. Results are given in Table 4.

The methods we used for ΔG_h calculations[26,27] considered only electrostatic interactions. For the PCM model, ΔG_h are differences between energies calculated for the molecule in water and in vacuo. For Table 4 experimental data are from Raevsky and Shaper.[28] PCM[a] and PCM[b] - polarizable continuum method results with ionic radii by Wells[29] (rCa^{2+} = 1.00Å, rMg^{2+} = 0.72Å) and Sutton[30] (rCa^{2+} = 1.14Å, rMg^{2+} = 0.86Å).

Table 4

Calculated versus experimental hydration ΔG_h -values (kcal/mol). Adapted from references 26,27

Species	Exp	PCM[a]	PCM[b]	BEM
Na^+	–81.00	–132.60	–116.00	-
Ca^{2+}	–381.20	–552.50	–481.90	–360.60
Mg^{2+}	–455.50	–761.50	–634.80	–437.40
Cl^-	–77.00	–69.40	–73.40	-
CH_3COO^-	-79.90	-67.55	-	-77.40
$(CH_3COO)_2Ca$	-	1.66	–74.81	-
$(CH_3COO)_2Mg$	-	–5.91	–6.11	-
MeOH	–5.10	0.13	-	–6.60
EtOH	–5.01	2.47	-	–6.20
n-PrOH	–4.74	4.57	-	-
i-PrOH	–4.74	4.84	-	-

For a and b see Table 3.

From the methods considered, BEM calculations[7,8] gave the best results, probably due to parameterization of van der Waals contact distances. PCM with its cyclic procedure, implying also effects of polarizations on charge distribution, may viciate the results of good parameterization obtained for molecules/ions with different electronic structure.

Calculations of dissociation ΔG_d by PCM. Some results[27] are listed in Table 5. ΔG_d and ΔG_h are dissociation, respectively hydration free enthalpies calculated by PCM[4] with Wells atomic radii; ΔG_d contains the ZPE correction. ZPE denotes zero point energy correction, D denotes1,4 dioxane. There are no corrections for increase in entropy by dissociation (creation of new particles).

Table 5

Calculated ΔG_d dissociation energies for the $ML_2 \rightarrow M + 2L$ process in water, by the PCM-method (kcal/mol) (data from reference 27)

ML_2 complex	Total free energy in vacuum for ML_2	ΔG_d for dissociation	ΔG_h for ML_2 complexes	ZPE for ML_2 complexes	ZPE correction for the dissociation process
$(CH_3COO)_2Ca$	–709,872.54	–84.99	+1.66	68.86	–3.30
$(CH_3COO)_2Mg$	–410,469.94	–283.57	–5.91	69.76	–4.30
$[(ethanediol)_2Ca]^{2+}$	–711,679.57	–18.14	–303.22	118.84	–2.70
$[(ethanediol)_2Mg]^{2+}$	–412,212.44	–323.55	–169.51	119.86	–4.40
$(CH_3OSO_3)_2Ca$	–1,349,204.05	–118.14	–109.50	78.48	–2.50
$(CH_3OSO_3)_2Mg$	–1,049,698.34	–333.78	–12.52	79.20	–3.20
D_2Mg	–508619.51	–379.30	–149.85	169.76	–3.54

As can be seen ΔG_d values indicate Mg^{2+} complexes more prone to dissociation than Ca^{2+} complexes, in agreement with experiment, but all complexes should be unstable towards dissociation; ΔG_d's, in absolute values, are at least one order of magnitude above what is to be expected for this type of interaction. The reason is that total free energies for ML_2 complexes are of the order of 10^6 kcal/mol; to have a precision of at least 1kcal/mol for ΔG_d's, the calculated total free energies should have an accuracy of at least 10^{-6}, probably too much for the methods used by us. ZPE corrections for the dissociation process are significant, about 3 kcal/mol (as compared to 5-10 kcal/mol, experimentally[18] for dissociation process); the range of these ZPE corrections is of about -3.4 ± 1.0 kcal/mol, which is rather narrow.

WATSON-CRICK BASE PAIRS IN AQUEOUS SOLUTION

The influence of solvent on structural assembly and properties of DNA base pairs investigated by *ab initio*[31a,b] and molecular dynamics[32] revealed the importance of solvent bulk property changes and the hydrophilic interactions of base pairs with water molecules on the base pairs stability. *Ab initio* calculations[31b] demonstrated that in aqueous solution the canonical AT is the only dominant pair, but in the case of CG there are few tautomeric forms that are populated significantly. Kabeláč and Hobza[31a] showed that the solvent influence markedly the tautomeric equilibrium and spatial arrangement of base pairs. More recently, Fonseca Guerra *et al.*[33] succeeded to describe bond energies and distances of hydrogen bonds in adenine-thymine (AT) and cytosine-guanine (CG) pairs in water solution with the help of dispersion corrected functional BLYP-D.

Fidelity of nucleotide insertion by DNA-polymerase. This is a first step in DNA replication, and the hydrogen bonding between base-pairs (of the templating nucleotide and of the incoming nucleotide-triphosphate) was considered by Watson and Crick to be responsible for the fidelity of the replication process. Error rates for single base substitutions due to proofreading deficient DNA polymerase (thereafter corrected by proofreading, mismatch repair) are reported in the 10^{-3}-10^{-6} range[34]. If hydrogen bonding is responsible for this primary error-process, then thymine-1H-enol guanine (TG_{enol}) false pairing should be the main responsible for error (3 hydrogen bonds in the TG_{enol} pair, G to G_{enol} ratio in water solution is near one). Michaelis-Menten at low concentration kinetics suggest that quotients in nucleotide incorporation rates should thus be determined by the ΔΔG difference between stabilities in aqueous solution of the false TG_{enol} and the correct AT and CG base pairs. Thus, for the base pair stabilities in water solution, experimental data suggest:[35]

$$\Delta G_{AT} \approx \Delta G_{CG}; \qquad \Delta\Delta G = \Delta G_{TGenol} - \Delta G_{AT} \approx 4\text{-}8\text{kcal/mol} \tag{6}$$

We calculated the stability of ten base pairs in aqueous solution (Fig. 1), more exactly between 9-CH_3-purines and 5-CH_3-pyrimidines[36]. These base pairs were selected such as to present at least two hydrogen bonds and to respect the condition of 11 ± 0.5 Å distances between the C atoms of CH_3 of the 9-purine and 5-pyrimidine base pairs.

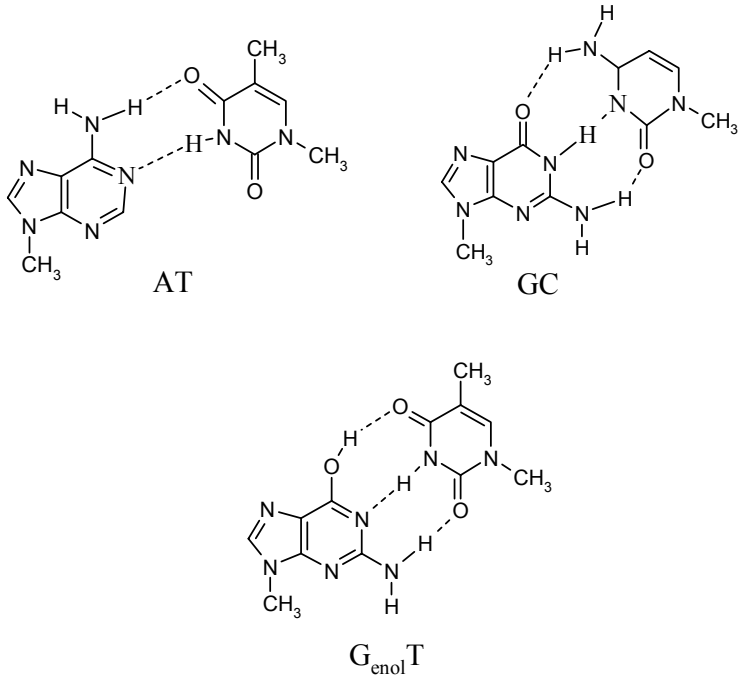

Fig. 1 – DNA base pairs presenting at least 2 hydrogen bonds and which observe the steric condition of 11 ± 0.5Å distance between the 9-CH_3 of purine and 5-CH_3 of pyrimidine.

Quantum chemistry calculations and results. Calculations were performed with various *ab initio* and DFT methods[37] coupled with PCM[4] for hydration. The base pair geometries and charge distributions for start were then provided by DFT B3LYP/6-31G*, with PCM (but also the more primitive BEM-method) being used for hydration ΔG_h calculation. The Gaussian 03W package[6] was used in quantum chemistry calculations. For hydration ΔG_h the PCM method was used, but also the simpler BEM-variant.

Table 6 gives the results of calculations for hydration energies and dipole moments for bases and base pairs compared with experimental data, where these were available.[38,40] Hydration energies ΔG_h^a, ΔG_h^b are calculated with DFT/B3LYP/6-31G* and respectively single point HF/6-31G*; ΔG_h^c-with BEM; μ calculations *in vacuo* with DFT/B3LYP/6-31G* (μ^a_{calc}) and single point *ab initio* HF/6-31G* (μ^b_{calc}) – data from reference 36.[36] Calculations are for 9-CH3 purine and 5-CH_3 pyrimidine derivatives. Experimental data were available for Watson-Crick bases and base-pairs in organic solvents (dioxane) for dipole moments of methylated bases.

Table 6

Calculated and experimental hydration energies and dipole moments (from ref. 36)

Base or base pair	Hydration energies (kcal/mol)				Dipole moments μ (D)		
	ΔG^a_h	ΔG^b_h	ΔG^c_h	ΔG^{exp}_h	μ^a_{calc}	μ^b_{calc}	μ_{exp}
A	-10.90	-12.74	-7.30	-13.6[38]	2.57	2.78	2.75[39]
T	-8.69	-11.29	-8.28		4.52	5.17	4.10[39]
C	-15.57	-19.13	-11.09		6.28	7.26	6.5[39]
G	-20.25	-23.89	-12.27		7.20	7.75	
AT	-9.05	-12.24	-10.84	-7.32[40]	1.86	2.42	
CG	-16.39	-20.45	-14.23		6.92	7.35	
TG_{enol}	-11.74	-15.38	-12.57		6.68	6.88	

We must mention that Bakalarsky *et al.*[41] reported DFT calculations with extended basis set for dipole moments of A, T, C, G and their methylated derivatives in good agreement with experiment, with a difference of only 0.1-0.3 D.[39] The HF/6-31G* calculations are single point calculations for geometries calculated by DFT B3LYP/6-31G*.

Results of calculations of total energies in aqueous solution (E_{tot}^{aq}), of ZPE - corrections and of entropy correction ($-T\Delta S_{TRV}$ at 298 K) for bases and base pairs are listed in Table 7. The entropy corrections calculated for isolated molecules (gas phase) are certainly not realistic for solutions.

Table 7

Calculated total energies, ZPE corrections and entropy corrections (data from reference 36)

Base or base pair	E_{tot}	ZPE		$-T\Delta S_{TRV}$	E^{aq}_{tot}
	kcal/mol (a)	kcal/mol (a)	kcal/mol (b)	(kcal/mol) (b)	(kcal/mol) (b)
A	–317928.32	79.65	80.48	94.59	–317916.59
T	–309653.98	81.57	82.17	95.73	–309643.17
C	–272506.04	71.85	72.71	88.10	–272490.05
G	–365145.90	82.09	83.15	98.97	–365125.07
G_{enol}	–365138.26	81.75	83.18	96.35	–365121.69
AT	–627591.08	162.52	163.82	152.92	–627576.00
CG	–637668.88	155.90	157.28	147.62	–637644.95
$G_{enol}T$	–674804.89	165.14	166.25	155.02	–674786.87

Results are in kcal/mol, with (a) B3LYP/6-31G* with PCM; (b) B3LYP/6-31G* with BEM.

As can be seen from Table 6 and Table 7 for total energies, hydration energies and various corrections, the order of these values for various bases and base pairs is the same for different methods but there are appreciable differences between the methods. For example, total energies are in the -272506 to -674804 kcal/mol range, and differences between methods are in the 10 to 20 kcal/mol range. For hydration ΔG_h, the intervals are from -9 to -20, respectively 2 to 4 kcal/mol.

Results for dissociation ΔG_d's and other data related to the dissociation process of the base pairs in water are given in Table 8 for B3LYP/6-31G* with PCM and respectively with BEM (for hydration energies). For base pairs, see Fig. 1.

Table 8

Calculated association - ΔG_d and other figures for base pairs in water solution (kcal/mol) (adapted from reference 36)

	ΔG_d		ΔE^{aq}_{tot}		ΔG_h		ΔZPE	
Base Pair	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
AA_{im}	–0.32	-	–5.71	-	4.06	-	1.33	-
AC_{im}	–6.32	14.50	–14.49	–1.84	6.81	16.34	1.36	1.93
AG	–11.16	-	–18.20	-	5.79	-	1.25	-
AT	–10.52	4.44	–16.24	–7.35	4.74	11.79	0.97	1.43
CC_{im}	–4.13	17.63	–12.28	0.78	6.90	16.85	1.24	2.04
CG	–19.27	7.75	–29.82	–14.79	9.14	22.54	1.42	2.15
CT	-8.78	7.98	–14.51	–4.60	4.81	12.58	0.91	1.67
GG_{enol}	–9.30	-	–19.20	-	8.72	-	1.18	-
$G_{enol}T$	–9.72	16.41	–18.63	–3.39	7.98	19.80	0.93	1.63
TT_{enol}	1.00	13.20	–3.18	4.22	3.69	8.98	0.49	1.08

Calculated with B3LYP/6-31G* and (a) BEM, respectively (b) PCM for hydration ΔG_h. Entropy term $T\Delta S_{TRV}$ is not included.

If base pairing by hydrogen bonds is the differentiating factor for nucleotide inclusion rates,[35] ΔG_d's should be approximately equal for the AT and CG-pairs, about 6 ± 2 kcal/mol higher for $G_{enol}T$ and sensibly higher for the other base pairs. As can be seen, neither of the two methods gives, from this point of view a correct order for the ΔG_d values. Other factors could also intervene: interactions of the DNA polymerase with various atomic groups of correct and false base pairs, interactions of incoming base pairs with proteins preexistent at the growing region of DNA double helix, etc.

The precision of the methods used here to calculate ΔG_d values could be too low. As seen from data of Table 6, differences among two methods in calculating individual E^{aq}_{tot} - values are in the 10÷20 kcal/mol range, and for the calculated E^{aq}_{tot} values in the range 0.3 to $0.7 \cdot 10^6$ kcal/mol, somewhat better for hydration ΔG_h (Table 6), while calculated ΔG_h values are in the –7 ÷ –20 kcal/mol. From Table 7, differences between ΔG_d - values calculated by two methods are in the range 5÷10 kcal/mol. The precision required to interpret experimental data (point mutations rates), as seen at the beginning of this paragraph dedicated to calculations for nitrogen bases and base-pairs, is of at least 1.0 kcal/mol.

Concerning the ΔZPE corrections for base pair formation, for the ten base pairs of Table 8, the values are within narrow ranges, 0.49 to 1.42 kcal/mol and 1.08 to 2.15 kcal/mol for B3LYP/6-31G* calculations with PCM, and 1.08 to 2.15 kcal/mol for those using BEM for hydration ΔG_h - calculations.

SOME INTERESTING RESULTS WITH MORE PRIMITIVE CALCULATION METHODS

A preliminary study with BEM was performed to calculate dissociation energies in aqueous solution for $(CH_3COO)_2Ca$ and $(CH_3COO)_2Mg$;[42] the results, ΔG_d of +8.54 kcal/mol and -3.12 kcal/mol, respectively are in agreement with the experimentally higher stability of saline bridges mediated by Ca^{2+} than of those mediated by Mg^{2+} and also with the stability inferred for such saline bonds.[18] Computations for the stability of the cobalt hexammine ion in aqueous solutions, *i.e.* for the process:

$$Co(NH_3)_6^{3+} + 6H_2O \rightleftharpoons Co(OH_2)_6^{3+} + 6NH_3 \quad (7)$$

were performed using an *ab initio* 3-21G - method for quantum chemistry calculations (vacuum) and the BEM method[7,8] for hydration energies.[43] From the experimental stability constant at 25 ^{o}C of this complex, ΔG_d = +31.8 kcal/mol[44] is inferred for equilibrium (5). The calculated[43] value is ΔG_d = 21.25 kcal/mol. For the unstable ferihexammine the lower, calculated ΔG_d = 18.03 kcal/mol is obtained.

Concerning stabilities of base pairs, ΔG_d, values in aqueous solution were calculated by the semiempirical MNDO/PM3 methods plus the BEM method for hydration ΔG_h for the AT, CG and TG_{enol} pairs.[35] The calculated ΔG_d for the base pairs were –4.32, –4.57 and 9.73 kcal/mol, respectively. Certainly, for the whole nucleotide incorporation process, several other interactions are present, but if these are equal for all three pairs and only the ΔG_d's are different, the experimental requirements (eq. 4) are met.

Semiempirical methods may give good results for the type of molecules or processes for which they were calibrated – see BEM calculations for usual intermolecular interactions, hydration energies, PM3 for hydrogen bonding. The ZPE corrections for dissociation of saline ML_2 bonds (Table 5) and are within a narrow range (2.5 to 4.4 kcal/mol). For equilibrium (5) molecules on both sides are of similar complexity, probably ZPE ≈ 0 for this process; for base pair dissociation processes (Table 8). ZPE values are also within a narrow range (0.5 to 1.4 kcal/mol).

CONCLUSIONS

The work here described is an attempt to contribute to the problem of universal, fast and accurate prediction of biological activities. To quote the title of Moitessier *et al.*[45] "Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go", this is true also for the general problem of prediction of biological activities.

Our studies concerning saline bond and hydrogen bond complexes in water solution indicate that, at least for the quantum chemical methods used by us, their precision is not sufficient for the about 1 kcal/mol value required by the weak forces implied in ligand-receptor complexes. Free enthalpies for complex formation appear as calculated small differences between large numbers (formation enthalpies in the 10^5-10^6 kcal/mol range).

REFERENCES

1. C. Hansch, *Drug Dev. Res.,* **1981**, *1*, 267-309.
2. R.D. Cramer III, D.E. Patterson and J.D. Bunce, *J. Am. Chem. Soc.,* **1988**, *110*, 5959-5967.
3. D.C. Clary, *Science,* **2008**, *301*, 789-791.
4. J. Tomasi, M. Persico, *Chem. Rev.,* **1994**, *94*, 2027 - 2094.
5. A. A. Granovsky, PC GAMESS version 7.0, http://classic.chemmser/gran/games/ index.html.
6. Gaussian 03, Revision B.01, 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, J. 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. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A.Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2003.
7. E. O. Purisima, *J. Comput. Chem.,* **1998**, *19*, 1494-1504.
8. T. Sulea, E. O. Purisima, *J. Phys. Chem. B,* **2001**, *105*, 889-899.
9. J. J. P. Stewart, *J. Comput. Chem.,* **1989**, *10*, 207 - 220.
10. HyperChem 5.11 Pro, ChemPlus 1.6, HyperCube Inc., Gainesville, FL, www.hyper.com.
11. T. Dudev, C. J. Lim, *Chem. Rev.,* **2003**, *103*, 773-787.
12. W. G. Richards, *Mol. Phys.,* **2009**, *107*, 819-822.
13. T. Dudev, C. Lim, *J. Phys. Chem. B* 2000, *104*, 3692-3694.
14. T. Dudev, C. Lim, *Annu. Rev. Biophys.,* **2008**, *37*, 97-116.
15. T. Dudev, C. Lim, *J. Am. Chem. Soc.,* **2006**, *128*, 1553–1561.
16. U. Dammer, O. Popescu, P. Wagner, H. I. Anselmetti, H. I. Guntherodt, G. M. Misevich, *Science,* **1995**, *267*, 1173-1175.
17. O. Popescu, I. Checiu, P. Gherghel, Z. Simon, G. N. Misevics, *Biochimie(Biochemistry),* **2003**, *85*, 181-188.
18. M. J. Hernais, J. M. de la Fuente, A. G. Barrientos, S. A Penades, *Angew. Chem. Int. Ed.,* **2002**, *41*, 1554-1557.
19. V. Careja, M. Mracec, L. Sayti, E. Sisu, R. Tudose, Z. Simon, *Rev. Roum. Chim.,* **2006**, *51*, 379-380.
20. V. Careja, S. Muntean, M. Mracec, L. Sayti, Z. Simon, *Int. J. Quantum Chem.,* **2007**, *107*, 1714-1718.
21. R. B. Helmholdt, E. J. Sonneveld, J. Schenk, *Z. Kristallogr.,* **1988**, *213*, 696-601.
22. J. Buschmann, E. Muller, P. Luger, *Acta. Cryst.,* **1986**, *C42*, 873-896.
23. A. L. Debbaudt, M. L. Ferreira, M. E. Gschaider, *Carbohydr. Polym.,* **2004**, *56*, 321-332.
24. J. Florian, A. Warshel, *J. Phys. Chem. B,* **1999**, *103*, 10282-10288.
25. W. L. Jorgensen, J. P. Ulmschneider, J. Tirado-Rives, *J. Phys. Chem. B,* **2004**, *108*, 16264-16270.

26. S. Muntean, L. Kurunczi, C. Bologa, G. Ilia, Z. Simon, *Macromol. Symp.,* **2006**, *235*, 215-219.
27. S. Muntean, L. Kurunczi, V. Careja, Z. Simon, *Rev. Roum. Chim.,* **2007**, *52*, 1111-1114.
28. O. A. Raevsky, R. J. Schaper, *QSAR Comb. Sci.,* **2004**, *22*, 926-942.
29. A. F. Wells, "Structural Innorganic Chemistry", 5th ed., Clarendon Press, Oxford, 1984, p. 506-507.
30. L. Sutton, "Tables of Interatomic Distances and Configuration in Molecules and Ions", Special Publication No 11 and 18, The Chemical Society, London, 1965, p. 450.
31. a) M. Kabélác, P. Hobza, *Phys. Chem. Chem. Phys.,* **2007**, *9*, 903-917; b) J. Rejnek, P. Hobza, *J. Phys. Chem. B,* **2007**, *111*, 641-645.
32. V. I. Danilov, T. van Mourik, *Mol. Phys.*, **2008**, 1-8.
33. C. Fonseca Guerra, T. van der Wijst, J. Poater, M. Swart, F. M. Bickelhaupt, *Theor. Chem. Acc.,* **2010**, *125*, 245-252.
34. T. A. Kunkel, K. Bebnek, *Annu. Rev. Biochem.,* **2000**, *69*, 497-536.
35. E. Seclaman, L. Kurunczi, Z. Simon, *Biochemistry (Moscow),* **2007**, *72*, 328-331.
36. L. Pacureanu, Z. Simon, *Int. J. Quantum Chem*., **2010**, *110*, 1295-1305.
37. W.Koch, M.C. Holthemsen, "A Chemists Guide to Density Functional Theory", Wiley/VCH, Weinheim, 1999, p. 213-234.
38. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.,* **1983**, *4*, 187-189.
39. I. Kulakowska, M. Geller, B. Lesyng, K. Bolewska, K.L. Wierzchowski, *Biochim. Biophys. Acta,* **1975**, *407*, 420-429.
40. S. J. Gill, D. B. Martin, M. Downing, *J. Am. Chem. Soc.,* **1963**, *85*, 706-708.
41. G. Bakalarski, P. Grochowski, J. S. Kwiatkowski, B. Lesyng, J. Leszczyński, *Chem. Phys.,* **1996**, *204*, 301-311.
42. T. Sulea, Z. Simon, *Internet Electron. J. Mol. Des.,* **2002**, *1*, 59-69.
43. L. Sayti, V. Careja, S. Muntean, R. Tudose, O. Costisor, Z. Simon, *Rev. Roum. Chim.,* **2007**, *52*, 299-300.
44. J. A. Plambeck, Chemical Data Tables. Stability Constants of aqueous Complex Ions, 1996, http://www.ualberta.ca/~jplambeck/che/data/p00408.htm.
45. N. Moitessier, P. Englebienne, D. Lee, J. Lawandi, C. R. Corbeil, *Br. J. Pharmacol.,* **2008**, *153*, S7-26.