Predicting hydrophobic alkyl-alkyl and perfluoroalkyl-perfluoroalkyl van der Waals dispersion interactions and their potential involvement in protein folding and the binding of drugs in hydrophobic pockets

Clifford W Fong

Predicting hydrophobic alkyl-alkyl and perfluoroalkyl-perfluoroalkyl van der Waals dispersion interactions and their potential involvement in protein folding and the binding of drugs in hydrophobic pockets

Clifford W. Fong
Eigenenergy, Adelaide, South Australia.

Email: cwfong@internode.on.net

Abstract

It has been shown that a model compound which has previously allowed the experimental determination of the hydrophobic van der Waals dispersion alkyl-alkyl folding free energy interaction in a wide range of polar and non-polar solvents is strongly correlated with the cavitation, dispersion, solvent-structure (CDS) solvation free energy and to a lesser extent the molecular polarizability of the foldamers.
The predicted folding alkyl-alkyl free energy interaction in water calculated from the unfolded and folded foldamers by dispersion corrected quantum mechanical modeling is in agreement with the calculated values derived from solution experiments. Electrostatic potential surfaces in solution can differentiate amongst foldamers.
The solute-solvent dispersion interaction overrides the solvent free intramolecular alkyl-alkyl dispersion interaction, but the latter may still be a significant effect deep within the hydrophobic interior of a folded protein.
The analogous perfluoroalkyl-perfluoroalkyl foldamer interaction shows different behavior in water compared to n-hexane where the unfolded state is more stable in water but the folded state is more stable in n-hexane.
The use of the SMD solvent model using water and non-polar solvents to examine protein folding processes may have advantages over more intensive explicit solvation studies.

Keywords: Protein folding, alkyl-alkyl dispersion interactions, perfluoroalkyl-perfluoroalkyl interactions, foldamers, solvophobicity, hydrophobicity, CDS, electrostatic potential surfaces, polarizability, solvent-foldamer dispersion, cavitation, repulsion, quantum mechanics

Abbreviations

Total free energy of water solvation ΔG_{Sol}, SMD cavity dispersion solvent structure of the first solvation shell CDS, free energy CDS $\Delta G_{CDS,}$ isotropic polarizability Pol, electrostatic potential surfaces EPS, solute-solvent interaction: ΔG_{Dis} dispersion, ΔG_{Cav} cavity and ΔG_{Rep} repulsive free energies, free energy of dispersion interaction in solvent $\Delta G_{Interact, sol,}$ multiple correlation coefficient R^2, the F test of significance, standards errors for the estimate (SEE) and standard errors of the variables ΔG_{CDS} $SE(\Delta G_{CDS})$, molecular polarizability $SE(\Delta G_{Pol})$

Introduction

The hydrophobic effect is important in understanding protein folding since minimizing the number of hydrophobic side chains exposed to water is the principal driving force behind the folding process, although formation of hydrogen bonds within the protein also stabilizes protein structure. Water-soluble proteins have a hydrophobic core in which side chains are buried away from water, and charged and polar side chains are situated on the solvent-exposed surface where they interact with surrounding water molecules.

Molecular dynamics modeling of protein folding and foldamers is well advanced, but one area which is still problematical is the role of solvation in water which is highly intensive using explicit water models. Implicit salvation models based on sound physical properties can reduce the calculational intensity. The incorporation of hydrophobicity into force fields would also advantageous.

Solute−solvent interactions can be divided into (i) the bulk electrostatic contribution associated with the polarization free energy of the solute−solvent system, and (ii) the non-bulk-electrostatic contribution, called the CDS which includes cavitation, dispersion, and solvent-structure effects where the solvent structure contributions include effects such as hydrogen bonding solvent structure breaking, exchange repulsion, solute−solvent charge transfer, dielectric saturation, and electrostriction effects. The lumped term non-electrostatic CDS describes the first solvent shell and this shell rather than the bulk solvent properties would be expected to be the dominant solvatochromic environment influencing the interior regions of a folded protein, or for a drug accessing a hydrophobic pocket of a biological protein or enzyme. The widely used SMD solvent model allows the calculation of the CDS for many solvents.

The solvophobic effect applies to interactions between polar solvents (particularly water in biological systems) and non-polar solutes. Pure solvents have strong cohesive forces between the solvent molecules due to hydrogen bonding or other polar interactions, so non-polar solutes tend to be insoluble in polar solvents because these solvent-solvent binding interactions must be overcome first. The classical hydrophobic effect involved in protein folding and drug protein binding in hydrophobic pockets includes solvophobic effects.

A recent comprehensive study of the hydrophobic alkyl-alkyl interaction using a molecular model (see Figure 1) which allows the folding and unfolding of two alkyl chains of the molecule in a wide variety of non-aqueous solvents has shown that the intramolecular hydrophobic van der Waals alkyl-alkyl interaction is smaller than the solute-solvent dispersion interaction. The crystal structure of the model compound has shown that the folded conformation is the most stable state. The analogous perfluoroalkyl-perfluoroalkyl interaction has been compared to the alkyl-alkyl interaction using the same molecular scaffold has also been studied in solvents, and it has been concluded that solvophobic effects were most important in aqueous and the highly polar organic

solvents, but differences in van der Waals dispersion forces were important in apolar organic and fluorous solvents. Dispersion interactions in apolar solvents were found to be strongly attenuated in solution compared to the gas phase, but still play a major role in fluorous and organic solvents. [1,2] Mecinovic [3] found that alkyl and perfluoroalkyl chains had similar hydrophobicities in carbonic anhydrase binding.

The importance of hydrophobic alkyl-alkyl interactions when drugs bind to proteins (protein-ligand interactions) is well known. [4-7] The hydrophobic effect related to a single methyl group of a drug initially exposed to water that gets buried within a hydrophobic binding pocket of a protein has a predicted stabilizing free energy gain of ca 21-25 cal/mole per $Å^2$ which translates into an affinity enhancement of ca 50 fold. [8] Fluorinated drugs are known to confer major pharmacological effects compared to their alkyl analogues particularly influencing potency, selectivity, absorption and metabolism. The impact of fluorine substitution on lipophilicity (log P and log D), cell membrane permeability and pKa can be large but variable. These effects can be attributed to the size of the fluorine atom (van der Waals radius of 1.47Å), compared to hydrogen atom (van der Waals radius of 1.20Å), the high electron withdrawing ability of fluorine, the greater stability of the C-F bond compared to the C-H bond, and the greater lipophilicity of fluorine compared to hydrogen. [9]

It is noted that lipophilicity and hydrophobicity are often used interchangeably in the pharmaceutical industry, but they are not exactly the same, and one definition used in the drug industry is that *lipophilicity = hydrophobicity – polarity*. [10] Hence molecular polarizability of the drug in solution might be expected to influence solvophobicity.

Non-covalent dispersion interactions (commonly called van der Waals or London interactions) can arise from induced dipoles between neutral molecules that bear no permanent dipolar attraction or from an intramolecular mutual recognition folding processes in proteins for example. Dispersive interactions originate from quantum mechanical correlated motions of electrons in approaching molecules or from intramolecular moieties in close proximity. These distance related forces are electrostatic in nature. [7]

Study Objectives:

To determine if a relationship can be found between the free energy of alkyl-alkyl interaction and free energy of solvation as it applies to the first solvent shell of the model molecule.
To determine whether a solvent mediated model can predict the folding alkyl-alkyl and perfluoroalkyl-perfluoroalkyl interaction in an aqueous environment as would be applicable in biological processes such as protein folding or drug binding in hydrophobic pockets of enzymes or related biological entities.

Results

Solvation studies of alkyl-alkyl interaction

Table 1 shows the calculated free energies of solvation for the model compound (Figure 1) in the folded state with the experimentally determined free energies of alkyl-alkyl interaction (ΔG_{DMC} and ΔG_{Exp}) from reference 1 in 15 solvents. ΔG_{DMC} and ΔG_{Exp} are two different methods used by Yang et al to experimentally determine the free energies of alkyl-alkyl interaction [1,2]. Correlations were sought between the experimental free energy alkyl-alkyl interactions (ΔG_{DMC} and ΔG_{Expt}) and the total free energy of solvation (ΔG_{Solv}), the CDS free energy (ΔG_{CDS}) the using the SMD solvent model, the dipole moment and molecular volume of the model compound. The only significant relationships found are shown in equations 1 and 2 which are very similar, with eq 1 being the most precise and significant.

Equation 1 Relationship between the alkyl-alkyl interaction $\mathbf{\Delta G_{DMC}}$ and solvent $\mathbf{\Delta G_{CDS}}$ of the model compound for 15 solvents

$\mathbf{\Delta G_{DMC}} = -1.43\mathbf{\Delta G_{CDS}} - 19.17$
Where $R^2 = 0.764$, SEE = 0.168, SE(ΔG_{CDS}) = 0.009, F=42.15, Significance=0.0000

Equation 2 Relationship between the alkyl-alkyl interaction $\mathbf{\Delta G_{Exp}}$ and solvent $\mathbf{\Delta G_{CDS}}$ of the model compound for 15 solvents

$\mathbf{\Delta G_{Exp}} = -1.47\mathbf{\Delta G_{CDS}} - 22.64$
Where $R^2 = 0.668$, SEE = 0.220, SE(ΔG_{CDS}) = 0.012, F=26.16, Significance=0.0002

Polarizability is the ability to induce instant electric dipole moments in a molecule, which can apply to solvents and solutes. Polarizability plays an important role in determining the hydrophobic force acting between weakly polar surfaces. [11, 12] Linear correlations were found with ΔG_{DMC} and ΔG_{Expt} with the isotropic molecular polarizability of the model compound in the various solvents. Equations 3 and 4 illustrate the relative contributions of ΔG_{CDS} and the polarizability (raw values have been scaled to allow direct comparison of the relative magnitudes of the CDS salvation and the polarizability). It can be seen that the solvation effect is about 4 to 10 times as great as the polarization effect.

Equation 3 Relationship between the alkyl-alkyl interaction ΔG_{DMC} and isotropic polarizability Pol of the model compound in 15 solvents

$\mathbf{\Delta G_{DMC}} = -1.56\mathbf{\Delta G_{CDS}} + 0.16\mathbf{Pol} - 29.71$
Where $R^2 = 0.771$, SEE= 4.13, SE(ΔG_{CDS})=0.31, SE(Pol)= 0.26, F=20.19, Significance=0.00014

Equation 4 Relationship between the alkyl-alkyl interaction ΔG_{Expt} and isotropic polarizability Pol of the model compound in 15 solvents

$\mathbf{\Delta G_{Expt}} = -1.88\mathbf{\Delta G_{CDS}} + 0.51\mathbf{Pol} - 56.81$
Where R^2 =0.725, SEE= 4.98, SE(ΔG_{CDS})=0.37, SE(Pol)= 0.32, F=15.83, Significance=0.00043

Dispersion studies of alkyl-alkyl and analogous perfluoroalkyl-perfluoroalkyl interactions

Solvent effect on alkyl-alkyl and perfluoroalkyl-perfluoroalkyl interactions ($\Delta G_{Interact, sol}$)

The alkyl-alkyl and perfluoroalkyl-perfluoroalkyl interactions of the folded model compound (Figure 1) in water were determined by (1) optimizing the model compounds in water or other solvent, (2) then rotating around the aryl-aryl bond as shown in Figure 1(b) to unfold the model molecule as shown in Figure 2, (3) then creating fragments (X=Y=H) for both the folded and unfolded states, (4) then calculating the gas phase ΔG values for all entities using the dispersion corrected wB97XD functional with the ONIOM lower level 6-31G(d) for the fragment as per (3) and 6-311++G(d,p) for the high level alkyl-alkyl or perfluoroalkyl-perfluoroalkyl fragments. Finally (5) using thermodynamic cycles, the solvation energy (ΔG_{CDS}) was added to derive the $\Delta G_{Interact, sol}$.

The following $\Delta G_{Interact, sol}$ values were obtained:

Alkyl-alkyl interaction: $\Delta G_{Interact, water}$ -1.11 kcal/mol, $\Delta G_{Interact, hexane}$ 0.27 kcal/mol

The alkyl-alkyl calculated values for water and n-hexane can be compared to the extrapolated experimental value for water of -1.01 kcal/mol and experimental value 0.03-0.14 kcal/mol for n-hexane [1,2] The extrapolated value for water from equations 1 and 2 are 0.14-0.29 kcal/mol

Figures 2-5 show the calculated electrostatic potential surfaces (EPS) for the alkyl-alkyl model in the unfolded and folded states in gas, water and n-hexane solvents. EPS were calculated in two modes, firstly from the total SCF electron density, and secondly from electron densities from CHELPG partial atomic charges. The CHELPG EPS's show a region of electron density around the end of the alkyl chains, in the unfolded state, and with greater density in the folded alkyl-alkyl chains in the gas, water (Figures 3 and 4) and particularly showing increased density for the n-hexane solvent (Figure 5). The electrostatic potential surface EPS from the total SCF electron density for example is shown in Figure 2 as $-7.425e^{-2}$ to $+7.245e^{-2}$ in water, or alternatively by showing the EPS at different electron densities from CHELPG partial atomic charges, (green $-1.662e^{-2}$ to $+1.662e^{-2}$. The wide variability of the ranges of electron density in the EPS shown in Figures 2-11 illustrates the extent of differentiation that can be resolved by using EPS.

Fluoroalkyl-fluoroalkyl interaction: $\Delta G_{Interact, water}$ not applicable, $\Delta G_{Interact, hexane}$ -2.71 kcal/mol

The experimental value for the perfluoroalkyl model in n-hexane was -0.01 [1,2]. It was found that the non-folded perfluoroalkyl foldamer was more stable in the gas phase (and water) than the folded conformation by a range of 11.3 – 26.29 kcal/mol depending on whether the foldamer is fully relaxed or transitioning from the less stable folded state. This result is undoubtedly due to the higher polarity of the perfluoroalkyl compound compared to the alkyl model compound. Inspection of Figures 6 to11 which show the unfolded and folded perfluoroalkyl foldamers which clearly illustrates the higher electron density on the perfluoroalkyl moieties compared to that shown for the alkyl chains in Figures 2-5. Figure 7 for the perfluoroalkyl foldamer in water shows the influence of solvated water on the EPS compared to Figures 6 or 8 in the gas phase or n-hexane solvent.

Solvent-solute dispersion, solvent-solute cavity formation, solvent-solvent repulsion

It was previously found by Cockroft [1,2] that solvophobic effects were most important in aqueous and the highly polar organic solvents, but differences in van der Waals dispersion forces were important in apolar organic solvents. Dispersion interactions in apolar solvents were found to be strongly attenuated in solution compared to the gas phase, but still play a major role in organic solvents. In light of the strong correlations found in eqs 1 and 2 between the ΔG_{CDS} and the experimental dispersion interactions in solution, it is instructive to examine water-foldamer dispersion interaction separately from the SMD ΔG_{CDS} correlation. The solvent PCM model (which is the basis for the SMD solvent model) allows the calculation of the solvent-foldamer interaction in water and n-hexane using the method of Floris and Pierotti (see experimental) which separates the solute-solvent interaction into the ΔG_{Dis} dispersion, ΔG_{Cav} cavity and ΔG_{Rep} repulsion free energies.

G_{Dis} values for alkyl-alkyl foldamers folded in water were -50.20, unfolded in water -59.21, and folded in hexane -43.18 and unfolded in hexane -50.72 kcal/mol. ΔG_{Dis} for the unfolded and folded perfluoroalkyl-perfluoroalkyl foldamers in water were -56.94 and -57.14 kal/mol. It is clear that the difference in ΔG_{Dis} |-9.01| kcal/mol between the unfolded and folded alkyl-alkyl foldamers in water (offset by the $+\Delta G_{Cav}$ and $+\Delta G_{Rep}$) is a significant factor in stabilizing the folded conformation in water, compared to the negligible difference in ΔG_{Dis} |~ 0| kcal/mol between the unfolded and folded perfluoroalkyl-perfluoroalkyl foldamers in water. This result is consistent with the finding above in calculating $\Delta G_{Interact,\ water}$ that the non-folded perfluoroalkyl foldamer was more stable in the gas phase and water than the folded conformation. These findings confirm and clarify earlier findings [1,2] that solvent-foldamer dispersion interactions override intramolecular alkyl-alkyl and especially fluoroalkyl-fluoroalkyl interactions. The difference in ΔG_{Dis} |-7.54| kcal/mol between the unfolded and folded alkyl-alkyl foldamers in the more solvophilic hexane compared to the value of |-9.01| in solvophobic water may be indicative of the interactions experienced in a more hydrophilic interior of a protein compared to the more hydrophobic water-exposed exterior.

Discussion

We have previously used the ΔG_{CDS} value to determine the free energy of desolvation in water (defined as $\Delta G_{desolv,CDS}$) and the lipophilicity of drugs (ΔG_{CDS} in octane or octanol, defined as $\Delta G_{lipo,CDS}$) in a number of quantitative structure activity studies of various biological activities involving drug-cell receptor binding and drug-cell membrane transport activities. [13-28] Lipophilicity as determined by $\Delta G_{lipo,CDS}$ is linearly related to logP (or logD) and other experimental measures of lipophilicity [21]. ΔG_{CDS} is a lumped term description of the non-electrostatic first solvent shell which incorporates the dielectric constant, refraction index, hydrogen bonding, surface tension, the carbon aromaticity, and fraction of non-hydrogenic solvent atoms that are F, Cl, or Br.[29]

A major difficulty exist in differentiating the intramolecular dispersion interaction between folded alkyl groups in solvents, and the solvent-solute dispersion of foldamers. The highly parameterized ΔG_{CDS} appears to be a good measure of most factors which can influence solvent-solute interactions, particularly in the critical first solvent shell. The relationships between the experimental alkyl-alkyl dispersion interactions in solution and ΔG_{CDS} shown in eq 1 and 2 indicate that ΔG_{CDS} is a good measure of solvent-solute interactions. Eq 3 and 4 show that the molecular polarizability of the foldamers is also influenced by the polarizability of the solvents used to determine the experimental ΔG_{DMC} and ΔG_{Expt} as expected. Cockcroft showed that ΔG_{DMC} and ΔG_{Expt} were linearly related to the cohesive energy density and polarizability of the solvents, both of which are bulk solvent factors, rather than first solvation shell factors. [1,2]

The calculated $\Delta G_{Interact,\ water}$ and $\Delta G_{Interact,\ hexane}$ values derived from the unfolded and folded foldamers are in good agreement with the experimental ΔG_{DMC} values derived from Wilcox molecular balances double-mutant cycle approach, and the alternative ΔG_{Expt} approach. [1,2] This congruence gives added confidence that all approaches are fundamentally sound.

The independent derivation of the solvent-foldamer ΔG_{Dis} interactions in water and hexane clearly shows that solvent-foldamer dispersion interactions outweigh intramolecular alkyl-alkyl van der Waals dispersion interactions, particularly in water the medium most applicable to biological systems. This aspect could only be inferred from the $\Delta G_{DMC\ and}$ ΔG_{Expt} studies. In so far that a hexane solvent environment may approximate a hydrophobic pocket deep inside a protein, and the dominance of solvent-solute dispersion interactions over that for an intramolecular alkyl-alkyl dispersion interaction, then it follows that the removal of water from a hydrophobic binding pocket is required before any hydrophobic binding can occur.

Conclusion

It has been shown that the hydrophobic van der Waals dispersion alkyl-alkyl folding free energy interaction in a wide range of polar and non-polar solvents is strongly correlated with the cavitation, dispersion, solvent-structure (CDS) solvation free energy and to a lesser extent the molecular polarizability of the foldamer.

The predicted folding alkyl-alkyl free energy interaction in water calculated from the unfolded and folded foldamers by dispersion corrected quantum mechanical modeling is in agreement with the calculated values derived from solution experiments. Electrostatic potential surfaces in aqueous and hexane solution are very useful in differentiating amongst foldamers.

The solute-solvent dispersion interaction overrides the solvent free intramolecular alkyl-alkyl dispersion interaction, but the latter may still be a significant effect deep within the hydrophobic interior of a folded protein.

The analogous perfluoroalkyl-perfluoroalkyl foldamer interaction shows different behavior in water compared to n-hexane where the unfolded state is more stable in water but the folded state is more stable in n-hexane.

The use of the continuum SMD solvent model using water and non-polar solvents to examine protein folding processes may have advantages over more intensive explicit solvation studies. Quantum mechanical ONIOM methods using the SMD solvent model can be used to examine fine details of dispersion interactions at a high level of theory, while describing less critical molecular interactions at lower levels of theory such as using molecular mechanics.

Experimental

All calculations were carried out using the Gaussian 09 package. Energy optimizations were at the DFT/wB97XD/6-31G(d) (6d, 7f) levels of theory for all atoms in all solvents. Selected optimizations at the DFT/wB97XD/6-31^{+}G(d,p) (6d, 7f) level of theory gave very similar results to those at the lower level. Optimized structures were checked to ensure energy minima were located, with no negative frequencies. Single point solvent and gas calculations were at the DFT/wB97XD/6-31G(d) (6d, 7f) or DFT/wB97XD/6-311++G(d,p) (6d, 7f) levels of theory.

Frequency ONIOM calculations of the folded and unfolded conformations as shown in Figure 1(a) and (b) were conducted at the low level at DFT/B3LYP/6-31G(d,p) (6d, 7f) and at the high level of theory at DFT/wB97XD/6-311G++(d,p) (6d, 7f) with optimized geometries in solvents, using the SMD solvent model. ONIOM calculations at the low level at DFT/wB97XD/6-31G(d,p) (6d, 7f) and at the high level of theory at DFT/wB97XD/6-311G++(d,p) (6d, 7f) gave very similar results. The high level ONIOM calculations were applied to the n-heptyl, n-hexyl and perfluorohexyl side chains of the folded and unfolded conformations as shown in Figure 1(a) and (b). Partial atomic charges were calculated using the CHELPG method with energy calculations where atomic charges are fitted to reproduce the molecular electrostatic potential at a number of grid points around the molecule.

With the 6-31G(d) basis set, the SMD model achieves mean unsigned errors of 0.6 - 1.0 kcal/mol in the solvation free energies of tested neutrals and mean unsigned errors of 4 kcal/mol on

average for ions. [29] The 6-31G(d,p) basis set has been used to calculate absolute free energies of solvation and compare these data with experimental results for more than 500 neutral and charged compounds. The calculated values were in good agreement with experimental results across a wide range of compounds. [30,31] Adding diffuse functions to the 6-31G(d) basis set (ie 6-31^{+}G(d,p)) had no significant effect on the solvation energies with a difference of less than 1% observed in solvents, which is within the literature error range for the IEFPCM/SMD solvent model.

Solvents in the SMD solvent model are defined using seven physical factors that determine the non-electrostatic CDS term: the dielectric constant, the square of the refraction index, hydrogen bond acidity and basicity, the macroscopic surface tension at a liquid-air interface, the carbon aromaticity or fraction of non-hydrogenic solvent atoms that are aromatic carbon atoms, and the electronegative halogenicity or fraction of non-hydrogenic solvent atoms that are F, Cl, or Br. PCM solvent ΔG_{Cav} cavitation, ΔG_{Dis} dispersion and ΔG_{Rep} repulsive energies were calculated according to the methods of Floris [32] and Pierotti [33]

For *alkyl-alkyl foldamers* (kcal/mol) ΔG_{Cav} values folded in water were 93.34, unfolded in water 101.39, folded in hexane 65.28, unfolded in hexane 69.67. ΔG_{Dis} values for alkyl-alkyl foldamers folded in water were -50.20, unfolded in water -59.21, folded in hexane -43.18, unfolded in hexane -50.72. ΔG_{Rep} values for alkyl-alkyl foldamers folded in water were 2.77, unfolded in water 3.43, folded in hexane 2.21, unfolded in hexane 2.70.

For *perfluoroalkyl-perfluoroalkyl foldamers* (kcal/mol) ΔG_{Cav}, ΔG_{Dis} and ΔG_{Rep} values in water were 100.01, -57.14 and 3.42 for the folded conformation, and 108.66, -56.94 and 4.09 for the unfolded conformation respectively.

Dipole isotropic polarizabilities (W= 0.000000) were calculated in Bohr^3 in water.

It is noted that high computational accuracy for each species in different environments is not the focus of this study, but comparative differences between various related species is the aim of the study. The literature values for ΔG_{DMC} and ΔG_{Exp} used to develop the multiple regression LFER equations have much higher experimental uncertainties than the calculated molecular properties. The statistical analyses include the multiple correlation coefficient R^2, the F test of significance, standards errors for the estimates (SEE) and each of the variables ΔG_{CDS} SE(ΔG_{CDS}), molecular polarizability Pol SE(ΔG_{Pol}) as calculated from “t” distribution statistics. Residual analysis was used to identify outliers.

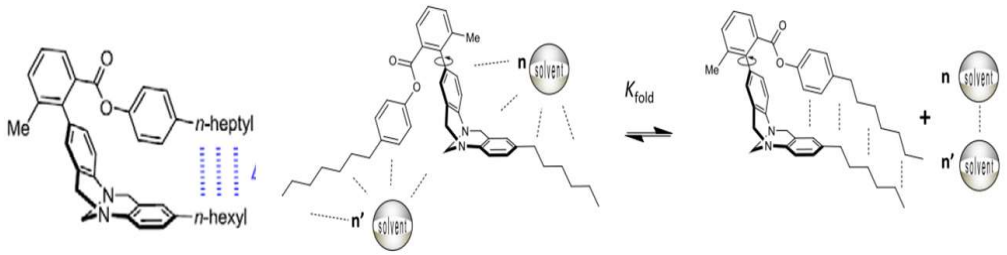

Figure 1(a) Model compound used to evaluate the folded n-heptyl-n-hexyl interaction (left) and representation of folding process showing the unfolded and folded interchange with solvent mediation (right) adapted from reference 1

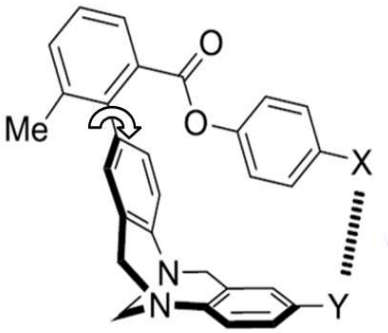

Figure 1(b) Model compound used to evaluate the folded alkyl-alkyl interaction (X=n-heptyl, Y=n-hexyl), and folded perfluoralkyl-perfluoroalkyl interaction (X=Y=perfluoro-n-hexyl); also X=Y=H fragment used to calculate van der Waals alkyl-alkyl and perfluoroalkyl-perfluoroalkyl interactions, with the circular arrow showing aryl-aryl bond rotation used to unfold the molecule

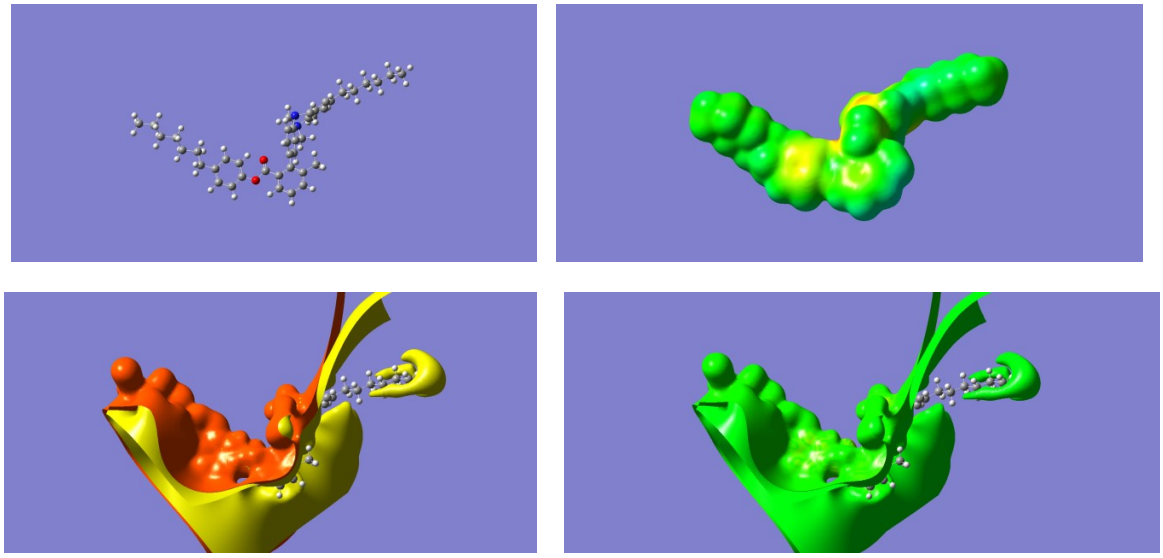

Figure 2. Unfolded conformation of model alkyl-alkyl compound (top left) and electrostatic potential surface EPS (from total SCF electron density, -7.425e^{-2} to +7.245e-^{2}) (top right) in water, lower images showing EPS at different electron density from CHELPG partial atomic charges, (green -1.662e^{-2} to +1.662e^{-2}. (high electron density to low density: red-yellow-green-light blue-dark blue)

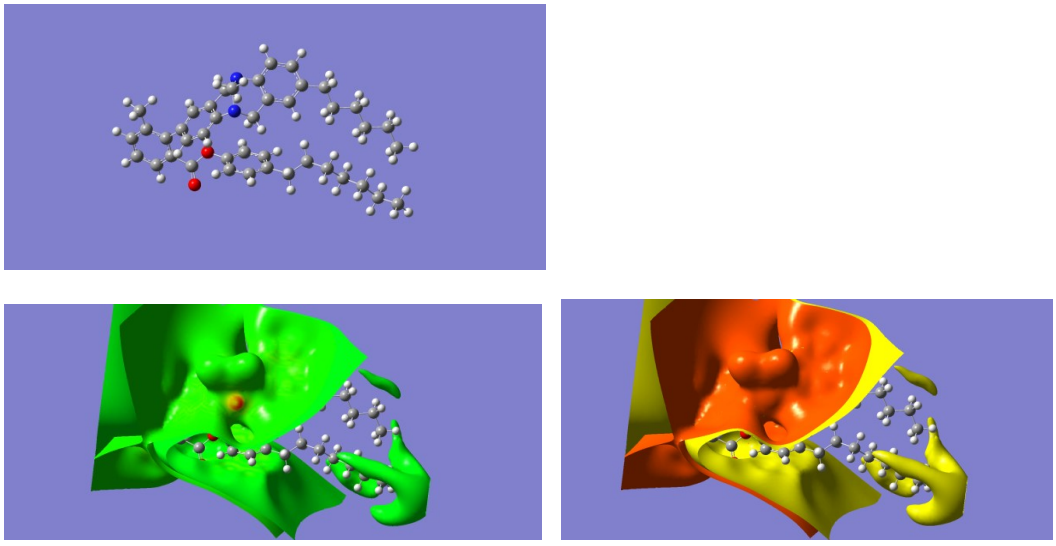

Figure 3. Folded conformation of model alkyl-alkyl compound gas phase (top) and two images (lower) showing two EPS, at different electron densities from CHELPG partial atomic charges, (green image from -1.286e^{-2} to +1.286e^{-2})

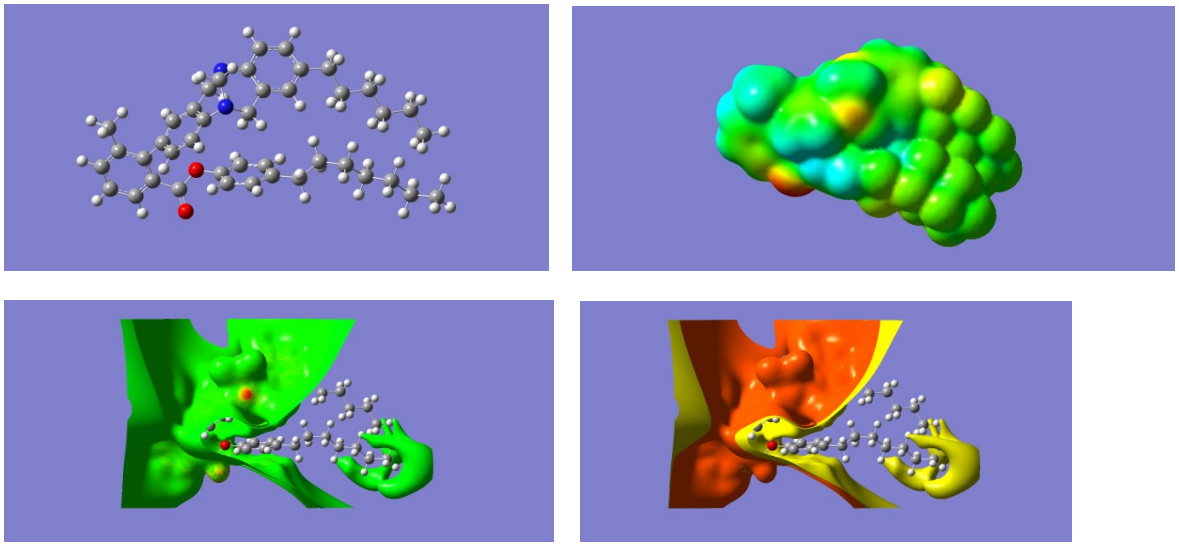

Figure 4. Folded conformation of model alkyl-alkyl compound in water and EPS from total SCF electron density (top) and two images (lower) showing two EPS at different electron densities, generated from CHELPG partial atomic charges (green image from $-1.756e^{-2}$ to $+1.756e^{-2}$)

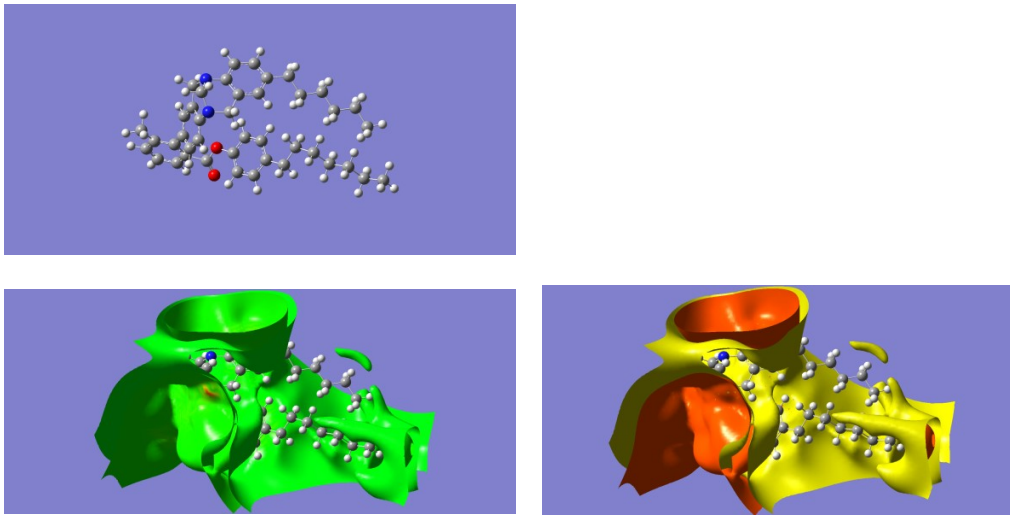

Figure 5. Folded conformation of model alkyl-alkyl compound in n-hexane (top) and lower two images showing two different EPS, at different electron density, generated from CHELPG partial atomic charges (green image from $-1.527e^{-2}$ to $+1.527e^{-2}$)

Figure 6. Unfolded conformation of model perfluoro-compound gas phase optimized in n-hexane (top left), EPS from total SCF electron density, ($-4.509e^{-2}$ to $+4.509e^{-2}$), and two images (lower left and right) showing two EPS at different electron densities from CHELPG partial atomic charges, (green image from $-1.329e^{-2}$ to $+1.329e^{-2}$) Fluorine atoms are light blue

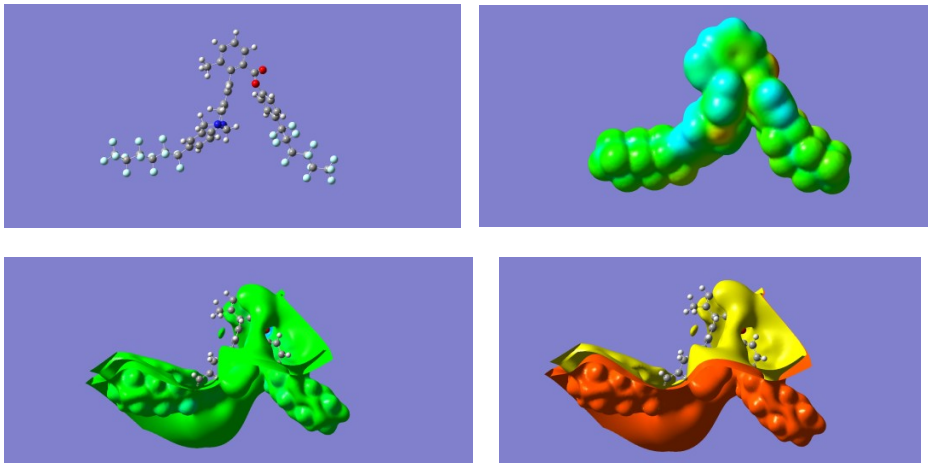

Figure 7. Unfolded conformation of model perfluoro-compound water phase optimized in water (top left), EPS from total SCF electron density, ($-6.023e^{-2}$ to $+6.023e^{-2}$), and two images (lower left and right) showing two EPS at different electron densities from CHELPG partial atomic charges, (green image from $-9.179e^{-2}$ to $+9.179e^{-2}$)

Figure 8. Unfolded conformation of model perfluoro-compound optimized in n-hexane, EPS of two images (left and right) showing two EPS at different electron densities from CHELPG partial atomic charges, (green image from $-1.400e^{-2}$ to $+1.400e^{-2}$

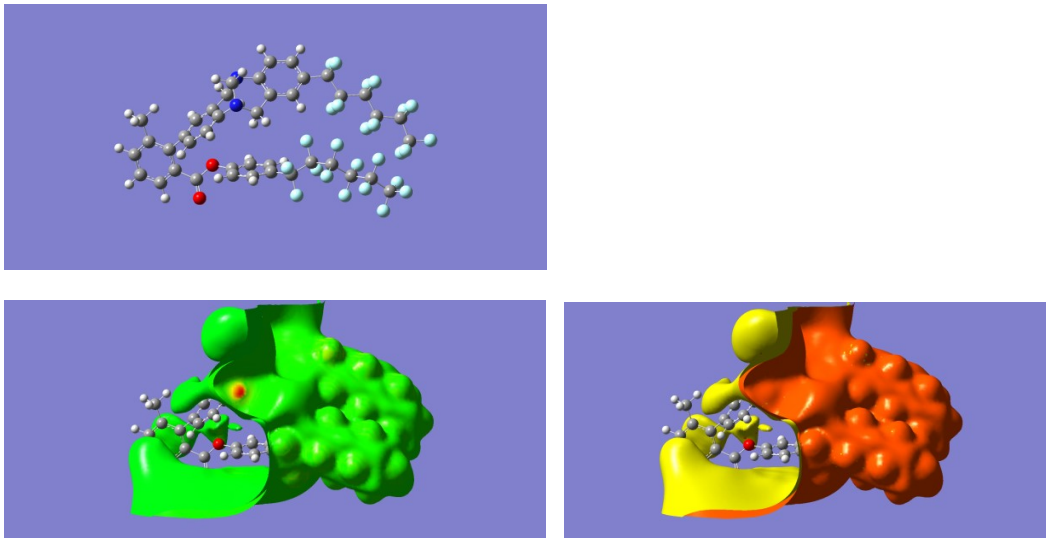

Figure 9. Folded conformation of model perfluoro-compound gas phase (top) and two images (lower) showing two EPS at different electron densities from CHELPG partial atomic charges, (green image from -1.385e^{-2} to +1.385e^{-2})

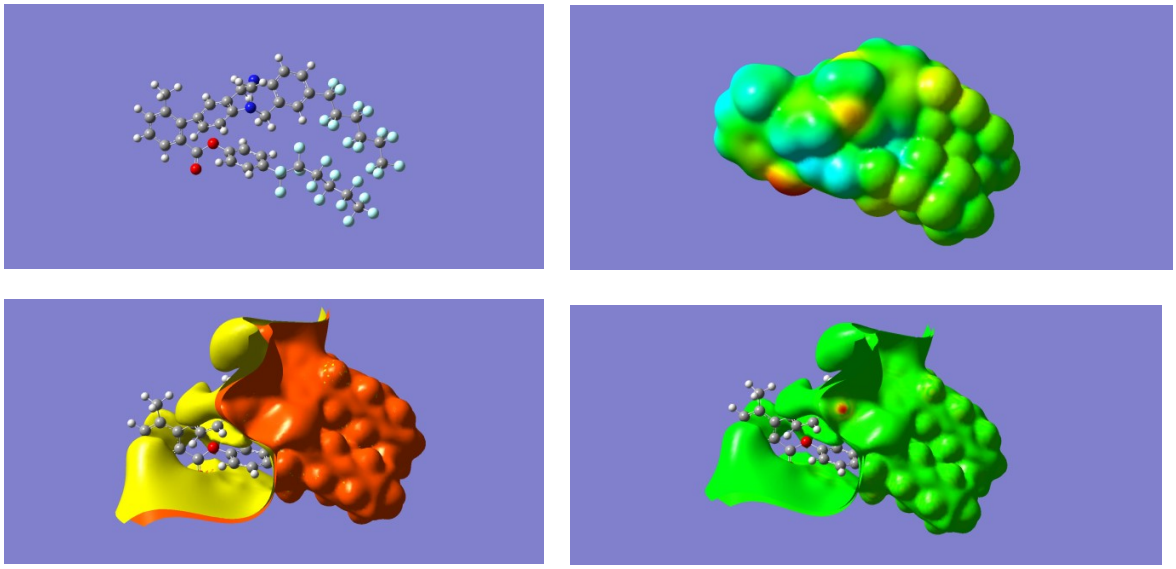

Figure 10. Folded conformation of model perfluoro-compound in water (top left), EPS from total SCF electron density (-5.838e^{-2} to +5.838e^{-2}), and two images (lower left and right) showing two EPS at different electron densities from CHELPG partial atomic charges, (green image from -1.815e^{-2} to +1.815e^{-2})

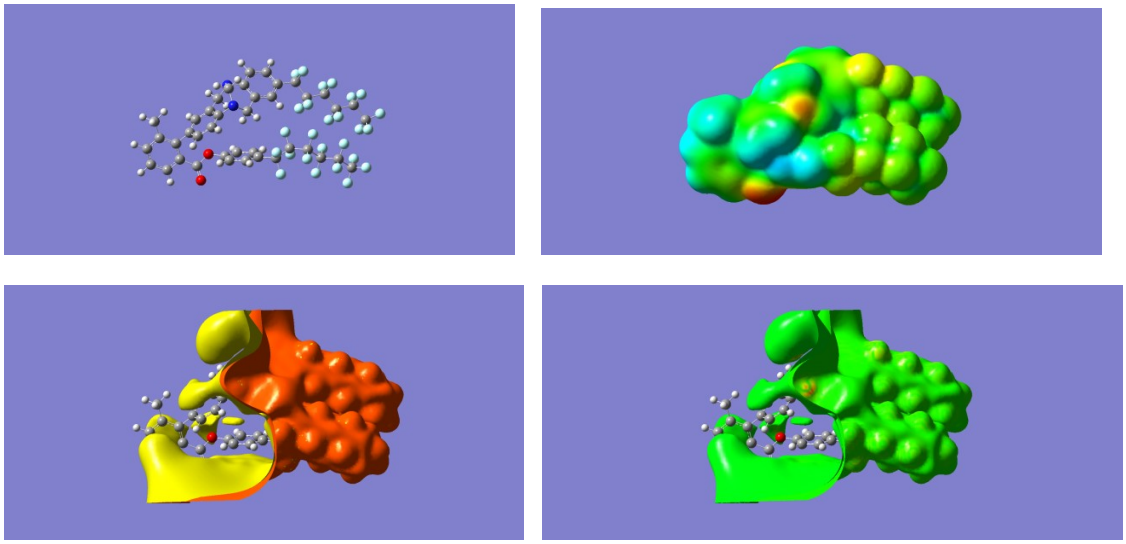

Figure 11. Folded conformation of model perfluoro-compound in n-hexane (top left), EPS from total SCF electron density, and two images (lower left and right) showing two EPS at different electron densities from CHELPG partial atomic charges, (green image from $-1.345e^{-2}$ to $+1.345e^{-2}$)

Table 1.

Solvent	ΔG_{DMC}	ΔG_{Exp}	ΔG_{Solv}	ΔG_{CDS}	Molec Dipole	**Molec Volume**	**Molec Polariz -ability**
Water	-28.68*	-13.15*	-15.52	6.55	2.27	502	73.5
Acetonitrile	3.35	-5.02	-33.87	-10.58	2.4	432	62.3
Acetone	-3.35	-9.56	-45.84	-12.16	2.67	542	44.9
Methanol	-5.74	-11.00	-26.9	-7.3	2.62	403	61.8
Ethanol	-9.86	-13.62	-51.52	-8.32	3	482	61.7
Dimethyl Sulphoxide	-19.36	-19.84	-32.24	-4.39	2.23	530	62.5
Carbon Disulphide	4.54	9.32	-51.04	-17.86	2.43	434	52.1
Carbon Tetrachloride	5.98	6.21	-50.7	-20.4	2.3	520	51.4
Chloroform	4.06	2.63	-33.33	-16.69	2.25	540	56.2

Dichloromethane	2.87	3.11	-36.54	-16.51	2.43	524	59.1
Benzene	4.54	8.60	-35.57	-19.26	2.23	536	51.5
Perfluorobenzene	9.08	-7,65	-29.73	-19.07	2.2	510	50.6
Pyridine	0.96	5.02	-52.78	-12.77	2.49	502	60.2
Tetrahydrofuran	1.67	-0.72	-39.26	-12.52	1.93	509	58.3
Cyclohexane	3.35	2.15	-29.21	-16.32	1.98	522	50.5
Hexane	13.86	3.35	-38.22	-18.3	1.96	495	50.1

Footnotes: ΔG_{DMC} and ΔG_{DMC} from reference 1, scaled to 100 times kcal/mol. ΔG_{Solv} and ΔG_{CDS} in kcal/mol. Molecular dipole in D, molecular volume in cm^3/mol, isotropic molecular polarizability scaled to 0.1 times $Bohr^3$. Values in red for water are calculated values from equations 1 and 2.

References

[1] L Yang, C Adam, GS Nichol, SL Cockroft, How much do van der Waals forces contribute to molecular recognition in solution, Nature Chem, 2013, 5, 1006-1010, and supplementary section DOI: 10.1038/NCHEM.1779

[2] C Adam, L Yang, SL Cockroft, Partitioning Solvophobic and Dispersion Forces in Alkyl and Perfluoroalkyl Cohesion Angewandte Chemie International Edition, 2015, 54, 1164-1167.

[3] J Mecinović, PW Snyder, KA Mirica, S Bai, ET Mack, RL Kwant, DT Moustakas, A Heroux, GM Whitesides, Fluoroalkyl and Alkyl Chains Have Similar Hydrophobicities in Binding to the "Hydrophobic Wall" of Carbonic Anhydrase, J Am Chem Soc. 2011, 133, 14017–14026.

[4] C Camilloni, D Bonetti, A Morrone, R Giri2, CM. Dobson, M Brunori, S Gianni, M Vendruscolo, Scientific Reports, 2016, 6, 28285, DOI: 10.1038/srep28285
Towards a structural biology of the hydrophobic effect in protein folding

[5] HJ Dyson, PE Wright, HA Scheraga, The role of hydrophobic interactions in initiation and propagation of protein folding, PNAS, 2006, 103, 13057–13061

[6] R Patil, S Das, A Stanley, L Yadav, A Sudhakar, AK Varma Optimized Hydrophobic Interactions and Hydrogen Bonding at the Target-Ligand Interface Leads the Pathways of Drug-Designing, PLOS One. 2010, https://doi.org/10.1371/journal.pone.0012029

[6] N Murshid, X Wang, Hydrophobic Effect of Alkyl Groups Stabilizing Self-Assembled Colloids in Water, J. Phys. Chem. B 2017, 121, 6280-6285

[7] K Daze, S Hof, Molecular interaction and recognition, Encyclopedia of Physical Organic Chemistry, First Edition. Edited by Zerong Wang, 2017 John Wiley & Sons, Inc. ISBN 978-1-

118-46858-6, ISBN: 9781118470459, Online ISBN: 9781118468586, DOI: 10.1002/9781118468586
[8] R Milo, R Phillips, Cell Biology by the Numbers, Garland Science, Taylor and Francis, 2015, Ch 3, page 204.
[9] P Shah, AD Westwell, The role of fluorine in medicinal chemistry, J Enzyme Inhib Med Chem, 2007, 22, 527-540,
[10] B Testa, G Caron, P Crivoria, S Reya, M Reist, PA Carrupt, Lipophilicity and Related Molecular Properties as Determinants of Pharmacokinetic Behaviour, Chimia, 2000, 54, 672-677
[11] F Bresme, A Wynveen, On the influence of solute polarizability on the hydrophobic interaction, J. Chem. Phys. 2007, 126, 044501
[12] MH Newt, BJ Berne, Molecular Dynamics Calculation of the Effect of Solvent Polarizability on the Hydrophobic Interaction, J Am Chem Soc.1995, 117, 7172-7179
[13] CW Fong, Free radical anticancer drugs and oxidative stress: ORAC and CellROX- colorectal cancer cells by quantum chemical determinations, HAL archives, hal-01859315v1, 2018
[14] CW Fong, Role of stable free radicals in conjugated antioxidant and cytotoxicity treatment of triple negative breast cancer, HAL archives, hal-01803297v2, 2018
[15] CW Fong, The role of free radicals in the effectiveness of anti-cancer chemotherapy in hypoxic ovarian cells and tumours, hal archives, hal-01659879v2
[16] CW Fong, Free radicals in chemotherapy induced cytotoxicity and oxidative stress in triple negative breast and ovarian cancers under hypoxic and normoxic conditions, HAL archives, hal-01815246v1
[17] CW Fong, Role of free radicals in PARP inhibition of cervical carcinogenesis: PARylation of eLa cells, and inhibition of hPARP-1 enzyme**,** HAL archives, hal-01859316v11
[18] Clifford Fong. Platinum anti-cancer drugs: free radical mechanism of Pt-DNA adduct formation and anti-neoplastic effect**,** Free Radical Biology and Medicine, 2016, 95:216-229.
[19] Clifford Fong. Platinum based radiochemotherapies: Free radical mechanisms and radiotherapy sensitizers**,** Free Radical Biology and Medicine, 2016, 99:99 - 109.
[20] CW Fong, The extravascular penetration of tirapazamine into tumours: a predictive model of the transport and efficacy of hypoxia specific cytotoxic analogues and the potential use of cucurbiturils to facilitate delivery, Int J Comput Biol Drug Design. 2017, 10:343-373
[21] CW Fong, Permeability of the Blood–Brain Barrier: Molecular Mechanism of Transport of Drugs and Physiologically Important Compounds, J. Membr. Biol. 2015, 248:651-69.
[22] CW Fong, The effect of desolvation on the binding of inhibitors to HIV-1 protease and cyclin-dependent kinases: Causes of resistance, Bioorg Med Chem Lett. 2016, 26:3705–3713.
[23] CW Fong, Physiology of ionophore transport of potassium and sodium ions across cell membranes: Valinomycin and 18-Crown-6 Ether. Int J Comput Biol Drug Design 2016, 9: 228-246.

[24] CW Fong, Statins in therapy: Understanding their hydrophilicity, lipophilicity, binding to 3-hydroxy-3-methylglutaryl-CoA reductase, ability to cross the blood brain barrier and metabolic stability based on electrostatic molecular orbital studies. Eur J Med Chem. 2014, 85:661-674
[25] CW Fong, Predicting PARP inhibitory activity – A novel quantum mechanical based model. HAL Archives. 2016, https://hal.archives-ouvertes.fr/hal-01367894v1.
[26] CW Fong, A novel predictive model for the anti-bacterial, anti-malarial and hERG cardiac QT prolongation properties of fluoroquinolones, HAL Archives. 2016, https://hal.archives-ouvertes.fr/hal-01363812v1.
[27] CW Fong, Statins in therapy: Cellular transport, side effects, drug-drug interactions and cytotoxicity - the unrecognized role of lactones, HAL Archives, 2016, https://hal.archives-ouvertes.fr/hal-01185910v1.
[28] CW Fong, Drug discovery model using molecular orbital computations: tyrosine kinase inhibitors. HAL Archives, 2016, https://hal.archives-ouvertes.fr/hal-01350862v1
[29] AV Marenich, CJ Cramer, DJ Truhlar, Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions, J Phys Chem B, 2009, 113:6378 -96
[30] S Rayne, K Forest, Accuracy of computational solvation free energies for neutral and ionic compounds: Dependence on level of theory and solvent model, Nature Proceedings, 2010, http://dx.doi.org/10.1038/npre.2010.4864.1.
[31] RC Rizzo, T Aynechi, DA Case, ID Kuntz, Estimation of Absolute Free Energies of Hydration Using Continuum Methods: Accuracy of Partial Charge Models and Optimization of Nonpolar Contributions, J Chem Theory Comput. 2006, 2:128-139.
[32] F Floris, J Tomasi, JL Pascual-Ahuir, Dispersion and repulsion contributions to the solvation energy: Refinements to a simple computational model in the continuum approximation, J. Comp. Chem., 1991, 12, 784.
[33] RA Pierotti, A scaled particle theory of aqueous and nonaqueous solutions, Chem. Rev., 1976, 717.