UNIVERSITY OF CALGARY

Bond Dissociation Energies of Small Biomolecules from Quantum Mechanical
and Monte Carlo Calculations

by

David Block

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF CHEMISTRY

CALGARY, ALBERTA

JANUARY, 2001

0-612-64943-1

Canada

Abstract

Two studies are made into the bond dissociation energies (BDEs) of C-H bonds of small, biologically relevant molecules. The first determined that proline has a higher BDE on its α-carbon than all other amino acids, approximately 370 kJ mol^{-1} in a protein environment. A peptide model was developed to mimic a β-turn environment, where proline is often found in nature. The constraints of β-turn secondary structure may serve to protect the proline $^{\alpha}$C-H bond from attack by weak oxidizing species, such as glutathione radical. The second study attempted to determine the effect of solvent (water) on the BDEs of several model systems. A hybrid Monte Carlo/ Quantum Mechanics approach was developed. The results were impressive for a series of simple alcohols, matching experiment within 2 kJ mol^{-1}, but less so for a similar set of simple amines and glycine. The lack of polarizability is the primary weakness in the method.

Acknowledgements

Dedication

To Debbra, for waiting

Table of Contents

List of Tables

List of Figures

I. Introduction

Determining the thermodynamics of free radical reactions in proteins is important in order to understand the radical mediated reactions occurring in biology. Some of these are part of normal biological function. Several enzymes have been found which utilize radical reactions as part of their function.[1] Radicals are also used for long-range intramolecular electron transfer.[2] However, the high reactivity and relative non-selectivity of radicals inevitably leads to undesirable side reactions and damage to biological components.[3,4] Inadequate cellular response to oxidative stress can lead to disease states. Proteins damaged by oxidation are normally marked for degradation, but buildup of such proteins has been shown to be a mark of aging.[5] Oxidative damage to proteins has also been implicated in various disorders, including inflammatory diseases, atherosclerosis, ischemia-reperfusion tissue damage, and various neurologic disorders.[5,6]

The damage to proteins is a poorly understood process, both in its mechanism and in its selectivity. It often occurs on the side chains of amino acids in the protein, but it has been found that the hydrogen on the α-carbon of amino acids is especially vulnerable.[7] The strength of this αC-H bond, which is a measure of the ease of radical formation, will be influenced by the chemical nature of the side chains of the amino acids, and by the constraints on the peptide backbone imposed by local secondary structure. The role of solvent is also poorly understood in this context,[8] but differential free energies of solvation may influence radical reactions as well.

The goal of this thesis is to determine the bond dissociation energies (BDEs) of some C-H bonds in model systems. The information gleaned from these models may allow inferences to be drawn about radical reactions in biology. The ability to determine BDEs theoretically has come about fairly recently.[7] The increase in computer power has allowed the examination of ever-larger systems at higher levels of detail. At the same time, new algorithms have decreased the amount of computation that must be performed to get experimental accuracy in calculation.[9] Understanding of the thermodynamics of radical systems has increased,[9] allowing for more accurate treatment of calculated data.

Experimental approaches have also advanced, allowing theoretical investigators a chance to validate their methodologies.[10,11] This work combines all of these factors in order to predict bond dissociation energies of biologically relevant small molecules.

Methodology

B3LYP/6-31G(D)

It has been found that the hybrid Hartree-Fock/density functional, B3LYP,[12] gives energies and geometries equivalent to or better than Moller-Plesset 2nd order perturbation results, at much less computational cost.[13] The majority of the calculations presented in this paper were therefore performed using B3LYP, as implemented in the suite of ab initio quantum mechanical programs, Gaussian 94.[14]

The split valence polarized 6-31G(D) basis set, with s-, p-, and d-type functions, has enough flexibility to describe most electronic wavefunctions and is suitable for accurate quantum mechanical calculations.[15] For greater accuracy, the number and types of functions can be expanded, i.e. 6-311+G(3df, 2p).[9] The determination of the wavefunction for a given system depends on the ability to solve complex integrals and diagonalize matrices. The size of the matrices, and the number and complexity of the integrals, increases with the fourth power of the number of functions required to describe the system, so large basis set calculations are used sparingly.

Bond Dissociation Energies

An important measure in radical thermodynamics is the bond dissociation energy (actually enthalpy) (BDE), or the strength of the bond that must be homolytically broken to create the radical. Knowledge of the BDE will give a good idea of the activation energy of the reaction, and thus of the likelihood of the reaction taking place.[16] The $^{\alpha}$C-H bond dissociation energies (BDEs), $D_{\alpha_{CH}}$, are defined as the heat of reaction (1), $\Delta H^0_{(1)}$:

$$RH \rightarrow R \cdot + H \cdot \qquad (1)$$

Isodesmic Reactions

In this work, as a means of reducing residual errors due to basis set and correlation effects, and for the sake of validation of the methodology by comparison with experimental data, the heats of reaction were derived from the heats of isodesmic reactions.[17,18] These reactions can be represented by equation (2):

$$R\cdot + AH \rightleftharpoons RH + A\cdot \tag{2}$$

AH is a reference molecule for which the BDE, $D_{CH}(AH)$, is known accurately. For each RH the heat of reaction (2), $\Delta H_{(2)}^{0}$, was evaluated from the energies obtained in *ab initio* calculations at the B3LYP/6-31G(D) level, which was shown to give reliable results (within 10 kJ mol^{-1}) for the glycine model peptides.[19] In the context of BDEs, $D_{CH}(RH)$ is then given by:

$$D_{CH}(RH) = D_{CH}(AH) - \Delta H_{(2)}^{0} \tag{3}$$

Glycine

In order to obtain the most effective cancellation of residual errors, the structures of the reference molecule and radical used in reaction (2) should be related as closely as possible to those of RH and $^aR\bullet$ respectively.

For peptides, AH should have both an adjacent amino group and a carbonyl group so that the special feature of the captodative effect can be taken into account.[20] Captodative stabilization arises when there is both an electron donating group and an electron accepting group interacting in a π fashion with an unpaired electron. It can be seen in Figure 1.

The interaction between the filled and the half-filled orbitals lowers the energy of the pair of electrons, and the interaction between the unpaired electron and the empty accepting orbital lowers the energy of the unpaired electron. In effect, captodative stabilization lowers the energy of the donated electron pair without
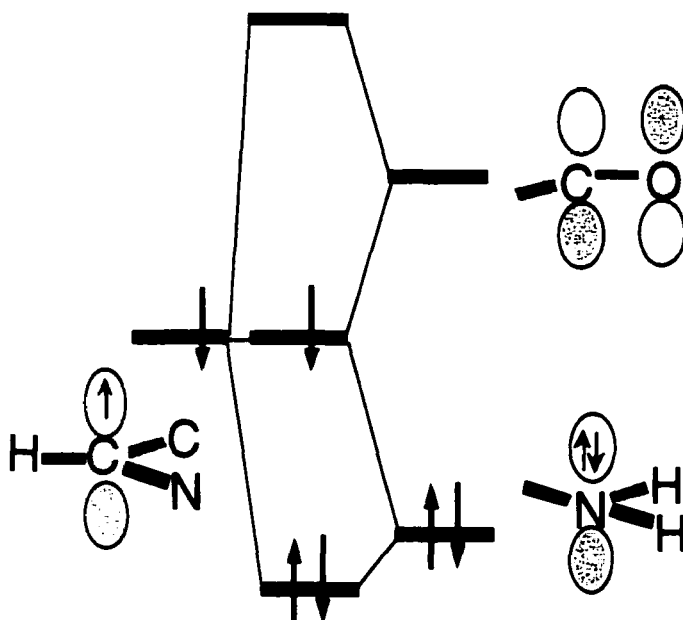
Figure 1: Illustration of Captodative Stabilization

The molecule pictured is a C-centred free radical. This orbital interaction diagram illustrates the interaction between the radical centre and the two groups neighbouring it. The captodative effect occurs because the lone pair on the amine is delocalized and stabilized, without energetic cost to the unpaired electron, because it is able to delocalize into the unoccupied π bond in the carbonyl.

cost to the unpaired electron.[20] Previous studies have shown[15,19] that H_2NCH_2COOH (glycine) [Structure 1]* is the most suitable reference molecule to give reliable values of $D_{\alpha_{CH}}$. The magnitude of $D_{\alpha_{CH}}$ for glycine itself (331.0 kJ mol^{-1}),[21] was not directly available from experiment. However, it has been derived from a number of isodesmic reactions with heats of reaction based on G2(MP2) calculations.[21]

---

* The structures can be seen in Appendix A, The Molecular Bestiary.

II. Proline

Importance

Proline (Structure 3) is a unique amino acid because it contains a five-membered ring on the peptide backbone. The presence of the ring creates a tertiary amide as opposed to the secondary amides in the other amino acids. This means that there is no proton on the amide nitrogen to form hydrogen bonds. It constrains the Ramachandran Φ dihedral angle in proline peptides to around 60°, where in other amino acids there is much more flexibility. Whereas the peptide linkage of all other amino acid residues has the *trans* relationship between the ${}^\alpha C$ centres,[22] the proline ring allows the *cis* rotamer to exist as perhaps 10% of the population of proline residues in nature (see Equation 4). The barrier to rotation about a N,N-dialkyl amide bond is on the order of 60-80 kJ mol^{-1},[23] so interconversion is a rare event. Because the NCR angle in an amide is tighter than the NCO angle, it is advantageous to have the bulkier of the two nitrogen substituents *syn* to the carbonyl oxygen.[23] For the other amino acids, the two groups on the nitrogen are a hydrogen and the alpha carbon, which leads to the *trans* form dominating. With proline, the choice is between the alpha carbon and a ring carbon, so the *cis* form has less of a disadvantage - 18.9 kJ mol^{-1}, according to these calculations (Table 3).

trans proline   cis proline

(4)
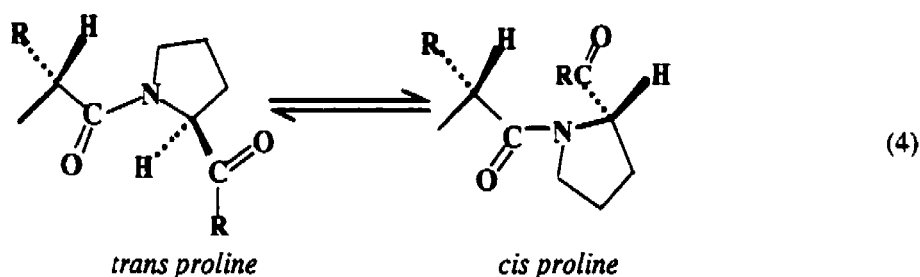
Uniqueness

Proline has a unique role in protein structure. Proline is the preferred residue at the second of four positions in β turns of type I, II and III.[24] It is also the preferred residue at the third of four positions in a β turn of type II'.[25] The *cis* rotamer is the third of four residues by definition in type VI turns.[24] Proline is also common on

the first turn of an α-helix, and is also seen just following the end of the helix. It is commonly found on the outside edge of β sheets, and is the cause of β–bulges in the interior of β sheets.[24] It is a crucial structural member of collagen, in which it is on the outside of a unique triple helical structure.[26]

The above properties make proline, although a hydrophobic residue, usually the most exposed residue in proteins.[22] Its position in β turns, on the edge of sheets, at the beginning of helices, and just following both helices and sheets, leaves it exposed to solvent and the environment. β turns, in particular, have been postulated to be important in molecular recognition because they are almost exclusively on the exterior of globular proteins, and they are highly conserved.[22,27]

Sensitivity to Damage

Proline has been found to be sensitive to oxidation, both by radiolysis[28] and in metal catalyzed systems.[29] The hydroxyl radical can produce hydroxyproline, or it can abstract the αC hydrogen, leaving a neutral radical. This radical can spontaneously cleave the peptide backbone (see Scheme 1a),[28] or it can be further attacked by O_2, also resulting in peptide cleavage (Scheme 1b).[30] Radiolysis can result in the cleavage of the C-N bond, leading to unnatural flexibility in the peptide backbone, and possibly further radical damage within the protein (Scheme 1c).[30] All of these results are expected to occur *in vivo*.

The combination of a crucial role in protein structure, general exposure of proline to the environment, and a αC-H bond that is easily oxidized leads to the importance of this study. We continue the work done earlier on the bond dissociation energies (BDEs) of the αC-H bond of peptide models of the amino acids.[18,19,21,31,32] The BDE of the αC-H bond of a proline peptide model, $D_{\alpha_{CH}}$ (pro), has been calculated, as well as the equivalent BDEs for a number of constrained systems, simulating the constraints placed on the proline residue in various secondary structural environments.

Constraints: model β-turns

It has been estimated that one third of amino acid residues in globular proteins are involved in turns.[25] As noted earlier, proline is frequently found in turns, the

Scheme 1.
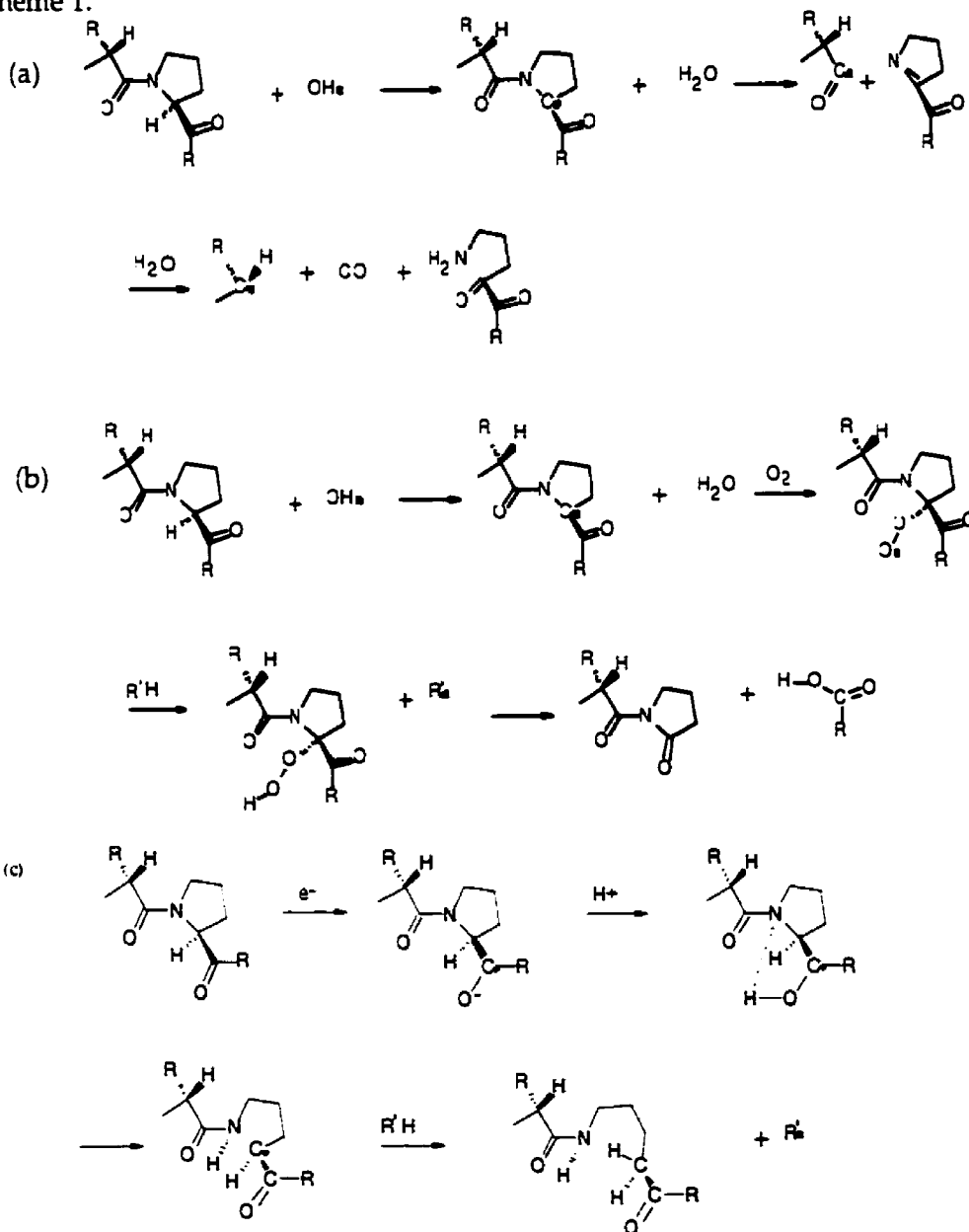
Table 1: Beta Turn Types[a] *(* indicate common proline positions)*

β turn type	Position i+1		Position i+2	
	Φ_{i+1} (°)	Ψ_{i+1} (°)	Φ_{i+2} (°)	Ψ_{i+2} (°)
I	*-60	*-30	-90	0
I'	60	30	90	0
II	*-60	*120	80	0
II'	60	-120	*-80	*0
VIa	-60	120	*-90	*0
VIb	-120	120	*-60	*0

[a]For a visual depiction of the locations of the dihedral angles, see Structure 4

most common of which is the β turn (structure 4), which allows the reversal of direction of the peptide chain within four residues. The convention is to number the residues relative to the first one, calling the first i, the second i+1, etc. The β turns have been classified in terms of the Φ and Ψ angles of the i+1 and i+2 residues (Structure 4). The major types are summarized, with asterisks indicating common proline positions, in Table 1.[24]

Type I and II turns contain proline most often in the i+1 position. It can be noted that type III turns are identical to type I turns at the i+1 position,[25] where proline is the most common residue.[24] Proline is the most common residue at the i+2 position of type II' β turns,[25] where the prime indicates a mirror image of the unprimed β turn types in terms of the Ramachandran dihedral angles Φ and Ψ of the peptides in the i+1 and i+2 positions. Both types VIa and VIb are defined as having a cis proline residue in the i+2 position.[25]

Proline residues in the first turn of an α–helix will have structural constraints very similar to type II' β turns.[22] Prolines are proportionately rare in β sheets due to their lack of an N-H for hydrogen bonding;[22] typical Ψ values in beta sheets are not favorable for proline's five-membered ring.

Collagens form the most common protein type by mass in the vertebrate body. They are made up primarily of repeats of the three peptides (Gly-X-Y) where X and Y are predominantly proline or lysine. The proline and lysine residues may or may not be oxidatively modified by the addition of a hydroxyl group to one of the side chain carbons. Proline's (and hydroxyproline's) rigidity is crucial in stabilizing the unique left-handed helical structure of collagen. Because three strands wind together, the small glycine unit is needed in the middle of the triple helix. This forces proline in collagen to be exposed to the solvent.[33]

Proline's conformation in collagen is very similar to proline in residue $i+1$ in a type II β turn, with Φ and Ψ angles of about -60 and +160 degrees, respectively,[34] as compared with -60 and +120 degrees for a type II turn. The energetics of radical damage to collagen at the proline α-centre is modelled adequately by our β turn type II model. Damage to proline at this site can be expected to lead to cleavage of the peptide bond[28] (see Scheme 1b).

In this study, proline was constrained in geometries corresponding to β turns of type I, II, II', VIa, and VIb (Table 2). Proline was assumed to be the $i+1$ residue for turn types I and II, and the $i+2$ residue for types II', VIa and VIb. In types VIa and VIb, the proline model was constrained to be a cis residue.

Table 2: Proline conformations optimized in this study[a]

Structure Label	β turn type	position	rotamer	Φ (°)	Ψ (°)
5	I	$i+1$	trans	-60	-30
6	II	$i+1$	trans	-60	120
7	II'	$i+2$	trans	-80	0
8	VIa	$i+2$	cis	-90	0
9	VIb	$i+2$	cis	-60	0

[a]*For a visual depiction of the locations of the dihedral angles, see Structure 4*

For each natural constraint the minimum energy geometry was calculated. The α-radical was created by removing the hydrogen attached to the α-carbon. This structure was then optimized as well, maintaining the existing constraints on the Ramachandran angles. The radical [°]C centre thus created was nearly planar and

sp^2-hybridized. A hydrogen atom was then added to the opposite side of the α-carbon to create the unnatural R isomer of proline. This was also optimized, while maintaining the constraints on the system. This procedure allows the examination of the thermodynamics of unnatural repair of an amino acid radical in the peptide backbone.

Alterations to previous methodology

Frequency calculations were not undertaken in the proline study, as the presence of constraints in nearly every structure leads to imaginary frequencies and difficulty in calculating the zero point energy. Fortuitously, the zero point energies and thermal corrections ($H_{298}^\circ - H_0^\circ$) approximately cancel each other out, as seen in this group's previous work.[18] The error introduced by not adding the zero point energy and the thermal correction is estimated to be less than 2 kJ mol^{-1}, whereas the estimated error in the BDEs calculated by the isodesmic reaction procedure is approximately ±10 kJ mol^{-1}. This cancellation, and the approximations involved in doing frequency analysis of constrained systems, led us to decide that frequency calculations were not justified. The structures presented are all converged geometries according to Gaussian 94's default convergence criteria.[14]

In previous studies of $^\alpha$C-H BDEs, an N-formyl amino acid amide was adopted as the model of the mid-chain residue.[18,19] In this model, the next $^\alpha$C site is replaced by the formyl hydrogen atom. It was shown that use of N-acetyl glycine amine rather than the N-formyl analogue resulted in only a 3 kJ mol^{-1} difference in the calculated $^\alpha$C-H BDE.[19] However, the higher steric requirements imposed by the five membered ring of proline, particularly for the *cis* rotamer, dictated the use of the larger N-acetyl proline amide {structures 5 and 7} as the peptide model for the present work since the steric requirements of the next $^\alpha$C site are better modelled by a methyl group than by a hydrogen atom. The N-formyl proline amide model {structure 24} will be briefly touched upon in the discussion.

Table 3: Calculated Energies and BDEs of various proline model conformations

Structure Label	Molecule description	Energy (hartrees)	Bond Dissociation Energy (kJ/mol)
1	Glycine	-284.42345[a]	
2	Glycine radical	-283.79039[a]	**331.0[a]**
5	Proline model *trans*	-533.95938	**368.6**
6	Proline model *trans* radical	-533.31201	
7	Proline model *cis*	-533.95219	**357.7**
8	Proline model *cis* radical	-533.30896	
9	Proline model β turn type I	-533.95108	**380.7**
10	Proline model β turn type I radical	-533.29911	
11	Proline model β turn type I R isomer	-533.93413	336.2
12	Proline model β turn type II	-533.95493	**397.8**
13	Proline model β turn type II radical	-533.29641	
14	Proline model β turn type II R isomer	-533.93234	338.5
15	Proline model β turn type II'	-533.95347	**385.4**
16	Proline model β turn type II' radical	-533.29967	
17	Proline model β turn type II' R isomer	-533.92413	308.4
18	Proline model β turn type VIa	-533.95153	**374.0**
19	Proline model β turn type VIa radical	-533.30210	
10	Proline model β turn type VIa R isomer	-533.91676	282.7
21	Proline model β turn type VIb	-533.94869	**355.0**
22	Proline model β turn type VIb radical	-533.30646	
23	Proline model β turn type VIb R isomer	-533.92328	288.3

[a]Data from Ref. 19.

Results

The results of our calculations are shown in Table 3. The corresponding structures of the natural (S)-isomers, the unnatural (R)-isomers, and the radicals are shown in the appendix. All represent minimum energy geometries subject

only to secondary structural constraints (on Φ and Ψ). All Φ and Ψ angle measurements are indicated. The bond dissociation energies were calculated using the isodesmic reaction described above, with glycine and the glycine α-radical as the known components of the reaction, and the value of glycine's BDE taken as 331.0 kJ mol^{-1}.[18]

Discussion

Stabilizing β turns is the most significant role for proline in protein secondary structure. Damage to the proline α-carbon can result in peptide bond scission, modification into glutamic acid, or formation of other carbonyl-containing derivatives (Scheme 1). Any of these results could lead to the loss of local secondary structure about the proline residue, and perhaps loss of tertiary structure as well. The oxidation of proline would thus initiate a cascade of events that may lead to pathological processes.[28,29,35] The most common pathway for the oxidation of amino acids generally is the abstraction of the hydrogen from the α-carbon by a hydroxy-radical.[30] Understanding the impact of the structural environment on the strength of the $^{\alpha}$C-H bond is important in order to predict the likelihood of oxidative damage to proline.

$^{\alpha}$C-H Bonds: The BDE of the $^{\alpha}$C-H bond in fully optimized trans- and cis-proline residues (structure 5, 368.6 kJ mol^{-1}, and structure 7, 357.7 kJ mol^{-1}, respectively, and Table 3, column four) are higher than the corresponding bonds of glycine (348 kJ mol^{-1}), alanine (344 kJ mol^{-1}), serine (348 kJ mol^{-1}), threonine (356 kJ mol^{-1}),[18] or any of the other residues studied thus far.[16] This is because structural requirements of the five membered ring result in increased steric repulsion from the groups of neighbouring residues and the $^{\alpha}$C-centered radical cannot achieve the planar geometry required for maximum captodative stabilization. That this is the case is dramatically illustrated in the N-formyl model (structure 24) in which a hydrogen atom replaces the next α-carbon. There the $^{\alpha}$C-centered radical is able to achieve a planar geometry (structure 25) and the BDE is calculated to be 321 kJ mol^{-1}. The additional steric hindrance manifests itself primarily in the parent system- the optimized trans conformer is 18.9 kJ mol^{-1} more stable than the cis conformer (structure 24), while the difference in the respective radicals is 8.0 kJ mol^{-1}.

The high $^\alpha$C-H BDEs are significant because they are close to the BDE of an alkyl sulfhydryl bond, 370 kJ mol^{-1}, as occurs in cysteine, cysteine residues,[36] or glutathione (GSH).[37] $^\alpha$C-H bonds which are weaker than 370 kJ mol^{-1} are liable to be damaged by thiyl radicals, such as GS·, or conversely, the corresponding $^\alpha$C-centered radicals would not be repairable by, e.g. GSH.

The constraints imposed by secondary structure on this system prevent any of the α-carbon radicals from achieving planarity, and cause a concomitant rise in BDE. Other than β turn type VIb, (structure 21), all of the constrained systems have BDEs higher than 370 kJ mol^{-1}. This indicates that the sulfur radical of glutathione, GS·, is not able to abstract hydrogens from proline's α-carbon, and glutathione would be able to repair damage at this site. Given proline's position at the 'hinge' of turns, the loss of a proline residue's $^\alpha$C-H bond may be expected to result in the destruction of the protein's tertiary structure and therefore its function. The high $^\alpha$C-H BDEs confer some protection against weaker oxidizing species such as thiyl radicals, ROO·, or superoxide (O_2^{-}·).

The $^\alpha$C-H BDE of proline residues in the type II beta turn (structure 12) is higher than the corresponding BDE of proline in type I turns (structure 9). The natural type II conformation is lower in energy, and the type II radical higher in energy, than the type I equivalents. The difference is in the rotation of the C-terminal amide group: the type II conformation does not allow the α-carbon radical to link its π orbitals with those of the amide system, which are perpendicular to it. The type I system allows some limited mixing of the π systems, which accounts for the lower BDE.

The BDE of proline residues in type II′ turns (structure 15) is similar to that of residues in type I turns. This is to be expected, as their structures differ by only 20 and 30 degrees at the Φ and Ψ angles, respectively. The molecule's attempt at planarity is blocked by the angle constraints, but the α-carbon and the nitrogen are each individually almost planar. Because proline residues in the first turn of an α-helix will have similar structural constraints, these results can be also applied to proline in that environment. Proline residues in an α-helix are thus protected against GS· attack.

Turn types VIa (structure 18) and VIb (structure 21) are similar, and have similar BDEs for proline residues. Type VIb (structure 21) has, as mentioned previously, the weakest °C-H bond among the constrained systems. It appears that the α-carbon is able to achieve planarity in these systems, at the expense of the planarity of the neighbouring nitrogen. This means that the amide system including the proline nitrogen is unable to delocalize electrons, but the amide system on the other side is able to delocalize electrons all the way to the radical centre. This strategy appears to be the most successful in stabilizing the radical.

It is evident from Table 3 that the BDEs of the unnatural (R)-isomers are much lower than those of the natural (S)-isomers. Loss of the hydrogen from the (S)-isomer does not induce much change in conformation outside of the α-carbon. Addition of a hydrogen to the radical on the pro-R face of the α-carbon forces the proline ring to pucker the opposite way, leading to a higher energy conformation. Thermodynamics will thus favor formation of the natural (S)-isomers when hydrogen transfer agents such as glutathione repair prolyl radical residues.

C-H Bonds of the Proline Ring: We have not explicitly examined the BDEs of the remaining six C-H bonds of the proline ring. The two C-H bonds adjacent to the N atom will have BDEs similar to the corresponding bonds in pyrrolidone, 377 kJ mol^{-1}.[38] The remaining four C-H bonds should have BDEs similar to those of cyclopentane, 398±5 kJ mol^{-1}.[39] Thus C-H BDEs of the methylene groups of proline are higher than the °C-H BDE of unconstrained proline, similar to the highest °C-H BDEs among the proline β-turn conformations, and substantially higher than the parent BDEs of the weaker oxidizing radicals, ROO· and RS·.

Enzyme-controlled oxidation of the 3- and 4- carbon of the 5-membered ring occurs only in collagen and collagen-like proteins. Prolyl 4-hydroxylase selectively and stereospecifically hydroxylates the inactivated carbon β to the nitrogen on the pyrrole ring. The peptide must contain an X-Pro-Gly triplet for the hydroxylation to proceed. The reaction consumes one O_2 molecule and one 2-oxoglutarate, and occurs at an active site containing an Fe^{2+} atom. Prolyl 3-hydroxylase uses a similar active site and set of cosubstrates, but requires a Pro-4-hydroxyproline-Gly triplet. The purpose of 3-hydroxyproline is unknown, but

the hydroxyl group of 4-hydroxyproline is important in hydrogen-bonding to other peptide residues in the formation of the unique collagen triple helix.[40]

Ascorbate is required to regenerate the Fe^{2+} after an uncoupled decarboxylation of the 2-oxoglutarate. Without ascorbic acid (Vitamin C), the enzyme will decarboxylate a 2-oxoglutarate molecule and be deactivated, leading to collagen without hydroxyproline. This reduces the strength of the collagen helices, leading to scurvy (the breakdown of the body's connective tissues).[40]

Significance for Oxidative Damage: The most striking result from the present work is that the $^{\alpha}$C-H BDEs of all the natural proline residues in Table 3 are appreciably larger than those of glycine and the residues of the other common amino acids thus far examined. While they are much lower than the O-H BDE of water (499 ± 1 kJ mol^{-1} [39]), and thus still susceptible to oxidative damage by the exothermic hydrogen abstraction reaction (5)

$$RH_{(res)} + OH\cdot \; \overline{}\!\!\!\longrightarrow \; ^{\alpha}R\cdot_{(res)} + H_2O \qquad (5)$$

the proline $^{\alpha}$C-centres will not be preferred sites of attack by OH·; they are less exposed than the C-H bonds on the ring. Nor will they be susceptible to oxidative damage by the weaker oxidants like peroxy (ROO·) and thiyl (RS·) radicals, which may be present in cellular systems and whose parent precursors have BDEs near 370 kJ mol^{-1}.[37,41] At the same time, the repair reaction (6)

$$^{\alpha}R\cdot_{(res)} + GSH \; \overline{}\!\!\!\longrightarrow \; RH_{(res)} + GS\cdot \qquad (6)$$

would be exothermic and they should be repairable by the natural protector glutathione, GSH, which has a BDE of ~370 kJ mol^{-1}.[37] As noted above, this repair will give preference to the natural isomer. The only radical potentially not repairable by GSH is the proline β turn type VIb (structure 21), which has a relatively low BDE (355 kJ mol^{-1}) near to that of the glycine residue.

Inspection of the β-turn conformation (structure 4) indicates that the ring carbons are more exposed to solvent than the α-carbon. The comparable C-H BDEs of the ring carbons indicate that accidental oxidation of these C-H bonds is perhaps more likely than attack at the $^{\alpha}$C-H bond, given the structural constraints of the

protein environment. Accidental oxidation of these bonds will create a C-centred radical that could abstract a hydrogen from a nearby carbon, propagating damage into the peptide backbone or onto another amino acid side chain. However, oxidation of the ring carbon is itself not likely to be damaging to a protein's secondary or tertiary structure, or its function.

Conclusion

Proline is vulnerable to oxidative damage at the $^\alpha$C-H bond when it is unconstrained; rigid constraints, such as would exist in nature, serve to protect the bond from damage. The degree of planarity at the radical centre seems to be a good visual indicator of the stabilization of the radical; better stabilization indicates a lower BDE and a weaker $^\alpha$C-H bond. Proline residues in β turns of type VIb are expected to have weak enough $^\alpha$C-H bonds to allow attack by weaker oxidizing agents, like ROO· and thiyl radicals, whereas other turn types should protect the $^\alpha$C-H bond from all but the strongest oxidizers.

Damage to proline's $^\alpha$C-H bond may be expected to lead to the degradation of the parent protein. This is a common event for many proteins, but may lead to pathological states if the damaged protein is not removed and replaced quickly. Thus pathology will result if damage occurs in structural proteins. This long-lasting damage can be seen in the breakdown of the crystallins of the vitreous humour of the eye during aging, for example.[42]

This work extends our investigation of the BDEs of the amino acids in proteins,[18,19,21] both unconstrained and in specific secondary structural environments. These calculations support earlier experimental work[28,29,35,43] and will be crucial in the goal of predicting the site specificity of oxidative damage to proteins.

III. The Role of Solvent in Radical Thermochemistry

While the thermodynamic properties of C-centered free radicals are readily described in the gas phase by modern computational techniques,[9] accurate methods for predicting their reduction potentials and other thermochemical properties in solution have not been developed as highly. Solution properties are required for an understanding of redox mechanisms and numerous other processes.[8,31]

Scarce Experimental Data

The problems faced included the scarcity of accurate experimentally derived BDEs in solution. Most calculations, including those from this group, assume that the radical is solvated to the same extent as the parent, sidestepping the matter completely.[15,18,19,21,31,32,36,38,44] This has some experimental support.[8] It was realized, however, that in order to tackle the problem of BDEs of proteins in solution, model systems that could be compared with experiment would have to be examined, and a procedure developed that could be extended to larger model systems, including amino acids in peptides.

Experimental values of ΔG_{soln} for radicals cannot be obtained by direct methods, such as vapor pressure measurements. One must therefore use radicals for which values of $\Delta_f G_{(aq)}$ and $\Delta_f G_{(g)}$ are known and obtain ΔG_{soln} from the fundamental relation:

$$\Delta G_{soln} = \Delta_f G_{(aq)} - \Delta_f G_{(g)} \tag{7}$$

Quantitative experimental data are difficult to find.

Early Theoretical Arguments

A π-donor stabilized C-centered radical necessarily has dipolar character since there is a net charge transfer from the doubly occupied π-donor orbital to the singly occupied 2p orbital at C.[6] It is reasonable that such polarity would be enhanced in a medium of high dielectric constant, such as water, and that the free energy of solution would increase with dielectric constant. The situation is compounded in the case of captodatively stabilized free radicals, for which polar

resonance forms can be written. Indeed, early SCRF studies (INDO-UHF-SCRF) found that the stabilization of such radicals, relative to the gas phase, was in the range 30 - 120 kJ mol^{-1}.[46] The conclusion of such theoretical studies is that the radical, R·, should have a more negative free energy of solution than the parent, R-H, and that the difference, $\Delta\Delta G_{soln}$, would increase with polarizability of the free radical and polarity of solvent. The difference, $\Delta\Delta G_{soln}$, would be greatest in aqueous solution. Internal factors not included in these considerations, such as preferential interaction of the solvent in the form of increased hydrogen bonding and bonding of the three- or one-electron kind, would appear to exacerbate the difference, whereas external factors associated with rearranged solution structure (a decrease of solvent entropy, for example) may work in the opposite direction. In contrast, the relatively few experimental studies that address this point lead to the conclusion that $\Delta\Delta G_{soln}$ is close to zero.

Thus, no effect of solvent polarity was found for the stabilization of captodative free radicals in a variety of polar organic solvents,[47] and a photoacoustic calorimetry study concluded that "moderately large organic molecules and their corresponding radicals are solvated to the same extent - even in water".[8] Subsequently, in the calculation of reduction potentials, $E^{o'}$, for the process:

$$R· + e^- + H^+ \rightarrow R-H \qquad (8)$$

with glycine radicals, it[31,36] has been assumed that the solvation free energy of the radical is the same as that of the parent species, and experimental results on the oxidation of the parent by RS· radicals have borne out the predictions.[48,49]

Despite the above experimental evidence, the theoretical predictions relating to $\Delta\Delta G_{soln}$ are compelling, and further examination of the problem is obviously required. A theoretical approach with a discrete molecular modeling of the solvent would potentially avoid the difficulties with the SCRF approach. Here we have applied the BOSS (Biochemical and Organic Simulation System) program package[50], combined with results derived from quantum mechanics, to investigate the solution properties of a biologically important class of free radicals. We evaluate the potential of this method, which does not require the introduction of new empirical parameters, and compare the free energies of solution (ΔG_{soln}) values calculated by BOSS with results derived from experimental information.

IV. Solvation of Radicals I: Alcohols

The $^\alpha$C-centered radicals of methanol, ethanol, 1-propanol, and 2-propanol were chosen for this study because values of $\Delta_f G_{(aq)}$ are available for three of them and for all of their parent alcohols. A study by Schwarz and Dodson was found in which aldehydes were reduced to alcohol α–radicals.[51] The authors gave free energies of formation in solution of the radicals. Values of $\Delta_f G_{(g)}$ were not directly available for the radicals, but they can be derived from experimental values of $\Delta_f H_{(g)}$ and standard *ab initio* calculations of entropies.

Oxidative damage of glycoproteins and carbohydrates occurs primarily by hydrogen abstraction at a site adjacent to an OH group. This is because the resultant C-centered free radical is stabilized by the π-donor ability of the dicoordinated oxygen atom. The (deoxy)ribose moiety of nucleic acids and the side chain of serine residues of proteins are among the sites susceptible to oxidation for the same reason.

Solvation Methods

One of the great enabling assumptions of computational chemistry is that gas phase molecules are effectively isolated, and therefore can be modelled as the only molecules in a system. Specifically, gas phase molecules are treated as if they exist in a pure vacuum at 0 K. As long as no collisions take place, that assumption is valid. This assumption has been extended to molecules in non-polar solvents, since the presence of such solvent molecules has very little effect on the electronic structure of the solute molecule. This assumption is faulty, however, when polar solvents are considered. Polar solvents polarize the electron distribution of solutes in them, in order to maximize dipole-dipole and other electrostatic interactions.

Water is one of the most polar solvents, with a dielectric constant of ~78.5.[52] Biology occurs in an aqueous environment, so many calculations of biological significance must take the solvent into account in order to reflect reality accurately. Unfortunately, the interaction with solvent is complex, with many

transient interactions combining to give a large set of possible structures for the molecule-solvent complex.

Solvent as Continuum: Self-Consistent Reaction Field, Self-Consistent Isodensity (surface) Polarizable Continuum Model

One way to model the molecule-solvent complex is to select the most important feature of the solvent and model only that feature in an average way. One can then smear the individual solvent molecules together in order to make an isotropic continuum that is parameterized to reproduce this average behaviour.[53] The most important descriptor of a solvent, electrostatically speaking, is its dielectric constant. Parameterizing the continuum to reproduce the dielectric constant of the solvent produces a dielectric continuum. The advantage to this approach is that it is computationally very efficient. The disadvantage is that anisotropic effects cannot be modelled without added parameterization.[53] In water, hydrogen-bonding is very important and completely dependent on the orientation of individual solvent molecules; as such, it is hard to model using a continuum. Nevertheless, continuum models have been used successfully in many applications,[54] and much work has gone into improving them.[53,55,56]

Continuum modelling has been built upon solutions to the Poisson-Boltzmann equation, which describes the potential energy of a spherically shaped charge placed in a dielectric continuum. The equation has been extended to describe general shapes. The question of what shape to use to describe a molecule has been debated extensively, and there are several alternatives.[53] This work uses the most physically meaningful approach available, the Self-Consistent Isodensity (surface) Polarizable Continuum Model (SCIPCM),[53] as implemented in Gaussian 94.[14] This defines the surface of the molecule as the isodensity surface of the electron probability density at some empirical value, which in this case is 0.0004. The surface is modified as part of the self-consistent field calculations within optimizations- as the electronic structure is modified, the shape of the molecule used to define its interaction with solvent is also modified. The solvent influence is thus included in the process of optimizing the structure. It was found that this approach produces realistic geometries and polarized wavefunctions, but not very reliable energetics for molecules that are expected to be involved in

hydrogen bonding. This approach was therefore not suitable for accurate thermodynamic work in aqueous solvent, with polarizable solutes such as free radicals. The geometries and wavefunctions were used, however, as a starting point for the next step in the calculations.

Solvent as Discrete Molecules: Monte Carlo

A fundamentally different approach to modelling solvation can be characterized as bottom-up. The continuum model can be described as top-down, since it starts from the assumption that the most important parts of the solute-solvent interaction are understood, and parameterizes a continuum to reproduce only those parts. In contrast, one can choose to model enough individual solvent molecules, interacting with the solute in enough random orientations, to reproduce bulk behaviour. There are several disadvantages to this scheme, but its major advantage is that it gives valuable structural information about solute-solvent geometries, since no assumptions are made about the physical placement of solvent molecules.

This approach was taken in this study. The solvent model used was TIP4P water, a well-characterized and understood water model.[57] It consists of three Lennard-Jones atoms, an oxygen and two hydrogens, with charges of +0.52 e on the hydrogens, and a balancing charge of −1.04 electrons displaced from the oxygen 0.15 Å on the bisector of the H-O-H angle. This structure produces realistic radial distribution functions, reasonable densities, and reliable solvation free energies.[58,59]

The Biochemical and Organic Simulation System (BOSS),[50] due to Dr. W. L. Jorgensen, was used to perform Monte Carlo simulations of the solvation of the solutes of interest. The Monte Carlo method[60] randomly rotates and translates a randomly-chosen molecule and evaluates the energy of the whole sample after each change. The resultant orientation is either accepted (i.e. included in the sample of orientation space) or rejected (i.e. ignored) by a stochastic process that results in the energy profile of the sample space following a Boltzmann distribution. If the energy of the system is lowered, the new state is accepted. If the energy of the new state has been raised, the probability of acceptance is

directly proportional to the difference in energy between the previous state
and the present state. The average properties of the orientational sample space
as a whole can thus be used as a good estimate of the properties of the real
system, which would include many molecules and many orientations due to
both the number of molecules in a real sample, and the finite measuring time.
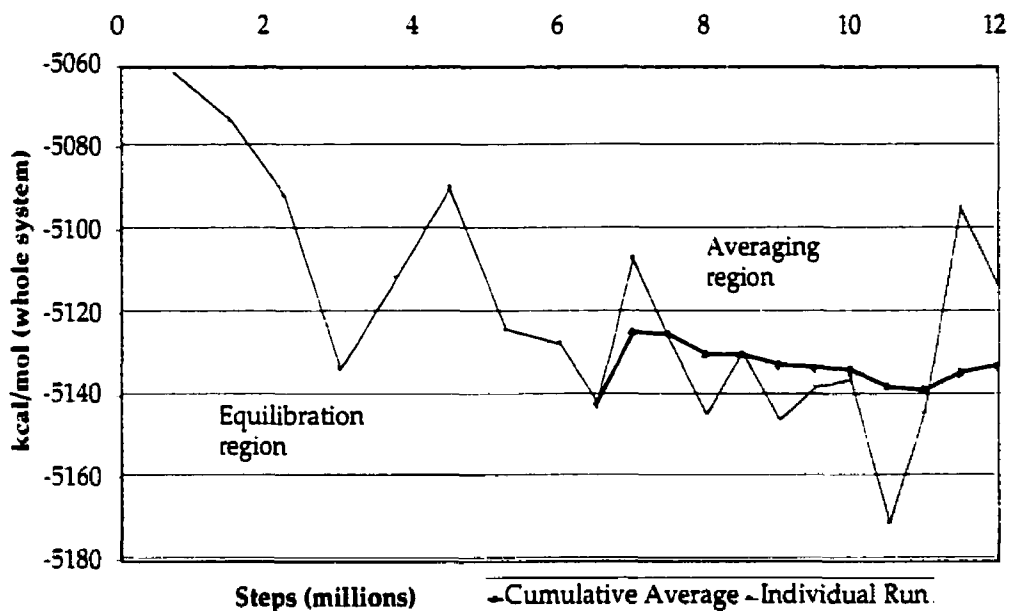
Figure 2: Total Energy- Glycine Sample Run

The energy of the solute+solvent system for glycine, surrounded by ~500 TIP4P water
molecules. The total energy stabilizes after 6 million steps, but reliable results require
averaging over a further 6 million configurations.

In order to ensure that the model included enough solvent molecules, a box of
about 500 water molecules was used, with periodic boundary conditions, where
the box is repeated in all directions.[61] Since there were no unbalanced charges in
the solutes studied, there were not expected to be long-range effects beyond this
box size.[62] Others have determined that thermodynamic quantities determined
through Monte Carlo simulations are surprisingly insensitive to the size of the
box beyond some minimum value.[62]

Because BOSS uses a Monte Carlo technique, there is no minimization procedure.
In order to ensure that the model sampled enough of orientational space,

extensive benchmarking (Figure 2) was done with glycine (structure 1) and the glycine zwitterion (structure 26). It was found that for a given structure, the radial distribution functions and total energies fluctuated about a stable mean after a simulation of 6,000,000 steps, where a step is the random movement and reorientation of one molecule in the box. The simulations were all run at 25° C. The first 750,000 steps were performed without allowing volume changes, in order to discourage non-physical expansion which could alleviate initial close contacts between the solvent and the solute. After the first 750,000 steps, the simulation was run in the NPT ensemble[63], at 1 atmosphere pressure. The system was found to be equilibrated after 6,000,000 steps. Averaging was performed over a further 6,000,000 steps, divided into 12 segments of 500,000 steps each. Errors were estimated as the root mean square of the standard deviations of each of the segments[61]. A cutoff distance for solute-solvent interactions of 12 Å (half the width of the box) is used, with quadratic feathering of the intermolecular interactions to zero in the last 0.5 Å. For solvent-solvent interactions, the cutoff is reduced to 10 Å.

Molecular Mechanics and OPLS

The task of sampling 6,000,000 configurations of ~500 water molecules plus a solute molecule at an *ab initio* level is beyond the present computational resources at the University of Calgary[64]. The Monte Carlo calculations were therefore performed using molecular mechanics[65], a gross approximation that has been parameterized to reproduce experimental results rather closely.[58]

Molecular mechanics treats molecules as systems of atom-centred point charges that interact with each other classically. Interactions between molecules are modelled by the Coulomb interaction of the point charges, q_i, and by a Lennard-Jones type[60] interatomic potential between each pair of atoms:

$$F_{(Lennard-Jones)} = \sum_{non-bonded} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\varepsilon r_{ij}} \right] \tag{9}$$

The potential function is parameterized for each atom type, and the interaction potential between two atoms is treated as the geometric mean of the two atomic potentials, i. e.:

$$A_{ij} = (A_{ii}A_{jj})^{1/2} \text{ and } C_{ij} = (C_{ii}C_{jj})^{1/2} \tag{10}$$

Because all of the MC calculations use an empirical force field, defining the atom types is the crucial step in setting up the simulation. The atom types are completely defined by three parameters (more are required when including flexibility in the model). The three are σ and ε from the Lennard-Jones parameters, and q, the charge on the atom. One may define σ and ε in terms of the previous A and C parameters through equations (11) and (12):

$$A_{ii} = 4\varepsilon_i\sigma_i^{12} \tag{11}$$

$$C_{ii} = 4\varepsilon_i\sigma_i^{6} \tag{12}$$

σ is then the Lennard-Jones radius term, while ε is the energy term[61]. Others have developed Lennard-Jones parameters for use in molecular mechanics calculations; we have adopted the OPLS set[58] as the source of our Lennard-Jones parameters. The model is not very sensitive to their exact values, as long as they are reasonable. This is demonstrated by the fact that there is essentially no difference in ΔG_{soln} between methanol radical with the α-carbon treated as an sp^3 carbon and as a tertiary carbocation (sp^2). Following the recommendation within BOSS, any hydrogen attached to a heteroatom has zeroes for its Lennard-Jones parameters, so that the only sign of its presence is its charge.

In order to simplify the calculations further, the assumption is made that intramolecular vibrations and torsional rotations are irrelevant to the free energy state function. The solutes can then be frozen in their optimized position. This assumption is valid if those vibrations and torsions are included in the total free energy some other way, and if all relevant conformations are included in the calculations. *Ab initio* frequency calculations allow the vibrational and torsional energy levels to be included in the thermodynamic free energy calculation. With this approach, it remains the researcher's responsibility to include the different conformations present either in gas phase or in solution in the appropriate calculations. The entropy of mixing of the different conformations must then be included in this free energy.

CHELPG

It can be seen by the above description that the point charges on the atoms in the molecular mechanics representation of the solute molecule are crucial. The approach adopted is to calculate atom-centred point charges so as to reproduce both the dipole moment and the electrostatic potential surrounding the molecule. This is done using the CHELPG (Charges from ELectrostatic Potentials, Grid-oriented) procedure[66], as implemented in Gaussian[14]. This procedure consists of evaluating the electrostatic potential at hundreds of grid points around the molecule, and optimizing a set of point charges to reproduce the calculated potential and the dipole moment. This calculation takes a few seconds and produces a set of charges that adequately represents the electrostatics of a solute molecule in the given conformation.

The question arises: at which stage in the *ab initio* calculations should one perform the CHELPG calculation? Jorgensen *et al* have settled on HF/6-31G(D) wavefunctions in the gas phase as the best starting point for CHELPG charges[67], as HF calculations overestimate the polarization to some extent. This overestimate approximates the polarization present in aqueous solution.

In this work, B3LYP/6-31G(D) calculations provide the starting wavefunctions, and tests were performed with methanol and the methanol radical to determine what combination of theoretical level and basis set produces the best CHELPG charges with the most reasonable amount of effort. The results of those calculations will be discussed later.

Free Energy Perturbation

The influence of the solvent on the total free energy of a system is captured in the term ΔG_{soln}, which expresses the change in free energy gained by bringing the solute from some specified standard state in the gas phase into solution at a concentration of 1 mol L^{-1}. For purposes of comparison with the BOSS results, ΔG_{soln} is the difference between the free energy of formation of the substance in the aqueous phase and the gaseous phase under the same conditions of temperature (298.15 K), and concentration (1 M). (The standard state for reporting $\Delta_f G_{(g)}$ is 1 atm, so most values must be corrected by -7.9 kJ mol^{-1} in
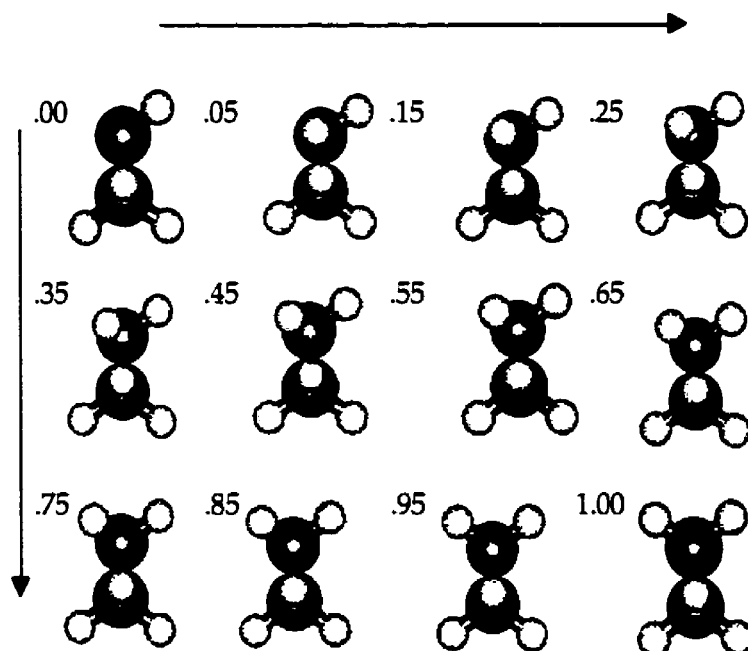
Figure 3: A Sample Mutation from Methanol to Methylamine

The numbers indicate the extent along the linear mutation path.

order to reflect a concentration of 1 M.) It is the difference in ΔG_{soln} which MC methods are able to determine.

The free energy perturbation (FEP) technique[68,69] is used to derive accurate *relative* free energies of solution of two species, say A and B; A is converted to B in 10 steps by linearly scaling geometry, charges, and Lennard-Jones parameters (see Figure 3). The relative free energy of solution of the two species is the sum of the changes of the 10 steps.

An individual mutation involves very small changes in the solute molecule, and so after the first run, less equilibration time is required if the solvent structure is carried over from one run to the next. In order to take advantage of the pseudo-equilibrated state of the solvent box, the equilibration step is reduced from 6,000,000 to 3,000,000 steps on subsequent runs. Again the first 750,000 steps are without volume changes. Averaging is then performed over 6,000,000 steps.

If there is no flexibility allowed in the solute molecule (i.e. the internal energy of the solute is ignored), the only energy differences observed will be in the
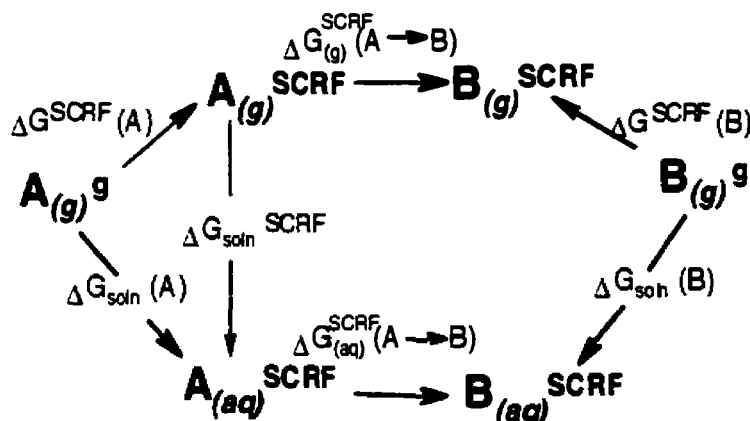
Figure 4: Thermodynamic Cycles and definition of Symbols used in the text and Tables.

solvent's changing response to the solute. This difference is the change in the free energy of solvation, which is what is sought.

Free Energy Perturbation. The left and right hand side of Figure 4 show the free energy changes occurring between species in solution with SCRF-optimized geometries ($A_{(aq)}^{SCRF}$, $B_{(aq)}^{SCRF}$) and in the gas phase with the same geometries ($A_{(g)}^{SCRF}$, $B_{(g)}^{SCRF}$) or optimized gas phase geometries ($A_{(g)}^g$, $B_{(g)}^g$). The arrows in the middle show the changes occurring in the conversion of A into B by the FEP technique. The relative free energy of solution of species A and B, $\Delta\Delta G_{soln}$ (= $\Delta G_{soln}(B)$ - $\Delta G_{soln}(A)$), is given by equation (13):

$$\Delta\Delta G_{soln} = \Delta G_{(aq)}^{SCRF}(A \rightarrow B) - \Delta G_{(g)}^{SCRF}(A \rightarrow B) + \Delta G^{SCRF}(B) - \Delta G^{SCRF}(A)$$

$$(13)$$

The last two terms correspond to the free energy change associated with the distortion of the optimized gas phase structures of the individual species to the corresponding solution structures (approximated as the SCRF-optimized structures) *in the absence of the reaction field.*[67] These are discussed below. The quantity actually calculated by BOSS is the solvent response part of the permutation of A to B, $\Delta G_{(aq)}^{BOSS}(A \rightarrow B)$ i.e.

$$\Delta G_{(aq)}^{BOSS}(A \rightarrow B) = \Delta G_{(aq)}^{SCRF}(A \rightarrow B) - \Delta G_{(g)}^{SCRF}(A \rightarrow B)$$

$$(14)$$

Therefore,

$$\Delta\Delta G_{\text{soln}} = \Delta G_{(aq)}^{\text{BOSS}}(\text{A} \rightarrow \text{B}) + \Delta G^{\text{SCRF}}(\text{B}) - \Delta G^{\text{SCRF}}(\text{A}) \tag{15}$$

The quantity, $\Delta G^{\text{SCRF}}(\text{A})$, may be represented as in equation (16):[67]

$$\Delta G^{\text{SCRF}}(\text{A}) = \Delta H^{\text{SCRF}}(\text{A}) - T\left(S\left(\text{A}_{(g)}^{\text{SCRF}}\right) - S\left(\text{A}_{(g)}^{\text{opt}}\right)\right) \tag{16}$$

We define the enthalpy change following Lim and Jorgensen:[67]

$$\Delta H^{\text{SCRF}}(\text{A}) = \left\langle \Phi^{\text{SCRF}} \middle| H^{\text{B3LYP}} \middle| \Phi^{\text{SCRF}} \right\rangle_{(g)} - \left\langle \Phi^{\text{opt}} \middle| H^{\text{B3LYP}} \middle| \Phi^{\text{opt}} \right\rangle_{(g)} \tag{17}$$

In practice, in order to determine the energy required to modify the geometries of these species from their gas-phase optimum to their solution optimum, the solution phase wavefunction was analyzed in the following way. A single-point energy calculation was performed on the SCRF wavefunction with the SCRF removed. The nuclear repulsion term is fixed in a single point calculation, and it was recorded. The very first SCF cycle electronic interaction energy was also recorded. This did not allow the electronic wavefunction to relax to a gas-phase optimum. The sum of the two terms gives the enthalpy of formation of the solution-phase geometry in the gas phase. Combined with the solution-phase entropy, this gives the free energy of the solution-phase structure in the gas phase. The difference between this value and the free energy of the gas phase structure is the free energy required to modify the structure from gas- to solution-phase, or ΔG^{SCRF}(Equation 16).

G2(MP2-B3LYP)

In order to determine the absolute energies of various species, a standard algorithm has been developed. The G2(MP2) method[70] has been shown to minimize basis set and correlation errors. It is based on HF frequencies and MP2 geometries. These are both inferior to their B3LYP equivalents for radicals[13], and so the procedure has been modified to reflect this improvement. The new procedure is referred to as G2(MP2-B3LYP). It gives a very accurate estimate of the absolute enthalpy of a species, which can be compared directly to experiment.

The algorithm consists of optimizing the wavefunction of a molecule at B3-LYP/6-31G(D) level, and performing a frequency analysis on that wavefunction. The zero point energy from that analysis is recorded. The wavefunction is recalculated at the optimized geometry, using MP2/6-31G(D), MP2/6-311+G(3df,2p), and QCISD(T)/6-31G(D). The MP2 calculation with the small basis set provides an energy that is subtracted from the MP2 large basis set energy. This difference is added to the QCISD(T) energy in order to correct for the small basis set. The zero point energy is added, scaled by 0.98^{71}. Finally the numbers of valence electron pairs and unpaired valence electrons are counted, and the totals scaled by empirical factors. The subtraction of these numbers from the total produces a very reliable enthalpy. The procedure can be summarized as in equation (18):

$E(G2(MP2-B3LYP)) =$

$$E(QCISD(T)) + E(MP2/6-311+G(3df,2p)) - E(MP2/6-31G(D))$$
$$- 0.005 \, (NVEP) - 0.00019 \, (NUE) + 0.98 \, (ZPE) \tag{18}$$

The following table (4) illustrates the procedure for propanol.

Table 4: Propanol Energy (G2(MP2-B3LYP) (Hartrees)

QCISD(T) energy				-193.81481
MP2/6-31G(D) energy				- (-193.75041)
MP2/6-311+G(3df,2p) energy				+ (-193.95526)
Number of Valence Electron Pairs (NVEP)	13	x	0.005 =	- 0.065
Number of Unpaired Electrons (NUE)	0	x	0.00019 =	- 0
Zero Point Energy (ZPE)	0.10903	x	0.98 =	+ 0.10685
Energy (G2(MP2-B3LYP)				-193.97781

Values for the enthalpy change determined at the B3LYP/6-31G(D)//B3LYP/6-31G(D) and B3LYP/6-311+G(3DF,2P)//B3LYP/6-31G(D) levels are discussed below. The entropy is derived from B3LYP/6-31G(D) frequencies with and without the reaction field. The difference represents changes to rotational and

vibrational terms as a consequence of the presence of the solvent modeled as a dielectric continuum. We shall see that the ΔS term is small and may be set to zero, in effect avoiding the tedious computation of vibrational frequencies in the presence of the reaction field.

The combination of Lennard-Jones parameters from similar atoms in the OPLS parameter set, and charges from CHELPG calculations on our wavefunction, provides a simple technique for converting a molecule from an *ab initio* structure to its equivalent molecular mechanical representation. Keeping a rigid solute keeps the vibrational and conformational information from the high-level calculation without relying on force fields for thermodynamic data.

In order to determine *absolute* solvation free energies, it is necessary to mutate a solute from being present in the solvent to being absent, to note the free energy difference in the system. This is accomplished by mutating a solvent molecule to nothing. This is obviously a drastic change, and requires some intermediate
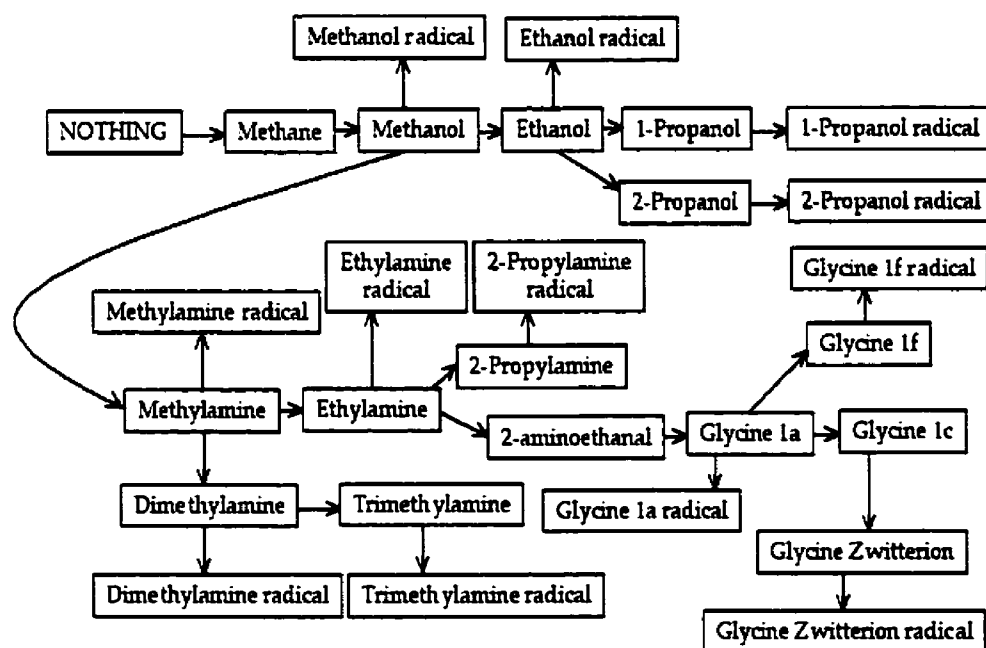

Figure 5: Solvation Tree
Each of the species in this figure was created by mutation from the one previous to it, following the paths indicated by the arrows. All species were derived from methane, which was created from NOTHING. BOSS provides a ΔG_{soln} for each of the mutations (each arrow).

perturbations in order to keep any individual perturbation from being too large. In order to accomplish this, functional groups from a solute are removed until one is left with methane (structure 27). At this point methane is mutated into a species with the identical geometry and Lennard-Jones parameters, but without any charges on the atoms, an approach known as electrostatic decoupling. This species only interacts with the solvent by taking up a certain volume. In a second perturbation, the C-H bonds are shrunk down to 0.1 Å and the Lennard-Jones parameters are removed, until the solute is tiny and has no effect on its surroundings. The sum of these two perturbations gives an absolute value for the solvation free energy of methane of 8.4 kJ mol^{-1}, exactly the same as experiment, and essentially the same as in the pioneering work of Jorgensen and coworkers.[59,72] Simply adding relative solvation free energies together gives the absolute solvation free energy of any molecule reachable through a series of mutations. This provides the absolute free energy of solvation values for the alcohols in this study, through the relationships pictured in Figure 5.

In order to avoid singularities due to the close approach of solvent molecules to charged species, electrostatic decoupling is more appropriate than direct mutation when aliphatic functional groups are being added or removed. Therefore, a similar procedure is used whenever a methyl group is being added. This procedure sees first, the growth of the methyl group out of a hydrogen, while maintaining the hydrogen's charge on the central carbon, and keeping all other charges constant. Second, the charges on all the atoms are mutated to their final values. This procedure increases the accuracy of the simulations without increasing the run times.

Thermodynamics

The thermodynamic results and data sources used in this study have been summarized in Table 5, on page 34. The quantities shown in bold were used to derive other dependent quantities in Table 5 and later in Table 6. For some of the parent compounds reliable values of $\Delta_f G_{(aq)}$ and $\Delta_f G_{(g)}$ exist in established databases. However, for all of the radicals and 1- and 2-propanol $\Delta_f G_{(g)}$ values had to be derived from $\Delta_f H_{(g)}$ and $\Delta_f S_{(g)}$ using the relation: $\Delta_f G_{(g)} = \Delta_f H_{(g)} - T\Delta_f S_{(g)}$. Literature values of $\Delta_f H(g)$ for the radicals were assessed carefully, and in two

cases values obtained by ab initio calculations with isodesmic reactions were preferred. The procedures used are described below.

Free Energies of Formation in the Gas Phase

$\Delta_f S_{(g)}$: The entropy of formation of a species, $\Delta_f S^O_{(g)}$, was calculated from the computed molar entropy, $S^O_{(g)}$, and the entropies of the elements. $S^O_{(g)}$ was calculated for all species, because the molar entropies of the parent alcohols in the literature serve as useful checks on the validity of the procedure used. Most species of interest here exist in the gas phase as equilibrium mixtures of conformers. Conformations arise from torsions around C-C and C-O bonds, and from the fact that the radical center is not planar, i.e., the out of plane bending potential has a double well. Relative energies, entropies, S_i, and populations (mole fractions), x_i, for each conformer of each species were therefore required. B3LYP calculations were used to obtain the energies, and S_i values were computed in the rigid rotator-harmonic oscillator model from the B3LYP structures and vibrational frequencies.

The total molar entropy is given by equation (19), where n is the number of

$$S^o_{(g)} = \sum_i^n x_i S_i - R \sum_i^n x_i \ln x_i \tag{19}$$

conformers and R is the ideal gas constant. The second term corresponds to the entropy of mixing. In Table 5, the second column gives S_i for the conformer used in the subsequent BOSS calculation. The third column gives the number of conformers and the fourth the molar entropy. For each of the species, the variation in entropy among the various conformers is very small, the largest range (2.8 J K^{-1} mol^{-1}) occurring in the case of 1-propanol. One might have assumed, as is usually done, that ΔS for a conformational change is negligibly small. However, ΔS_{mix}, the entropy of mixing term (which is approximately the difference between the 2nd and 4th columns of Table 5) is not. For instance, reasonable agreement between the calculated and experimental gas phase entropies of ethanol, and 1- and 2-propanol, could only be achieved by including ΔS_{mix}. The magnitude of ΔS_{mix} is not very sensitive to the actual proportions of the components, however. It is readily verified that R $\ln(n)$, the entropy of mixing of n equally populated conformations, provides a reasonable

approximation to ΔS_{mix}. As expected, the largest numbers of components occur for the most flexible species, 1-propanol (n =9) and its radical (n =12).

It may be noted that the double well nature of the radical bending potential is counted as giving rise to two conformations, even though in the case of CH_2OH (structure 29), the planar structure is lower in energy after the (harmonically approximated) ZPEs are taken into account. One may regard the "mixing" entropy of the two pyramidal forms of the radical center as compensating for the anharmonicity of the umbrella vibrational mode. The two component treatment of CH_2OH yields a value for the entropy, 245.4 J K^{-1} mol^{-1}, in close agreement with that obtained from a detailed analysis of the bending-torsion potential energy surface, 244.2 J K^{-1} mol^{-1},[75] (Table 5). The previously tabulated value[78] of 255.6 J K^{-1} mol^{-1} appears to be too large. The agreement between the calculated and literature values of $S^{\circ}_{(g)}$ for the parent alcohols is generally within 2 J K^{-1} mol^{-1} and satisfactory.

$\Delta_f H_{(g)}$: For the parent alcohols, experimental heats of formation are available (methanol,[73] ethanol,[73] 1-propanol,[80] 2-propanol[80]) and were adopted. For two of the radicals also there exist recent experimental $\Delta_f H_{(g)}$ values. In the case of $\cdot CH_2OH$, a combined theoretical and experimental (spectroscopic) investigation has established the heat of formation to high accuracy, -17.8 ± 1.3 kJ mol^{-1},[75] and this is consistent with independent experimental measurements of the BDE of methanol (401.9 kJ mol^{-1}.[11]). A UV-PES investigation[10] has established a value for the ethanol radical (structure 31), -56.9 ± 3.8 kJ mol^{-1}, which is almost 7 kJ mol^{-1} higher than previously reported values. It corresponds to an $^{\alpha}$C-H BDE of 396.2 kJ mol^{-1} for ethanol. The isodesmic reaction (20) with ab $initio$ energies computed at the G2MP2'-B3LYP level of theory,[70] yields a value of 396.3 kJ mol^{-1}
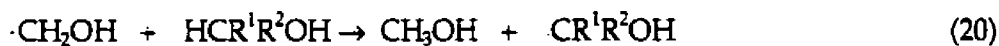
$$\cdot CH_2OH + HCR^1R^2OH \rightarrow CH_3OH + \cdot CR^1R^2OH \qquad (20)$$

Table 5: Thermodynamic data at 298.15 K: gas phase 1 atm; aqueous phase 1M.[a]

Molecule	$S^O_{(g)}$ [b]	n	$S^O_{(g)}$ [c]	$\Delta_f H^O_{(g)}$	$BDE_{(g)}$	$\Delta_f G^O_{(g)}$	$\Delta_f G^O_{(aq)}$
	J K⁻¹mol⁻¹		J K⁻¹mol⁻¹	kJ mol⁻¹	kJ mol⁻¹	kJ mol⁻¹	kJ mol⁻¹
Methanol	237.7	1	237.7	-200.7[d]	401.9[e]	-162.0[d]	-173.3[d]
(Structure 28)			239.8[d]	-201.6[f,g]	393.7[h]		
Methanol radical	239.6	2	245.4				-9.2[i]
(Structure 29)			244.2[j]	17.8±1.3[i,j]		0.1	
			255.6[k]	17.1[e]			
Ethanol	270.0	3	278.5	-235.1[d,g]	396.3[l]	-168.5[d]	-181.6[d]
(Structure 30)			282.7[d]	-234.8[l,m]	386.8[h]	-167.9[m]	
			281.6[m]				
Ethanol radical	276.3	4	287.6	-56.8[n]		-11.2	-15.1[i]
(Structure 31)				-56.9±3.8[o]			
				-63.6±4[i,p]			
1-Propanol	302.7	9	319.1	-255.4[q]	397.3[l]		
(Structure 32)			322.6[m]	-255.1[g,m]		-159.7	
1-Propanol radical	309.5	12	328.9	-75.8[n]		-1.8	
(Structure 33)							
2-Propanol	298.0	3	307.1	-272.8[g,q]	393.2[l]		
(Structure 34)			309.2[m]	-272.6[l,m]		-173.2	
2-Propanol radical	322.8	2	328.6	-97.4[n]		-23.2	-27.6[i]
(Structure 35)				-111.3±4.6[p]			

Definitions of table heading: $S^o_{(g)}$ entropy in the gas phase; n number of conformers; $\Delta_f H^o_{(g)}$, enthalpy of formation in the gas phase; $BDE_{(g)}$ gas phase bond dissociation enthalpy; $\Delta_f G^o_{(g)}$, $\Delta_f G^o_{(aq)}$ Gibbs free energy of formation in the gas and aqueous phase, respectively. [b] Calculated by rigid rotator-harmonic oscillator model. [c] Includes average entropy of n conformers and entropy of mixing. [d] Ref. 73 [e] Ref. 11; see also Ref. 74 (402.3 kJ mol⁻¹) and Ref. 75 (401.1 kJ mol⁻¹). [f] Ref 76. [g] Ref. 77. [h] Ref. 39. [i] Ref. 51. [j] Ref. 75. [k] Ref. 78. [l] Calculated from $HCR^1R^2OH + \cdot CH_2OH = \cdot CR^1R^2OH + CH_3OH$ with $\Delta_f H^o_{(g)}(H\cdot)=218$ kJ mol⁻¹. [m] Ref. 52. [n] Calculated using BDE and $\Delta_f H^o_{(g)}$ of parent. [o] Ref. 10. [p] Ref. 79. [q] Ref. 80.

(see Table 5). The level of agreement is excellent. Therefore the BDEs of 1-
and 2-propanol in Table 5 were calculated with this isodesmic reaction using the
same level of theory. The heats of formation of 1- and 2-propanol-derived
radicals were calculated from them and the $\Delta_f H_{(g)}$ values of the parents. The
enthalpies of the most stable parent and radical were used in equation (20).
Experimentally derived BDEs *are* available for ethanol and 1- and 2-propanol
(Table 5). However, these show a greater stabilization with increased methyl
substitution at the radical site than is supported by the theory. In light of earlier
experience with C-H BDEs in alkyl amines,[38] the values from the isodesmic
reactions were considered to be more reliable.

$\Delta G_{(g)}$: $\Delta G_{(g)}$ of CH_3OH and CH_3CH_2OH were taken from reference 73. For other
species, $S^\circ_{(g)}$, was converted to an entropy of formation from the elements,
$\Delta S^\circ_{(g)}$,[73] and combined with the best value of $\Delta_f H^\circ_{(g)}$ to obtain the $\Delta G^\circ_{(g)}$ values
shown in column 6 of Table 5.

Free Energies of Formation in Solution

The values of $\Delta_f G_{(aq)}$ for three of the radicals, $\cdot CH_2OH$, $\cdot CH(CH_3)OH$ {structure
31}, and $\cdot C(CH_3)_2OH$ {structure 35}, have been calculated from the measurement
of the reduction potentials of CH_2O, CH_3CHO and $(CH_3)_2CO$.[51] For CH_3OH and
CH_3CH_2OH $\Delta_f G_{(aq)}$ and $\Delta_f G_{(g)}$ were taken from reference 73.

Experimental Free Energies of Solution

For purposes of comparison with the BOSS results, ΔG_{soln} is the difference
between the free energy of formation of the substance in the aqueous phase and
the gaseous phase under the same conditions of temperature (298.15 K), and
concentration (1 M). Since the standard state for $\Delta_f G_{(g)}$ in Table 5 was 1 atm
rather than 1M, the values obtained from equation (1) must be corrected by -7.9
kJ mol^{-1}, as explained previously. ΔG_{soln} values calculated for the radicals by that
method are given in column five of Table 6, on page 39. Although the values of
$\Delta_f G_{(g)}$ for the radicals involved *ab initio* calculations of the entropies, the values of
ΔG_{soln} for all species obtained in this section are referred to as *Experimental Free
Energies of Solution*. The values for the four alcohols are from a review by Cramer
and Truhlar.[54] For methanol and ethanol, ΔG_{soln} values derived from $\Delta_f G_{(aq)}$ and

$\Delta_fG_{(g)}$ listed in Wagman, et al.[73] are in agreement with these. In the case of 2-propanol, a value derived from vapor pressure data is cited by Schwartz and Dodson.[51]

BOSS Calculations

Ab initio Methods. The structures of the parent alcohols and corresponding $^\alpha$C radicals were determined by complete geometry optimization using the B3LYP hybrid HF-DFT procedure as implemented in the Gaussian 94 suite of programs,[14] and the 6-31G(D) basis set. Vibrational frequencies were calculated and scaled by 0.98 for the purpose of deriving zero point energies and thermodynamic data in the rigid rotator-harmonic oscillator approximation.[63]

One conformation of each species was the "solute" of the BOSS calculations. Its gas phase entropy is shown separately and in no case differs by more than 1 e.u. from the average entropy of the components of the mixture (first term in equation (18)). In order to simulate the state of the solute as it exists in solution, the geometry optimization, frequency analysis, and thermodynamic calculations were repeated with SCRF=SCIPCM. A low torsional vibrational mode in the $^\alpha$C radical of 2-propanol in the SCRF calculation was approximated as a free internal rotor. All other internal rotations were treated as vibrations. Single point calculations at the B3LYP/6-311+G(3DF,2P) level were carried out in the presence and absence of the reaction field in order to test the effect of the basis set on the results. CHELPG charges to simulate the electrostatic potential,[66] were calculated with the large and small basis sets.

BOSS (Monte Carlo) Calculations. Absolute free energies of solution were derived using the BOSS Monte Carlo package,[50] following a modification of a procedure suggested by Lim and Jorgensen.[67] The SCRF structures, with charges calculated using the CHELPG procedure, were transferred to a periodic solvent box containing about 500 TIP4P[57] water molecules.

In the present work, the permutation tree shown in Figure 5 was applied. The quantities, $\Delta G_{(aq)}$BOSS(A$\ddagger$B), for each permutation of B3LYP-SCRF/6-31G(D) structures and CHELPG charges are indicated in Table 7. The absolute free energy of solution of a species is the sum of all of the permutation free energy

changes back to "NOTHING", and the associated error is the RMS of the individual errors. The largest single statistical error is for the permutation of methane to NOTHING.

Explanations

The computed results are summarized in Table 6. The first column of numbers corresponds to ΔG^{SCRF} of Figure 4, the free energy of distortion in the gas phase of the gas phase optimized structure to the geometry and the wave function of the solution structure (as modeled by SCRF=SCIPCM). The second column lists absolute free energies of solution which would be obtained if the gas and solution phase structures were identical and equal to the SCRF-derived species. These correspond to the vertical dashed arrow in Figure 4. The best calculated and "experimental" ΔG_{soln} values are listed in the last two columns. In Figure 6, are plotted the $\Delta\Delta G_{soln}$ values, namely the differences between the free energies of solution of the $^{\alpha}C$-centered radical and its parent alcohol by all three measures (columns 3 - 5 of Table 6). Table 7 and Figure 8 detail tests on the effects of basis set size and medium for optimization on methanol and its radical.

Comparison of Calc vs "Expt" (Table 6)

From Table 6 it is immediately apparent that there is agreement between calculated and best experimental values of ΔG_{soln} (columns 4 and 5 of Table 5) to within experimental accuracy of ~4 kJ mol^{-1} in every case.[81] The average deviation is 1.2 kJ mol^{-1}, the largest being 3.6 kJ mol^{-1} in the case of the ethanol parent. Within this small sample, the radical species are as well described as the parents. In Figure 6, the differential solvation, $\Delta\Delta G_{soln}$, of the $^{\alpha}C$ radical and its parent alcohol is displayed. Again, there is good agreement between theory and experiment in the three cases for which comparison is possible. While the $\Delta\Delta G_{soln}$ values are small, they are significant in the context of the statistical error of the BOSS FEP calculation. They indicate that the free radical is *less* solvated than the parent. It is of considerable interest that SCRF calculations (not shown) predict higher solvation (albeit to within 1 kJ mol^{-1}) for the radical relative to the parent. These results are contrary to the present discrete solvent model results
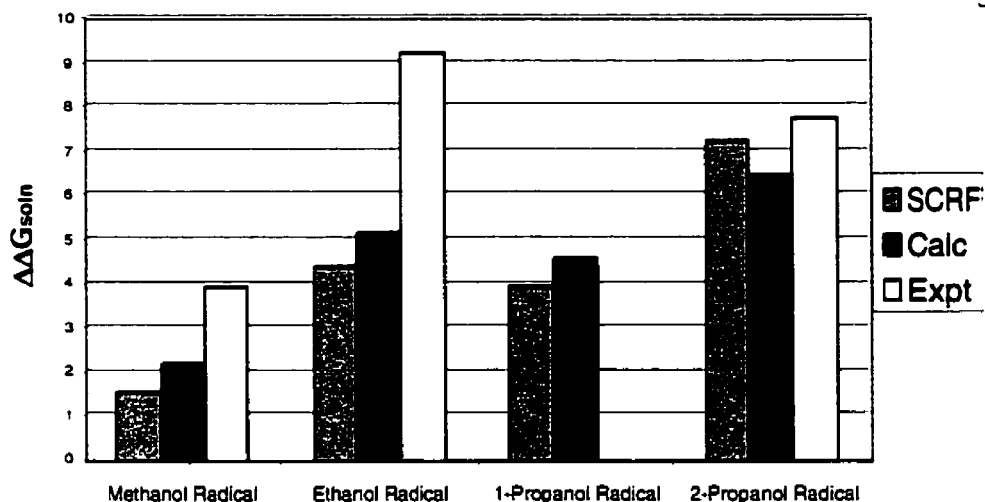
Figure 6: The difference in ΔG_{soln} of the C-centred radicals and their parent alcohols

$\Delta\Delta G_{soln}=\Delta G_{soln}(radical)-\Delta G_{soln}(parent)$ in kJ mol-1. SCRF, Calc, and Expt refer to columns 3, 4, and 5 of Table 6.
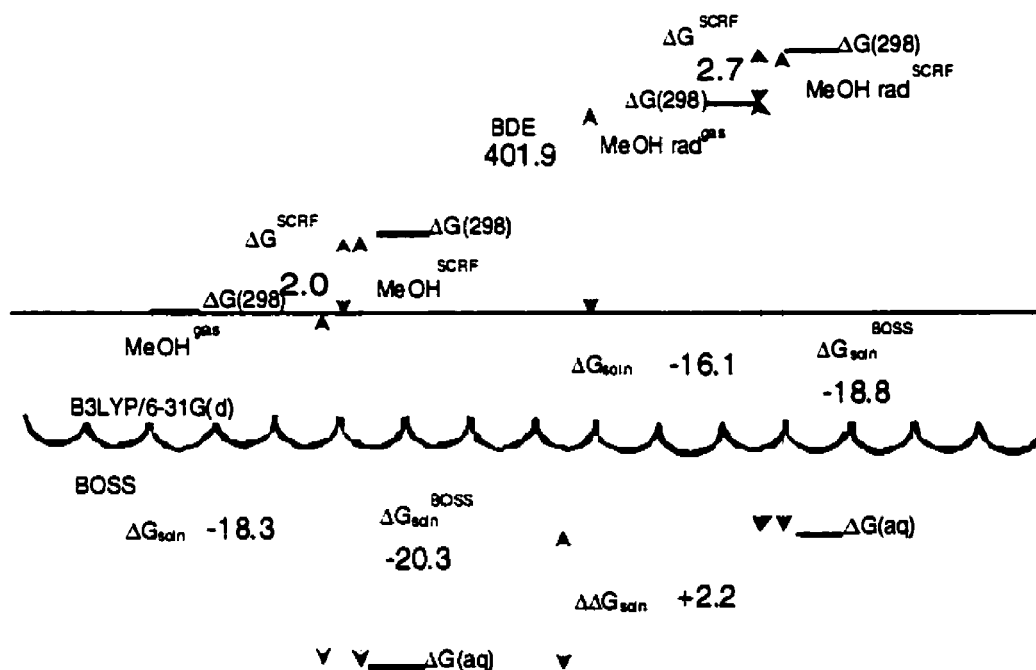

Figure 7: Schematic of the Methanol/ Methanol radical system

All energies in kJ mol^{-1}. This figure is not drawn to scale. $\Delta\Delta G_{soln}$ refers to the difference in solvation energies between methanol and the methanol radical ($\cdot CH_2OH$).

Table 6: Absolute and Relative Free Energies of solution (kJ mol⁻¹).[a]

Molecule	ΔG^{SCRF}	ΔG_{soln}^{SCRF} [b]	ΔG_{soln} Calc	ΔG_{soln} "Expt"
Methane	-0.38	8.4 (1.2)	8.0	8.4[c]
Methanol	1.97	-20.4 (0.6₅)	-18.4	-21.3[c,d]
Methanol radical	2.67	-18.8 (0.7)	-16.2	-17.3[e]
Ethanol	1.37	-18.9 (0.8)	-17.5	-21.1[c,d]
Ethanol radical	2.16	-14.5 (0.8)	-12.3	-11.8[e]
1-Propanol	1.65	-21.7 (0.9)	-20.0	-20.1[d]
1-Propanol radical	2.30	-17.7 (0.9₅)	-15.4	
2-Propanol	1.44	-21.3 (0.9)	-19.8	-20.1[d]
2-Propanol radical	0.67	-14.0 (1.0)	-13.3	-12.3[e]

[a] See Figure 4 for definition of symbols. [b] Numbers in parentheses are cumulated statistical errors relative to methane. The error for methane is relative to NOTHING. [c] Ref. 73. [d] Ref. 54. [e] Derived from data in Table 5.

and experiment. Furthermore, the present results suggest a general *decrease* in the magnitude of the free energy of solvation of the radical with increasing size or substitution at the $^\alpha C$ center, while the solvation of the parent is essentially constant (in this small sample). As the interaction of the solvent with the solute, whether radical or parent, is primarily electrostatic in nature, and the electrostatic potential of each is described approximately as an atom-centered monopole expansion via the CHELPG charges, one may compare the CHELPG charges of the two solutes for insight into the reason for the lower solvation of the radical. CHELPG charges of OH and methyl/methylene groups of the methanol radical and parent in both the gas phase and solution (SCRF) are representative of the series and are listed in columns 2 and 5 of Table 8. The principal difference between radical and parent in either phase is the considerably greater charge separation across the C-O bond in the case of the parent compared to the radical. This is shown graphically in Figure 8. The lower bond polarity of the radical reflects the delocalization of the oxygen lone pair of

Table 7: Relative Free Energies of Solvation calculated in this study

From...	To...	ΔG_{soln} (kJ mol^{-1})	error ($\pm$)(kJ mol^{-1})
NOTHING	Methane	8.4	1.20
Methane	Methanol	-28.7	0.65
Methanol	Methanol radical	1.5	0.18
Methanol	Ethanol	1.5	0.49
Ethanol	Ethanol radical	4.4	0.21
Ethanol	1-Propanol	-2.8	0.41
Ethanol	2-Propanol	-2.4	0.43
1-Propanol	1-Propanol radical	4.0	0.24
2-Propanol	2-Propanol radical	7.3	0.44

electrons into the half empty 2p orbital of the $^\alpha$C atom. As anticipated, the presence of the reaction field leads to increased charge separation compared to the gas phase in each case but the change is smaller in the case of the radical and the lower electrostatic potential around the O of the radical remains. In the parent, the C-O bond becomes more polar, making the oxygen more basic, and the polarity change increases as the $^\alpha$C center goes from primary to secondary to tertiary (not shown). The associated increase in H-bonding from the water must compensate for the hydrocarbon part of the parent, resulting in approximately constant solvation free energy over the series. On the other hand, the reaction field induces a smaller change in the radical, and the oxygen remains a poorer H-bond acceptor. The inability to account for such important solute-solvent interactions as H-bonding is an inherent limitation of all continuum models.
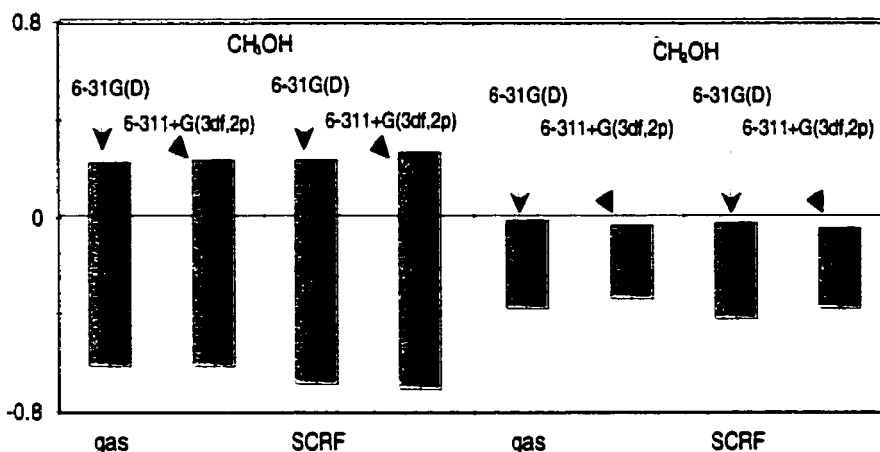
calculations into others which differ in wave function (gas phase or SCRF) with small (6-31G(D)) or large (6-311+G(3DF,2P)) basis set, and geometry (gas phase or SCRF). $\Delta\Delta G_{soln}$ describes the direct result of the FEP, and ΔG_{soln}^{SCRF} is the corresponding absolute free energy of solution (relative to NOTHING). The ΔG_{soln}^{SCRF} values may be directly compared with the experimental ΔG_{soln} values from Table 6: CH_3OH -21.3 kJ mol^{-1}; CH_2OH -17.3 kJ mol^{-1}. It is immediately apparent that all three SCRF wave function entries (columns 4-6) are in good agreement with experiment while the two entries where the gas phase wave function was used to derive the CHELPG charges (columns 2 and 3) show too low free energies of solution for both the parent and radical. These results are independent of whether a small or large basis set was used. Comparison of columns 4 and 5 confirms that essentially the same results are obtained with either gas phase- or SCRF-optimized geometries, provided the SCRF wavefunction (from a single point calculation in the first case) is used. The larger difference occurs in the case of the radical for which the free energy of solution is

Table 8: SCRF and Basis Set Dependence of ΔG_{soln} at 298.15 K for CH_3OH and $\cdot CH_2OH$.

Wave Function	gas	gas	SCRF	SCRF	SCRF
Basis Set	small	large	small	small	large
Geometry	gas	gas	gas	SCRF	SCRF
CH_3OH					
$\Delta\Delta G_{soln}$	7.7	7.8	0.3	0.0	-0.2
ΔG_{soln}^{SCRF}	-12.6	-12.6	-20.1	-20.4	-20.6
CHELPG of H(O)	+0.393	+0.381	+0.440	+0.440	+0.432
CHELPG of O	-0.611	-0.619	-0.677	-0.678	-0.700
CHELPG of CH_3	+0.218	+0.238	+0.237	+0.238	+0.268
$\cdot CH_2OH$					
$\Delta\Delta G_{soln}$	7.9	9.2	1.0	0.0	1.2
ΔG_{soln}^{SCRF}	-11.0	-9.6	-17.8	-18.8	-17.7
CHELPG of H(O)	+0.391	+0.369	+0.441	+0.445	+0.426
CHELPG of O	-0.371	-0.336	-0.414	-0.418	-0.380
CHELPG of $\cdot CH_2$	-0.020	-0.033	-0.027	-0.027	-0.046

calculated to be lower by 1.0 kJ mol^{-1} with the gas phase geometry compared to the SCRF-optimized geometry.

Sensitivity to Choice of Lennard-Jones Parameters. Because calculations on free radicals by BOSS have not previously been reported, one further test was carried out, namely the consequences of the choice of Lennard-Jones parameters for the tricoordinated radical center. The geometry at this carbon atom is intermediate between planar (sp^2 hybridized) and one with tetrahedral angles (sp^3 hybridized). All of the other results in this work use the same parameters as internally defined for an sp^2 hybridized carbon atom. An FEP calculation on the methanol radical was carried out in which the standard state (column 5 of Table 8) was changed to one with sp^3 hybridized carbon parameters. The result (not shown) was a $\Delta\Delta G_{soln}$ value of -0.04 kJ mol^{-1}. Thus, either choice would produce equivalent results. It was found that larger basis sets produce more polarized wavefunctions, which in most cases produce calculated free energies of solvation that are close to experiment. However, it was also found that the difference between small (6-31G(D)) and large (6-311+G(3df, 2p) basis sets was not as significant as the effect of continuum solvation calculations. Without the polarization introduced by the continuum solvation model, solute point charges were significantly smaller in magnitude, leading to smaller solvation free energy values. The optimal combination was found to be B3LYP/6-31G(D) SCIPCM wavefunctions- the computational expense imposed by the large basis set did not justify the insignificant improvement in calculated solvation free energy.

Conclusions

Methodology. The BOSS Monte Carlo discrete solution simulation package, combined with quantum mechanical (QM+BOSS) calculations, is capable of yielding accurate free energies of aqueous solution for $^\alpha$C-centered free radicals derived from alcohols, and for the parent alcohols themselves. The results are not sensitive to the choice of Lennard-Jones parameters for the radical center. The recommended procedure involves geometry optimization and frequencies by QM methods (B3LYP/6-31G(D)) in the gas phase, followed by a single point SCRF=SCIPCM calculation to obtain CHELPG charges. The SCRF-derived CHELPG charges model the electrostatic potential of the substance as it exists in

solution and is seen by the solvent water molecules. They are not sensitive to the size of the basis set. Omission of the SCRF step yields free energies of solution which are too low compared to experiment.

Gas phase free energies of formation of $^\alpha$*C-centered free radicals.* The gas phase free energies of formation of the $^\alpha$C-centered radicals from the lower alcohols were derived from a combination of experimental data and theoretical procedures. Enthalpies of formation were taken from experiment (methanol radical and ethanol radical) or derived from $\Delta_f H^\circ_{(g)}$ of the parent alcohols and calculated BDEs (radicals of 1- and 2-propanol). Entropies were obtained from the rigid rotator harmonic oscillator approximation, taking account of the conformational mix of the free radicals. The derived quantities are listed in Table 5.

Solvation of $^\alpha$*C-centered free radicals.* The absolute free energies of solution, ΔG_{soln}, of the $^\alpha$C-centered radicals from the lower alcohols in water are quantitatively described by QM+BOSS calculations: methanol radical, expt -17.3 kJ mol^{-1}, calc -16.2 kJ mol^{-1}; ethanol radical, expt -11.8 kJ mol^{-1}, calc -12.3 kJ mol^{-1}; 2-propanol radical, expt -12.3 kJ mol^{-1}, calc -13.3 kJ mol^{-1}. A value is predicted for the 1-propanol radical, ΔG_{soln} = -15.4 kJ mol^{-1}. The radicals are less solvated than the parent alcohols. Examination of the CHELPG charges suggests that the reason lies in the lower polarity of the C-O bond and lower H-bond acceptor ability of the oxygen atom. The latter factor is not modeled by continuum models. The SCRF=SCIPCM procedure actually yields slightly higher solvation energy for the radicals in contradiction with the experimental and the present QM+BOSS results.

V. Solvation of Radicals II: Amines and Glycine

Introduction

Given the success of the alcohol study, it was seen to be useful to extend the work to a different set of systems that had already been thoroughly examined in the gas phase. This research group had previously published an extensive study of the thermodynamics of bond dissociation of amines[38]. That study indicated that the presence of aqueous solvent had no effect on the thermodynamics of the process. In order to test whether that holds in the case of simple aliphatic amines and their radicals, QM+BOSS solvation calculations were performed on methylamine, ethylamine, 2-propylamine, dimethylamine, and trimethylamine, and their radicals.

Amines

Previous Work

In earlier work, the BDEs of the carbon-hydrogen bond alpha to the nitrogen in methylamine, ethylamine, dimethylamine, trimethylamine, and 2-propylamine were calculated[38]. This was done in the gas phase, at 298 K, at the G2(MP2) level, with HF and MP2/6-31G(D) geometries. These values were confirmed with photo-acoustic calorimetry measurements[38]. The excellent agreement between the calculated gas phase results and the experimental solution measurements implied that the relative free energies of solution between parent and radical are close to zero. This can be tested using the newly-developed solvation calculation technique described previously. Predictions can then be made about the BDEs of these amines in aqueous solution.

Present Calculations

Recognizing that in the alcohol calculations, the differences in geometry between the structures optimized with and without SCIPCM were minimal, the amines' geometries were optimized in the gas phase only, at the B3LYP/6-31G(D) level. The CHELPG charges depend on a polarized wavefunction, however, so single-

point calculations with the SCIPCM model were performed on the gas-phase structures, and CHELPG charges taken from the resultant polarized wavefunctions. This was tested by performing full SCIPCM optimizations for both methylamine {structure 36} and ethylamine {structure 38}; the difference in both geometry and ΔG^{SCRF} between the approaches was minimal (results not shown).

Again, the Lennard-Jones parameters were taken from OPLS parameters available within the BOSS package for similar systems, with the radical centre modelled as an sp^2 carbon. Each species was modified from its predecessor over 10 sampling windows with double-wide sampling, and each window's energy was averaged over 6,000,000 configurations, as previously described.

Solvation Thermodynamics of Amines

Results: The results of the QM+BOSS calculations on the amines and their αC-centred radicals are summarized in Table 9.

Table 9: Solvation Energies for Amines and Derived Radicals

Molecule	Structure	ΔG_{soln} [a]	ΔG_{soln} [b]	ΔG_{soln} [b,c]	$\Delta \Delta G_{soln}$ [b,d]
Methylamine	36	-19.1	-20.6	-10.8	9.8
Ethylamine	38	-18.8	-16.5	-6.4	10.1
2-Propylamine	40	N/A	-20.8	-8.8	12.1
Dimethylamine	42	-17.9	-6.9	-3.0	3.9
Trimethylamine	44	13.6	+5.9	+5.5	0.4

All values in kJ mol^{-1} [a] Experimental (Ref. 82) [b] Present calculation [c] Radical [d]Parent -> Radical

Experimental ΔG_{soln} values for the parent amines are available for all but 2-propylamine. Agreement with these values would validate the extension of the technique to the amines. The results of the solvation calculations are presented in Table 9. The agreement for methylamine and ethylamine is excellent, with a difference of only 1.5 kJ mol^{-1} between the calculated and experimental values for methylamine (-20.6 kJ mol^{-1} (calculated), -19.1 kJ mol^{-1} (experiment) [38]), and a difference of 2.3 kJ mol^{-1} for ethylamine (-16.5 kJ mol^{-1}

(calculated), -18.8 kJ mol^{-1} (experiment)[38]). The calculations are expected to have errors of less than 10 kJ mol^{-1}. Unfortunately, the agreement for dimethylamine {structure **42**} and trimethylamine {structure **44**} is very poor, indicating that something is wrong with the assumptions underlying the calculation (see Discussion).

Because there are no experimental sources to compare with regarding ΔG_{soln} for the amine radicals, no conclusions can be drawn as to the accuracy of the calculations.

Discussion (parent species): Molecular mechanics[65] and CHELPG[66] both assume solvation is dominated by electrostatics, that is, the interaction of point charges. The addition of a methyl group to an amine causes the absolute value of the charge on the nitrogen to be lowered, as observed by CHELPG. The overall charge on the methyl group is less positive than the charge on an amine hydrogen, so to compensate, the charge on the nitrogen is made less negative. Reducing the charge on the nitrogen reduces the electrostatic interaction between the amine and water, which reduces the magnitude of the solvation free energy. In addition, the methyl group is hydrophobic in itself. However, in reality, the solvation energy of dimethyl- and trimethylamine is dominated by the hydrogen-bonding ability of the nitrogen, which arises from its *basicity*, not its excess negative charge. The basicity of the nitrogen is *increased* by the electron-donating methyl group, which allows the strengthened hydrogen bond to water to overcome the hydrophobic effects of the extra methyl group. This electronic effect is not allowed for in any standard solvation code, including continuum models.[55] Therefore, any prediction of the absolute solvation energy of secondary or tertiary amine species using molecular mechanics is suspect at this time.
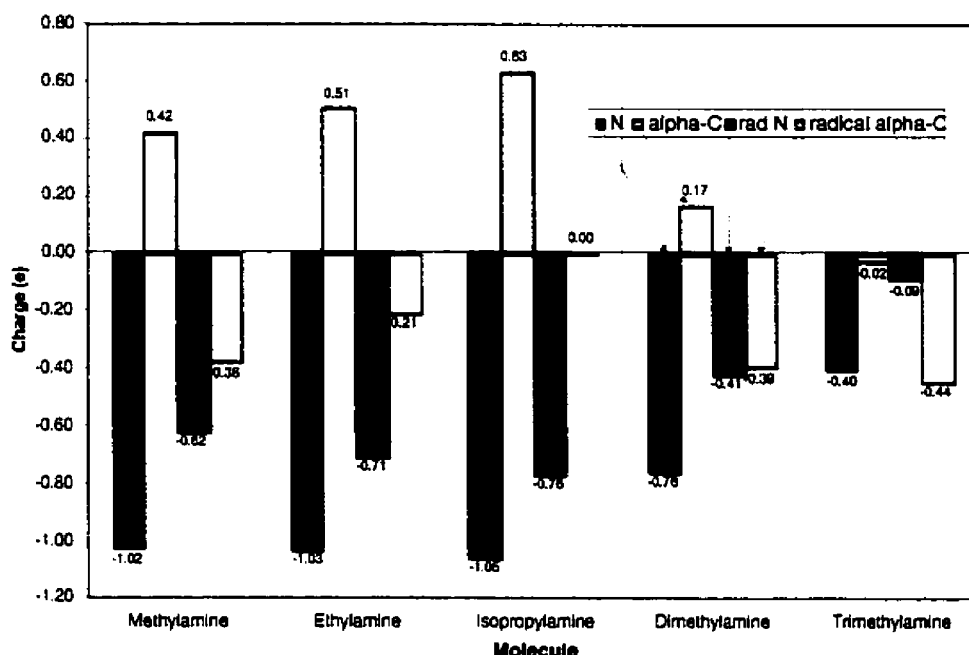
48

Figure 9: Variation of Charge across the Amines
N: Nitrogen on parent amine   a-C: alpha carbon on parent amine
rad N: Nitrogen on amine radical   rad a-C: alpha carbon radical centre

Discussion (radical species): Since the electrostatic component of the solute-solvent interaction is the significant variable over the course of the mutation from parent to radical, the calculated solvation energy of the amines and their radicals can be correlated with the difference between the CHELPG charges on the nitrogen and the α-carbon. Primary amines allow delocalization of charge from the nitrogen to the radical centre, significantly decreasing the difference between the two (Figure 9), and therefore also decreasing the solvation energy (Figure 10).

There are two issues to be concerned with: first, since the agreement with experiment for secondary and tertiary amines is poor, the absolute solvation energies calculated here for their radicals are almost certainly unreliable. Second, the BOSS package only calculates relative solvation free energies, and that data may or may not be useful. The relative calculations between very similar species have less potential for error than calculations between very different molecules, since there are fewer variables in the former case. However,

if the technique itself is unable to accurately measure solvation free energies for amines at all, then the usefulness of even the relative measurements is suspect.
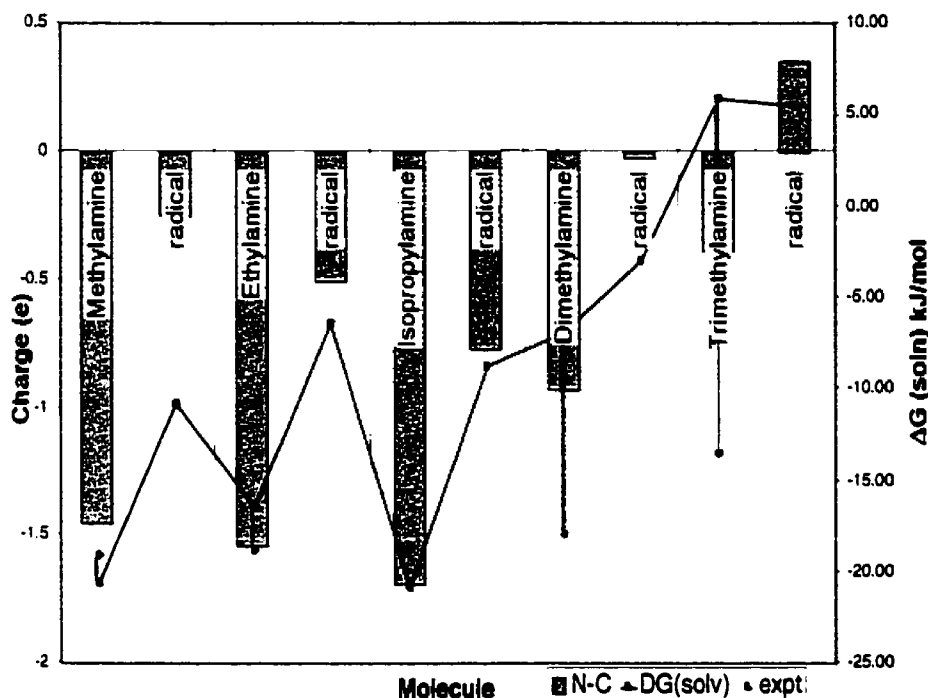

Figure 10: Variation of Charge and Solvation Energies
The wide grey bars show the charge separation between the nitrogen and carbon. The thin line connects the calculated free energies of solvation of the compounds. The grey vertical lines connect the calculated values of ΔG_{soln} to the experimental values, from Ref. 38.

Possible solutions to the 'amine solvation problem': This problem was recognized in recent work by Rizzo and Jorgensen.[33] It was noted in that work that the solvation of amines and amides is an area where classical force fields have failed. The solution presented was to re-parameterize the force fields, taking into account the failures of the previous parameters. The effort began with the adjustment of the point charges on the atoms in amines and amides, away from CHELPG-like values. Further optimization yielded new Lennard-Jones parameters as well. This approach, while useful for its intended purpose (inclusion into a database of suitable parameters for organic molecules) fails for our purposes.

The alcohol study was successful in part because it was proved to be a useful assumption that the CHELPG charges were a good representation of the charge distribution in alcohol radicals. It is apparent that this was the case because the alcohol-water interaction is dominated by electrostatics more than the amine-water interaction is. The approach presented in this work attempts to use as physically reasonable a model as possible. Thus CHELPG charges, calculated *ab initio*, were used in all cases for solute parameterization. Rizzo *et al* adjusted the charge distribution in order to reproduce desired chemical behaviour. This does not allow another researcher to extend the model using any physically reasonable assumptions. Rizzo's study rejects CHELPG charges explicitly because of their failure to reproduce experimental results. Therefore, successful calculation of amine radical solvation energy using Rizzo's approach must await successful experimental measurements, which would provide a target for optimization. Prediction is thus ruled out.

A physically reasonable solution to the prediction of amine radical solvation may be found if the theory underpinning BOSS can be adjusted to include polarizability, the influence of a third body on the interaction between two. Promising work has been reported in the latter direction for several years.[84,85,86] Indeed, Jorgensen has published work toward that end,[87,88] but an agreed-on standard with a modified parameter set has not been established at this point. Many workers, including Rizzo and Jorgensen,[83] have abandoned polarizability as being not as productive as optimization of the parameters used in current methods. Cases where experimental results are not available are rare, so it may fall to groups studying hard-to-measure systems to extend the model for their own purposes.

The more promising alternative, given the continuing increase in computational power available, is inclusion of quantum-mechanical calculations into solvation calculations. The solution would entail a Monte Carlo calculation of a system, with the solvent being treated using molecular mechanics, and the solute, with perhaps the nearest neighbour solvent molecules, being treated with QM. This has been done using semi-empirical code.[89] The nature of MC requires the recalculation of the energy of the QM system thousands or millions of times,

which is not a trivial task, but its achievement awaits only the processing power to make it practical.

Results: Bond Dissociation Energies

The inability of this solvation model to reproduce the known solvation free energies of the secondary and tertiary amines casts doubt on the whole methodology. However, because this approach relies on calculating *relative* free energies, the $\Delta\Delta G_{soln}$ data (Table 10, column three) may prove useful. It is clear that the model predicts that primary amine radicals are solvated to a lesser extent than their parents, by about 10 kJ mol^{-1}. Secondary and tertiary amines do not have as great a charge difference as the primary amines, so that the formation of a radical cannot have as great an effect, in electrostatic terms. This means the solvation of the radicals is calculated to be about the same as that of their parents. This prediction is the result of the change in bond polarity of the C-N bond, and depends on this electrostatic effect overcoming any change in the basicity of the radical.

Table 10: Bond Dissociation Energies for Amines

Molecule	BDE$_{(g)}$ [a]	$\Delta\Delta G_{soln}$ [b]	BDE$_{(expt)}$ [c]
Methylamine	388	+10	393
Ethylamine	384	+10	377
2-Propylamine	388	+12	372
Dimethylamine	386	-1	364
Trimethylamine	387	0	351

All values in kJ mol^{-1} [a] Calculated (Ref. 38) [b] Present calculation [c] Experimental (Ref. 90)

Table 10 contains the calculated BDEs of the five amines studied. The calculated BDEs in the gas phase were taken from the previous study, and only the $\Delta\Delta G_{soln}$ was contributed by this present study. Gas phase experimental BDEs from reference 90 are given in the last column. These experimental BDEs show a steady trend to lower values with increasing substitution, which was questioned in reference 38, both because of the calculated values (Table 10, column two), and

new photo-acoustic experimental data.[38] In addition, the reduction potential of the trimethylamine radical cation was recently measured in aqueous solution. The BDE may be derived from that measurement, and was determined to be 389 kJ mol^{-1}.[38] This supports the calculated value of 387 kJ mol^{-1} in column two. The experimental value depended upon the assumption that $\Delta\Delta G_{soln}$ was negligible for trimethylamine and its radical (Me$_2$NC·H$_2$). That assumption is entirely in agreement with the value in column two.

Conclusion

Any conclusions about amines from this study are premature, but it does seem likely that *primary* amine radicals, like those of alcohols, will not be solvated to the same extent as their parents. Some adjustments, on the order of +10 kJ mol^{-1}, may be required when estimating BDEs in aqueous solution as opposed to the gas phase or non-polar media.

The poor results for the secondary and tertiary amine parents throw into question the results obtained for their radicals. If the radicals are expected to exhibit differing basicities relative to their parents, and if basicity is the source of the poor agreement, then the $\Delta\Delta G_{soln}$ results presented here will almost certainly be in error. This problem awaits the development of a more accurate technique.

Glycine

Glycine is a very important problem in its own right. Glycine {structure 1} is the simplest of the amino acids and as such is usually the first amino acid studied with any new approach.[91,92,93] Glycine is particularly significant for a solvation study, as its structure is very different depending on its environment. In the gas phase, glycine exists as a neutral molecule with an intramolecular hydrogen bond. Once glycine is placed in aqueous solution, the acid hydrogen is transferred to the nitrogen, forming a zwitterion {structure 26}. This charge-separated species is considerably higher in energy in terms of its electronic structure, but the very favourable interaction of the charged species with the solvent more than balances out the difference in electronic energies. The result is that the zwitterion is overwhelmingly favoured in solution.

Previous Calculations

This research group has long experience with glycine, leading up to it being chosen as the isodesmic partner for (gas-phase) BDE calculations with other amino acids.[18] The previously calculated BDE is 331 kJ mol^{-1},[21] using G2(MP2) theory. A previous study[19] has determined the relative energetics of various conformations of glycine in the gas phase. The notation of that paper has been adopted here. Six of these conformations (1a-1f) were re-optimized at the B3LYP/6-31G(D) level of theory, and their energies recalculated at B3LYP/6-311+G(3df, 2p) to determine the relative energies of the conformers. B3LYP/6-31(D) frequencies were calculated, and transformed into entropies using standard methods. Three of these conformations were placed in solution, using both SCIPCM with frequency calculations, and BOSS mutations, to determine their ΔG_{soln}. The free energies of solvation are hard to determine experimentally, as the zwitterion (zwit) is effectively the only species present in solution, but is absent in the gas phase. Indeed, the zwitterion is not a stable species in the gas phase according to high-level calculations. However, the (gas phase) single-point energy of the optimized (in solution) zwitterion geometry and wavefunction were calculated.

Table 11: Relative Energetics and BDEs for the lowest-energy conformers of Glycine

Glycine Conformer	ΔG relative to 1a (gas phase)	ΔG relative to 1a (solution)	BDE (gas)	BDE (soln)
1a (structure 1)	0.0	0.0	331.0	334.6
1b (structure 47)	5.6			
1c (structure 48)	5.6		377.6	
1d (structure 50)	6.1			
1e (structure 51)	11.4			
1f (structure 52)	19.5	12.2	332.8	354.2
Zwitterion (structure 26)	N/A	-18.9		407.2

The gas phase results elicited few surprises; the 1a conformation is the most stable, as expected. It exhibits Cs symmetry, with the carboxylic acid hydrogen in *cis* form, attached to the oxygen furthest from the nitrogen. About 6 kJ mol^{-1} higher lie conformations 1b, 1c, and 1d, {47,48,50} almost identical in energy. Some 5 kJ mol^{-1} higher lies conformation 1e {51}, with 1f {52} another 8 kJ mol^{-1} above that. In solution, glycine is almost entirely a zwitterion {26}. The experimental estimate of the difference in $\Delta_f G^o_{(aq)}$ between the zwitterion and the next lowest conformation is 31 kJ mol^{-1}. This calculation produced a difference of only 19 kJ mol^{-1}. The difference will be discussed below.

The question of which neutral form is favoured in solution is somewhat complex: the solvation energy is dominated by the number of hydrogen bonds that can be formed. The most favourable *inter*molecular interaction is between the solvent and 1f {structure 52}, which has the acid hydrogen in *anti* form. Conformer 1a, with a *cis* hydrogen, contains an intramolecular hydrogen bond, and is slightly favoured overall.

The study of the radicals in the gas phase also held few surprises. The radicals were formed by removing one of the hydrogens from the central carbon and reoptimizing at B3LYP/6-31G(D). It was found that the 1a form was again the lowest in energy.

The differences between the parent and radical are instructive, and have to do with captodative stabilization. In the glycine radical, in order to maximize the interactions of the available π bonds, the C--H becomes trigonal planar, as does the NH$_2$ group. The favoured form is therefore all planar.

Like the parent zwitterion, the zwitterion radical was not found to be a stable species in the gas phase. The hydrogen that completes the intramolecular hydrogen bond between the nitrogen and the carboxyl oxygen would swing back to the oxygen, forming the 1c conformation.

It was found that for radicals in both gas phase and solution, and for the closed shell population in solution, the dominant species exists as greater than 99.95% of the population. For the closed shell in the gas phase, there is a contribution to the total energy from the entropy of mixing, because the six lowest energy

conformers are relatively close. This provides ~2 kJ mol^{-1} of energy to the free energy of the mixture. However, the energy of the most stable conformer (1a) is a good approximation (within 1 kJ mol^{-1}) of the corrected energy of the mixture.
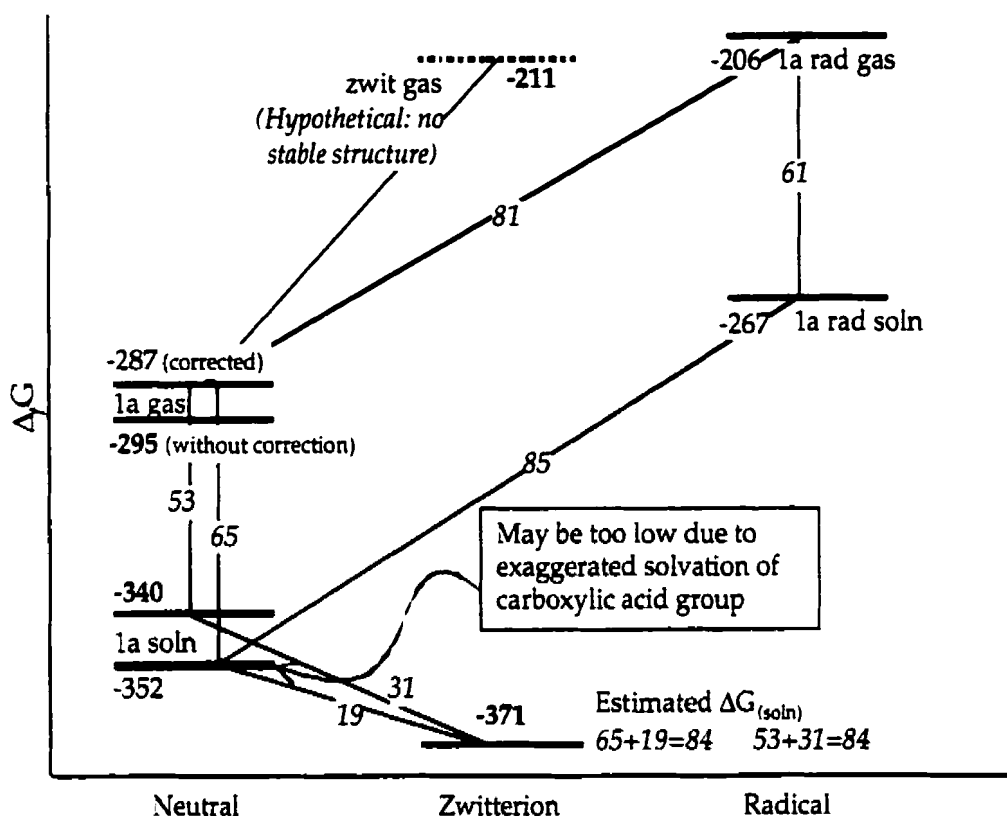

Figure 11: Schematic of Glycine Energetics

All energies in kJ mol^{-1}, at 1 M. Bold numbers indicate results from Ref. 31. Non-bold numbers are from the present study. Italic numbers indicate relative measurements. Correction is 7.9 kJ mol^{-1}, the difference between 1 atm and 1 M standard state.

Glycine Radicals in Solution

The solvation energy of the closed shell zwitterion was able to overcome its higher internal energy. Is this the case with the zwitterion radical? The present calculations indicate that the answer is negative. It was found that even though the closed-shell zwitterion is favoured in solution, the 1a radical (structure 2) is

lower in energy than the zwitterionic radical {structure **54**} by 53 kJ mol^{-1} when solvated. This is due to the extremely high internal energy of the zwitterion radical. Captodative stabilization, which depends on the presence of an electron donating group and an electron withdrawing group (Figure 1), is not present in the radical because of the extra hydrogen on the amine. The NH_2 group in the neutral species is an electron-donator, but NH_3^+, with a net positive charge, withdraws electrons from the alpha carbon. The acid carbon is also electron withdrawing, due to the electronegativity of the oxygens. The radical centre is thus surrounded by two electron-withdrawing groups. Despite the considerable solvation advantage the zwitterion radical enjoys, it cannot overcome the resulting loss of captodative stabilization. Therefore, the captodatively stabilized 1a radical is the preferred conformer in solution. This was confirmed through QM determination of the BDE for the 1c conformation, and QM+BOSS determination of the BDE in solution for the 1f conformation (Table 11).

Limitations of BOSS Calculations

The gas-phase calculations of glycine's conformers are expected to be of experimental accuracy. The solvation calculations, on the other hand, disagree with experiment on the crucial measure of the relative $\Delta_f G^o_{(aq)}$ of the glycine neutral/zwitterion pair (Figure 11). This shortcoming may arise from errors in determining the solvation of either of the two functional groups present in glycine.

It has been seen that BOSS calculations on amines are not entirely trustworthy, and that the OPLS parameter set is being modified in un-physical ways to accommodate that fact.[83] BOSS also has difficulty calculating *acid* solvation free energies.

Tests done on formic and acetic acid do not give appropriate solvation free energies, regardless of whether the acidic hydrogen is *cis* or *anti* to the carbonyl group (unpublished results). This weakness does not seem to be unique. Judging from the review by Cramer and Truhlar,[54] it is seen in both continuum models and in molecular mechanics force-field-based calculations.

Analysis of these calculations shows that the problem may lie in the assumption made by classical force fields that three-body effects may be ignored. [83,94] The carboxylic acid group is perfectly arranged for multiple hydrogen bonds, especially at the double-bonded oxygen. This one oxygen can be H-bonded to two different water molecules at the same time. In that case, the average electron density available to each of the water molecules is reduced. In effect, the first H-bond reduces the 'effective charge' of the oxygen atom that the second H-bond experiences. The H-bonds are thus expected to be weaker than they would be if they were the only H-bond the oxygen was participating in.

Because of the lack of three-body interactions, the BOSS simulation does not reflect this. Instead, the full 'effective charge' of the oxygen is available to both hydrogens, allowing both to form strong H-bonds. The only effective counterbalancing interaction between the two participating hydrogens is their electrostatic repulsion, as they will rarely be in Lennard-Jones contact. This is not expected to make up for the artificially strong hydrogen bonds produced as an artifact of the program. This error results in carboxylic acids being too strongly hydrated, relative to experiment.

A polarizable force field would not be expected to have this problem. Supermolecular QM calculations would also be expected to display more accuracy in this situation. Modification of the CHELPG parameters to correct for this error, á la Rizzo's amine optimization,[83] is not compatible with the goal of a physically reasonable methodology.

The problem is not as obvious in the case of the zwitterion in solution, since plain electrostatics dominate the solvation free energy, due to the presence of two charged groups both available to the solvent. In these calculations, the greater (in magnitude) solvation free energy of neutral glycine was counter-balanced by less of a difference between the neutral and zwitterionic forms, so that the total solvation free energy of glycine was calculated to be the same as in previous work, 84 kJ mol[-1].

Summary

The present work is of value in giving relative energies between conformations of glycine, and between parent and radical. Previously calculated gas phase energetics have been confirmed, using hybrid density functional theory, and with the use of much less computer time. QM+BOSS calculations were performed to determine relative free energies of solvation between conformers, and across the bond dissociation reaction. The zwitterion radical was found to be not as stable in solution as the neutral radical (conformation 1a).

VI. Conclusion

Two very different projects were embarked upon, in order to further the understanding of the strength of C-H bonds in biological molecules. The αC-H BDEs of proline and several proline model peptides were determined. The models simulated several of proline's most commonly found roles in protein secondary structure. It was found that the BDEs of the proline models' αC-H bonds are actually stronger than those of the other amino acid model peptides, This confers some protection against oxidative damage from weak oxidizers such as S· at proline. These results, in contrast with the experimental results which indicate that proline is vulnerable to oxidative damage, show that thermochemistry is not always the deciding factor in protein radical chemistry.

Does $\Delta\Delta G_{soln}=0$? The second part of this work attempted to answer that question for some simple alcohols and amines, and glycine. Unfortunately, one cannot make a blanket generalization about the solvation behaviour of radicals relative to their parents. In order to predict the relative free energies of solvation, one must examine the polarizability of the system. In radicals where an interior heavy atom is the site of the bond dissociation, delocalization will increase the electron density at the radical centre, in the interior of the molecule, where it is unavailable to solvent. This often results in lowered charge at the exterior, so that solvent is not bound as strongly as in the parent.

The magnitude of the effect is not very large. In most cases studied the previous assumption, that $\Delta\Delta G_{soln} = 0$, is reasonable within the expected error of the calculation. In some systems, however, the effect is significant enough that it

must be taken into account. In all systems we were able to check against experiment, the hybrid QM+BOSS Monte-Carlo/quantum mechanics approach is superior to a simple continuum model in determining the magnitude of this effect.

It is apparent, however, that there are significant defects within the current approach to solvation. The reduction of solvation to two-centred electrostatics removes both polarizability and multi-body effects from the calculation. This does not allow for such things as the basicity effect of amines in solution, or the distribution of charge around a solvated carboxylic acid. These shortcomings strike directly at the biological structures this project was intended to study, leaving no option but to declare that the present approach, while promising in theory, is inadequate in practice. It is expected, however, that the addition of a better description of electron polarizability, either within the classical force-field approximation, or via increased QM calculations, would significantly improve the quality of these results.

Endnotes

1.   Stubbe, J; van der Donk, W A; *Chem Rev.* **1998**, *98*, p 705.

2.   Klapper, M H; Farragi, M; Mishra, A K; Chandrasekar, R; *J Am Chem Soc.* **1994**, *116*, p 1414.

3.   Simic, M G; Taylor, K A; Ward, J F; von Sonntag, C; *Oxygen Radicals in Biology and Medicine,* **Plenum Press**: *New York,* (1988).

4.   Sies, H; *Oxidative Stress: Oxidants and Anti-Oxidants,* **Academic Press**: *Toronto,* (1991).

5.   Stadman, E R; Oliver, C N; *J Biol Chem.* **1991**, *266*, p 2005.

6.   Stadman, E R; *Annu Rev Biochem.* **1993**, *62*, p 797.

7.   Leroy, G; Sana, M; Wilante, C; *J Mol Struct.* **1991**, *228*, p 37. See also Ref. 3.

8.   Kanabus Kaminska, J M; Gilbert, B C; Griller, D; *J Am Chem Soc.* **1989**, *111*, p 3311.

9.   Parkinson, C J; Mayer, P M; Radom, L; *Theoret Chem Acc.* **1999**, *102*, p 92.

10.   Dyke, J M; Groves, A P; Lee, E P F; Zamanpour Niavaran, M H; *J Phys Chem.* **1997**, *101*, p 373.

11.   Berkowitz, J; Ellison, G B; Gutman, D; *J Phys Chem.* **1994**, *98*, p 2744.

12.   Becke, A D; *J Chem Phys.* **1993**, *98*, p 5648.

13.   Jursic, B S; *Chem Phys Lett.* **1996**, *256*, p 603.

14.   Binkley, J S; Frisch, M J; Trucks, G W; Schlegel, H B; Gill, P M W; Montgomery, J A; Johnson, B G; Raghavachari, K; Peng, C Y; Robb, M A; Al Laham, M A; Ayala, P Y; Cheeseman, J R; Zakrzewski, V G; Defrees, D J; Chen, W; Keith, T A; Ortiz, J V; Baker, J; Wong, M W; Petersson, G A; Foresman, J B; Stewart, J P; Andres, J L; Cioslowski, J; Head Gordon, M; Replogle, E S; Stefanov, B B; Gonzalez, C; Gomperts, R; Nanayakkara, A; Pople, J A; Martin, R L; Challacombe, M; Fox, D J; *Gaussian 94 Revision D2,* **Gaussian Inc**: *Pittsburgh PA,* 1995.

15.   Jonsson, M; Wayner, D D M; Armstrong, D A; Yu, D; Rauk, A; *J Chem Soc Perkin Trans 2.* **1998**, p 1967.

16. Rauk, A; Yu, D; Taylor, J; Shustov, G V; Block, D A; Armstrong, D A; *Biochemistry*. **1999**, *38*, p 9089.

17. Hehre, W J; Ditchfield, R; Radom, L; Pople, J A; *J Am Chem Soc*. **1970**, *92*, p 4796.

18. Rauk, A; Yu, D; Armstrong, D A; *J Am Chem Soc*. **1997**, *119*, p 208.

19. Armstrong, D A; Yu, D; Rauk, A; *Can J Chem*. **1996**, *74*, p 1192.

20. Viehe, H G; Janousek, Z; Merenyi, R; Stella, L; *Acc Chem Res*. **1985**, *18*, p 148.

21. Yu, D; Rauk, A; Armstrong, D A; *J Am Chem Soc*. **1995**, *117*, p 1789.

22. MacArthur, M W; Thornton, J M; *J Mol Biol*. **1991**, *218*, p 397.

23. Dixon, D A; Dobbs, K D; Valentini, J J; *J Phys Chem*. **1994**, *98*, p 13435, referred to in Rauk, A; Glover, S A; *J Org Chem*. **1996**, *61*, p 2337.

24. Wilmot, C M; Thornton, J M; *J Mol Biol*. **1988**, *203*, p 221.

25. Richardson, J S; *Adv Protein Chem*. **1981**, *34*, p 116.

26. Moran, L A; Scrimgeour, K G; Horton, H R; Ochs, R S; Rawn, J D; *Biochemistry 2nd ed*, **Neil Patterson Publ Prentice**: *Englewood Cliffs NJ*, 1994.

27. Rose, G D; Gierasch, L M; Smith, J A; *Adv Protein Chem*. **1985**, *37*, p 1.

28. Uchida, K; Kato, Y; Kawakishi, S; *Biochem Biophys Res Commun*. **1990**, *169*, p 265.

29. Amici, A; Levine, R L; Tsai, L; Stadman, E R; *J Biol Chem*. **1989**, *264*, p 3341.

30. Sharpatyi, V A; *High Energy Chem*. **1995**, *29*, p 2.

31. Yu, D; Rauk, A; Armstrong, D A; *J Chem Soc Perkin Trans 2*. **1995**, p 553.

32. Yu, D; Armstrong, D A; Rauk, A; *Can J Chem*. **1992**, *70*, p 1762.

33. Wolfe, S L; *Molecular and Cellular Biology*, **Wadsworth**: *Belmont Calif*, 1993.

34. Lehninger, A L; Nelson, D L; Cox, M M; *Principles of biochemistry 2nd ed*, **Worth**: *New York*,

1993.
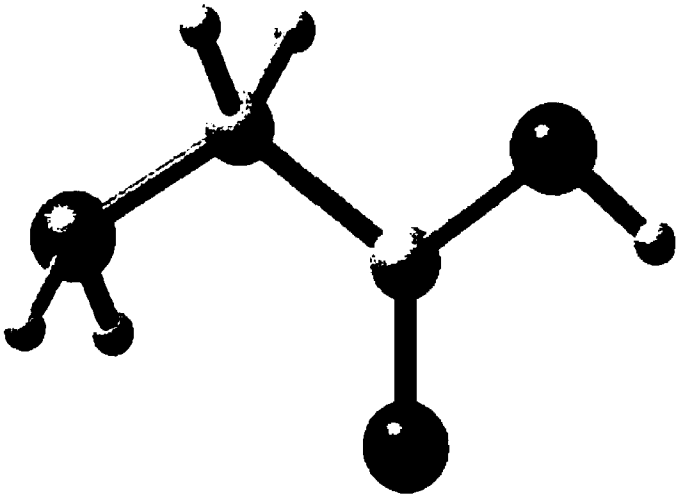
35.   Creeth, J M; Cooper, B; Donald, A S R; Clamp, J R; *Biochem J.***1983**, *211*, p 323.

36.   Rauk, A; Yu, D; Armstrong, D A; *J Am Chem Soc.***1998**, *120*, p 8848.

37.   Griller, D; Simoes, J A M; Wayner, D D M; Chatgilialoglu, C; Asmus, K D; *Sulfur-centered reactive intermediates in chemistry and biology*, **NATO ASI Series Plenum Press**: *New York*, 1990.

38.   Wayner, D D M; Clark, K B; Yu, D; Rauk, A; Armstrong, D A; *J Am Chem Soc.***1997**, *119*, p 8925.

39.   Lias, S G; Bartmess, J E; Liebman, J F; Holmes, J L; Levin, R D; Mallard, W G; *J Phys Chem Ref Data.***1988**, *17*.

40.   Kivirkko, K I; Myllylo, R; *Proc NY Acad Sci.***1985**, *460*, p 187.

41.   Holmes, J L; Lossing, F P; Mayer, P M; *J Am Chem Soc.***1991**, *113*, p 9723.

42.   Goss Sampson, M; Vivian, A J; Kelly, F J; Blake, D; Winyard, P G; *Immunopharmacology of free radical species*, **Academic Press**: *Toronto*, 1995.

43.   Cooper, B; Creeth, J M; Donald, A S R; *Biochem J.***1985**, *228*, p 615.

44.   Block, D A; Yu, D; Armstrong, D A; Rauk, A; *Can J Chem.***1998**, *76*, p 1042.

45.   Rauk, A; *Orbital Interaction Theory of Organic Chemistry*, **Wiley Interscience**: *New York*, 1994.

46.   Katritsky, A R; Zerner, M C; Karelson, M M; *J Am Chem Soc.***1986**, *108*, p 7213.

47.   Beckhaus, H D; Richardt, C; *Angew Chem Int Ed Engl.***1987**, *26*, p 770.

48.   Zhao, R; Merenyi, G; Lind, J; Eriksen, T E; *J Am Chem Soc.***1994**, *116*, p 12010.

49.   Zhao, R; Lind, J; Merenyi, G; Eriksen, T E; *J Chem Soc Perkin Trans 2.***1997**, p 569.

50.   Jorgensen, W L; *BOSS Version 3.8*, **Yale University**, 1997.

51.   Schwarz, H A; Dodson, R W; *J Phys Chem.***1989**, *93*, p 409.
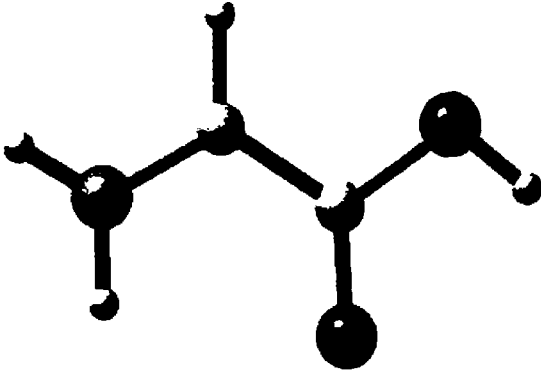
52.   Lide, D R; *CRC Handbook of Chemistry and Physics 76th ed.* **CRC Press:** *Boca Raton FL,* 1998.

53.   Foresman, J B; Keith, T A; Wiberg, K B; Snoonian, J; Frisch, M J; *J Phys Chem.* **1996,** *100,* p 16098.

54.   Cramer, C J; Truhlar, D G; *Rev Comput Chem.* **1995,** *6,* p 1.

55.   Cortis, C; Kim, K; Marten, B; Murphy, R B; Friesner, R A; Honig, B; Sitkoff, D; Ringnalda, M N; *J Phys Chem.* **1996,** *100,* p 11775.

56.   Klamt, A; Schuurmann, G; *J Chem Soc Perkin Trans* 2. **1993,** p 799.

57.   Jorgensen, W L; Chandrasekhar, J; Madura, J D; Impey, R W; Klein, M L; *J Chem Phys.* **1983,** *79,* p 926.

58.   Jorgensen, W L; Tirado Rives, J; *Perspectives Drug Discovery Design.* **1995,** *3,* p 123.

59.   Jorgensen, W L; Blake, J F; Buckner, J K; *Chem Phys.* **1989,** *129,* p 193.

60.   Allen, M P; Tildesley, D J; *Computer Simulations of Liquids,* **Clarendon:** *Oxford,* 1987.

61.   Jorgensen, W L; *J Phys Chem.* **1983,** *87,* p 5304.

62.   Chandrasekhar, J; Spellmeyer, D C; Jorgensen, W L; *J Am Chem Soc.* **1984,** *106,* p 903.

63.   McQuarrie, D A; *Statistical Thermodynamics,* **Harper & Row:** *New York,* 1973.

63.   A Silicon Graphics Power Challenge supercomputer and a group of Compaq Alpha workstations were the main computational workhorses for this project.

65.   Burkert, U; Allinger, L; *Molecular Mechanics,* **American Chemical Society:** *Washington DC,* 1982.

66.   Breneman, C M; Wiberg, K B; *J Comput Chem.* **1990,** *11,* p 361.

67.   Lim, D; Jorgensen, W L; *J Phys Chem.* **1996,** *100,* p 17490.

68.   Zwanzig, R W; *J Chem Phys.* **1954,** 22, p 1420.
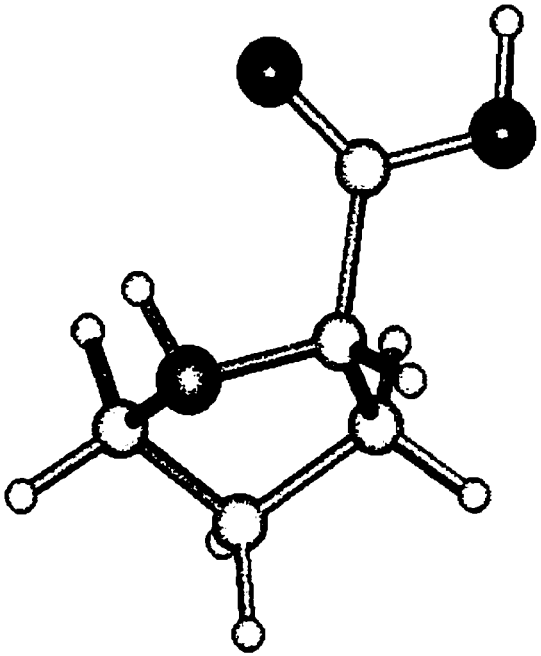
69.   Jorgensen, W L; Ravimohan, C; *J Chem Phys.* **1985**, *83*, p 3050.

70.   Curtiss, L A; Raghavachari, K; Pople, J A; *J Chem Phys.* **1993**, *98*, p 1293.

71.   Scott, A P; Radom, L; *J Phys Chem.* **1996**, *100*, p 16502.

72.   Kaminski, G; Duffy, E M; Matsui, T; Jorgensen, W L; *J Phys Chem.* **1994**, *98*, p 13077.

73.   Wagman, D D; Evans, W H; Parker, V B; Schumm, R H; Halow, I; Bailey, S M; Churney, K L; Nuttal, R L; *J Phys Chem Ref Data.* **1982**, *11*.

74.   Dobe, S; Berces, T; Turanyi, T; Marta, F; Grussdorf, J; Temps, F; Wagner, H Gg; *J Phys Chem.* **1996**, *100*, p 19864.

75.   Johnson, R D III; Hudgens, J W; *J Phys Chem.* **1996**, *100*, p 19874.

76.   Atkinson, R; Baulch, D L; Cox, R A; Hampson, R F Jr; Kerr, J A; Rossi, M J; Troe, J; *J Phys Chem Ref Data.* **1997**, *26*, p 1329.

77.   Cohen, N; *J Phys Chem Ref Data.* **1995**, *25*, p 1411.

78.   Tsang, W; *J Phys Chem Ref Data.* **1987**, *16*, p 471.

79.   McMillan, D F; Golden, D M; *Annu Rev Phys Chem.* **1982**, *33*, p 493.

80.   Afeefy, H Y; Liebman, J F; Stein, S E; *http://webbooknistgov/*, **NIST Webbook.**

81.   It should be noted that the experimental data as well as the calculated gas phase thermodynamic data are for equilibrium mixtures of conformers while the BOSS calculation pertains only to a single conformer. Direct comparison is valid to the extent that there is no preferential solvation of any of the conformers, i.e., that the population mix does not change significantly from gas phase to solution.

82.   Cabani, S; Gianni, P; Mollica, V; Lepori, L; *J Sol Chem.* **1981**, *10*, p 563.

83.   Rizzo, R C; Jorgensen, W L; *J Am Chem Soc.* **1999**, *121*, p 4827.

84.   Rick, S W; Stuart, S J; Berne, B J; *J Chem Phys.* **1994**, *101*, p 6141.

85.   Van Belle, D; Wodak, S J; *Comp Phys Comm.* **1995**, *91*, p 253.

65

86. Caldwell, J W; Kollman, P; *J Am Chem Soc.* **1995**, *117*, p 4177.

87. Jorgensen, W L; Severance, D L; *J Chem Phys.* **1993**, *99*, p 4233.

88. Jorgensen, W L; McDonald, N A; Selmi, M; Rablen, P R; *J Am Chem Soc.* **1995**, *117*, p 11809.

89. Gao, J; *J Am Chem Soc.* **1995**, *117*, p 8600.

90. Griller, D; Lossing, F P; *J Am Chem Soc.* **1981**, *103*, p 1586. Also Burkey, T J; Castelhano, A L; Griller, D; Lossing, F P; *J Am Chem Soc.* **1983**, *105*, p 4701.

91. Romano, S; Clementi, E; *Int J Quantum Chem.* **1978**, *14*, p 839.

92. Mezei, M; Mehrotra, P K; Beveridge, D L; *J Biomolecular Struct Dyn.* **1984**, *2*, p 1.

93. Alagona, G; Ghio, C; Kollman, P; *J Mol Struct.* **1988**, *166*, p 385.

94. Jorgensen, W L; Madura, J D; Swenson, C J;; *J Am Chem Soc.* **1984**, *106*, p 6638.
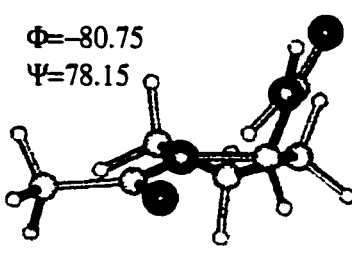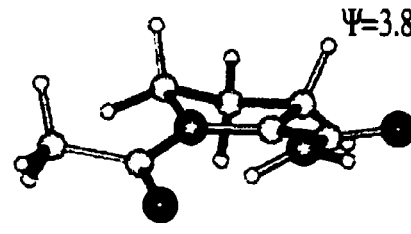
Appendix A: Molecular Bestiary
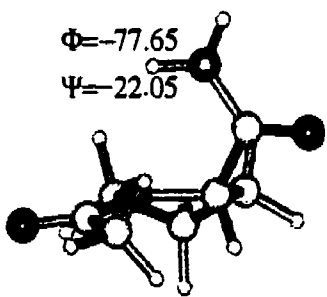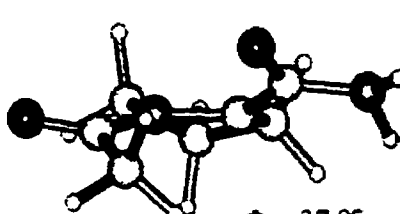
Molecule label	Glycine 1a
Structure 1	
Related Species	2-Aminoethanal, All Glycine conformers, esp. 1c and 1f
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from 2-Aminoethanal.
ΔG_{soln}	-64.7 kJ mol^{-1} (this calc), -53 kJ mol^{-1} (previous estimate)
Comments	Most stable form of glycine in the gas phase. In solution, the most stable *non-zwitterionic* form (zwitterion dominates, however). Cs symmetry.
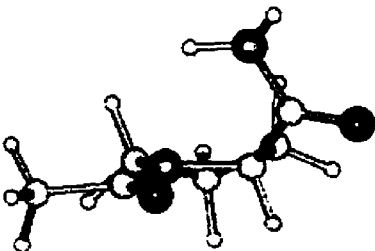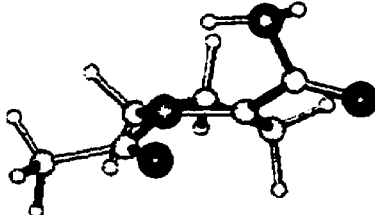
Radical	Glycine 1a radical
Structure 2	
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Glycine 1a.
Comments	Planar structure. This is the most stable radical structure in the gas phase; more surprisingly, it is the most stable radical structure in solution, being much more stable internally than the zwitterion radical, which makes up for the zwitterion radical's greater solvation energy .
BDE	331 kJ mol^{-1} (gas phase calc, partner for isodesmic reactions), 334.6 kJ mol^{-1} (aqueous solution calc)
ΔG_{soln}	-60.7 kJ mol^{-1} (calc)
$\Delta\Delta G_{soln}$	-4.0 kJ mol^{-1} (calc)
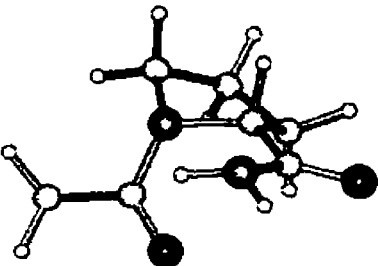
Molecule label	Proline
Structure 3	
Related Species	Proline peptide models, Glycine
Comments	Most stable conformation of amino acid proline. All proline calculations performed at B3LYP/6-31G* level only.

Molecule label	Sample β Turn (Type II *gly-pro-gly-gly*)
Structure 4	
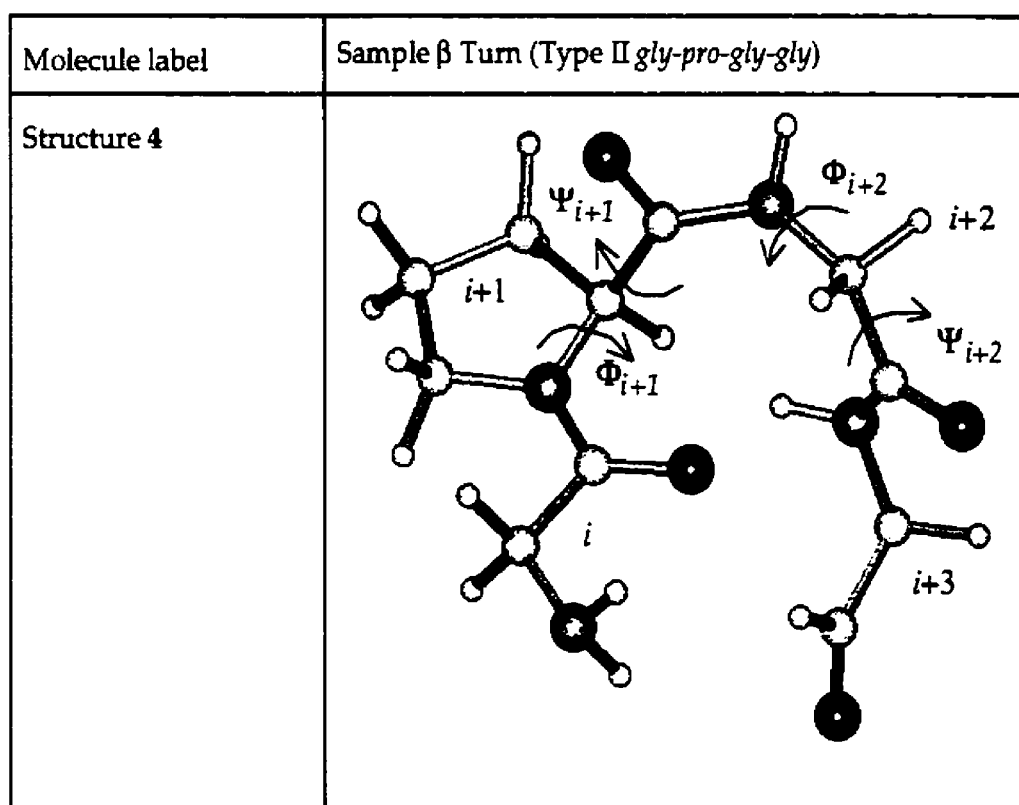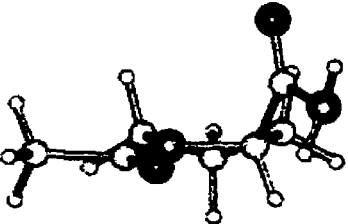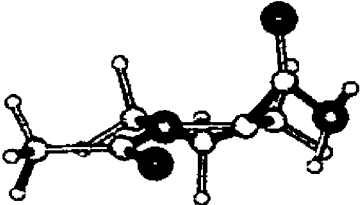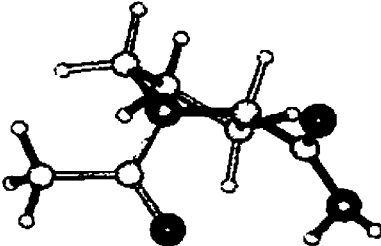
Molecule label	Proline Model 2 (N-acetyl) *trans*
Structure 5	$\Phi = -80.75$ $\Psi = 78.15$
Related Species	Proline, Proline model 1, Proline model 2 *cis*
Comments	Unconstrained N-acetyl model shows the proper *trans* form of the peptide bond, with the methyl group where the next α-carbon would be.
Radical	Proline model 2 *trans* radical
Structure 6	$\Phi = -29.15$ $\Psi = 3.85$
Comments	The almost planar structure is not easily formed in a protein environment. The unpaired electron freely delocalizes across the amide bond on the C-terminus, but the N-terminus is not included.
BDE	368.6 kJ mol^{-1}
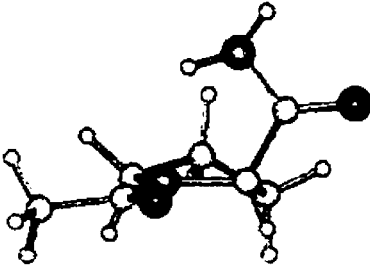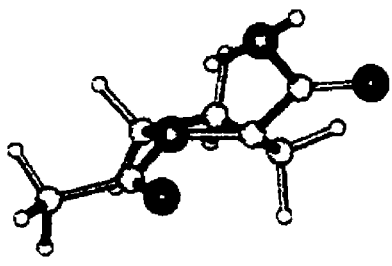
Molecule label	Proline Model 2 *cis* form
Structure 7	$\Phi=-77.65$ $\Psi=-22.05$
Related Species	Proline, Proline model 1, Proline model 2 *trans*
Comments	The *cis* form causes steric interference between the methyl group on the acetyl and the amide group, pushing its energy up to the point that only roughly 10% of proline residues in proteins adopt this conformation.
Radical	Proline model 2 *cis* radical
Structure 8	$\Phi=-37.95$ $\Psi=178.25$
Comments	Lower BDE due to higher energy parent, delocalization on radical
BDE	357.7 kJ mol^{-1}
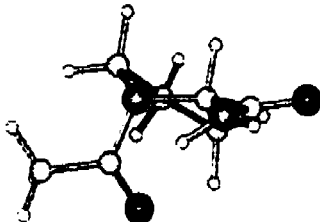
Molecule label	Beta Turn Type I: $\Phi=-60°$, $\Psi=-30°$
Structure 9	
Comments	*trans* form, on residue *i+1*
BDE	380.7
Radical	Type I radical
Structure 10	
Comments	Almost planar structure, but unpaired electron is not able to delocalize across both amide moieties.
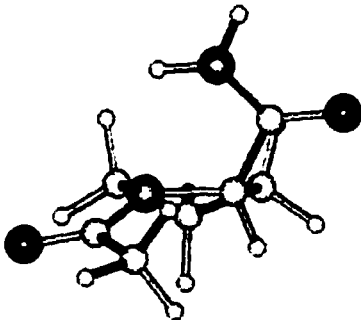
Molecule label	Beta Turn Type I **R** isomer: Φ=-60°, Ψ=-30°
Structure 11	
Comments	R form not very stable with Φ and Ψ constraints, therefore BDE low, allowing easy return to radical, and more likely adoption of S form.
BDE	336.2
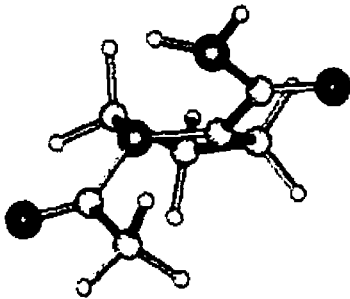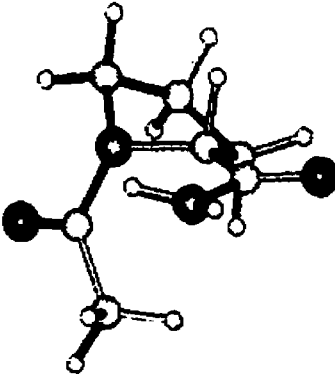
Molecule label	Beta Turn Type II: Φ=-60°, Ψ=120°
Structure 12	
Comments	*trans* form, on residue *i+1*
BDE	397.8
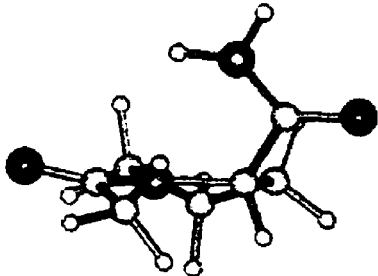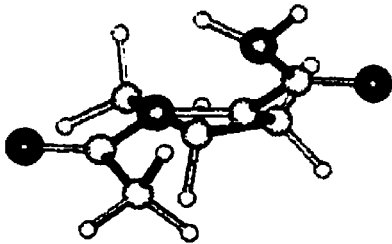
Radical	Type II radical
Structure 13	
Comments	Not planar at the radical centre, so very little captodative stabilization.
Molecule label	Beta Turn Type II R isomer: Φ=-60°, Ψ=120°
Structure 14	
Comments	High energy conformer encourages the S conformer to be re-formed.
BDE	338.5
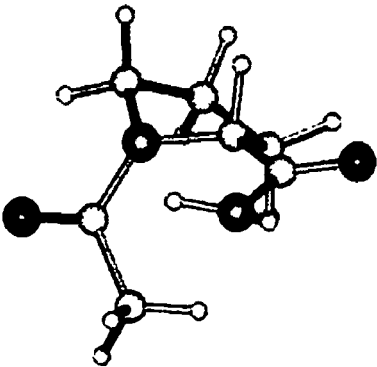
Molecule label	Beta Turn Type II': $\Phi=-80°$, $\Psi=0°$
Structure **15**	
Comments	*trans* form, on residue $i+2$
BDE	385.4
Radical	Type II' radical
Structure **16**	
Comments	Planar amide, but the radical centre is again excluded from linking to the ring nitrogen.
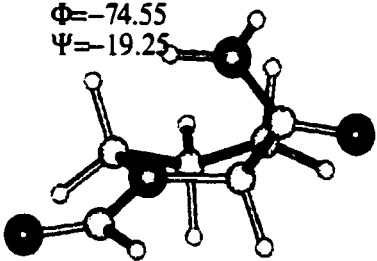
Molecule label	Beta Turn Type II' **R** isomer: Φ=-80°, Ψ=0°
Structure **17**	
Comments	
BDE	308.4
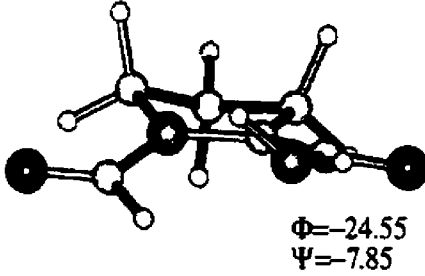
Molecule label	Beta Turn Type VIa: Φ=-90°, Ψ=0°
Structure **18**	
Comments	*cis* form, on residue *i+2*
BDE	374.0
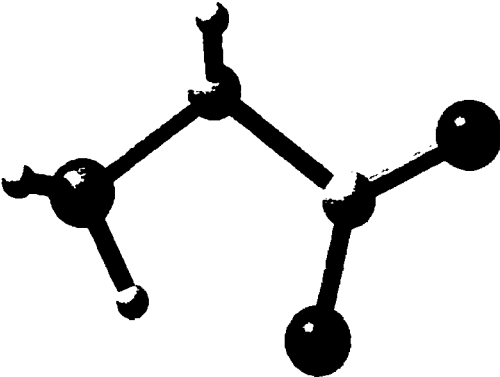
Radical	Type VIa radical
Structure 19	
Comments	Here the bend is in the nitrogen, so the radical centre is in a more planar environment, lowering the BDE.
Molecule label	Beta Turn Type VIa **R** isomer: $\Phi=-90°$, $\Psi=0°$
Structure 20	
Comments	
BDE	282.7
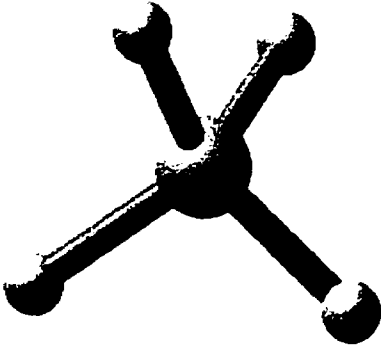
Molecule label	Beta Turn Type VIb: Φ=-60°, Ψ=0°
Structure 21	
Comments	*cis* form, on residue *i+2*
BDE	355.0
Radical	Type VIb radical
Structure 22	
Comments	The unpaired electron is able to delocalize into the amide bond in this conformation, resulting in the lowest BDE of any of the models studied (excluding **R** isomers).
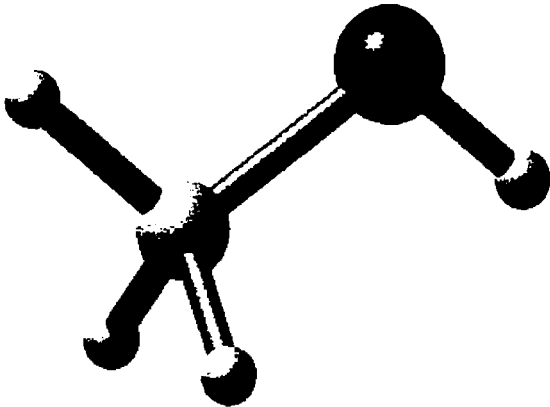
Molecule label	Beta Turn Type VIb **R** isomer: $\Phi=-60°$, $\Psi=0°$
Structure 23	
Comments	
BDE	288.3
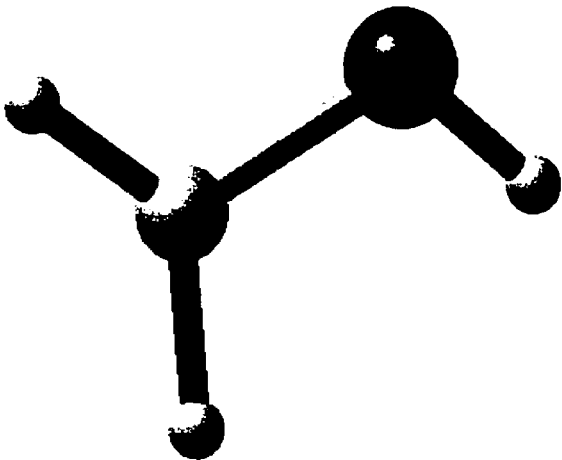
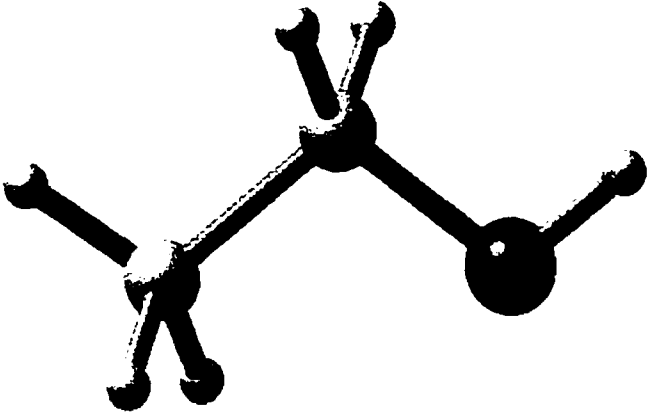| Molecule label	Proline Model 1 (N-formyl) *cis* form
Structure 24	$\Phi=-74.55$ $\Psi=-19.25$
Related Species	Proline, Glycine model peptide, Proline model 2
Comments	N-formyl model allows planar structure too easily, and does not model steric constraints on *cis* form. Model 2 therefore used (N-acetyl). Model 1 formed by converting carboxylic acid to amide, and adding a formyl group to the nitrogen on the ring.

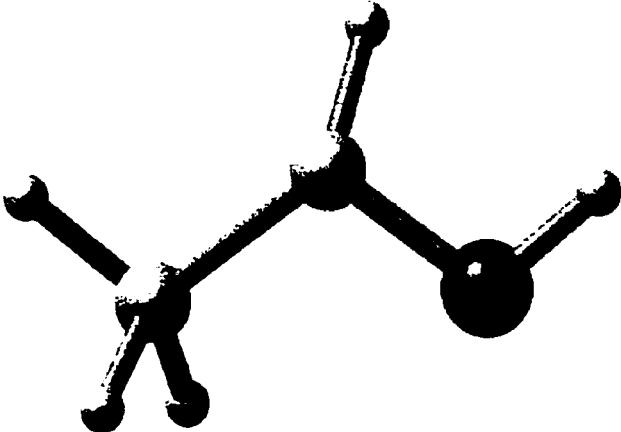| Radical	Proline model 1 radical
Structure **25**	$\Phi=-24.55$ $\Psi=-7.85$
Comments	Parent structure too stable compared to N-acetyl, so BDE unnaturally high.
BDE	393.2

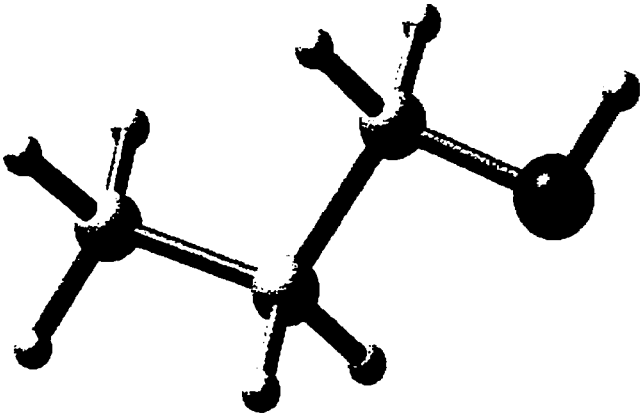| Molecule label	Glycine Zwitterion
Structure 26	
Related Species	All Glycine conformers, esp. 1c
Calculations	B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Glycine 1c
ΔG_{soln}	-198.9 kJ mol-1 (this calc)
Comments	Not a stable structure in the gas phase. Structure optimized only with continuum model. There is a considerable energy penalty in taking this structure into the gas phase, even without alteration (133 kJ mol-1). In solution, this is effectively the only glycine species, due to the enormous charge separation and the resulting excellent solvation.

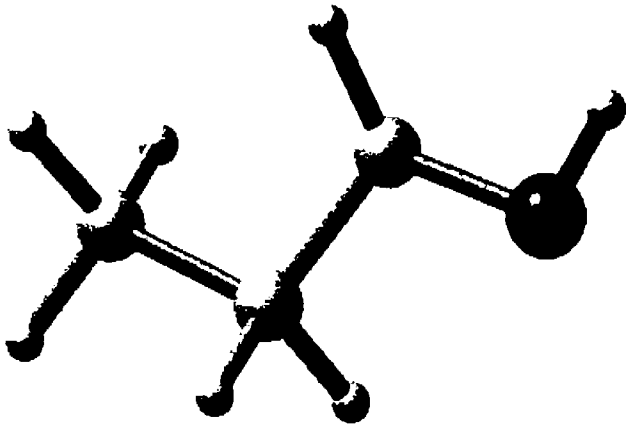| Molecule label	Methane
Structure 27	
Related Species	Methanol, Nothing
Calculations	B3LYP/6-31G* in gas phase and with SCRF=SCIPCM. ΔG^{SCRF}=0.4 kJ mol^{-1}. BOSS mutation from NOTHING.
ΔG_{soln}	8.0 ± 1.2 kJ mol^{-1} (calc), 8.4 kJ mol^{-1} (expt)
Comments	ΔG^{BOSS}=8.4 kJ mol^{-1} (same as expt). Electrostatic decoupling used to mutate to nothing, but that mutation still the source of most of the statistical error in further BOSS calculations.

83

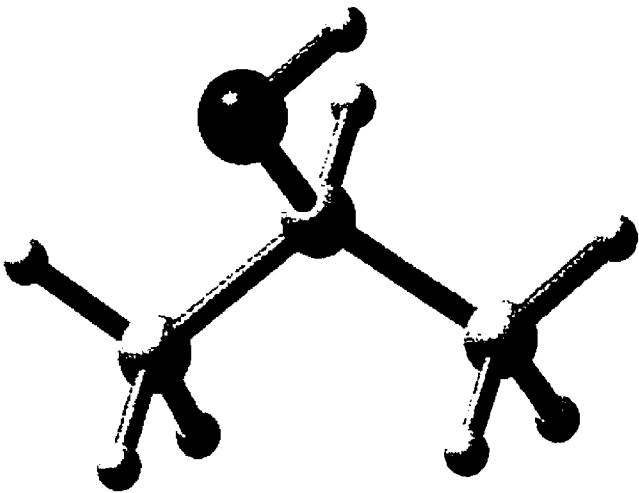| Molecule label	Methanol
Structure 28	
Related Species	Methane, Methanol radical, Methylamine
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM B3LYP/6-311+G(3df, 2p) B3LYP/6-311+G(3df, 2p) SCRF=SCIPCM BOSS mutation from Methane.
ΔG_{soln}	−18.3 kJ mol^{-1} (calc), -21.3 kJ mol^{-1} (expt)
Comments	Used with radical as test case for basis set calculations- found that SCRF step was vital for CHELPG charges that give accurate solvation energies.

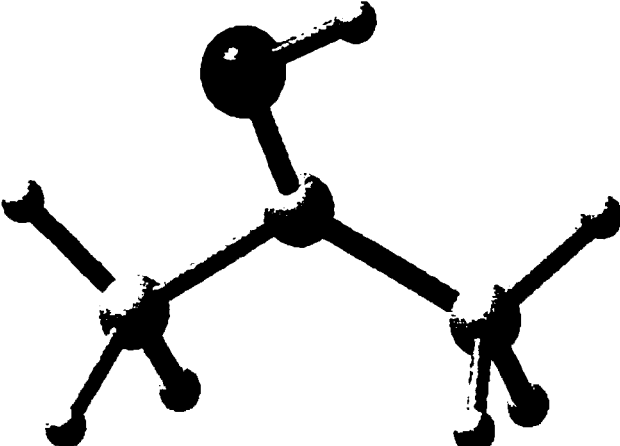| Radical	Methanol Radical
Structure **29**	
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM B3LYP/6-311+G(3df, 2p) single point at previous geometry B3LYP/6-311+G(3df, 2p) SCRF=SCIPCM single point BOSS mutation from Methanol.
Comments	Tested for OPLS dependence- redid solvation calculation with sp^3 carbon centre. 0.04 kJ mol^{-1} difference (negligible) from sp^2 carbon parameters. 2 conformers- symmetric because of radical double well
BDE: Experimental	401.9
ΔG_{soln}	-16.2 kJ mol^{-1} (calc),-17.3 kJ mol^{-1} (expt)
$\Delta\Delta G_{soln}$	2.2 kJ mol^{-1} (calc), 4.0 kJ mol^{-1} (expt)

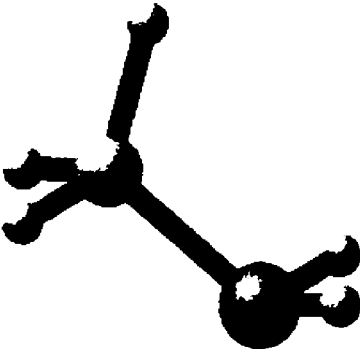| Molecule label	Ethanol
Structure 30	
Related Species	Methanol, Ethanol radical, Ethylamine
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Methanol.
ΔG_{soln}	-17.4 kJ mol^{-1} (calc), -21.1 kJ mol^{-1} (expt)
Comments	Ethanol produced the worst agreement with experiment of the alcohols. 3 conformers present

Radical	Ethanol Radical
Structure 31	
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Ethanol.
Comments	Poor agreement with experiment for $\Delta\Delta G_{soln}$ comes from poor ethanol parent ΔG_{soln} value 4 conformers present
BDE	396.3 kJ mol^{-1}
ΔG_{soln}	-12.3 kJ mol^{-1} (calc), -11.8 kJ mol^{-1} (expt)
$\Delta\Delta G_{soln}$	5.2 kJ mol-1 (calc), 9.3 kJ mol-1 (expt)

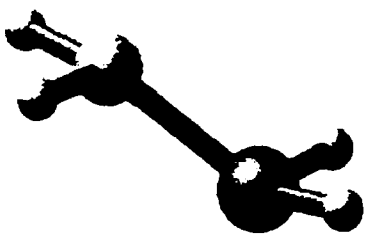| Molecule label	1-Propanol
Structure 32	
Related Species	Ethanol, 1-Propanol radical, 2-Propanol
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Ethanol.
ΔG_{soln}	-20.0 kJ mol-1 (calc), -20.1 kJ mol-1 (expt)
Comments	9 conformers present
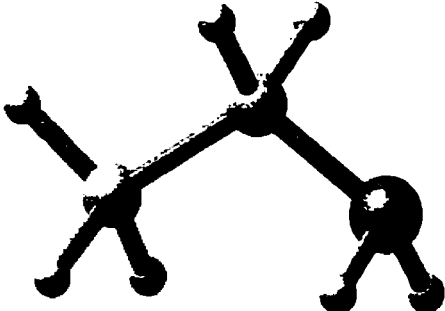
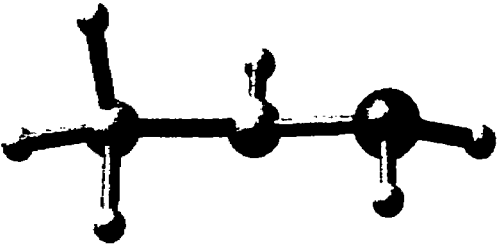| Radical	1-Propanol Radical
Structure 33	
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from 1-Propanol.
Comments	no experimental ΔG_{soln} value for radical 12 conformers present
BDE	397.3 kJ mol-1
ΔG_{soln}	-15.4 kJ mol-1 (calc)
$\Delta\Delta G_{soln}$	4.7 kJ mol-1 (calc)

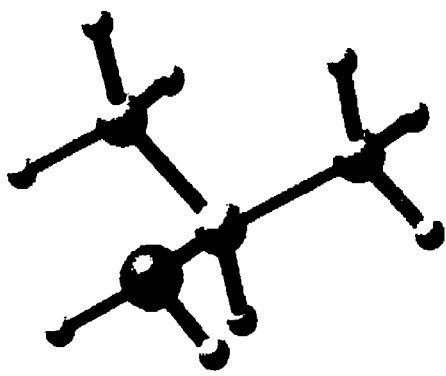| Molecule label	2-Propanol
Structure **34**	
Related Species	Ethanol, 1-Propanol, 2-Propanol radical, 2-Propylamine
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Ethanol.
ΔG_{soln}	-19.8 kJ mol-1 (calc), -20.1 kJ mol-1 (expt)
Comments	3 conformers present

Radical	2-Propanol Radical
Structure 35	
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from 2-Propanol.
Comments	2 conformers present
BDE	393.2
ΔG_{soln}	-13.3 kJ mol-1 (calc), -12.3 kJ mol-1 (expt)
$\Delta\Delta G_{soln}$	6.5 kJ mol-1 (calc), 7.8 kJ mol-1 (expt)
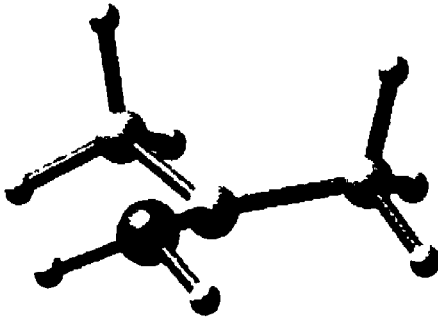
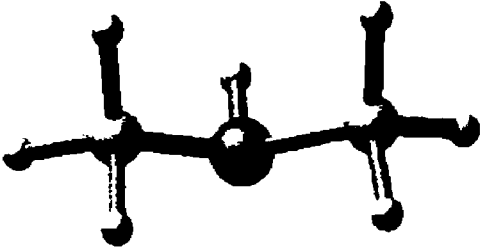| Molecule label	Methylamine
Structure **36**	
Related Species	Methanol, Methylamine radical, Ethylamine
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. Comparison with B3LYP/6-31G* SCRF=SCIPCM optimization gives negligible difference in geometry or energy. BOSS mutation from Methanol.
ΔG_{soln}	-20.6 kJ mol-1 (calc), -19.1 kJ mol-1 (expt)
Comments	Assumed for amines (based on alcohol results) that geometry modification in solution is negligible, but CHELPG needs SCRF wavefunction

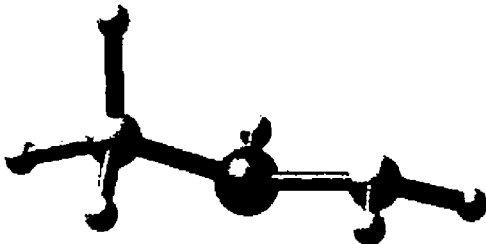| Radical	Methylamine Radical
Structure 37	
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. BOSS mutation from Methylamine.
Comments	2 conformers present
BDE	388 kJ mol^{-1} (prev. calc- 25 °C, gas phase), 398 kJ mol^{-1} (present calc- 25 °C, in aqueous solution)
ΔG_{soln}	-10.8 kJ mol^{-1} (calc)
$\Delta\Delta G_{soln}$	9.8 kJ mol^{-1} (calc)

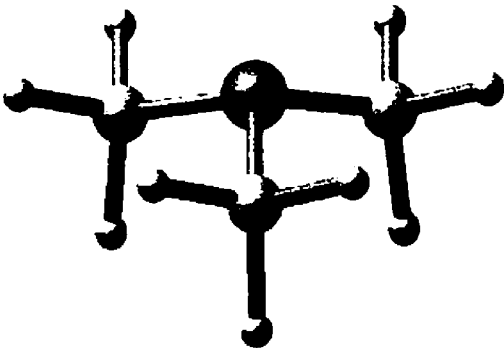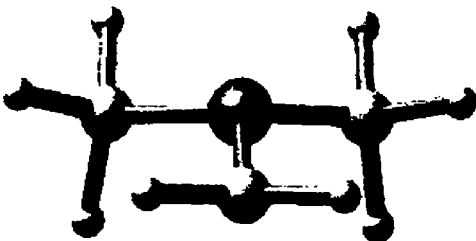| Molecule label	Ethylamine
Structure 38	
Related Species	Methylamine, Ethylamine radical, 2-Propylamine, 2-aminoethanal, Ethanol, Dimethylamine
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. Comparison with B3LYP/6-31G* SCRF=SCIPCM optimization gives negligible difference in geometry or energy. BOSS mutation from Methylamine.
ΔG_{soln}	-16.5 kJ mol^{-1} (calc), -18.8 kJ mol^{-1} (expt)
Comments	Substitution at the α-carbon position does not affect the accuracy of our calculations on the parent- not so on the nitrogen (see dimethylamine)

Radical	Ethylamine Radical
Structure 39	
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. BOSS mutation from Ethylamine.
Comments	Loss of charge density on the nitrogen leads to less negative solvation energy.
BDE	384 kJ mol^{-1} (prev. calc), 394 kJ mol^{-1} (present calc)
ΔG_{soln}	-6.4 kJ mol^{-1} (calc)
$\Delta\Delta G_{soln}$	10.1 kJ mol^{-1} (calc)

95

Molecule Label	2-Propylamine
Structure 40	
Related Species	Ethylamine, 2-Propylamine radical, 1-Propylamine, Trimethylamine, 2-Propanol
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. BOSS mutation from Ethylamine.
ΔG_{soln}	-20.8 kJ mol^{-1}
Comments	Addition of methyl group does not affect solvation energy as much as might be expected- solvation apparently dominated by amine group

Radical	2-Propylamine radical
Structure 41	
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. BOSS mutation from 2-Propylamine.
Comments	In radical form, amine has much lower charge density, and is unable to counter the hydrophobic propyl group, hence large positive $\Delta\Delta G_{soln}$
BDE	388 kJ mol^{-1} (prev. calc), 400 kJ mol^{-1} (present calc)
ΔG_{soln}	-8.8 kJ mol^{-1} (calc)
$\Delta\Delta G_{soln}$	12.1 kJ mol^{-1} (calc)

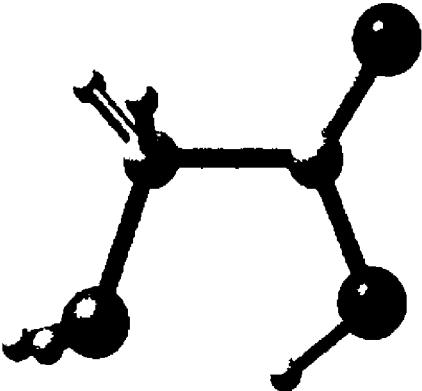| Molecule Label	Dimethylamine
Structure 42	
Related Species	Methylamine, Dimethylamine radical, Trimethylamine, Ethylamine
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. BOSS mutation from Methylamine.
ΔG_{soln}	-6.9 kJ mol^{-1} (calc), -17.9 kJ mol^{-1} (expt)
Comments	Molecular Mechanics and CHELPG both assume solvation is dominated by electrostatics. The CHELPG response to adding a methyl group is to lower the charge on the nitrogen, thus decreasing the solvation of the molecule. However, in dimethylamine, the solvation energy is dominated by the hydrogen-bonding ability of the nitrogen, which arises from its basicity, not its excess charge. The basicity of the nitrogen is increased by the electron-donating methyl group, allowing the strengthened hydrogen bond to water to overcome the extra methyl group. This effect is not allowed for in any standard solvation code.[55]

Radical	Dimethylamine radical
Structure 43	
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. BOSS mutation from Dimethylamine.
Comments	Dimethylamine radical may be intermediate between the primary amines and Trimethylamine, as may be expected. The radical has less solvation, probably because of even lower charge density at the nitrogen, but basicity concerns (see parent) may make the whole calculation moot.
BDE	386 kJ mol^{-1} (prev. calc), 385 kJ mol^{-1} (present calc)
ΔG_{soln}	-3.0 kJ mol^{-1} (calc)
$\Delta\Delta G_{soln}$	3.9 kJ mol^{-1} (calc)
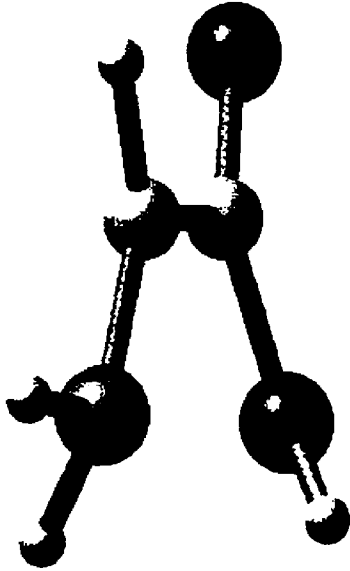
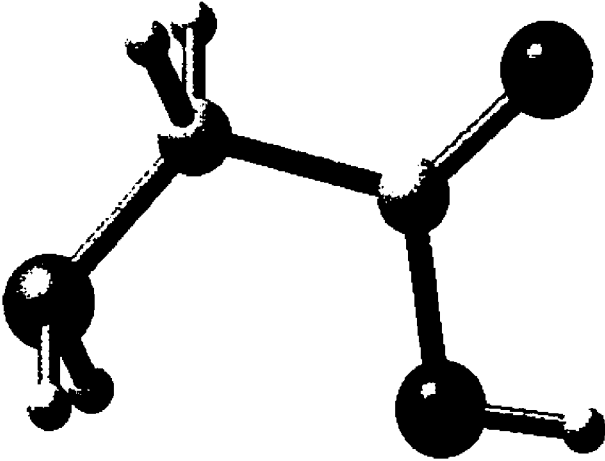| Molecule Label	Trimethylamine
Structure **44**	
Related Species	Dimethylamine, Trimethylamine radical, 2-Propylamine
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction. BOSS mutation from Dimethylamine.
ΔG_{soln}	+5.9 kJ mol^{-1} (calc), -13.6 kJ mol^{-1} (expt)
Comments	See discussion of Dimethylamine.
Radical	Trimethylamine radical
Structure **45**	
Calculations	B3LYP/6-31G* in gas phase Single point SCRF=SCIPCM calculation to get wavefunction.

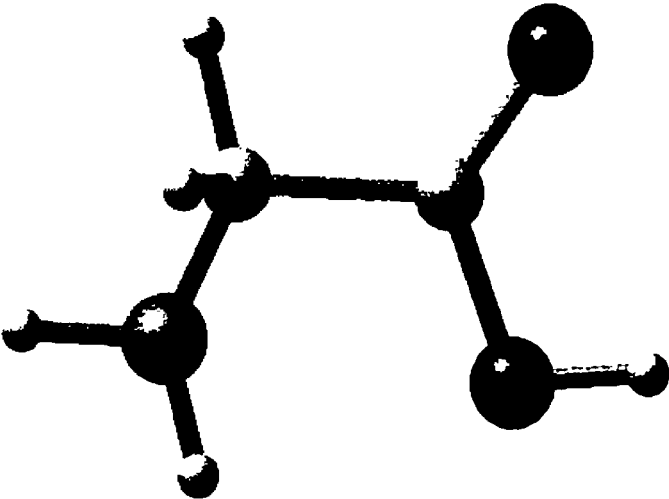|	BOSS mutation from Trimethylamine.
Comments	Radical solvation prediction is unsure because of inaccuracy of parent. With no evidence to the contrary, it appears that this radical is just as well solvated as its parent. However, if the formation of the radical makes the amine less basic, this assumption may be faulty.
BDE	387 kJ mol^{-1} (prev. calc), 387 kJ mol^{-1} (present calc)
ΔG_{soln}	+5.5 kJ mol^{-1} (calc)
$\Delta\Delta G_{soln}$	-0.4 kJ mol^{-1} (calc)
Molecule label	2-Aminoethanal
Structure 46	
Related Species	Ethylamine, Glycine
Calculations	B3LYP/6-31G* B3LYP/6-31G* SCRF=SCIPCM

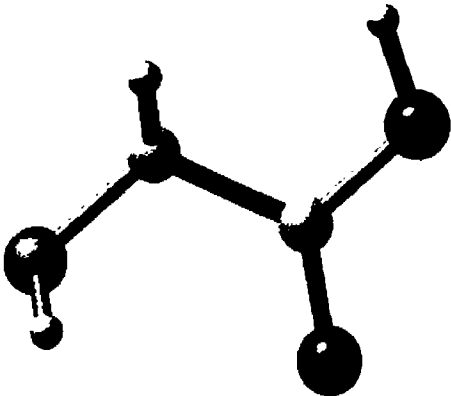|	BOSS mutation from Ethylamine.
ΔG_{soln}	-32.7 kJ mol^{-1} (calc)
Comments	Calculated as a transitional molecule between ethylamine and glycine. Introduced the carbonyl, mutated from a methyl group.

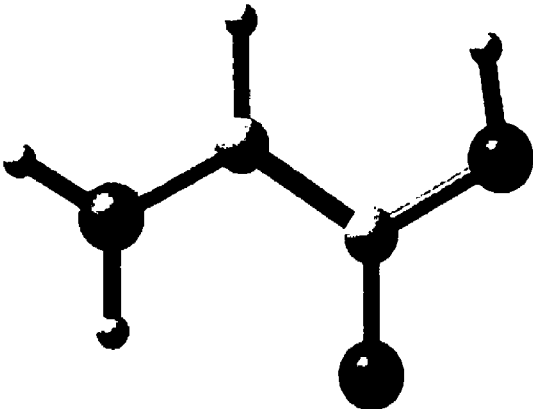| Molecule Label	Glycine 1b
Structure 47	
Related Species	All Glycine conformers
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM
$\Delta G_{(g)}$ relative to conformer 1a	5.6 kJ mol^{-1}
Comments	Minor contributor to glycine population in the gas phase- insignificant in solution due to domination by zwitterion.

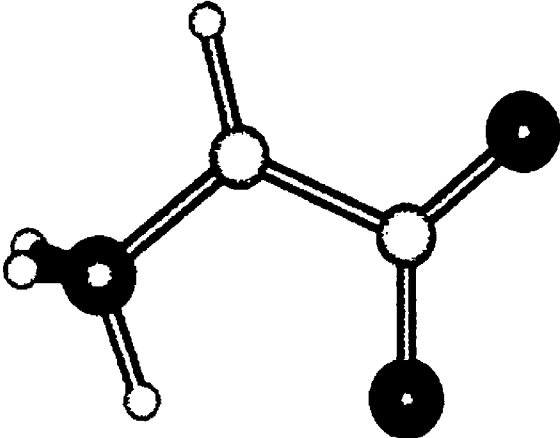| Molecule label	Glycine 1c
Structure **48**	
Related Species	All Glycine conformers, esp. 1a and Zwitterion
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Glycine 1a.
$\Delta_f G_{(g)}$ relative to conformer 1a	5.6 kJ mol^{-1}
ΔG_{soln}	-60.8 kJ mol^{-1} (this calc)
Comments	Second most stable conformer in the gas phase. Slightly higher in internal energy in SCIPCM optimized form, which raises the ΔG_{soln} higher than 1a.

Radical	Glycine 1c Radical
Structure **49**	
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM
Comments	Non-planar structure, which reduces the ability of the π orbitals to overlap, lowering the captodative stabilization energy. This radical is very unstable in the gas phase, and even more so in solution, so no BOSS mutation was done.
BDE	377.6 kJ mol^{-1} (gas phase calc)

Molecule label	Glycine 1d
Structure **50**	
Related Species	All Glycine conformers
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM
$\Delta_r G_{(g)}$ relative to conformer 1a	6.1 kJ mol^{-1}
Comments	Minor contributor to glycine population in the gas phase- insignificant in solution due to domination by zwitterion.

Molecule label	Glycine 1e
Structure **51**	
Related Species	All Glycine conformers
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM
$\Delta_f G_{(g)}$ relative to conformer 1a	11.4 kJ mol^{-1}
Comments	Negligible contributor to glycine population in the gas phase- insignificant in solution due to domination by zwitterion.

Molecule label	Glycine 1f
Structure **52**	
Related Species	All Glycine conformers, esp. 1a
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Glycine 1a.
$\Delta_f G_{(g)}$ relative to conformer 1a	19.5 kJ mol^{-1}
ΔG_{soln}	-72.0 kJ mol^{-1} (this calc)
Comments	Quite unstable in gas phase, but *trans* conformation almost makes up for it in solution. Still higher than 1a in solution, however.

Radical	Glycine 1f radical
Structure 53	
Calculations	B3LYP/6-311+G(3df,2p) // B3LYP/6-31G* B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Glycine 1f.
Comments	Higher internal energy of the radical in solution relative to the gas phase combined with less negative ΔG_{soln} leads to negligible population of this conformation of radical in solution. Gas phase BDE very similar to 1a. Radical appears to neutralize advantage of *trans* conformation.
BDE	332.8 kJ mol^{-1} (gas phase calc), 354.2 kJ mol^{-1} (aqueous solution calc)
ΔG_{soln}	-50.6 kJ mol^{-1} (calc)
$\Delta\Delta G_{soln}$	21.4 kJ mol^{-1} (calc)

Radical	Glycine Zwitterion Radical
Structure **54**	
Calculations	B3LYP/6-311+G(3df,2p) SCRF=SCIPCM // B3LYP/6-31G* SCRF=SCIPCM BOSS mutation from Glycine Zwitterion.
Comments	In contrast to the zwitterion, the zwitterion radical has enormous internal energetic disadvantages, due to the loss of captodative stabilization from the amine. This lost energy is not recovered enough by the admittedly strong solvation energy, so that the zwitterion radical is not a significantly populated conformer in solution.
BDE	407.2 kJ mol-1 (aqueous solution calc)
ΔG_{soln}	-225.5 kJ mol-1 (calc)
$\Delta\Delta G_{soln}$	-26.7 kJ mol-1 (calc)