The QCFF/SOL Manual

Contents

1   Introduction

The QCFF/SOL program is a molecular modeling package for determining the properties of medium sized molecules and biological cofactors in solution and in protein microenvironments. The QCFF/SOL program is based on the heteroatom version of the widely used QCFF/PI with an extremely efficient incoporation of microscopic solvent effects and a direct interface with the electrostatic potential generated by POLARIS or ENZYMIX. This includes the ability to evaluate the solvent effect on quantum mechanical calculations of ground and excited states.

The QCFF/SOL program has been used extensively in studies of spectroscopic and conformational properties of molecules in solution and in proteins with an emphasis on reliability and efficiency. The program is particularily effective in exploring the actual effect of the proteins microenvironment on bound cofactors including such crucial effects as changes in redox potential, photochemical activity and color control. Also, this program allows the user to study properties of aromatic drugs where quantum mechanical delocalization effects might be significant, to explore solvent effects on electronic and vibrational calculations and to perform molecular mechanics calculations on solvated molecules.

Features

- Quantum mechanical calculations of solvated conjugated molecules.

- Energy minimization and normal mode analysis of solvated molecules.

- Electronic spectra (UV and CD) of molecules in solution and in proteins.

- Resonance Raman of biological cofactors in protein sites.

- Charge transfer states in solution and molecular crystals.

- Photochemical and redox properties of biological factors.

- Conducting polymers.

2   Some Background

2.1   The Philosophy of the QCFF Approach

The study of the properties of small molecules by ab-initio quantum mechanical methods has advanced significantly in recent years[1]. Yet such methods are still far from being able to provide a practical way for modeling medium and large molecules. Semiempirical all-valencece electron methods (e.g. INDO[2a], MINDO[2b], AM1[2c], etc.) can be used to study medium size molecules. However, although very promising and quite useful in many cases, these methods still require very extensive computer time for convergent minimization and normal mode analysis of medium size molecules and their reliability in treating various properties of large molecules (e.g. torsional barriers of conjugated molecules) still need improvement. On the other hand it has become clear in recent years that many properties of large molecules can be simulated effectively and reliably by the so called "Classical Force Fields" methods[3,4]. Such methods exploit the fact that large molecules are built from small fragments whose properties can be transferred from small molecules. In this way one can study the dependance of the bond energy of the bond length in small molecules (by either ab-initio calculations or by analysing the relevant experimental information) and then use the same functional dependance for the same type of bonds in large molecules. Such an approach allows one to study large molecules without facing the problem of treating them quantum mechanically. In fact, force field methods are expected to be more reliable than semiempirical all-valence electron methods as long as one deals with localized electrons and with regions of the potential surface which do not involve large deformations of bond lengths and bond angles. A systematic approach of obtaining reliable force fields is the Consistent Force Field (CFF) method[3] which is the basis of many of the the current popular molecular modeling programs (e.g. AMBER[5a] or CHARM[5b]). This method represents molecular potential surfaces by analytical empirical potential functions in the form which includes the bond lengths, b_i, bond angles, θ_i, torsional angles, ϕ_i, and non-bonded distance r_{ij} (these coordinates are treated as analytical functions of the molecular

cartesian coordinate vector **r**).

$$V(\boldsymbol{S}(\boldsymbol{r})) \;=\; \frac{1}{2}\sum_i K_b(b_i - b_o)^2 + \frac{1}{2}\sum K_\theta(\theta_i - \theta_0)^2 \qquad (1)$$

$$+\frac{1}{2}\sum_i K_\phi(1 + \cos(n\phi_i + \delta_i)) + \sum_{i>j} f(r_{ij}) + \sum_{i>j} q_i q_j / r_{ij}$$

where **S** is the vector representing the internal coordinates of the given molecule The non-boned potential function $f(r_{ij})$ is either a 12-6 or an exp-6 function and the q's represent the residual atomic changes. The parameters in these potential functions (e.g. K_b, K_θ, K_ϕ, etc.) are determined by a systematic fitting of various calculated and observed properties of the given class of molecules (see ref. 3) Unfortunately, the CFF approach is restricted to molecules with localized electrons (and transferable bonding properties) and cannot be used in studies of molecules with delocalized electrons. However, even in such cases there is no need for a full quantum mechanical treatement of the entire molecule. One can clearly separate the molecular system to fragments with localized and delocalized electrons, respectively. While the delocalized electrons should be treated quantum mechanically, the molecular fragments with localized electrons could be treated by force field approaches. This is the basic philosophy of the QCFF (Quantum Mechanical Extension of the Consistent Force Field) approach[6a]. For example, the QCFF method treats conjugated molecules by assuming a formal separation of the σ and π electrons, with the former represented by analytical empirical functions and the latter by a second-order analytical representation of a semiempirical mode of the Dariser-Parr-Pople type corrected for orbital overlap. The method can treat ground electronic states and excited $\pi \to \pi^*$ states representing the potential surface of the $N^{\underline{th}}$ electronic state by

$$V^N(\boldsymbol{r}) = V_\sigma^0(\boldsymbol{r}) + V_\pi^0(\boldsymbol{r}) + \Delta V_\pi^N(\boldsymbol{r}) \qquad (2)$$

where **r** is the Cartesian coordinate vector, V_σ^0 (**r**)+V_π^0(**r**) is the sum of the $\sigma-$ and π-electron energies for the ground state and $\Delta V_\pi^N(\boldsymbol{r})$ is the π-electron excitation energy.

The σ-electron potential surface is given by

$$V_\sigma^0(\boldsymbol{r}) \;=\; \sum_i M(b_i) + \frac{1}{2}\sum_i [K_a(a_i - a_0)^2 + 2D_i] \qquad (3)$$

5

$$+\frac{1}{2}\sum_i K_\theta(\theta_i - \theta_0)^2 + \frac{1}{2}\sum_i F(q_i - q_0)^2 + \frac{1}{2}\sum_i K_\theta^{(1)}\cos\phi_i$$

$$+\frac{1}{2}\sum K_\theta^{(2)}\cos 2\phi_i + \frac{1}{2}\sum_i K_\chi(\chi_i - \chi_0)^2 + \sum_{k>l} f(r_{ij})$$

$$+\sum_{k>l}(Q_k^\sigma Q_l^\sigma)e^2/r_{kl}$$

The last term represents the electrostatic interaction between all atom pairs which have no π electrons.

The π-electron ground state is give by[5]

$$V_x^0(r) = \sum_A [P_A^\pi(W_\mu - \frac{1}{2}Q_A^\sigma\gamma_{AA}) + \frac{1}{4}(P_A^\pi)^2\gamma_{AB}] \qquad (4)$$

$$+ \sum_{A>B}\sum_\mu\sum_\nu [2P_{\mu\nu}^\pi\beta_{\mu\nu} - \frac{1}{2}(P_{\mu\nu}^\pi)^2\gamma_{AB}] + \sum_{A>B}\bar{Q}_A\bar{Q}_B\gamma AB$$

$$+ \sum_{A,C}[\bar{Q}_A Q_C^\sigma]e^2/r_{AC}$$

where A and B run over π atoms and C runs over σ atoms. The last term represents the electrostatic interaction between atoms with π electrons (A) and atoms without π electrons (C). The γ's and β's are quantum mechanical integrals which are approximated by analtyical functions. The Q's are the atomic charges and the P's are the so called bond orders, obtained from the π-electron molecular orbitals by

$$P_{\mu\nu} = \sum_i n_i v_{i\mu} v_{i\nu} \qquad (5)$$

where n_i is the number of electrons in the ith molecular orbital and where the molecular orbital vectors v_i are obtained by solving the SCF equation

$$\boldsymbol{F}v^i = \epsilon_i v^i \qquad (6)$$

The F matrix element are given by

$$F_{\mu\mu}^\pi = \bar{W}_\mu + (P_A^\pi - \frac{1}{2}P_{\mu\mu})\gamma_{AA} - \sum_B \bar{Q}_B\gamma_{AB} \qquad (7)$$

$$F_{\mu\nu}^\pi = \beta_{\mu\nu} - \frac{1}{2}P_{\mu\nu}\gamma_{\mu\nu}$$

where the W's γ's and β's are the same semiempirical integral used in eq. (4).With the bond orders evaluated quantum mechanically from the v's eq. (6) we obtain an analytical equation for the potential surface of eq. (4), which is correct to second order. Thus if we update the P's every iteration in the energy minimization procedure we can consider them to be constants and eq. (4) can be treated in the same way as the analytical function of the σ-potential (eq. (3)).

The excitation energy ΔV_π^N is also obtained in an analytical approximation in terms of the same semiempirical γ and β of eq. (4) but with other coefficients which depend on the Configuration Interaction (CI) vectors of the Nth excited state (see ref. 6 for details).

2.2 Consistent treatment of localized charges in the quantum Hamiltonian

One of the key problems associated with separation of σ and π electrons is the interrelation between the two subsystems. This problem which is not treated consistently in many π-electron treatements is solved in the Hetero atom version of the QCFF method[7]. This is done by formally considering the molecular system as an all-valence electron problem and keeping the σ-electron bond orders and charges unchanged in the π-electron calculations. This results in a modified F matrix that includeds the effect of the σ electrons in its diagonal in a way that accounts consistently for the effect of σ-core on the π-electron charge distribution (see ref. 7 for details). The same idea is used in treating solvent effects and is described below.

2.3 Incorporating the Solvent in the Calculations

Most practical questions about molecular conformations and spectra are related to molecules in solutions and not in the gas phase. The QCFF/SOL program addresses this issue, focusing on efficient calculations of molecular properties in solution. The solvent is represented by the very efficient Langevin dipoles (LD) model[8]. This model represents the time average polarization of the solvent molecules by a cubic grid of polarizable dipoles. The model is constructed in three steps (Fig.1) (i) A cubic grid (typically with 3 A spacing) is placed around the solute atoms (ii) Each grid point, which is

(a)   **(b)**   **(c)**
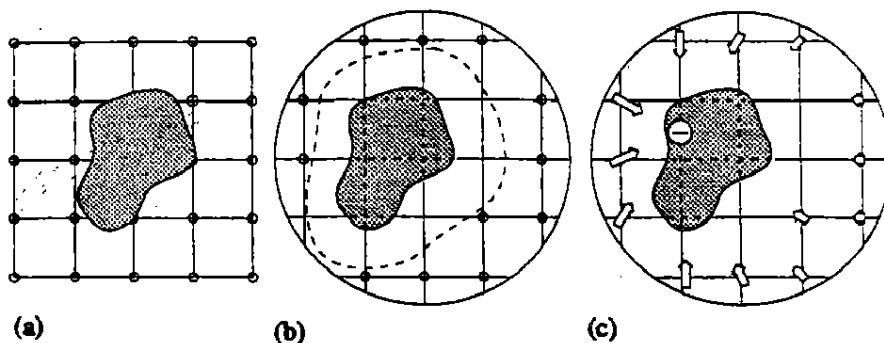
Figure 1: A schematic description of the langevin dipole model. The figure illustrates the three steps involved in constructing the model

within a van der Waals distance,r_w, from a solute atom is excluded (iii) The remaining grid points are then replaced by point dipoles whose polarization should mimic the average polarization of the solvent molecules at the same region in space. Each of this dipoles involve a permanent and induced moments, where the permanent component is represented by a Langevin type polarization law:

$$\mu_{per} = \mu_0(coth(\chi) - 1/\chi) \tag{8}$$

$$\chi = \mu_0 E/kT \tag{9}$$

while the induced dipoles are represented by

$$\mu_{ind} = \alpha\xi \tag{10}$$

where the α is the polarizability of the solvent molecule and ξ is the field generated by the solute charges, the solvent permanent dipoles and the solvent induced dipoles. This model allows one to obtain a very fast yet reliable estimate of solvation energies of medium size molecules and to include solvation free energy in energy minimization procedures.

8

Another key feature of the QCFF/SOL model is the unique ability to evaluate spectral properties of conjugated molecules (and most biological chromophores in solutions). This is done by incoperating the potential from the LD dipoles in the solute quantum mechanical Hamiltonian, using[6]

$$F_{\mu\mu} = F^0_{\mu\mu} - \sum_C Q_C / r_{AC} + \sum_D \mu_D \dot{r}_{AD} / r^3_{AD} \qquad (11)$$

where $\mathbf{F}^0$ is the SCF matrix of the isolated molecule (the gas phase molecule) the $\mathbf{Q}$ are the charges in the environment (e.g. counter ions) and μ are the solvent dipoles. These dipoles include permanent and induced components which are determined by the ground state charge distribution of the solute. In CI calculations of excitation energies we allow the solvent induced dipoles to repolarize themselves, while keeping the permanent dipoles at their ground state polarization.

This approach allows one to explore solvent effects on electronic spectra in a reliable way, reproducing experimentally observed shifts in polar and non polar solvents.

3   Getting the System Ready for QCFF/SOL

To familiarize ourself with the program lets consider a typical command file e.g. (Acrelein.COM) that looks like.

Run qcffsol	[1]
QCFF.parm	[2]
0 (1 if Chem3D)	[3]
Acrolein.mol	[4]
qcff.opt	[5]
sol.opt	[6]

The meaning of each of these lines are

- Line 1 - The command for running the program

- Line 2 - The parameters file

- Line 3 - A parameter that controls the source of the molecular geometry (0 for CHEM 3D and 1 for the standard QCFF input).

9

- Line 4 - The actual molecular information (geometry and topology).

- Line 5 - The control parameters.

- Line 6 - The parameters for solvation energy calculations.

As a first step just run this COM file and see that the program is working. Next lets familiarize ourselves with the three steps that the user must perform in order to prepare a run for a new molecules. These steps are:

1. Preparing molecular input file (e.g.mol.dat) that describes the molecular topology and initial structure.

2. Preparing (or updating) the force fields parameter file **qcff.parm**

3. Preparing the control file (e.g. **qcff.opt**)

The actual details of each of these steps are described below.

3.1   Preparing the Molecular Input

The information about the molecule under consideration can be prepared with our graphics program MOLARIS but in case you don't have access to this program here are some other options:

3.1.1   Using a CHEM3D file

If you have access to the popular CHEM3D program then just generate the output file for the given molecule and prepare a standard CHEM3D output file. To do this turn on your Chem-3D program and select 'Build' from the menu at the top of the screen. This option allows you to add/delete atoms, add/delete bonds. Depending upon the atom type standard bond lengths and bond angles are chosen. If you want to look at the values of bond lengths, bond angles etc use the option 'measurements' and specify 'show all measurements'. Another option is to first draw the molecule using the program Chem-draw and convert it to the 3D structure with Chem-3D. For saving the file, go to the 'file' menu and specify the option 'save as'. It will prompt for the filename and the format of the file. Choose the format 'cartesians' rather than 'Chem-3D' to save the cartesian coordinates.

```
10
   C    1     0.400528      1.155655      0.040039   63   2   5    6
   N    2    -1.090302      1.451965      0.040009   73   1   3    7
   C    3    -1.362274     -0.855042      0.040009   63   2   4    8
   C    4    -0.050888     -1.115646      0.040024   63   3   5    9
   C    5     0.830475     -0.110306      0.040039   63   1   4   10
   H    6     1.125687      1.982697      0.049576   11   1
   H    7    -1.375870      1.844788     -0.913422   11   2
   H    8    -2.017761     -1.331589      0.783768   11   3
   H    9     0.302841     -2.157166      0.030502   11   4
   H   10     1.909348     -0.324738      0.049576   11   5
```

To transfer the cartesian coordinates file return to your directory and type
'KERMIT', followed by 'receive [filename]' and follow the prompts to receive
the coordinates from Chem-3D into your directory.

A typical Mol file (for pyrolle) is given below
The QCFF/SOL program will translate the above information into its own
format.

3.1.2   Using Standard QCFF/SOL Format

In case you don't have access to MOLARIS or CHEM3D you should define
the molecular topology using the special symbolic formals described in ap-
pendix 1 (e.g. Benzene is (*AH AH AH AH AH *AH)). The input file will
now be arranged as:

11

Pyrolle	[1]
(*CHCHCHCH*NH)	[2]
	[3]
HO	[4]
(4F8.4)	[5]
1.796 -1.126 0.000 0.061	[6]
1.362 -2.109 0.000 0.042	
3.104 -0.716 0.000 -0.026	
3.967 -1.357 0.000 0.035	
3.104 0.713 0.000 -0.026	
3.967 1.354 0.000 0.035	
1.785 1.127 0.000 0.061	
1.351 2.110 0.000 0.042	
0.993 0.000 0.000 -0.371	
0.000 0.000 0.000 0.146	[15]

The meaning of each line is as follows:

- Line 1 - The title

- Line 2 - The symbolic formula (see appendix 1) - don't put more than 80 characters on a line.

- line 3 - A blank line reserved for a continuation of the symbolic formula.

- Line 5 - Side atoms (see appendix 1 for definition).

- Line 5 - The format for the molecular coordinates (which are given below).

- Line 6-15 - Molecular cartesian coordinates and atomic charges (X,Y,Z, and Q).

3.2 Potential Parameters

Although QCFF/SOL includes many default parameters it is sometimes necessary to add new parameters for some molecules. The parameter update and selection is done in qcff.parm. The current file and the program include a default parameter set and we strongly recommend that you run the program a few times without changing this file. Examination of the parameter set in

the output is extremely instructive for understanding the meaning of various parameters.

Once you have familiarized yourself with the types of potential parameters and their typical ranges, then you are ready to change them, if necessary. The arrangement of the parameters in the qcff.parm file and the way to modify them is decribed in Appendix 2.

3.3   Control Parameters

The options of the program are controlled by a set of parameters which are described below:

```
        Default Updated
```

	Default	Updated	
Nsteep	20	20	Nsteep is the number of steepest descent iterations
Nmin	4	0	Nmin is the number of modified Newton-Raphson iterations
Ivib	1	1	Ivib=1 for calculation of vibrational frequencies
Imode	0	0	Imode=1 for print out ofnormal mode vectors
Nao	3	3	Nao is the number of occupied orbitals
Nex	0	0	Nex is 0 for ground electronic state,
			N for N'th excited state,
			-1 for the reduced state,
			-2 for oxidized state
Fia	Sing	Trip	Fia='Sing' for calculation of singlet excitation Fia='Trip' for triplet excitation
Itest	0	0	Itest=1 for test comparison between numerical and analytical derivatives
Ipolr	1	1	Ipolr=1 for correction of vibrational frequencies by including the derivatives of the bond orders
Ntor	0	0	Ntor is the number of steepest descent iterations in the subspace of given torsional variables if Ntor.gt.0 the torsional variables should be defined
Ng	4	4	Ng is the number of levels in the C.I. calculations if Ng=0 then the configurations should be

```
                          read from input cards
Idip      0       0       Idip=1 for computation of dipole moment
                          if Idip=1 the data for computing the dipole
                          orientation have to be defined
Isig      0       1       Isig=1 for computing the sigma charge influence
                          the charges have to be read in or computed
                          by Del Re procedure
Idelre    0       1       Idelre=1 the sigma charges are computed by
                          the Del Re method
Isolv     0       1       Isolv=0 - solvation effects are not included
                          in the calculations,
                          Isolv=1 - solvation effects are considered
                          within the PDLD model
```

The default values of this parameters can be changed by including the new value in the file **qcffsol.opt** for example you may have the following file:

MOL	[1]
Nsteep=20,	[2]
Isig=1,	[3]
Idip=0,	[4]
Imode=0,	[5]
Itest=0,	[6]
Ivib=1,	[7]
Nmin=0,	[8]
Idelre=1,	[9]
Isolv=1,	[10]
Fia='trip',	[11]
end	[12]

In this specific case we have modified the specified default parameters and left the other unchanged. If you like to change any other control parameters just add it to the list.

4 Running the QCFF

4.1 Standard runs with minimization, normal mode analysis and electronic spectrum

A standard QCFF job minimizes the given molecule in its ground electronic state, calculates the $\pi \rightarrow \pi^*$ electronic spectrum of this molecule (if it has π atoms) and performs normal mode analysis at the equilibrium geometry. To examine this mode of operation run **pyrolle.com**(with isol=0). The results of such a run are given below.

```
final calculations of gas MOs
```

```
final iteration count      1
final p-matrix
```

	A 1	A 3	A 5	A 7	N 9
A 1	1.0711	0.7703	-0.1053	-0.4119	0.4699
A 3	0.7703	1.0936	0.5766	-0.1053	-0.2331
A 5	-0.1053	0.5766	1.0937	0.7703	-0.2331
A 7	-0.4119	-0.1053	0.7703	1.0710	0.4700
N 9	0.4699	-0.2331	-0.2331	0.4700	1.6706

```
final f-matrix
```

	A 1	A 3	A 5	A 7	N 9
A 1	-4.4389	-4.9401	0.2960	1.1732	-4.1840
A 3	-4.9401	-4.0105	-4.1749	0.2959	0.5605
A 5	0.2960	-4.1749	-4.0111	-4.9400	0.5605
A 7	1.1732	0.2959	-4.9400	-4.4385	-4.1842

N 9   -4.1840   0.5605   0.5605   -4.1842   -10.7603

final eigenvalues

| 1 | -14.499490 | 2 | -9.682799 | 3 | -8.703386 | 4 |

final eigenvectors

	A 1	A 3	A 5	A 7	N 9
1	-0.3806	-0.2099	-0.2099	-0.3806	-0.7888
2	-0.1413	-0.6111	-0.6112	-0.1413	0.4616
3	0.6089	0.3595	-0.3595	-0.6089	0.0000
4	-0.5789	0.2871	0.2871	-0.5790	0.4058
5	0.3595	-0.6089	0.6089	-0.3595	0.0000

h-matrix

	A 1	A 3	A 5	A 7	N 9
A 1	-13.0250	-2.4512	0.0000	0.0000	-2.6919
A 3	-2.4512	-11.3897	-2.3146	0.0000	0.0000
A 5	0.0000	-2.3146	-11.3910	-2.4512	0.0000
A 7	0.0000	0.0000	-2.4512	-13.0240	-2.6920
N 9	-2.6919	0.0000	0.0000	-2.6920	-25.2335

c.i. matrix:

| 5.798350 | -0.000028 | -0.000009 | -0.087542 |
| -0.000028 | 6.761863 | 0.702374 | -0.000021 |

16

```
      -0.000009        0.702374        6.836688        0.000006
      -0.087542       -0.000021        0.000006        7.857754
```

configurations:
```
   3   4            3   5            2   4            2   5
```

c.i. eigen values and eigen vectors

```
      5.794690
      0.999101        0.000046       -0.000023        0.042394
      6.095860
     -0.000050        0.725669       -0.688044        0.000009
      7.502639
      0.000016       -0.688044       -0.725669       -0.000024
      7.861417
      0.042394        0.000025        0.000010       -0.999101
```

TOPDLD output

Solvent ground state potential (vq) at the solute atoms
(vqperm and vqind are the corresponding permanent and induced components)
and total atomic charges and their changes in each excited state
----

atom	vq	vqperm	vqind	q(gr)	dq:	1	2	3	4
1	0.000	0.000	0.000	0.037		0.035	-0.022	-0.052	-0.109
2	0.000	0.000	0.000	0.064		0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	-0.132		0.047	0.011	0.039	0.003
4	0.000	0.000	0.000	0.054		0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	-0.131		0.047	0.011	0.039	0.003
6	0.000	0.000	0.000	0.054		0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.037		0.035	-0.022	-0.052	-0.109
8	0.000	0.000	0.000	0.064		0.000	0.000	0.000	0.000

```
 9    0.000  0.000  0.000    -0.263      -0.164 0.023 0.025 0.212
10    0.000  0.000  0.000     0.214       0.000 0.000 0.000 0.000
```
----

excitation energies, transition dipoles, oscillator strength and
quantum-chemical and classical estimates of solvent shifts contributions

----

exc en, cm-1 (in solvent)	mu-x	mu-y	mu-z	osc str	solv-gas shift quantum total	classical contributions	
						perm	ind
46740.0	0.00	0.60	0.00	0.36	0.0	0.0	0.0
49169.3	0.00	0.00	0.00	0.00	0.0	0.0	0.0
60516.3	0.96	0.00	0.00	0.93	0.0	0.0	0.0
63410.2	0.00	0.52	0.00	0.27	0.0	0.0	0.0

Final Results

  Pyrolle

internal coordinates


```
Bond length  type  atomi  atomj
      1.0804  AH      1      2      1
      1.3938  AA      1      3      2
      1.3350  AN      1      9      3
      1.0719  AH      3      4      4
      1.4289  AA      3      5      5
      1.0719  AH      5      6      6
      1.3938  AA      5      7      7
```

1.0804	AH	7	8	8
1.3349	AN	7	9	9
1.0077	NH	9	10	10

Theta angle	type	atomi	atomj	atomk	
125.5412	HAA	2	1	3	11
125.3092	HAN	2	1	9	12
109.1499	AAN	3	1	9	13
127.4130	AAH	1	3	4	14
105.8058	AAA	1	3	5	15
110.0887	ANA	1	9	7	16
124.9557	ANH	1	9	10	17
126.7816	HAA	4	3	5	18
126.7815	AAH	3	5	6	19
105.8059	AAA	3	5	7	20
127.4128	HAA	6	5	7	21
125.5409	AAH	5	7	8	22
109.1502	AAN	5	7	9	23
125.3093	HAN	8	7	9	24
124.9559	ANH	7	9	10	25

Phi angle	type	atomi	atomj	atomk	atoml	
0.0000	HAAH	2	1	3	4	26
180.0000	HAAA	2	1	3	5	27
180.0000	NAAH	9	1	3	4	28
0.0000	NAAA	9	1	3	5	29
180.0000	HANA	2	1	9	7	30
360.0000	HANH	2	1	9	10	31
0.0000	AANA	3	1	9	7	32
180.0000	AANH	3	1	9	10	33
180.0000	AAAH	1	3	5	6	34
360.0000	AAAA	1	3	5	7	35

0.0000	HAAH	4	3	5	6	36
180.0000	HAAA	4	3	5	7	37
180.0000	AAAH	3	5	7	8	38
0.0000	AAAN	3	5	7	9	39
360.0000	HAAH	6	5	7	8	40
180.0000	HAAN	6	5	7	9	41
360.0000	ANAA	1	9	7	5	42
180.0000	ANAH	1	9	7	8	43
180.0000	AANH	5	7	9	10	44
360.0000	HANH	8	7	9	10	45
180.0000	HAAN	2	3	1	9	46
180.0000	AHAA	1	4	3	5	47
180.0000	AANH	1	7	9	10	48
180.0000	AHAA	3	6	5	7	49
180.0000	AHAN	5	8	7	9	50

atomic   coordinates

		x	y	z
1	A	1.762	-1.095	0.000
2	H	1.398	-2.112	0.000
3	A	3.103	-0.714	0.000
4	H	3.962	-1.355	0.000
5	A	3.102	0.715	0.000
6	H	3.960	1.357	0.000
7	A	1.761	1.093	0.000
8	H	1.395	2.110	0.000
9	N	0.997	-0.001	0.000
10	H	-0.011	-0.002	0.000

energy contributions

total energy	=	-902.92	kcal
diagonal core energy	=	-114.39	kcal
bond energy	=	-895.50	kcal
non-bond energy	=	-15.86	kcal
repulse energy	=	356.23	kcal

```
theta energy          =        34.29 kcal
phi energy            =      -267.69 kcal
solvation energy      =         0.00 kcal
```

first derivation with respect to

x	y	z
0.0002	-0.0010	0.0000
0.0000	0.0000	0.0000
0.0002	0.0007	0.0000
0.0000	0.0000	0.0000
0.0001	-0.0005	0.0000
0.0000	0.0000	0.0000
-0.0006	-0.0008	0.0000
0.0000	0.0000	0.0000
0.0001	0.0016	0.0000
0.0000	-0.0001	0.0000

iteration history

iter	energy	step
1	-900.391410	0.002000
2	-901.188173	0.002400
3	-901.885634	0.002880
4	-902.390304	0.003456
5	-902.675941	0.004147
6	-902.740845	0.004977
7	-902.519239	0.002488
8	-902.858384	0.002986
9	-902.738102	0.001493
10	-902.886897	0.001792
11	-902.863533	0.000896
12	-902.905419	0.001075
13	-902.898986	0.000537
14	-902.912881	0.000645
15	-902.913675	0.000774
16	-902.910451	0.000387

17	-902.917019	0.000464
18	-902.916489	0.000232
19	-902.917449	0.000279
20	-902.917107	0.000139

	vibrational frequency	I.R. intensity
1	3366.17	1.43
2	3108.88	0.09
3	3105.07	0.80
4	3103.16	1.40
5	3101.67	3.06
6	1554.86	1.39
7	1476.39	1.17
8	1450.91	0.00
9	1401.21	1.20
10	1366.54	0.00
11	1155.51	0.04
12	1130.21	0.08
13	1129.10	0.81
14	1114.01	0.00
15	1089.83	0.18
16	1061.63	1.20
17	988.70	1.30
18	938.04	0.06
19	900.75	0.00
20	791.95	0.00
21	786.53	0.08
22	680.25	6.52
23	530.79	0.41
24	518.18	0.00
25	11.45	0.58
26	8.91	0.67
27	8.91	0.58
28 (*)	8.91	0.67
29	8.91	0.58
30 (*)	8.91	0.67

vibrational enthalpy at room temperature =   52.44
   zero point energy =   51.23
   atomization energy =   -851.69

   ground and 1-st excited states charges

atom	qtot(gr)	qpi(gr)	qsig(gr)	qtot(ex)
1	0.04	-0.07	0.11	0.07
2	0.06	0.00	0.06	0.06
3	-0.13	-0.09	-0.04	-0.08
4	0.05	0.00	0.05	0.05
5	-0.13	-0.09	-0.04	-0.08
6	0.05	0.00	0.05	0.05
7	0.04	-0.07	0.11	0.07
8	0.06	0.00	0.06	0.06
9	-0.26	0.33	-0.59	-0.43
10	0.21	0.00	0.21	0.21

sum of total, pi, sig crgs: 0.000 0.000 0.000

ground and excited state dipole moments

	dipx	dipy	dipz	dipole
ground -	-1.617	-0.002	0.000	1.617
excited -	-0.420	-0.001	0.000	0.420

To familiarize yourself with the program change the number of steepest descent minimization steps (nsteep) and the number of the Newton Raphson steps. See how the minimization changes. Note that a good minimum has zero first derivatives (examine the output for this) and have six zero frequencies (examine the calculated vibrational frequencies).

Next examine the calculated electronic excitation energies and compare them to corresponding experimental results listed in ref. 7. Similarily examine the calculated vibrational spectra.

Now try to prepare an input for some other molecules (preferably molecules which are listed in ref. 7, and examine how well you reproduce the observed geometry, vibrations, electronic spectra and dipole moments.

4.2   Excited State Calculations

The program can calculate equilibrium geometries and vibrational spectra in excited electronic states (singlets and triplets) and in oxidized and reduced molecules. To examine this option take for example the butadien molecule and run it with nex=1 compare the change in equilibrium geometry and vibrational frequencies to the analysis given in ref. 6.

The ability to calculate excited state surfaces is important for example in calculations of Resonance Raman spectra.

4.3   Incorporating the Solvent in the QCFF Calculations

As explained in section 2.3, you can use the QCFF/SOL program to evaluate solvent effects on structure, energy and spectra. In particular we emphasize the unique ability of the program to evaluate solvent effects on electronic spectra.

In order to activate the solvent part of the QCFF calculations you have to change the value of the parameter "isol " in the option file from 0 to 1 (see the file **qcff.opt**)

The calculations involve a formation of a grid of Langevin dipoles (LD) around the solute. This LD solvent model involves permanent and induces components and is controlled by the parameters in the file **sol.opt**. This file looks like:

```
6 3
   12.000      3.10      3.10      12.0      2.500      1.48      0.350
```

The meaning of each line is as follows:

- Line 1 - Format(2I2) - The first parameter gives the number of different grids of solvent dipoles on the solvation surface of the molecule. This

24

number is employed in getting the average solvent potential (numbers 2, 4, 6 can be used), when it equals 1 - only the cubic grid is built. the second parameter controls the calculation of the solvent potential: 1 - the field of permanent solvent dipoles is taken into account, 2 - the field of both permanent and induced dipoles is taken into account, 3 - the same as 2, but additionally the permanent dipoles are allowed to reorient according to the field from the excited state charges of the solute.

- Line 2 - These are, respectively, the radius of the langevin grid, distance between the langevin dipoles of the outer grid, distance between the langevin dipoles on the solvation surface, langevin dipole cutoff for selfconsistent procedure (sci), the cutoff for the solute-solvent distance, the polarizability of the solvent dipole, the magnitude of the solvent dipole moment in units of A and electron charge (for water use 0.35)

As an example for the QCFF/SOL calcualtions you may try to run the file **acrolein.com**. The results of this run are shown below:

```
final calculations of solv MOs



final iteration count      7
final p-matrix
```

	O 1	A 2	A 4	A 6
O 1	1.4689	0.8375	-0.0852	-0.2674
A 2	0.8375	0.6336	0.3463	0.2107
A 4	-0.0852	0.3463	1.0041	0.9342
A 6	-0.2674	0.2107	0.9342	0.8934

```
final f-matrix
```

	O 1	A 2	A 4	A 6
O 1	-9.1729	-5.8035	0.2324	0.5167
A 2	-5.8035	-3.6734	-3.2825	-0.5657
A 4	0.2324	-3.2825	-4.3838	-5.6918
A 6	0.5167	-0.5657	-5.6918	-4.2594

final eigenvalues

| 1 | -13.327101 | 2 | -10.269075 | 3 | -0.479522 | 4 |

final eigenvectors

	O 1	A 2	A 4	A 6
1	0.7465	0.5628	0.3038	0.1833
2	0.4209	-0.0034	-0.6401	-0.6427
3	0.4474	-0.6304	-0.2761	0.5712
4	-0.2558	0.5346	-0.6494	0.4765

h-matrix

	O 1	A 2	A 4	A 6
O 1	-19.1848	-2.4479	0.0000	0.0000
A 2	-2.4479	-11.7882	-2.1682	0.0000
A 4	0.0000	-2.1682	-10.2404	-2.6725
A 6	0.0000	0.0000	-2.6725	-9.2062

26

c.i. matrix:
```
     5.484671        0.392086        0.400266       -0.012131
     0.392086        7.740304        0.063524       -0.139323
     0.400266        0.063524        7.567775        0.477433
    -0.012131       -0.139323        0.477433       10.318841
```

configurations:
```
   2  3          2  4          1  3          1  4
```

c.i. eigen values and eigen vectors

```
     5.350613
     0.972448       -0.154056       -0.174332        0.014807
     7.484066
     0.088967       -0.444843        0.875036       -0.168855
     7.871241
    -0.215349       -0.881035       -0.420070        0.030722
    10.405675
     0.007349       -0.046460        0.165718        0.985051
```

TOPDLD output

estimation of bulk solvation using born equation

Born radius -   13.50 total charge -   0.00
Born energy -   0.00

LANGEVIN ENERGY -   -12.955
NUMBER OF DIPOLES-   235
TOTAL NUMBER OF DIPOLES =   235

DIPOLE PAIR LIST GENERATED
        TOTAL PAIRS   -   55225
        PAIRS EVALUATED -   11787

$(V*Q)/2=$ -8.4483 ELGVN=   -7.3416 EMUP=   -7.2920 EMUI=   -1.3619

stabilization by solvent permanent dipoles
 for ground and excited states   (in kcal/mol)

   -7.29   -10.22   -1.97   -7.10   1.71

stabilization of solute by solvent polarizability
for ground and excited states (in kcal/mol)

   -1.36   -2.91   -0.70   -1.34   -1.64

Solvent ground state potential (vq) at the solute atoms
(vqperm and vqind are the corresponding permanent and induced components)
and total atomic charges and their changes in each excited state
----

atom	vq	vqperm	vqind	q(gr)	dq: 1	2	3	4
1	-1.152	-0.959	-0.193	-0.576	-0.008	0.309	0.149	0.487
2	-0.183	-0.126	-0.058	0.401	-0.385	-0.121	-0.254	0.027
3	0.300	0.286	0.014	0.061	0.000	0.000	0.000	0.000
4	0.415	0.395	0.020	-0.046	0.316	0.003	0.009	-0.319
5	0.229	0.231	-0.002	0.054	0.000	0.000	0.000	0.000
6	1.036	0.942	0.094	-0.003	0.078	-0.192	0.097	-0.195
7	1.146	1.043	0.103	0.054	0.000	0.000	0.000	0.000

```
    8   1.301  1.175  0.126    0.054      0.000 0.000 0.000 0.000
```
----

solvent shift components:

dq is the change in charge upon excitation,
V is the potential from solvent permanent and induced dipoles

```
dq(ex-gr)*V(gr)=         -2349.894
dq(ex-gr)*Vperm(gr)=     -2049.198
dq(ex-gr)*Vind(gr)=       -300.696
```
----

excitation energies, transition dipoles, oscillator strength and
quantum-chemical and classical estimates of solvent shifts contributions

----

exc en, cm-1 (in solvent)	mu-x	mu-y	mu-z	osc str	solv-gas shift quantum total	classical contributions perm	ind
43158.0	-0.74	0.25	0.00	0.61	-3489.2	-2050.7	-541.1
60366.5	-0.29	0.04	0.00	0.09	669.1	3723.4	230.7
63489.4	0.55	0.09	0.00	0.31	946.7	134.3	8.0
83932.2	0.30	0.18	0.00	0.12	3732.7	6304.5	-98.1

Note the effect of the solvent on the calculated excitation energies. Here, the excitation energy (for the molecule in solution) is obtained as a sum of the excitation energy from the quantum calculations (including the solvent effect in the Hamiltonian) and the classical contribution from the induced dipoles energy. The total quantum solvent shift is obtained as the difference between the excitation energy of the given solvent and the excitation energy of an isolated molecule with the same geometry. The classical contributions from the permanent and induced solvent dipoles are calculated using the

classical formulas for solute-solvent interactions[10].

4.4   Special Feature

1. Controlling the CI calculations

2. Minimizing in torsional space

3. Changing the σ charges explain overwriting Del Re

4. n-π^* Transitions

5   QCFF Calculations With All Atoms Solvent (or Proteins) Models.

The standard option for QCFF/SOL involves the simplified and efficient LD model. Sometimes, however, you may be interested in more realistic (and much more expensive) approach which incorporates the effect of an all-atom solvent or protein model in the solute Hamiltonian. This approach which is integrated with a MD simulation program requies the package ENZYMIX.

5.1   Spectra of Chromophores in Protein Sites

If you have access to the combined package of ENZYMIX + QCFF/SOL you can perform calculations of the spectral properties of biological chromores inside their protein active site. This is done by the following procedure: After consulting the enzymix manual use the regular PREPARE procedure to create Enzymix input file. Then prepare a running com.file which is almost identical to the regular Enzymix com.file and look like:

100	Number of steps	[1]
5	Temperature	[2]
0.02	Stepsize	[3]
40	NB cycle rate	[4]
bptix2		[5]
0	Flag for restart option	[6]
rest1		[7]
gap1		[8]
3	option flag	[9]
nevb.dat		[10]
qcfenz tyr.poly		[11]
1.0	AC mapping parameter	[12]
15.0 15.0 11.0	Center of protein	[13]
10.0	Exclusion radius	[14]
10.0	Water radius	[15]
0	Protein constraints	[16]
0	Atom constraints	[17]
1 10 2.5	Induced dipole flag	[18]
0.01	Constraint for moving atom	[19]
10. 10. 10. 10. 10.	Cutoffs	[20]
qcff.parm		[21]
qcff.opt1		[22]
tyr.spectr		[23]

The meaning of each line is as follows:

- Line 1 - This is the number of steps of dynamics that the run will last. For a stepsize of 0.04 (Line 3), 500 steps will give one picosecond of dynamics.

- Line 2 - This is the temperature (in Kelvin) that the run will be done at.

- Line 3 - This is the stepsize (in arbitrary units). For normal runs this should be 0.04.

- Line 4 - This is the number of steps between updates of the nonbonded pair list. 40 is a good value, but if you find that your system blows up for no obvious reason during dynamics (AFTER it has been relaxed), set it to a lower value.

- **Line 5** - This is the name of the Big input file generated by PREPARE.

- **Line 6** - This flag is 0 if you are using the x-ray coordinates from the Brookhaven file and 1 if you want to read in the coordinates from a previous run. The flag is negative if you are generating downhill trajectories. The absolute value of the negative number is the number of restart files the program will genereate.

- **Line 7** - This is the name of the restart INPUT file if the flag on line 6 is 1 or is negative and it is the name of the restart OUTPUT file if the flag is 0.

- **Line 7a** - This line only exists if the flag on line 6 is 1 or is negative. In that event this line has the name of the restart OUTPUT file. (In the above example this line doesn't exist).

- **Line 8** - This is the name of the energy gap file for the enzymix run. This is the main output data from the program and is used to map the free energy profile of the reaction you are modelling.

- **Line 9** - This is the flag that controls the enzymix option and it is 1 - for EVB calculation, 2 - for AC calcu- lations and 3 - for QCFF/SOL.

- **Line 10** - The EVB option, it is not used, leave as it is.

- **Line 11** - This is the name of the AC file for our "quantum" chromophore (e.g. tyrosine in BPTI)

- **Line 12** - This option is not being used (leave 1.0)

- **Line 13** - These are the coordinates for the 'center' of the protein - pick a point approximately at the center of the EVB atoms that you have defined. This can be done by looking at the coordinates of the EVB atoms using the option 'resatm' when running PREPARE .

- **Line 14** - This is the radius of the exclusion region, beyond this region all protein atoms are constrained to lie at their x-ray coordinates.

- **Line 15** - This is the maximum radius (from the protein center) that the program will attempt to put waters at.

- **Line 16** - This is the number of constrained distances that the user wants

- **Line 17** - This is the number of atomic constraints that the user wants.

- **Line 18** - This is the flag that controls the calculations of induced-dipole effects. Set the flag to 0 if the effect of induced dipoles is not included. When the flag is set to 1 your results will include the effect of induced dipole but the calculation will be much longer. This line also contains values of step and scaling for induced dipoles.

- **Line 19** - This parameter is the prefactor of the constraint imposed on the whole protein system.

- **Line 20** - These are the cutoffs, respectively, for the intercations between: protein-protein, protein-water, water-water, AC/EVB-protein and AC/EVB-water. Typical parameters are given in this control file.

- **Line 21** - The QCFF potential parameters

- **Line 22** - The QCFF option file (see section 3.3)

- **Line 23** - An output file that collects the time dependent excitation energy for calculation of the spectral line shape

To understand what the program does it is useful to run first the demo job qcfenz_tyr1.com. This run evaluates the exitation energies of a tyrosine in BPTI during a molecular dynamics run of the enzyme (the time average of the excitation energies gives you the actual spectra). During the run you will see the calculated excitation energies fluctuating in response to the changing field from the protein + water surroundings at each time step. the values of the excitation energies are collected in the file tyr.spectr. When you complete the run analyse this file in the following way

1. The average of ΔE, will give you the the λ_{max} in the actual protein site

2. A furrier transform of $\Delta E(t)$ will give you the line shape of the given transition

To see this please run spectra.com. Further demonstration can be obtained by running acr_qcfenz.com which gives the time dependent spectra of acrolein in water.

6   Solvation calculations coupled with all valence electron programs

Methods of theoretical chemistry are widely used now in the studies of chemical mechanisms of drugs biological activity and corresponding Quantitative Structure-Activity Relationships (QSAR). Solvation properties of the molecules are usually incorporated in the physico-chemical QSAR model with some reactivity indices. Solvation energies in water can be obtained by Free Energy Pertubation (FEP) method using the program Enymix. However the FEP method and related approaches are very time consuming and in many cases cannot be integrated with standart standart quantum-chemical approaches without special modification. A powerful general alternative is provided by the PDLD solvent model that can be incorporated in quantum chemical hamiltonian (as demonstrated above with the QCFF hamiltonian). Apparently the PDLD approach can be integrated with any other quantum chemical program. This only requires to create the proper interface when the potential from the PDLD dipoles is consistently introduced into the given quantum chemical hamiltonian using eq.(11). and where the PDLD polarisation is determined iteratively by the solute quantum mechanical charges. This procedure is implemented here with the AMPAC quantum chemical package.

    The changes in AMPAC program have been minimal and involved only the subroutines, which control the input of data (READ, COMPFG), building of the Hamiltonian (FOCK1), calculation of electronic energy (HELECT), iterative process (ITER). The version of the PDLD program is the same one used in QCFF/SOL approach. The PDLD/AMPAC version (which is called here SOLMOL) include the modified subroutines READ, COMPFG, FOCK1, HELECT and ITER which can be incorporated in the user own AMPAC version and the PDLD program which is automatically called by subroutine FOCK1.

    To submit a job with the SOLMOL program one should type the following

line
<div align="center">**@solmol   name**</div>

where **name** is a PDLD/AMPAC input file name.dat. A sample **nh4.dat**
file which contains the data for the ammonium ion is included.

```
mndo CHARGE=1   1scF
ammonium ion, standart geometry
mndo ground state
N
H   1   1.023
H   1   1.023   2   109.5
H   1   1.023   2   109.5   3   120.0
H   1   1.023   2   109.5   3   -120.0

 1
   6   4   12.0
    2   3   4   5
```

The first part of this data file is a standart MNDO data ending with a blank
card. Note, that for the calculation of solvent effects in this variant of the
program the keyword **1SCF** should be used. The blank card is followed by
the PDLD input, where the meaning of each line is following:

Line 1 Format(1x,I1) - Control parameter for inclusion solvent effects:

        0 - no calculation of solvent effects,
        1 - calculation of the effect of polar solvent,
        2 - calculation of the effect of nonpolar solvent.

Line 2 Format(2I4,F7.1) - Three control parameters nrg, natsol, rg

      nrg - the number of different grids of solvent dipoles, employed
            in getting the average solvent potential

<div align="center">35</div>

(even numbers 2, 4 or 6 can be used)

natsol - number of atoms, which are specifically solvated. For these
         atoms the minima of electrostatic potential on the first
         solvation envelope are determined and the point dipoles are
         placed at this sites.

rg - radius of the solvent grid.

Line 3 Format(10I4) - Should be present if natsol > 0. Contains the
                       array iatsol (with the dimension natsol).

iatsol - the sequential numbers (according to the chemical formula)
         of atoms, which are specifically solvated

The results of PDLD/AMPAC calculations are placed in **name.out** file
in the same directory. To illustrate the perfomance of the SOLMOL program
we give below the results of the calculations with PDLD/AM1 method of the
solvation energies for a series of methyl substituted ammonium ions in water.

Table 1. Solvation energies (kcal/mole) of methylammonium ions in water.

Ion	$\Delta(\Delta H)$	E_{perm}	E_{ind}	E_{solv}^{total}	E_{solv}^{exp} [xx]
NH_4^+	0.0	-61.8	-14.5	-87.6	-84.0±3.0
$N(CH_3)H_3^+$	0.3	-55.7	-10.7	-77.3	-75.4
$N(CH_3)_2H_2^+$	0.3	-52.4	-8.3	-71.8	-69.7
$N(CH_3)_3H^+$	0.0	-48.1	-6.4	-65.8	-63.3

7   Appendix A: Using the symbolic formula to express molecular topology

If you like to use the standard QCFF format you should define the bonding
pattern in the molecule by a symbolic formula that looks, for example, like
(*AH AH AH AH AH *AH) for benzene. The roles of constructing this type
of formula are given below.

Atoms are represented by the following letters (HONCABDMQRTXYZW)
H-hydrogen, D-deuterium, C-sp3 carbon (CH2 group), B-sp3 carbon (CH3

group), A-sp2 (pi) carbon, O-oxygen in C=O group N-sp2 (pi) pyrrole type nitrogen, M-sp2 (pi) pyridine type nitrogen, the other letters can be used for defining any other atom. The connectivity is defined by the order of the atomic symbols and a few special symbols. Two classes of atoms are distinguished: "side atoms" which can never be connected to more than one atom, and "chain atoms" which can. The special symbols are left and right parenthesis "(", ")" the digits "2" and "3" and the period ".". Blanks are ignored and have no effect other than that 10 consecutive blanks mark the end of the formula. All other characters are known and "un-regionized symbols"; they are used to indicate bonds not implied by the ordering of atoms in the formula.

The use of this symbolic notation is best explained by the following steps (see also fig.2):

1. Draw the molecule in 2-dimensions, using the appropriate symbols (eq. H for hydrogen) for atoms and drawing lines between bonded atoms.

2. If there are closed rings of atoms choose any bond in the ring and write a particular un-recognized symbol before the symbol of each of the two atoms forming the bond. Now the line indicating this bond can be rubbed out because the connection between the bonded atoms is implied by the pair of unrecognized symbols. After doing this for one bond in every ring the remaining lines between the atoms form a simply connected network.

3. Select a chain atom at the end of one chain as the first atom, i.e. its index number will be 1. Write a left parenthesis and follow it with the atomic symbol of the chosen chain atom. If the chain atom has an un-recognized symbol next to it, this symbol must be written down before the chain atom symbol. Next write down the symbols of the side atoms connected to the proceeding chain atom, using the multiplicity symbols to indicate 2 or 3 side atoms of the same type. Continue in this way until reaching a chain atom that is connected to more than two other chain atoms (a branch point). One branch of the chain after this point is a continuation of the previous chain, the other branches are new chains. After writing down the side atoms connected to the branch point chain atom, write a left parenthesis (to mark the beginning of a new chain) followed by the first chain atom of the new chain. Continue

like this until all the atoms of the new chain have been written down.
Now write a right parenthesis (to mark the end of a chain) and return
to the previous chain. In this way all the atoms in a chain are enclosed
in parenthesis and follow the symbol of the branch point atom. If
there are two or more new chains connected to a branch point atom
then each new chain is enclosed in parenthesis and is written after side
atoms connected to the branch point atom. If there are two or more
new chains connected to a branch point atom then each new chain is
enclosed in parenthesis and is written after side atoms connected to the
branch point atom. At the end of the last chain write a left parenthesis
to enclose the whole formula. Any 'un-recognized symbol' can be used
only for one bond. '2' and '3' are multiplicity factors for side atoms, '.'
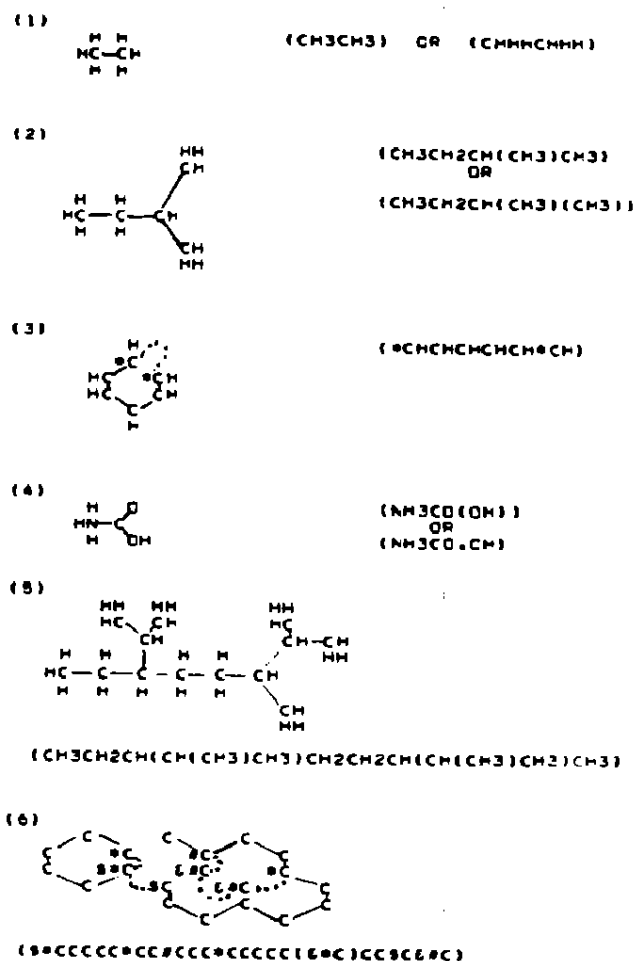converts the side atom it preceeds to a chain atom.

Fig.2 Examples of the QCFF symbolic formula.

7   Appendix B: The Potential Parameters and How to Modify them.

The parameter file **qcff.parm** contains the default parameters of QCFF/SOL. Sometimes you may encounter a molecule where some parameters are missing

or not well optimized. This will be indicated to you by a warning message from the program. In order to add or modify parameters just follow the description of the various parameters below and add the missing parameters in the same format used for other parameters of the same type.

8.1   Classical Bond Parameters

The functional form of the energies are:

- for sigma atoms (regular quadratic potential)

$$F(b) = \frac{1}{2}K_1(b - b_0)^2 + K_2 \qquad (12)$$

- for pi atoms (Morse potential)

$$F(b) = D[\exp\{-2a(b - b_0)\} - 2\exp(-a(b - b_0)] \qquad (13)$$

The parameters for each pair of bonded atoms are given (in a free format) in two lines as described below.

  AB

$K_1(D) \qquad b_0(b_0) \qquad K_2(a)$

where A and B are the letter codes of the atoms involved and the three relevant potential parameters are either K_1, b_0 and K_2 or D, b_0 and a depending on the atom type. The end of this group of parameters is marked by a blank line.

8.2   Classical Angle Parameters

The functional form of the classical angle bending energy term is

$$F(\theta) = K_1(\theta - \theta_0)^2 + K_2(s - s_0) + K_q(q - q_0)^2 \qquad (14)$$

where θ is the bond angle defined by the three atoms i, j, k (whose letter codes are A, B and C, respectively) and q is the distance between atom i and

atom j (this distance is called a 1-3 distance). The variable s is taken as θ for the CCC angle and q for all others.

The parameters for each type of angle are given in a free format in two lines as described below.

ABC

K_1 θ_0 s_0 K_q K_2 q_0

This group of parameters is ended by a blank line.

8.3   Classical Torsion Parameters

The classical contribution to the torsional potential is

$$V(\phi) = K_1 \cos(n\phi) + K_3 \cos \phi - \rho_1(\phi_1 - \theta_0)(\theta_2 - \theta_0) \cos \phi \qquad (15)$$

where ϕ is the torsional angle defined by the four atoms i, j, k, l (where the letter codes of the j and k atoms are A and B, respectively), θ_1 is the bond angle defined by the ijk atoms and θ_2 is the bond angle defined by the j, k, l atoms. The parameters for each type of torsional angle are given (in a free format) in two lines

XABX

K_1 K_2 K_3 N   IATG

where ρ_1 is taken as K_3 for the special case of HCCH and as $-\sqrt{K_1 K_3}$ for HCCC, while it is taken as K_2 for all other cases. The integer IATG specifies how many torsional angles can be defined around the AB atom pair (e.g. we can define nine angles around the CC bond in ethane).

8.4   Classical Nonbonded Parameters

The interactions between nonbonded atoms are described by

$$V(r_{ij}) = A \exp(-\mu r_{ij}) - B/r^6 + q_i q_j / r_{ij} \qquad (16)$$

The parameters for each atom pair are arranged in two lines as follows.

AB

A μ B

The atomic charges (the q's) are determined by the Del Re procedure.

8.5 Atomic Charges and Del Re Parameters

The atomic charges in the program are composed of sigma and pi contributions. The pi contributions are calculated by the quantum mechanical part of the program. The sigma charges are evaluated by the Del Re[9] procedure which appears to provide a practical and reliable way of reproducing of much more expensive quantum mechanical calculations.

The Del Re parameters, are described in ref. 7 and should be left unchanged. In case you prefere to use your own set of sigma charges please consult section 4.4.

8.6 Quantum Parameters

The quantum part of the program uses semiempirical integral (α, β and γ) for the π electron interactions. These include the following:

(i) The alpha integral is given by

$$\alpha = \alpha_0 - \alpha_1 \exp(-\mu(b - b_0)) \cos \phi^2 \qquad (17)$$

where each π atom is characterized by a given α_0 and the α_1 term is added (as a Lowdin correction) for any bond to a neighboring π atom. The α parameters are given in two lines

A

α_0 α_1 α_2

The parameter μ is taken from the corresponding β integral and the parameter α_2 is used in the γ integral below. Note that any atom for which you define an α parameter becomes automatically a π atom.

(ii) The β integral (the resonance integral) is given by

$$\beta = \beta_0 \exp\{-\mu(b-b_0)\}(1+K(b-b_0))\cos\phi(1+K'\cos\phi))/(1+K') \quad (18)$$

where b is the bond between a pair of π atoms. The parameters are given in two lines as

AB

$$\beta_0 \quad b_0 \quad \mu \quad K \quad K'$$

(iii) The γ integral (the repulsive integral) is given by

$$\gamma_{ii} = \alpha_2(i) + \gamma_3 \exp\{-\mu(b_{ij} - b_0))\cos\phi_{kijl} \quad (19)$$

where atom i and j determine a non zero β integral with the corresponding μ and b (see ref.6a) and

$$\gamma_{ij} = (\gamma_1 - \gamma_2)\exp\{-\mu'r\} + 14.39/(r + 14.39/\gamma_2) \quad (20)$$

but for atom whose β integral is not defined we use

$$\gamma_{ij} = 14.39/(r + 14.39 \times 2/(\alpha_2(i) + \alpha_2(j))) \quad (21)$$

The b term in eq. (21) runs on the bonds between the ith atom and other π atoms bonded to it. The γ parameters are given in two lines as

AB

$$\gamma_1 \quad \gamma_2 \quad \mu \quad \gamma_3$$

References

(1) W. J. Hehre, L. Radom, P. v. R. Schleyer, AB INITIO Molecular Orbital Theory, John Wiley'Sons, New York, 1988.

(2) a G. Klopman and R.C. Evans, in "Modern Theoretical Chemistry", Vol.7, G. Segal, ed., Plenum, New York, 1977. b R.C. Bingham, M.J.S. Dewar, and D.H.Lo, *J. Am. Chem. Soc.*, **97**, 1294, 1285, 1302, 1307, 1311 (1975). c M.J.S. Dewar, E.G.Zoebisch, E.F.Healy, J.J.P. Stewart, *J. AM. Chem. Soc.*, **107**,3902 (1985)

(3) Lifson, S. and Warshel, A., *J. Chem. Phys.* **49**, 5119 (1968). b Warshel, A. and Lifson, S.,*J. Chem. Phys.* **53**, 582 (1970). c Warshel, A., Levitt, M. and Lifson, S., *J. Mol. Spectrosc.*, **33**, 84 (1970).

(4) N.L. Allinger, *Adv. Phys. Org. Chem.*, **13**, 1 (1976).

(5) a P.K.Weiner, P.A.Kollman, *J.Comp.Chem.*, **2**, 287 (1981). b B.R.Brooks, R.E.Brrucoleri, M.Karplus, *J. Comp. Chem.*, **4**, 187 (1983).

(6) a A. Warshel and M. Karplus, *J. Am. Chem. Soc.*, **94**, 5612 (1972). b A. Warshel, in "Modern Theoretical Chemistry", Vol. 7, G. Segal, Ed.,Plenum, New York, 1977. c A. Warshel and M. Levitt, Quantum Chemistry Program Exchange, No. 247,Indiana University, 1974.

(7) A. Warshel and A. Lappicirella, *J. Am. Chem. Soc.* **103**, 4664 (1981).

(8) a A. Warshel, *J. Phys. Chem.* **83**, 1640 (1979). b A. Warshel and S.T. Russell, *Quart. Rev. Biophys.* **17** 283(1984). c S.T. Russell, and A. Warshel, *J. Mol. Biol.* **195**, 389 (1985).

(9) G. Del Re, *Theor. Chim. Acta.* **1**, 188 (1963).

(10) V.Luzhkov, A.Warshel, submitted.