JOM 23018

Regioselectivity of nucleophilic additions to cyclopentadienyliron complexes of substituted benzenes: a novel theoretical approach based on Boltzmann probabilities

Xincai Luo [a], Gustavo A. Arteca [a], Chunhao Zhang [b] and Paul G. Mezey [a,c]

[a] *Department of Chemistry, University of Saskatchewan, Saskatoon, Sask., S7N 0W0 (Canada)*
[b] *Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 1A1 (Canada)*
[c] *Department of Mathematics, University of Saskatchewan, Saskatoon, Sask., S7N 0W0 (Canada)*

(Received April 13, 1992; in revised form June 15, 1992)

Abstract

In this work, we compute the relative probabilities of the different locations a model nucleophile may occupy on the molecular surfaces of cyclopentadienyliron complexes of a series of substituted benzenes. To this end, we construct a Boltzmann probability function evaluated at various positions along a formal molecular surface.

The results show that for the addition of a hydride ion to these organometallic complexes, the experimental product distribution (regioselectivity) is a function of the changes of the classical statistical probability of the nucleophile to contact different regions of the molecular surface.

A correlation is found between the experimental relative product yields and the theoretical relative Boltzmann probabilities of a nucleophile attacking a given position on the reactant. The results provide a quantitative description of regioselectivity of the nucleophilic additions, with electron-donating substituents favoring addition to the *meta*-position and electron-withdrawing substituents favoring *ortho*-additions.

1. Introduction

The analysis of the regioselectivity of the nucleophilic additions to organometallic complexes is of key importance to organometallic chemistry [1–4].

Considerable amounts of experimental data are now available on the regioselectivity of the nucleophilic addition of hydride ion to cyclopentadienyliron (FeCp) complexes of substituted benzenes, $[C_6H_5XFeCp]^+$, [5–10]. It has been found that the nature of the substituents influences the hydride addition. Electron-donating substituents favor the addition at the *meta*-position, whereas electron-withdrawing ones favor *ortho*-addition. By contrast, there is no clear trend in the orientation towards the *para*-addition.

An INDO treatment of the FeCp complexes of toluene, anisole, and methylbenzoate was carried out by Clack and Kane-Maguire [11]. Their results suggest that the product distribution cannot be deduced easily from a simple study of charges on the molecule or on the frontier orbitals. The authors propose some correlations with other properties and characteristics of the nucleophile, but their description does not account for the actual values of regioselectivity. Some computational work has been done at semi-empirical and *ab initio* levels for other metal-pentadienyl compounds [12], but to our knowledge no further theoretical research has dealt with the $[C_6H_5XFeCp]^+$ species. In this work, we propose a simple, alternative theoretical approach to the study of regioselectivity in these compounds.

A molecule can be treated as a three-dimensional entity described by a formal molecular surface. The study of molecular surfaces has found many applications in studying the chemical and biological activity of molecules [13]. Molecular van der Waals surfaces [14,15] (built up by fused atomic spheres) are well

---

Correspondence to: Dr. P.G. Mezey.

0022-328X/93/$6.00

suited for studying regioselectivity of chemical reactions. Recently, graphic displays of molecular surfaces have been used as a theoretical tool for modeling nucleophilic and electrophilic reactions in organometallic chemistry [16]. Weber *et al.* [16] have placed a model nucleophile or electrophile on molecular surfaces and displayed these surfaces on a computer screen by colouring regions according to the total interaction energies. As the positions for nucleophilic and electrophilic attacks depend on the total interaction energies, these displays are quite useful in recognizing reactive positions. In this work, we move a step further and carry out a quantitative calculation of the distribution of the probability of a model nucleophile attacking the molecular surfaces of cyclopentadienyliron complexes of a series of substituted benzenes. Our procedure uses notions from statistical mechanics in that we consider a molecular surface as a formal, restricted phase space on which Boltzmann probabilities can be computed. From the practical point of view, we follow an approach which parallels closely the simple computational framework of Weber *et al.* [16] for the analysis of the reactivity in organometallic compounds.

2. The method and its practical implementation

According to statistical mechanics, the distribution of particles in a closed system in thermal equilibrium is given by the Boltzmann distribution [17]. Let ϵ_i be the energy associated to a physical state i. Then, the probability P_i of a particle to appear in such a state is given by

$$P_i = \frac{e^{-\epsilon_i/kT}}{\sum_i e^{-\epsilon_i/kT}} \leq 1 \tag{1}$$

where k is the Boltzmann constant and T is the absolute temperature. The sum in the denominator of eqn. (1) is a partition function and it runs over all the possible states in the system.

Similarly, one can define the probability $P_{[i_1,i_2]}$ of the system to appear in any state with energies between ϵ_{i_1} and ϵ_{i_2} (with $\epsilon_{i_2} \geq \epsilon_{i_1}$). In this case one has:

$$P_{[i_1,i_2]} = \frac{\sum_{i_1 \leq i \leq i_2} e^{-\epsilon_i/kT}}{\sum_i e^{-\epsilon_i/kT}} \tag{2}$$

From this equation, one can introduce the relative probabilities corresponding to the ratio of the probabilities of the system to appear in states associated to two different ranges of energies. If these ranges are defined by the indices $[i_1,i_2]$ and $[i_3,i_4]$, then

$$\frac{P_{[i_1,i_2]}}{P_{[i_3,i_4]}} = \frac{\sum_{i_1 \leq i \leq i_2} e^{-\epsilon_i/kT}}{\sum_{i_3 \leq i \leq i_4} e^{-\epsilon_i/kT}} \tag{3}$$

These equations are completely general and they describe the statistical behavior of a system comprising an arbitrary number of particles and energy levels $\{\epsilon_i\}$. In particular, they are valid for a system, defined over some general phase space, with a single particle populating the energy levels $\{\epsilon_i\}$. The above model allows one to introduce a Boltzmann probability measure that is useful in the study of regioselectivity for organic compounds.

In this work we are interested in modeling the product distribution in the nucleophilic addition to cyclopentadienyliron complexes of substituted benzenes. For this case, the *particle* in our statistical system is the *model nucleophile.* The *phase space* on which the "energy levels" are defined will be *the accessible molecular surface* [14,15]. This surface is the one generated by the center of a model spherical nucleophile being rolled on the van der Waals surface of the complex [14,15]. When the nucleophilic sphere (the "particle") is placed at a point on the accessible molecular surface, it is possible to associate an energy ϵ_i with that point, *i.e.* the interaction energy between the nucleophile and the molecule at this precise configuration. From these energies one can compute Boltzmann probability values and ranges, as shown above.

Our phase space (the accessible surface) is continuous. Accordingly, the summation in eqn. (3) should be replaced by an integral over the phase space. In other words, the summations in eqns. (1)–(3) become surface integrals. As a result, one can define probabilities associated either with ranges of energies or *with sections of the molecular surface.* Several schemes of surface partitioning can be followed [13]. Consider two different regions of the molecular surface with areas A_1 and A_2. The ratio of probabilities P_{A_1} and P_{A_2} of the model nucleophile appearing on these regions is thus given by (*cf.* eqn. (3)):

$$\frac{P_{A_1}}{P_{A_2}} = \frac{\int_{A_1} e^{-\epsilon(\mathbf{r})/kt}\, d\sigma(\mathbf{r})}{\int_{A_2} e^{-\epsilon(\mathbf{r})/kT}\, d\sigma(\mathbf{r})} \tag{4}$$

In eqn. (4), the integrals in the numerator and denominator are respectively over regions A_1 and A_2 of the surface, **r** is a vector specifying the location of the nucleophile on the surface, and $d\sigma(\mathbf{r})$ is the differential of area at **r**.

In this study, we are interested in describing relative product yields for nucleophilic additions to the cyclopentadienyliron complexes of a series of substituted benzenes. To this end, we shall calculate the ratios P_o/P_m of Boltzmann probabilities associated with the *ortho* (P_o) and *meta* (P_m) positions on the substituted benzenes. According to eqn. (4), these probabilities P_o and P_m can be evaluated by integrating on the regions of the accessible surface generated by the atomic spheres of *ortho* and *meta* carbons respectively. Note that the areas associated with each single atom can be computed rigorously and the points on the surface common to two or more atoms lie on spherical arcs and thus do not contribute to the area.

The integrals eqn. (4), are computed numerically. As we shall compare our results with relative product distributions, we have only computed the ratio of probabilities (eqn. 4). This saves one the computation of the integral (or sum) in the denominator of eqn. (2), which is over the whole molecular surface. This term, the computationally more demanding, cancels out in the ratio. The interaction energy $\epsilon(\mathbf{r})$ is taken as an electrostatic energy, which in a first approximation can be calculated by the interaction of atomic point charges,

$$\epsilon(\mathbf{r}) = \sum_i \frac{q_i q_m}{D r_i} \tag{5}$$

In eqn. (5), q_i is the atomic charge on the atom i of the complexes, q_m is the charge of the model nucleophile, D is the dielectric permittivity, and r_i is the distance between atom i and the model nucleophile. The summation runs over all atoms of the complex. Note that $\epsilon(\mathbf{r})$ changes only as a function of the point on the atomic sphere touched by the nucleophile. In turn, the changes in electrostatic potential $\epsilon(\mathbf{r})$ over the exposed atomic surface are the main effect on the actual P_i probability values.

The atomic charges associated with the atoms in the iron complexes have been computed by extended Hückel molecular orbital (EHMO) method [18]. We have decided to use this method in order to follow an approach in tune to the ones used to explore the reactivity of organometallic compounds [16]. The charge of the model nucleophile is taken as -0.242 a.u., which is the atomic charge of a H atom on the actual nucleophile $[BH_4]$-calculated with the RHF *ab initio* method using the 6–31 G* basis set [19]. Other quantum mechanical methods, other than EHMO, can of course be used to compute charges. However, at present, these methods may not be very practical for large organometallic complexes.

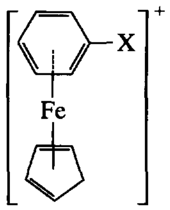
It must be mentioned that the EHMO atomic charges alone do not account for the correct distribution of products, as was already found by using INDO charges [11]. For example, EHMO charges would predict much larger yields for *meta*-derivatives than for *ortho*-derivatives for the chloro-substituted benzene cyclopentadienyliron complex. The experimental result is the other way around, as it would be expected from the fact that Cl^- is a mostly electron-donating group. Similarly, EHMO charges do not predict the correct order in the regioselectivity of derivatives of benzene cyclopentadinyliron complexes substituted by electron-withdrawing groups. As we shall see below, where point charges fail, the model based on the probability of attack seems to work. In this latter case, instead of the point charges, one considers the overall electrostatic interaction with the charged molecular surface.

The solvent effect is approximately accounted for by a distance-dependent dielectric permittivity. By adopting an approximation used in molecular dynamics [20], we have taken D to be 1.0 when the carbon atom has a direct contact with the model nucleophile, and $2r_i$ when the atom has no direct contact with the model nucleophile (with r_i measured in Å). In this work, we shall not study the effect of changes in the dielectric constant on the results, since we have so far experimental data available only for a single solvent (1,2-dimethoxyethane for all derivatives, except tetrahydrofuran for the chloro-derivative, ref. 9).

For $[C_6H_5FeCp]^{2+}$, we use the geometry estimated by Clack and Kane-Maguire [11]. The geometries for the substituents are estimated from average crystal structures [21].

In generating the points of the phase space (the accessible surface), the van der Waals radii for carbon, hydrogen, oxygen, nitrogen, and chlorine have been taken from ref. 22. For iron, we have taken into consideration the experimental fact that the complexes studied exhibit only exo-addition under nucleophilic attacks by hydride, that is, the hydride ion always attacks the complex from the "top" of the benzene ring [10]. Our model, based on point charges, is too simple to explain all reactivity features. For this reason, we have constrained it to describe only exo-addition. Accordingly, we estimated a radius of 2.9 Å, so that the endo-face of the surface of the carbon atoms in benzene is not accessible to the hydride ion. The radius for the model nucleophile sphere is taken as the van der Waals radius of a hydrogen atom, 1.17Å [22].

We have performed the EHMO calculations with the program FORTICON8 and the standard set of EHMO parameters [18]. For iron, we use a quadratic dependence on the atomic charge for the diagonal matrix elements of the Hamiltonian $\{H_{ii}\}$. For other atoms,



1: X = CN
2: X = COOCH$_3$
3: X = Cl
4: X = CH$_3$
5: X = OCH$_3$
6: X = N(CH$_3$)$_2$

Fig. 1. Cyclopentadienyliron complexes studied. The six substituents X considered are listed on the right-hand side. The regioselectivity of the nucleophilic addition to these compounds is described in terms of the relative location of the entrant group with respect to X (*ortho* or *meta*).

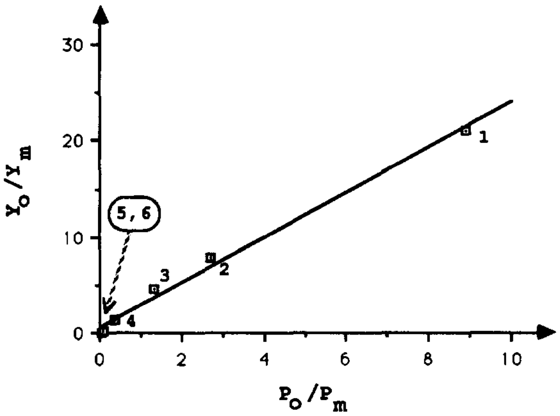


Fig. 2. Correlation between experimental relative product yields for the *ortho*- and *meta*-additions (Y_o/Y_m) and the theoretical relative Boltzmann probabilities of attack to the *ortho* and *meta* positions (P_o/P_m) by a model nucleophile. [The numbers 1,2,3,4,5, and 6 in the figure correspond to the substituents X labelled as in Fig. 1. Note that the substituents 5 and 6 (OCH$_3$ and N(CH$_3$)$_2$ respectively) are almost superimposed at the scale of the drawing. The ratio Y_o/Y_m is smaller for X = N(CH$_3$)$_2$.]

the $\{H_{ii}\}$ values are kept fixed. All integrals are calculated with a standard program.

As the experimental yields were obtained at room temperature, we have used $T = 298$ K in all calculations.

3. Correlation between theoretical and experimental results

We have carried out a calculation of the relative Boltzmann probabilities P_o/P_m for the cyclopentadienyliron complexes of the six substituted benzenes shown in Fig. 1.

These six complexes cover a wide range of values in the relative product yields [9]. The ratio of the experimental yields of the products of *ortho*-addition (Y_o) to those of *meta*-addition (Y_m) vary from $Y_o/Y_m = 21$ to 0 when moving from the primarily electron-withdrawing substituent CN to the predominantly electron-donating substituent N(CH$_3$)$_2$ respectively.

Our main results are presented in Table 1. The experimental relative yields Y_o/Y_m are taken from previous work by one of us [9], where all reactions have been measured under the same experimental conditions. Table 1 compares these relative yields to the ratio of Boltzmann probabilities, P_o/P_m, of a nucleophile attacking *ortho* and *meta* positions, respectively. The results show that our approach can be used successfully to model the regioselectivity of these additions. For the series of six substituents, the change in theoretical relative probabilities establishes the following decreasing order of preference of the *ortho*-addition over the *meta*:

$$P_o/P_m\text{: CN} > \text{COOCH}_3 > \text{Cl} > \text{CH}_3 > \text{OCH}_3 > \text{N(CH}_3)_2$$

which coincides with the experimental ordering. Comparing these probabilities with the experimental ratios Y_o/Y_m, one can immediately see that the relation between these two quantities is not only qualitative but also nearly quantitative. From the electron-withdrawing substituent CN to the electron-donating substituent N(CH$_3$)$_2$, the relative probability P_o/P_m decreases from 8.89 to 0.056, whereas the experimental relative product yield decreases from 21 to 0. Figure 2 summarizes the comparison, showing a good linear correlation between Y_o/Y_m and P_o/P_m.

TABLE 1. Comparison between relative Boltzmann probabilities and experimental relative *ortho*/*meta* product distributions for the nucleophilic additions on cyclopentadienyliron complexes of substituted benzenes

Complex	Substituent	Relative probabilities (P_o/P_m)	Relative product distributions (Y_o/Y_m)
1	CN	8.89	21
2	COOCH$_3$	2.71	7.9
3	Cl	1.32	4.5
4	CH$_3$	0.343	1.3
5	OCH$_3$	0.081	0.15
6	N(CH$_3$)$_2$	0.056	0.0

The results in Fig. 2 satisfy the following least-square linear equation (with a confidence level of 95%):

$$Y_o/Y_m = (2.46 \pm 0.62)\, P_o/P_m \qquad (6)$$

The correlation coefficient is 0.986. The fitting has been constrained to have zero intercept (only the slope is varied), since the condition $Y_o = 0$ if $P_o = 0$ is expected.

Note that the three electron-withdrawing substituents, Cl, COOCH$_3$, and CN, clearly favor *ortho*-additions since they have high values of P_o/P_m. In

TABLE 2. Comparison between relative Boltzmann probabilities and experimental relative *ortho/para* product distributions for the nucleophilic additions on cyclopentadienyliron complexes of substituted benzenes

Complex	Substituent	Relative probabilities (P_o/P_p)	Relative product distributions (Y_o/Y_p)
1	CN	8.4	68
2	$COOCH_3$	2.17	11.1
3	Cl	2.16	9.0
4	CH_3	1.13	1.57
5	OCH_3	0.54	0.30
6	$N(CH_3)_2$	0.37	0.0

contrast, the two electron-donating substituents $N(CH_3)_2$ and OCH_3 favor meta-additions (low values of P_o/P_m). Finally, the nearly electron-neutral substituent CH_3 shows no major preference of *ortho*-additions over *meta*-additions.

Our results predict correctly that *ortho*-additions are not favored if P_o/P_m approaches zero. For example, the probability of a model nucleophile to stay on the molecular surface of the $N(CH_3)_2$ derivative is nearly twenty times larger at the *meta*-position than at the *ortho* one. This observation agrees with the fact that no *ortho*-substituted products are experimentally detected.

In contrast, the experimental relative product distributions Y_p/Y_m vary over a much smaller range (from 0.31 to 0.83), and no clear trend is observed with respect to the type of substituent [9]. Table 2 shows our results for the correlation between the experimental relative product distributions Y_o/Y_p *vs.* P_o/P_p. The results agree qualitatively with the ones shown in Table 1 and Fig. 2 and the correct experimental order of reactivity is predicted:

$$P_o/P_p\text{: CN} > COOCH_3 > Cl > CH_3 > OCH_3 > N(CH_3)_2$$

This indicates that our procedure provides a reasonable model to represent the *ortho*-addition in relation to the other two possible attack positions. However, in this latter case, the agreement is not as good as for the the ratio of Y_o/Y_m, indicating that the ratio P_p/P_m does not show a clear trend. At this moment, we cannot explain this feature. The experimental results show the same, somewhat erratic behaviour, with some electron-withdrawing substituents producing the same Y_p/Y_m ratio as some electron-donating groups.

4. Discussion and conclusions

If a reaction is kinetically controlled, then the final product distribution will depend on the energies of all possible transition structures ("states"). Similarly, if a reaction is thermodynamically controlled, then the final product distribution will depend on the relative free energies of all possible products. In principle, the regioselectivity can be studied by calculating the energy barriers along all possible reaction paths and the free energies of all possible products.

However, the above approaches are not very practical for most actual chemical problems such as the present nucleophilic additions of the hydride ion (of $[BH_4]^-$) to cyclopentadienyliron complexes of substituted benzenes. It is not simple to calculate reaction paths and free energies by using the current theoretical methods and the present generation of computers. Consequently, it is important to have alternative, simple but reliable methods to treat reactions of interest.

In this study, we have proposed a method that combines notions of statistical and quantum mechanics to analyze the probability of nucleophilic attacks over different locations on a model molecular surface. Our results indicate that the approach can be useful in modelling the regioselectivity of the addition reactions on organometallic compounds.

The model consisting of a spherical nucleophile on the molecular surfaces of the cyclopentadienyliron complexes mimics the initial conditions of the reactions. The linear correlation obtained between the theoretical descriptor and the experimental results seems to indicate that the *final product distribution* is related to the *initial attack probabilities*. This may indicate that the reaction possesses an early transition structure, and that the reaction is thus kinetically controlled. The verification of this conclusion would require different models and computations.

The results provide a strong support for the use of molecular surfaces to describe quantitatively the regioselectivity of nucleophilic reactions. A modification of the proposed method, using a shape group analysis [23,24] of level sets (ranges) of local probabilities along molecular surfaces, can provide a new similarity concept for molecules [25]. The present model, though simplified, accounts for the main experimental features observed in the set of reactions. If desired, a number of refinements can be introduced. For instance, interaction terms other than the purely electrostatic one have been ignored. Other terms can be included by using the electron density distribution instead of point atomic charges. Moreover, one can take into account the polarizing effect of the charge of the nucleophile on the charge distribution on the molecule by means of a reaction field. This may be an observable effect in the case of hydride ion as nucleophile, which is quite polarizable.

These approaches above would lead to a more accurate description of the dynamics of the interaction.

However, in view of the present results, the effect of these corrections should not be dominant. Note that small changes in the point charges will be rather "smoothed out" by the Boltzmann exponentials in eqn. (4).

As stated before, other quantum mechanical methods, other than EHMO, can be used to provide charge distribution. Nevertheless, for large organometallic compounds other alternatives may still not be very practical.

References

1 S. G. Davies, M. L. H. Green and D. M. P. Mingos, *Tetrahedron, 34* (1978) 3047.
2 M. F. Semmelhack, G. R. Clark, J. L. García, J. J. Harrison, Y. Thebtaranonth, W. Wulff and A. Yamashita, *Tetrahedron, 37* (1981) 3957.
3 D. Astruc, *Tetrahedron, 39* (1983) 4027.
4 R. G. Sutherland, M. Iqbal and A. Piórko, *J. Organomet. Chem., 302* (1986) 307.
5 I. U. Khand, P. L. Pauson and W. E. Watts, *J. Chem. Soc. C*, (1968) 2257.
6 I. U. Khand, P. L. Pauson and W. E. Watts, *J. Chem. Soc. C*, (1968) 2261.
7 I. U. Khand, P. L. Pauson and W. E. Watts, *J. Chem. Soc. C*, (1969) 116.
8 J. F. McGreer and W. E. Watts, *J. Organomet. Chem., 110* (1976) 103.
9 R. G. Sutherland, C. H. Zhang, R. L. Chowdhury, A. Piórko and C. C. Lee, *J. Organomet. Chem., 333* (1987) 367.
10 C. H. Zhang, R. L. Chowdhury, A. Piórko, C. C. Lee and R. G. Sutherland *J. Organomet. Chem., 346* (1988) 367.
11 D. W. Clack and I. A. P. Kane-Maguire, *J. Organomet. Chem., 174* (1979) 199.
12 R. D. Ernst, *Struct. Bonding, 57* (1984) 1.
13 P. G. Mezey, in K. B. Lipkowitz and D. B. Boyd (eds.), *Reviews in Computational Chemistry*, VCH Publishers, Inc., New York, 1990.
14 F. M. Richards, *Annu. Rev. Biophys. Bioeng., 6* (1977) 151.
15 M. L. Connolly, *J. Appl. Crystallogr., 16* (1983) 548.
16 J. Weber, P. Y. Morgantini, J. Leresche and C. Daul, in R. Carbó (ed.), *Studies in Physical and Theoretical Chemistry*, Vol. 62, Elsevier, Amsterdam, 1989.
17 T. L. Hill, *An Introduction to Statistical Thermodynamics,* Addison-Wesley, Reading, 1960.
18 J. Howell, A. Rossi, D. Wallace, K. Haraki and R. Hoffmann, *QCPE, 11* (1979) 344.
19 M. J. Frisch, J. S. Binkley, H. B. Schlegel, K. Raghavachari, C. F. Melius, R. L. Martin, J. J. P. Stewart, F. W. Bobrowicz, C. M. Rohlfing, L. R. Kahn, D. J. Defrees, R. Seeger, R. A. Whiteside, D. J. Fox, E. M. Fleuder and J. A. Pople, *Gaussian 88*, Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburgh, 1988.
20 S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta Jr. and P. Weiner, *J. Am. Chem. Soc., 106* (1984) 765.
21 F. H. Allen, O. Kennard, D. G. Watson, L. Brammer, G. Orpen, and R. Taylor, *J. Chem. Soc.*, (1987) S1.
22 A. Gavezzotti, *J. Am. Chem. Soc., 105* (1983) 5220.
23 P. G. Mezey, *Int. J. Quantum Chem., Quantum Biol. Symp., 12,* (1986) 133.
24 P. G. Mezey, *J. Comput. Chem., 8* (1987) 462.
25 X. Luo, G. A. Arteca and P. G. Mezey, in preparation.