Theor Chem Acc (2002) 107:147–153
DOI 10.1007/s00214-001-0308-8

Theoretical
Chemistry Accounts

*Regular article*

# The search for stationary points on a quantum mechanical/molecular mechanical potential-energy surface

**Xavier Prat-Resina[1], Mireia Garcia-Viloca[1], Gerald Monard[2], Angels González-Lafont[1], José M. Lluch[1], Josep Maria Bofill[3], Josep Maria Anglada[4]**

[1] Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
[2] Groupede Chimie et Biochimie Théorique, Université Henri Poincaré – Nancy I, UMR 7565, Faculté des Sciences – B.P. 239, 54506 Vandoeuvre-lès-Nancy, France
[3] Departament de Química Orgànica and Centre Especial de Recerca en Química Teòrica, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain
[4] Institut d'Investigacions Químiques i Ambientals, C.I.D.–C.S.I.C., Jordi Girona Salgado 18–26, 08034 Barcelona, Spain

Received: 16 July 2001 / Accepted: 9 October 2001 / Published online: 9 January 2002


**Abstract.** We propose a methodology to locate stationary points on a quantum mechanical/molecular mechanical potential-energy surface. This algorithm is based on a suitable approximation of an initial full Hessian matrix, either a modified Broyden–Fletcher–Goldfarg–Shanno or a Powell update formula for the location of, respectively, a minimum or a transition state, and the so-called rational function optimization. The latter avoids the Hessian matrix inversion required by a quasi-Newton–Raphson method. Some examples are presented and analyzed.

**Key words:** Quantum mechanical/molecular mechanical stationary points location – Modified Broyden–Fletcher–Goldfarg–Shanno – Rational function optimization methods – Update Hessian matrix formula

## 1 Introduction

A challenging problem in computational chemistry is the optimization of a function of many degrees of freedom. The computational effort and performance of each optimization method basically depend on the number of variables and the evaluation cost of the function under optimization and its derivatives. In addition, the degradation of the optimization method is strongly related to the shape or behavior of the function. From the experience accumulated in the past years in the use of optimization algorithms [1] the main conclusion is that the efficiency of any optimization method is characterized by three factors; namely, the type of coordinates used to define the system, the well behavior of the Hessian matrix and the way to predict the next molecular geometry during the iterative process. In the last two decades much research has been done to propose and improve algorithms to optimize the geometry of large molecular systems. The main reasons for this effort are the recent developments of combined quantum mechanical (QM) and molecular mechanical (MM) approaches to evaluate the potential energy [2, 3] and linear scaling methodologies which both now allow the computation of very large reactive systems.

The algorithms used to find stationary points of large molecular systems can be divided into two groups. The first group contains the mixed algorithms which, in fact, combine a truncated Newton–Raphson optimizer with a preconditioned linear conjugate- gradient technique [4, 5]. In the second group, the algorithms are basically those employed to optimize molecules of medium size using QM potentials, but with some minor changes. The modifications of these algorithms account for the fact that the direct application of these techniques needs a big storage memory for large systems [6–13]. One of the most important factors in optimization is the selection of the coordinate system and owing to this fact much effort has been concentrated on the treatment of the **G** matrix [14]. For molecules of medium size the **G** matrix is fully diagonalized, which implies a cubic scaling of the computational effort with the number of redundant coordinates. This fact limits the optimization of large molecular systems. To avoid this important limitation, the transformation of gradient and coordinates and the computation of the delocalized internal coordinates is carried out in terms of iterative solutions of linear equations [1, 7, 8, 10–13].

In contrast to the attention on the treatment of the **G** matrix, very little effort has been made in the evaluation

*Correspondence to:* J.M. Lluch
e-mail: lluch@klingon.uab.es

and manipulation of the Hessian matrix, which is one of the factors that ensures efficiency in the optimization process. For molecules of medium size the storage of the Hessian matrix (analytic or approximated) is $O(n^2)$ memory locations, where $n$ is the number of variables. On the other hand, the solution of the standard Newton–Raphson equations implies $O(n^3)$ additions and multiplications [4]. This fact is again a serious limitation for the optimization of large molecular systems. In minimization problems the Hessian matrix is updated at each iteration by using the Broyden–Fletcher–Goldfarg–Shanno (BFGS) formula [15]. There already exist expressions which avoid storing the full Hessian matrix using the BFGS update formula for large molecular systems, the so-called limited BFGS (L-BFGS) [16–18]. Billeter et al. [11] used the L-BFGS formula to update the Hessian matrix when a minimization of a large molecular system is carried out. In the specific case of location and optimization of transition structures, both Turner et al. [10] and Billeter et al. [11] proposed a partition of the molecular system such that the "environment" part is minimized using L-BFGS and a first-order saddle point in the degrees of freedom of the rest of the molecule (the core) is searched for. In this way a Hessian matrix needs only to be stored and maintained for the core.

Recently Anglada et al. [9] reformulated the Powell formula to update Hessian matrices [19] with limited memory, L-Powell. Since the Powell formula is the standard update used in the optimization of transition structures [20, 21], the reformulation of Anglada et al. [9] enables the Powell formula [19] to be used to locate large transition structures without partition of the Hessian matrix. In addition Anglada et al. [9] used the augmented Hessian (AH) or rational function optimization (RFO) technique [20–26] to compute the next geometry. The Hessian matrix is diagonalized partially; only the first two eigenpairs are computed. These two eigenpairs are related to the transition vector and the geometry displacement vector [20, 21]. The diagonalization of the AH is carried out by using a Lanczos-type algorithm [27, 28] in the way that the full Hessian matrix is never stored in the memory of the computer. In general, the RFO technique predicts the best-improved molecular geometry for the next iteration. This fact contributes to the stability of the optimization process.

In this article we present a location of stationary points of relatively large molecules by using the QM, MM and QM/MM methodologies and the RFO technique. The update formula used in this iterative process for the location of a minimum is a modified form of the BGFS formula.

## 2 Method

The standard Newton–Raphson method is based on the optimization of a quadratic model of the energy with respect to the geometry parameters. By replacing this quadratic model by a rational function approximation we obtain the RFO method [20, 21]. This rational expansion at iteration k is

$$\mathrm{E}(\mathbf{q}_k+\Delta\mathbf{q}_k)-\mathrm{E}(\mathbf{q}_k)\cong \mathrm{q}(\Delta\mathbf{q}_k)=\frac{\mathrm{Q}(\Delta\mathbf{q}_k)}{1+\Delta\mathbf{q}_k^{\mathrm{T}}\mathbf{S}_k\Delta\mathbf{q}_k}$$
$$=\frac{\frac{1}{2}\left(1/\mathrm{T}\Delta\mathbf{q}_k^{\mathrm{T}}\right)\begin{pmatrix}0 & \mathbf{g}_k^{\mathrm{T}}\\ \mathbf{g}_k & \mathbf{B}_k\end{pmatrix}\begin{pmatrix}1\\ \Delta\mathbf{q}_k\end{pmatrix}}{\left(1/\mathrm{T}\Delta\mathbf{q}_k^{\mathrm{T}}\right)\begin{pmatrix}1 & \mathbf{0}^{\mathrm{T}}\\ \mathbf{0} & \mathbf{S}_k\end{pmatrix}\begin{pmatrix}1\\ \Delta\mathbf{q}_k\end{pmatrix}}\quad , \tag{1}$$

where the vector $\mathbf{q}_k$ contains the current molecular coordinates, the vector $\Delta\mathbf{q}_k$ gives the correction of the molecular geometry and $\mathbf{g}_k$ and $\mathbf{B}_k$ are the gradient vector and the Hessian matrix (analytic or approximated) of the energy, $E$, at $\mathbf{q}_k$, respectively. The $\mathbf{S}_k$ matrix is a symmetric matrix that has to be specified but normally is taken as the unit matrix $\mathbf{I}$. The matrix that appears in the numerator of Eq. (1) is the AH matrix. The advantage in using the RFO method rather than the standard Newton–Raphson method is due to the fact that the rational approximation by definition represents the energy values as conic isocontours which are not necessarily concentric and similar [24]. This conic behavior of the contours of the approximated energy function guides the $\Delta\mathbf{q}_k$ vector in the direction of the desired stationary point region [24]. The vector $\Delta\mathbf{q}_k$ that extremalizes $\mathrm{q}(\Delta\mathbf{q}_k)$, i.e., $\nabla_{\mathbf{q}}\mathrm{q}(\Delta\mathbf{q}_k)=\mathbf{0}$, is obtained by diagonalization of the AH matrix [20, 21]:

$$\begin{pmatrix}0 & \mathbf{g}_k^{\mathrm{T}}\\ \mathbf{g}_k & \mathbf{B}_k\end{pmatrix}\mathbf{v}_\upsilon^{(k)}=\begin{pmatrix}0 & \mathbf{g}_k^{\mathrm{T}}\\ \mathbf{g}_k & \mathbf{B}_k\end{pmatrix}\begin{pmatrix}\mathrm{v}_{1,\upsilon}^{(k)}\\ \mathbf{v}_\upsilon'^{(k)}\end{pmatrix}$$
$$=\lambda_\upsilon^{(k)}\begin{pmatrix}1 & \mathbf{0}^{\mathrm{T}}\\ \mathbf{0} & \mathbf{S}_k\end{pmatrix}\begin{pmatrix}\upsilon_{1,\upsilon}^{(k)}\\ \mathbf{v}_\upsilon'^{(k)}\end{pmatrix}$$
$$=\lambda_\upsilon^{(k)}\begin{pmatrix}1 & \mathbf{0}^{\mathrm{T}}\\ \mathbf{0} & \mathbf{S}_k\end{pmatrix}\mathbf{v}_\upsilon^{(k)}\quad \forall\,\upsilon=1,\ldots,n+1\quad , \tag{2}$$

where $(\mathbf{v}_\upsilon'^{(k)})^{\mathrm{T}}=(\upsilon_{2,\upsilon}^{(k)},\ldots,\upsilon_{n+1,\upsilon}^{(k)})$. The corresponding displacement vector is evaluated as

$$\Delta\mathbf{q}_k=\frac{1}{\mathrm{v}_{1,\upsilon}^{(k)}}\mathbf{v}_\upsilon'^{(k)}\quad . \tag{3}$$

In Eq. (3) if one is interested in locating a minimum then $\upsilon=1$, and for a transition structure (first-order saddle point) $\upsilon=2$ [20, 21, 25]. Finally, the quadratic variation energy, $Q(\Delta\mathbf{q}_\mathrm{k})$, is evaluated as

$$Q(\Delta\mathbf{q}_k)=\frac{1}{2}\frac{\lambda_\upsilon^{(k)}}{\left(\mathrm{v}_{1,\upsilon}^{(k)}\right)^2}\quad , \tag{4}$$

where $\upsilon$ is selected as explained in Eq. (3). Note that as the optimization process converges, $\mathrm{v}_{1,\upsilon}^{(k)}$ tends to 1 and $\lambda_\upsilon^{(k)}$ to 0.

When we are trying to locate a minimum, the Hessian matrix is updated by using a modified form of the BFGS formula [15]. The most general rank-two update Hessian matrix formula is [15, 29]

$$\mathbf{B}_{\mathrm{k}+1}=\mathbf{B}_0+\sum_{i=0}^{k}\left[\mathbf{j}_i\mathbf{u}_i^{\mathrm{T}}+\mathbf{u}_i\mathbf{j}_i^{\mathrm{T}}-\left(\mathbf{j}_i^{\mathrm{T}}\Delta\mathbf{q}_i\right)\mathbf{u}_i\mathbf{u}_i^{\mathrm{T}}\right]\quad k=0,1,\ldots, \tag{5}$$

where $\mathbf{j}_i=\mathbf{D}_i-\mathbf{A}_i$, $\mathbf{D}_i=\mathbf{g}_{i+1}-\mathbf{g}_i$, $\mathbf{A}_i=\mathbf{B}_i\,\Delta\mathbf{q}_i$, $\mathbf{u}_i=\mathbf{M}_i\,\Delta\mathbf{q}_i/(\Delta\mathbf{q}_i^{\mathrm{T}}\mathbf{M}_i\,\Delta\mathbf{q}_i)$ and $\mathbf{B}_\mathrm{k}$ is the approximated Hessian matrix. $\mathbf{M}_i$ represents a symmetric and positive-definite matrix. Different election of the $\mathbf{M}_\mathrm{i}$ matrix leads to a different update Hessian matrix formula, in particular for the BFGS update, $\mathbf{M}_i=\mathrm{a}_i\mathbf{B}_{i+1}+b_i\mathbf{B}_i$, for some selected positive-definite scalars $a_\mathrm{i}$ and $b_\mathrm{i}$ [30]. Before showing the structure of the modified BFGS formula we remember that $\mathbf{M}_i\Delta\mathbf{q}_i=\ a_i\mathbf{B}_{i+1}\,\Delta\mathbf{q}_i+b_i\mathbf{B}_i\,\Delta\mathbf{q}_i=a_i\mathbf{D}_i+\ b_i\mathbf{A}_i$, where the Newton–Raphson condition has been employed [15, 29]. Now the proposed modified form of the BFGS expression only differs with respect to the normal BFGS in the calculation of the two scalars $a_i$ and $b_i$. In this modified form these two scalars are evaluated as

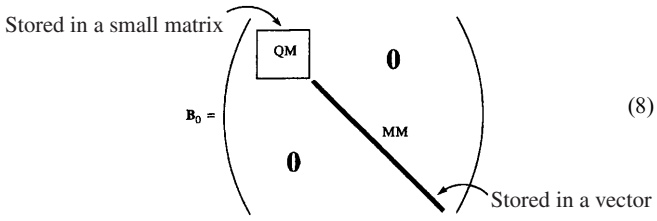
$$a_i=\frac{\left[\mathbf{A}_i^{\mathrm{T}}\mathbf{D}_i\right]^2}{\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{A}_i\right)\left(\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i\right)}\quad , \tag{6}$$

$$b_i=1-a_i\quad . \tag{7}$$

Note that both $a_i$ and $b_i$ are positive quantities. The resulting $\mathbf{B}_{k+1}$ updated Hessian will be positive-definite if both the $\mathbf{B}_k$ matrix and the scalar $\Delta\mathbf{q}_k^T\mathbf{D}_k$ are positive-definite as occurs in the normal BFGS formula.

For location of first-order saddle points the Powell formula is used, where in this case the **M** matrix of Eq. (5) is equal to the unit matrix, **I**.

The guess $\mathbf{B}_0$ matrix is selected in the following way




(8)

The general update formula given in Eq. (5) makes it possible to optimize any type of stationary point even for large molecular systems since the $\mathbf{B}_{k+1}$ matrix is evaluated by using the $\mathbf{B}_0$ matrix and the set of the pair vectors $\{\mathbf{j}_i, \mathbf{u}_i\}_{i=1}^{k}$ and the set of scalars $\{\Delta\mathbf{q}_i^T\mathbf{j}_i\}_{i=1}^{k}$. At each iteration of the optimization process only two vectors, $\mathbf{j}_k$ and $\mathbf{u}_k$, and the scalar $\Delta\mathbf{q}_k^T\mathbf{j}_k$ should be stored. Finally, the resulting $\Delta\mathbf{q}_k$ is scaled by a factor if the step length $|\Delta\mathbf{q}_i^T\Delta\mathbf{q}_i| > R$, where $R$ is the trust radius [15]. The trust radius is modified according to a dynamic algorithm used in Ref. [26].

When L-BFGS is used only a matrix vector product is employed. The starting vector is $v_{1,v}^{(k)} = 1$ and $\mathbf{v}_v^{(k)} = \mathbf{0}$. The matrix product between $\mathbf{B}_k$ and $\mathbf{v}'^{(k)}_v$ is evaluated according to the next formula:

$$\mathbf{B}_k\mathbf{v}'^{(k)}_v = \mathbf{B}_0\mathbf{v}'^{(k)}_v + \sum_{I=0}^{k-1}\left(c_I^{(k)}\mathbf{j}_i + d_i^{(k)}\mathbf{u}_i\right) \quad k = 1, 2, \ldots, \tag{9}$$

where $c_i^{(k)} = \mathbf{u}_i^T\mathbf{v}'^{(k)}_v$ and $d_i^{(k)} = \mathbf{j}_i^T\mathbf{v}'^{(k)}_v - \Delta\mathbf{q}_i^T\mathbf{j}_i\mathbf{u}_i^T\mathbf{v}'^{(k)}_v$.

## 3 Implementation

To obtain the QM/MM potential energy every molecular system was partitioned into a reactive part (the core) treated quantum mechanically with the PM3 [31] or the AM1 [32] semiempirical Hamiltonians and a nonreactive part (the environment) treated by means of molecular mechanics with the AMBER force field [33]. The QM atoms are influenced by the partial charges of the MM atoms, and, in addition, bonding and van der Waals interactions between the two regions are included consistently. We used link atoms to cap exposed valence sites due to bonds which cross the QM/MM boundary. The AMBER 5.0 (Roar-cp module) [34] program was used to carry out the QM/MM calculations. The Roar-cp module is the result of coupling Roar 2.0, i.e., a modified version from the Pennsylvania State University of Sander (the basic energy minimizer and molecular dynamics program of AMBER), and the semiempirical quantum mechanical program MOPAC 7.0 [35].

The search algorithm works in the following way:

1. An initial geometry $\mathbf{q}_0$ in Cartesian coordinates is chosen. The QM region, the MM region and the corresponding link atoms are defined. Then, three environment atoms are always fixed in order to get rid of the translations and rotations of the whole molecular system.
2. The QM/MM energy and the gradient at $\mathbf{q}_0$ are calculated using the Roar-cp module.
3. The guess $\mathbf{B}_0$ Hessian matrix is built up according to Eq. (8). To this aim, we implemented numerical second derivatives in the Roar-cp module as required here. Diagonalization of $\mathbf{B}_0$ allows us to test whether $\mathbf{q}_0$ lies on the suitable quadratic region of the potential-energy surface (i.e., zero or one negative eigenvalue for searching a minimum or a transition state, respectively). If this is not the case, the $\mathbf{B}_0$ matrix is forced to have the convenient number of negative eigenvalues.
4. The RFO method plus a procedure to solve the corresponding secular equations through the full diagonalization of the AH matrix are used in order to obtain the displacement vector at each iteration. This vector is scaled by a factor according to the procedure described in Sect. 2. Then, the new geometry $\mathbf{q}_{k+1}$ is obtained as a result of the current $k$ iteration.
5. If the square root of the gradient norm at $\mathbf{q}_{k+1}$ is smaller than a suitable convergence criterion, it is considered that the corresponding stationary point has been reached and the search ends. Otherwise, the algorithm proceeds to step 6.
6. For minima, the modified BFGS formula (m-BFGS) given in Eqs. (5), (6) and (7) is used to update the Hessian matrix. For transition states the Powell formula is employed. Owing to the update the QM and the MM parts of the resulting $\mathbf{B}_{k+1}$ Hessian matrix become coupled and the MM part is no longer diagonal. Diagonalization of $\mathbf{B}_{k+1}$ allows us to test whether $\mathbf{q}_{k+1}$ lies on the suitable quadratic region of the potential-energy surface. If this is not the case, the $\mathbf{B}_{k+1}$ matrix is forced to have the convenient number of negative eigenvalues. Then the algorithm proceeds to the step 4 to start a new iteration cycle.

All the calculations, except the evaluation of energies and gradients, were carried out with the algorithm just described. Hereafter this algorithm is called RFO-m-BFGS or RFO-Powell when a minimum or a transition state, respectively, is looked for.

## 4 Results and discussion

### *4.1 Description of the systems studied*

To test our algorithm we chose several chemical or biochemical systems, taken from published works, from small to medium size. We ran both geometry minimization and transition-state searches on every molecular system. In what follows we describe (see Fig. 1) the molecular systems we have chose as test systems.

The DHAP system is the deprotonated dihydroxyacetone phosphate (the dihydroxyacetone phosphate is the substrate of triosephosphate isomerase studied by Alagona et al. [36]). The reaction tested is a proton transfer between the hydroxy and the ketone group. DHAP has 14 atoms and we partitioned it into a reactive part of nine atoms treated with the PM3 Hamiltonian and a nonreactive part of five atoms treated with a MM

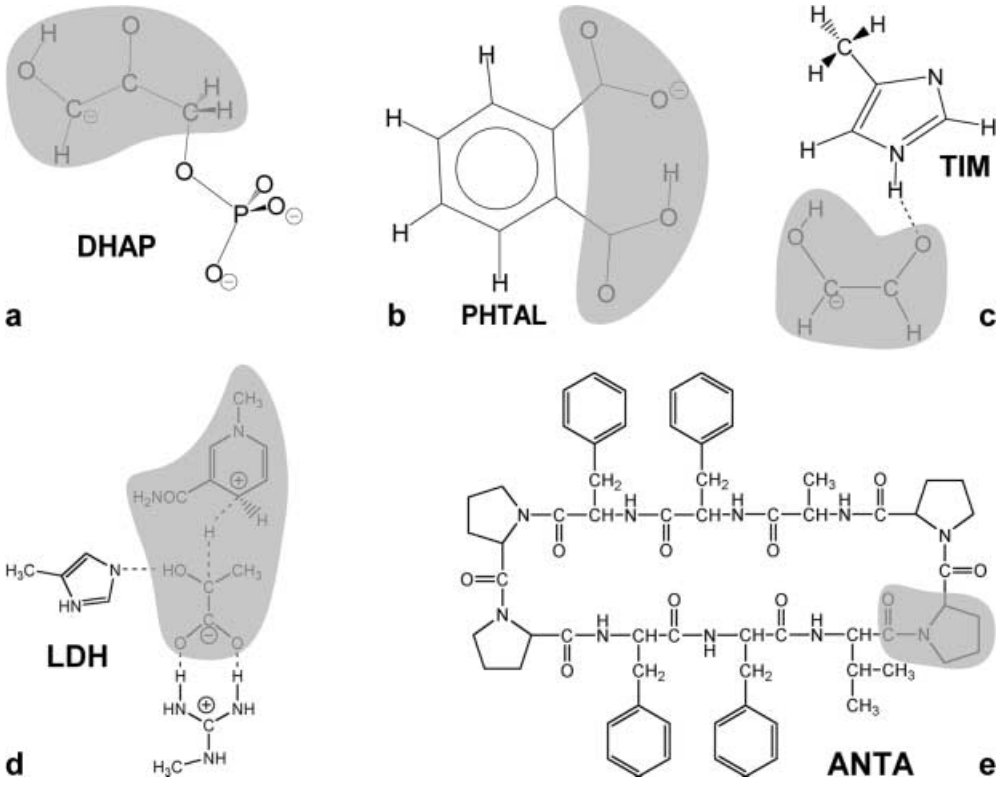



**Fig. 1a–e.** Schematic description of the molecular systems studied in this work. The *shaded zone* in each picture corresponds to the quantum mechanical region

potential. Note that in all the systems pictured in Fig. 1 the shaded zone corresponds to the QM region. A link atom is required for each covalent bond that joins any atom belonging to the QM region with any atom lying in the MM (nonreactive) region. Then, only one link atom has to be included in this case. The overall charge for this system is −3 au, while the total charge of the QM region is −1 au.

The PHTAL system is the phthalate anion. The reaction consists of a proton transfer between the two carboxylic groups. The whole system has 17 atoms. Seven of them, including both carboxylic groups, are treated with the AM1 semiempirical Hamiltonian, and the ten atoms of the phenyl ring are in the MM part. Two link atoms are added. The overall charge for this system is −1 au. The total charge of the QM region is also −1 au.

The TIM system is a model of the active site of triosephosphate isomerase studied by Cui and Karplus [37], where the deprotonated dihydroxyacetone phosphate is emulated by an enediolate. Nineteen atoms are involved here. Seven atoms are in the quantum part treated with the PM3 Hamiltonian, and a singly protonated imidazole ring modeling an histidine, i.e., 12 atoms, defines the MM part. No link atoms are required since there is no covalent bond crossing the QM/MM frontier. The reaction studied is a proton transfer between the two oxygen atoms, similar to that studied in DHAP. The overall charge for this system is −1 au. The total charge of the QM region is also −1 au. In Table 1 TIM1 and TIM2 stand for the reactant and the product of the reaction, respectively.

LDH is a model of the active site of lactate dehydrogenase enzyme studied by Andrés et al. [38]. The whole system is medium-sized and constituted by a total of 55 atoms. It includes a pyruvate and a nicotinamide ring involving 30 atoms in the reactive part treated with PM3, and a guanidino group and a methyl imidazole including 25 atoms for the MM part. As in the TIM system, no link atoms are required. The reaction studied is the transformation of pyruvate to lactate owing to a hydride transfer between pyruvate and nicotinamide. The overall charge for this system is 1 au, while the QM region is neutral. In order to reproduce the system studied by Andrés et al. [38] we also performed an all-QM calculation, including all 55 atoms in the quantum part.

The biggest system tested here is ANTA, which is a decapeptide called antamanide, studied by Fischer and Karplus [39]. In that work the reaction studied was a conformational change in a proline ring. The fact that no bond is broken during this reaction enables us to study the whole system with a MM force field. With the AMBER united atom model a total of 90 atoms are handled. When the whole system is treated classically the Hessian matrix does not have the form shown in Eq. (8), but the full Hessian matrix is built. In addition, a QM/MM partition was also done: the QM part includes the 14 atoms from the all-atom proline ring treated with the PM3 Hamiltonian (this partition takes into account that the frontier cannot be in the peptide bond owing to the fact that it has some double-bond character), whereas the MM part is the rest of the decapeptide involving 83 united atoms. Two link atoms are also included. Both the complete system and the QM region are neutral.

**Table 1.** Comparative results corresponding to the location of the minima for the different systems studied

| System | Number of quantum mechanical and molecular mechanical coordinates | Initial gradient norm (kcal $mol^{-1} \cdot Å^{-1}$) | ΔE (kcal/mol)[a] | qNR-L-BFGS[b] | qNR-BFGS[b] | RFO-m-BFGS(HI)[b] | RFO-m-BFGS(HF)[b] |
|---|---|---|---|---|---|---|---|
| DHAP | 30 + 15 | 10.811 | 11.2 | 96/104 | 34/77 | 46/77 | 10/11 |
| PHTAL | 27 + 30 | 11.739 | 3.48 | 35/41 | 39/83 | 63/108 | 28/43 |
| TIM1 | 21 + 36 | 5.191 | 0.55 | 121/129 | 68/146 | 81/128 | 37/48 |
| TIM2 | 21 + 36 | 5.429 | 0.71 | 123/130 | 56/117 | 90/142 | 57/66 |
| LDH | 90 + 75 | 15.39 | 76.04 | 4101/4196 | 479/965 | 1852/1935 | 1838/1959 |
| LDH | 165 + 0 | 18.47 | 26.86 | 972/1014 | 215/433 | 716/809 | 282/302 |
| ANTA | 0 + 270 | 8.044 | 10.2 | 403/423 | 268/539 | 321/544 | 58/60 |
| ANTA | 48 + 249 | 124.61 | 815.0 | 3215/3340 | 625/1258 | 1949/2054 | 1459/1467 |

[a]Energy difference between the initial geometry and the minimum
[b]Number of steps/number of gradient and energy evaluations

### *4.2 Location of the minima*

Although nowadays several cheap and efficient minimization algorithms exist, we present here the comparative results between our algorithm and two others that are widely used. To this aim we carried out geometry minimizations with our algorithm (RFO-m-BFGS) and two quasi-Newton–Raphson algorithms. These two quasi-Newton–Raphson algorithms use the BFGS and the L-BFGS update Hessian matrix formula. These two algorithms, which were also coupled to the Roar-cp module, are labeled qNR-BFGS and qNR-L-BFGS. The results corresponding to the minimizations with the L-BFGS update formula [18] were obtained using information from the five previous iterations. After several tests this number of iterations was proven to be the best compromise between the efficiency of the method and memory requirements.

The comparative results for the search for the minima are presented in Table 1. We are able to choose the minimization algorithm between qNR-BFGS, qNR-L-BFGS and RFO-m-BFGS. The unity matrix is taken as the initial guess Hessian matrix for qNR-BFGS and qNR-L-BFGS. For the sake of comparison we used two different initial guess Hessian matrices for RFO-m-BFGS. So, RFO-m-BFGS(HI) stands for the RFO-m-BFGS algorithm with the unity matrix as an initial guess, whereas the initial Hessian matrix for the RFO-m-BFGS(HF) algorithm was calculated numerically according to Eq. (8) (except for the ANTA system when the whole system is treated classically).

In each system the minimum reached for all of the algorithms is always the same. So we just specify the energy difference between the starting point and the minimum reached. A numerical Hessian calculation was done in order to characterize the stationary point. The convergence criterion on the root mean square (RMS) of the gradient is $10^{-3}$ kcal/(mol · Å), except for ANTA, which is $10^{-4}$ kcal/(mol · Å).We also present the number of steps and the number of energy and gradient evaluations (the energy and gradient calculations required to build up the numerical initial guess Hessian matrix are not counted). This information will give us the efficiency of every step. Note that the energy and the gradient are calculated only once each step unless the displacement vector needs to be corrected. This is why the number of steps is always smaller than the number of energy and gradient calculations, as seen in Table 1.

No global conclusions about the compared efficiency between the different algorithms can be drawn, we can just note a general tendency because the behavior of an optimization depends not only on the algorithm but also on the intrinsic characteristics of the system (size, starting point, fixed atoms and convergence criteria). Nevertheless it can be seen that qNR-L-BFGS tends to need more steps than the other algorithms. This is due to the fact that it only works with the information of the last five preceding steps, and in this way is able to treat big size systems avoiding the usual memory problems.

Comparing the results for the columns corresponding to the RFO-m-BFGS(HI) and RFO-m-BFGS(HF) algorithms, it is evident that when an initial Hessian matrix is calculated numerically the number of steps and the energy and gradient evaluations required decrease compared to when the starting Hessian matrix is a unity matrix. It can be seen that RFO-m-BFGS(HF) behaves reasonably well in comparison with qNR-BFGS. On the other hand, bearing in mind that the number of energy and gradient evaluations compared to the number of steps indicates the efficiency of the step, it is shown that the efficiency of an RFO-m-BFGS(HF) step is greater than that of a qNR-BFGS step because the ratio between those two numbers for qNR-BFGS is always more than 2, whereas the ratio for the RFO-m-BFGS(HF) algorithm is close to 1.

In addition, we also studied how the RMS of the gradient behaves during the minimization process. Although we present here only the QM/MM ANTA system as an illustrative example (Fig. 2) the comparative results are similar in all the systems studied. It can be seen that the RFO-m-BFGS(HF) algorithm reaches a low-gradient region faster, and it is in this quasi-converged region where it spends most of the steps. This is true even for the cases for which qNR-BFGS needs fewer steps to reach the minimum. The reason why

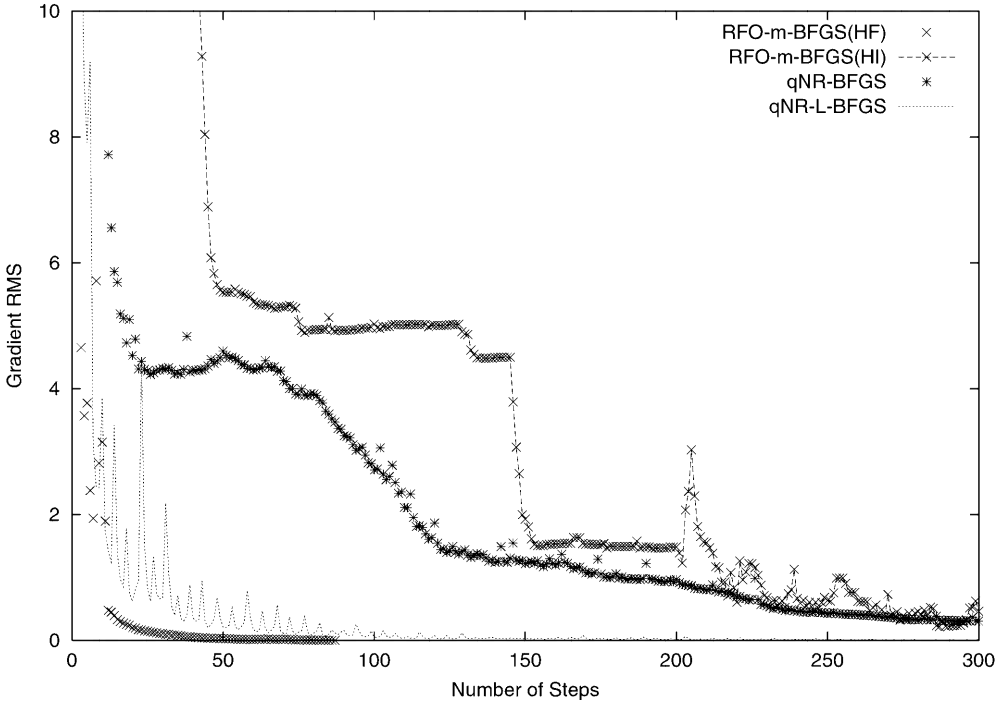



**Fig. 2.** Gradient root mean square (kcal $mol^{-1}\cdot Å^{-1}$), for the quantum mechanical/molecular mechanical ANTA system as a function of the number of steps of minimization with four algorithms

RFO-m-BFGS(HF) reaches a low-gradient zone faster can be due to the information that provides the initially calculated Hessian matrix and the RFO procedure efficiency. The reason why once in a quasi-converged RMS gradient region, RFO-m-BFGS(HF) can sometimes require many steps could be the intrinsic conic behavior of RFO because it does not give the correct shift [20, 21, 22, 23, 24, 26]. Keeping this in mind we think that it would be useful to take advantage of the RFO-m-BFGS(HF) efficient behavior for the initial steps and when a low gradient is reached switching the algorithm to qNR-BFGS or qNR-L-BFGS if a minimization is being performed.

### *4.3 Location of transition states*

We report the test of the RFO-Powell algorithm with the same reactive systems as for the minima in Table 2. We recall that for a transition-state search the BFGS formula cannot be used because in this case the **M** matrix involved in Eq. (5) is not positive-definite [40].

The initial structure for the transition-state search is usually the most energetic point in a few-point scanning along the approximated reaction path. During the search we have to ensure that the algorithm is following the correct direction. Then, at each step the eigenvector of the recently updated Hessian matrix, which has the maximum overlap with the eigenvector of the previous step augmented Hessian matrix associated with the negative eigenvalue ($v=1$), is looked for. Note that, according to Eqs. (2) and (3), the first component of the augmented Hessian eigenvector is not taken into account in order to perform the scalar products. Then, a negative eigenvalue is assigned to the eigenvector of the Hessian matrix found in that way, and the rest of the eigenvalues are forced to be positive. Once the transition state is reached we have characterized the structure found by a numerical calculation of the Hessian matrix.

**Table 2.** Results corresponding to the location of transition states for the different systems studied

| System | Number of quantum mechanical and molecular mechanical coordinates | Initial gradient norm (kcal $mol^{-1}\cdot Å^{-1}$) | $\Delta E$ (kcal/mol)[a] | RFO-Powell (HF)[b] |
|---|---|---|---|---|
| DHAP | 30 + 15 | 6.86 | 22.82 | 56/59 |
| PHTAL | 27 + 30 | 1.18 | 4.03 | 39/54 |
| TIM | 21 + 36 | 11.17 | 3.70 | 65/88 |
| LDH | 165 + 0 | 6.78 | 4.19 | 614/636 |
| LDH | 36 + 129 | 3.24 | 0.87 | 316/392 |
| ANTA | 0 + 270 | 8.04 | 7.78 | 291/315 |
| ANTA | 48 + 249 | 2.10 | 1.25 | 421/598 |

[a]Energy difference between the initial geometry and the transition state
[b]Number of steps/number of gradient and energy evaluations

Overall, the RFO-Powell algorithm performs well in locating transition states. The ratio between the number of steps and the number of gradient and energy evaluations is still close to 1 as previously found during the minimization process. Our implementation allows the location of transition-state structures even if they are quite far away from the starting structures as depicted in Table 2 for the DHAP system (i.e, $\Delta E = 22.82$ kcal/mol). Therefore, we can deduce from the previous results that the RFO-Powell algorithm is a solid algorithm to locate transition-state structures of systems from small to medium size, involving different ratios of QM and MM atoms,

described in Cartesian coordinates, including link atoms and representing several types of chemical reactions.

## 5 Final remarks

We have presented a robust algorithm able to locate mimima and transition-state structures on QM, MM and QM/MM potential-energy surfaces. It is based on a suitable approximation to an initial full Hessian matrix, a modified BFGS formula or a Powell update formula for the location of a minimum or a transition state, respectively and the RFO. This RFO avoids the Hessian matrix inversion required by a quasi-Newton–Raphson method. It also introduces in an automatic way a shift that preserves the current behavior of the optimization process. This algorithm has been successfully tested for a variety of chemical and biochemical systems from small to medium size.

The good behavior of the algorithm presented here has encouraged us to extend it to locate minima and transition states on QM/MM potential-energy surfaces corresponding to real reactive biochemical systems including thousands of atoms. We will modify this algorithm in order to take into account the special problems derived from the large size of those systems, but still handling the information contained in a full Hessian matrix. To this aim, the limited Powell algorithm will be used as update method to locate transition states, and a Lanczos-type algorithm associated with the RFO will be employed to solve the secular equation. This work is now in progress.

*Acknowledgements*. We are indebted to S. Olivella for valuable suggestions. This research was supported by the Spanish DGICYT (grants PB98-0915, PB98-1240-C02-01 and PB98-1240-C02-02).

## References

1. Paizs B, Fogarasi G, Pulay P (1998) J Chem Phys 109: 6571
2. Åqvist J, Warshel A (1993) Chem Rev 93: 2523
3. Monard G, Merz KM (1999) Acc Chem Res 32: 904
4. Schlick T, Overton M (1987) J Comput Chem 8: 1025
5. Derreumaux P, Guihua Z, Schlick T, Brooks B (1994) J Comput Chem 15: 532
6. Eckert F, Pulay P, Werner HJ (1997) J Comput Chem 18: 1473
7. von Armin M, Ahlrichs R (1999) J Chem Phys 111: 9183
8. Farkas O, Schlegel HB (1999) J Chem Phys 111: 10806
9. Anglada JM, Besalú E, Bofill JM, Rubio J (1999) J Math Chem 25: 85
10. Turner AJ, Moliner V, Williams IH (1999) Phys Chem Chem Phys 1: 1323
11. Billeter SR, Turner AJ, Thiel W (2000) Phys Chem Chem Phys 2: 2177
12. Németh K, Coulaud O, Monard G, Ángyán JG (2000) J Chem Phys 113: 5598
13. Paizs B, Baker J, Suhai S, Pulay P (2000) J Chem Phys 113: 6566
14. Wilson EB, Decius JC, Cross PC (1955) Molecular vibrations. McGraw-Hill, New York
15. Fletcher R (1987) Practical methods of optimization. Wiley, New York
16. Buckley A, LeNir A (1983) Math Programming 27: 103
17. Gilbert JC, Lemaréchal C (1989) Math Programming 45: 407
18. Liu DC, Nocedal J (1989) Math Programming 45: 503
19. Powell MJD (1971) Math Programming 1: 26
20. Simons J, Jørgensen P, Taylor H, Ozment J (1983) J Phys Chem 87: 2745
21. Banerjee A, Adams N, Simons J, Shepard R (1985) J Phys Chem 89: 52
22. Cerjan CJ, Miller WH (1981) J Chem Phys 75: 2800
23. Helgaker T (1991) Chem Phys Lett 182: 503
24. Culot P, Dive G, Nguyen VH, Ghuysen JM (1992) Theor Chim Acta 82: 189
25. Khait YG, Panin AY, Averyanov AS (1995) Int J Quantum Chem 54: 329
26. Besalú E, Bofill JM (1998) Theor Chem Acc 100:265
27. (a) Anglada JM, Besalú E, Bofill JM (1999) Theor Chem Acc 103: 163; (b) Anglada JM, Besalú E, Bofill JM (2000) Theor Chem Acc 104: 414
28. Bofill JM, Moreira I de PR, Anglada JM, Illas F (2000) J Comput Chem 21: 1375
29. Dennis JE Jr, Moré JJ (1977) SIAM Rev 19: 46
30. Anglada JM, Bofill JM (1998) J Comput Chem 19: 349
31. Stewart JJP (1989) J Comput Chem 10: 209
32. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107: 3902
33. (a) Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner W (1984) J Am Chem Soc 106: 765; (b) Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Fergusona D M, Spellmeyer D C, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117: 5179
34. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatman TE, DeBolt S, Ferguson D, Seibel G, Kollman PA (1995) Comput Phys Com 91: 1
35. Stewart JJP (1990) QCPE 10: 86
36. Alagona G, Ghio C, Kollman PA (1995) J Am Chem Soc 117: 9855
37. Cui Q, Karplus M (2000) J Chem Phys 112: 1133
38. Andrés J, Moliner V, Krechl J, Silla E (1995) J Chem Soc Perkin Trans 2: 1551
39. Fischer S, Karplus M (1992) Chem Phys Lett 194: 252
40. Bofill JM, Comajuan M (1995) J Comput Chem 16: 1326