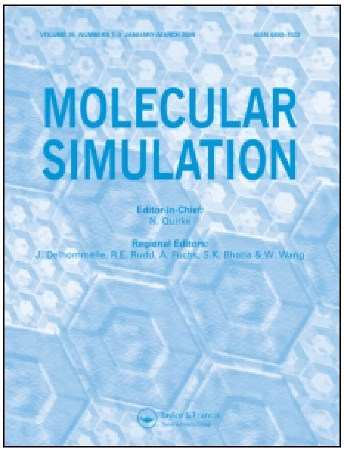

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

On the Simulation of Conformational Transitions: Smoothing Path Energy Minimization Results

Oliver S. Smart[a]; Julia M. Goodfellow[a]
[a] Department of Crystallography, Birkbeck College, University of London, London, England

To cite this Article Smart, Oliver S. and Goodfellow, Julia M.(1995) 'On the Simulation of Conformational Transitions: Smoothing Path Energy Minimization Results', Molecular Simulation, 14: 4, 291 — 302

To link to this Article: DOI: 10.1080/08927029508022024

URL: http://dx.doi.org/10.1080/08927029508022024

Molecular Simulation, 1995, Vol. 14, pp. 291–302
Reprints available directly from the publisher
Photocopying permitted by license only

ON THE SIMULATION OF CONFORMATIONAL TRANSITIONS: SMOOTHING PATH ENERGY MINIMIZATION RESULTS

OLIVER S. SMART and JULIA M. GOODFELLOW

Department of Crystallography, Birkbeck College, University of London, Malet Street, London WCIE 7HX, England

(Received September 1993, accepted July 1994)

Path energy minimization is a method for calculating a quasi-continuous reaction path linking two known conformers of a molecule. Although the method locates the transition state configuration lower energy regions of the path are not optimized. This work examines four methods of following routes downhill from the transition state. Results are compared for the test system of a pucker angle change for the sugar α-D-xylulofuranose. Restrained energy minimization is found to be the superior method. The steepest descents path obtained by this method is shown to be distinct from the route found by adiabatic mapping.

KEY WORDS: Molecular mechanics, sugar conformation, transition state, reaction cooridinate, restrained energy minimization.

1 INTRODUCTION

Conformational transitions play an important role in many biological processes. Examples include: substrate binding to enzymes; allosteric regulation of enzymes and transport proteins and ion translocation through membranes by channels and carriers. The path energy minimization (PEM) method is a new technique which shows an impressive ability to simulate major conformational transformations in large molecules [1,2]. Though the approach deals only with a non-dynamical representation of transitions, such calculations are a necessary preliminary in identifying a suitable reaction co-ordinate for potential of mean force calculations [3,4]. This paper outlines the procedure and compares methods for obtaining a reasonable overall route once PEM has located a transition state configuration.

The PEM method optimizes a quasi-continuous route through configuration space linking two known conformers a molecule or set of molecules. The method requires the specification of molecular configuration vectors $\mathbf{X}_0$ and $\mathbf{X}_{\mathrm{Nmove}+1}$ describing the two conformers, which can be in terms of Cartesian co-ordinates, torsions or some other description. In addition, a function of the vector: $E(\mathbf{X})$, giving reasonable representations for the potential energy of the molecule at $\mathbf{X}_0$ and $\mathbf{X}_{\mathrm{Nmove}+1}$ and intermediate positions in configuration space, must be supplied. The method as currently implemented is limited to simulations of simple conformational transitions because the standard macromolecular potential energy function used cannot adequately describe changes in which bonds are made or broken.

A path is defined as a series of molecular configurations $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \ldots \mathbf{X}_{N_{\text{move}}}\}$ and all points on the line sections through configuration space linking adjacent points $\{\mathbf{X}_0 \rightarrow \mathbf{X}_1, \mathbf{X}_1 \rightarrow \mathbf{X}_2, \ldots, \mathbf{X}_{N_{\text{move}}} \rightarrow \mathbf{X}_{N_{\text{move}}}\}$. The PEM method takes a discrete number (N_{inter}) of equally spaced sample points on each of the line sections. Transition state theory [5-7] states that the most favourable path would have the lowest peak energy. The PEM method finds the most favourable route local to some initial path by adjustment to the moving configurations. The method avoids the problem of the non-differentiability of the maximum function, which would preclude the use of optimization techniques that require the derivative vector to be supplied. This is achieved by noting that for any set of n positive numbers $\{\zeta_1, \zeta_2, \zeta_3 \cdots \zeta_n\}$ the function $\{\sum_{i=1}^{n} \zeta_i^M\}^{1/M}$ tends toward the maximum of $\{\zeta_1, \zeta_2, \zeta_3 \cdots \zeta_n\}$ as M tends to infinity. This expression is used to construct the PEM objective function:

$$S_{\text{PEM}}(\mathbf{X}_1, \mathbf{X}_2, \ldots \mathbf{X}_{N_{\text{move}}}) = \left\{ \sum_{i=0}^{N_{\text{move}}} \sum_{\phi=0}^{N_{\text{inter}}} E'\left(\frac{\{N_{\text{inter}} + 1 - \phi\}\mathbf{X}_i + \phi \mathbf{X}_{i+1}}{N_{\text{inter}} + 1} \right)^M \right\}^{1/M} \tag{1}$$

The function is continuous and tends toward the peak energy of the path as the number of M is increased. Importantly, the derivative vector of the function can be expressed analytically in terms of the molecular potential energies and force vectors at all sampled points on the path. The potential energies in the objective function must be offset so that they are greater than zero.

The objective function is dominated by the high energy regions of the path. On the application of minization, the procedure converges to the transition state local to the starting configuration. In a simulation of a large scale conformational transition of the ion channel forming peptide gramicidin A the procedure shows an impressive radius of convergence [2].

In addition, as the configurations immediately adjacent to the transition state contribute to the objective function, the optimal vector through the transition state will also be found. However, as the lower energy regions of the path make no effective contribution to the function these regions will not be optimized and will contain artefacts.

2 TEST SYSTEM: PUCKER ANGLE CHANGE IN α-D-XYLULOFURANOSE

The displacement of the atoms from the average plane in a five-membered ring can be completely described in terms of two parameters, a pucker angle and a displacement [8]. Using this description it is possible to construct adiabatic potential energy maps for pucker parameter changes in sugars with a furanose ring [9].

The system provides an excellent opportunity to assess the results of different procedures on a non-trivial system with a low dimensionality. The pentulofuranose sugar α-D-xylulofuranose was modelled in an arbitrary configuration and subjected to energy minimization, using the Polak-Ribere conjugate gradients procedure [10–12], to yield the configuration shown as Figure 1. The molecule was represented by the AMBER united atom energy function [13] as implemented in the program TIC [14].

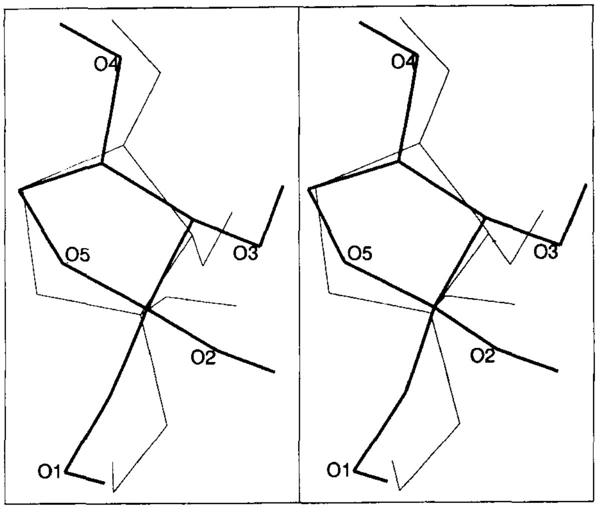

Figure 1 The two configurations of α-D-xylulofuranose used to test procedures: thick lines show the lower energy minimum (pucker angle 224°) and thin lines mark the other energy minimum (pucker angle 393.5°).

A dielectric constant of 3.0 was used and no non-bonded cutoff was applied. This and all subsequent runs were conducted with a Cartesian co-ordinate representation of the molecule.

To provide a bench mark against which to compare PEM and other techniques the "adiabatic surface" against pucker angle [8] was produced. This is the route produced by the "reaction co-ordinate" or "adiabatic mapping" method [15]. A variable or function of the variables judged to be important in the transition of interest (in this case the pucker angle) is controlled and energy minimization is applied changing all other variables. By judicious control of the "reaction co-ordinate" it is hoped that the transition of interest can be provoked. Often two co-ordinates are used and a two dimensional contour map is built up. The results are referred to as the adiabatic surface as the method gives a reasonable approximation to the potential of mean force (free energy profile) along the reaction co-ordinate if the time scale for the motion along the co-ordinate is much larger than for the other variables [15]. The method is widely used [9, 14–18], despite being prone to failure: even in the case of a two dimensional model function [19].

In this case, a restraint on the Cremer-Pople pucker angle was added to the molecular potential energy function:

$$E_{\mathrm{rest}} = K_\theta(\theta - \theta_{\mathrm{rest}})^2 \qquad (2)$$

where θ_{rest} is the target pucker angle and K_θ is the restraint constant (set to 50 kcal/(mol.rad^2)). Starting from the energy minimum configuration found above, which has a pucker angle of 224° (between the ^2E and $^2T_3$ forms) the target pucker angle was adjusted in 5° steps, through a full range of 360°. After each change energy minimization was applied. No large discontinuities in atomic positions were found along the path. The procedure results in a potential energy profile versus pucker angle (shown in Figure 2a). The profile shows a second, higher energy, minima with a pucker

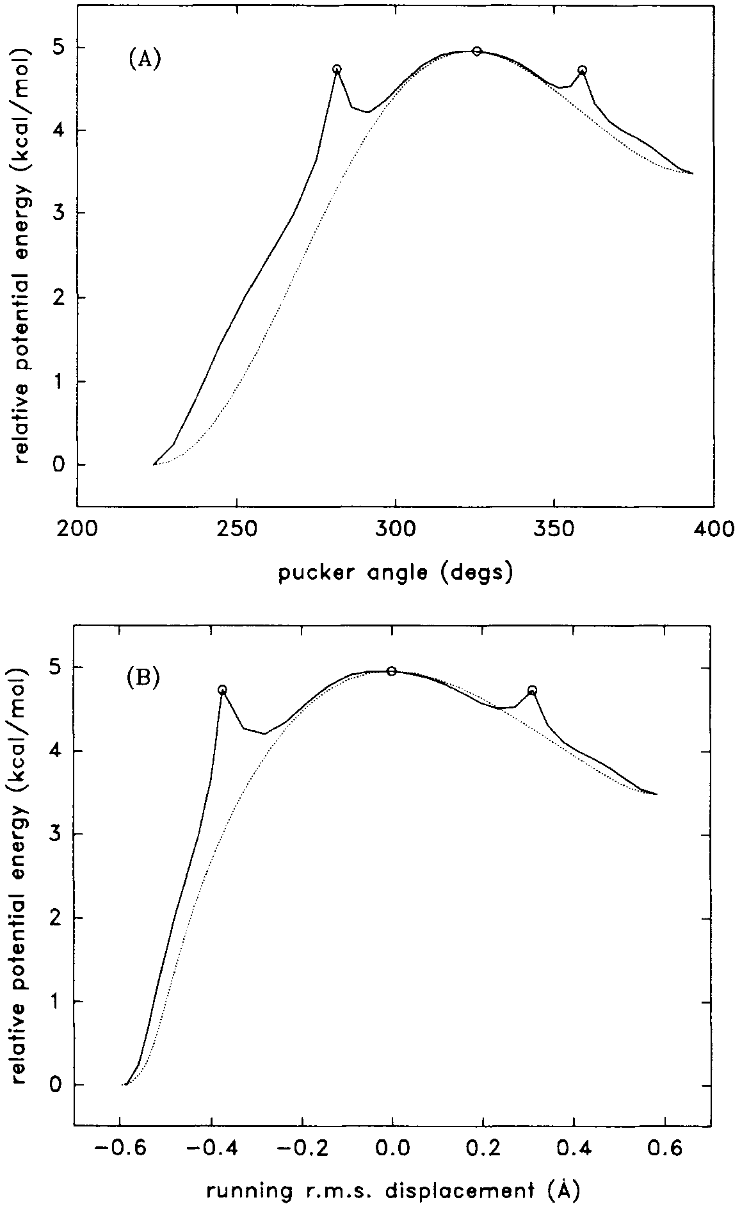

Figure 2 Comparison between the PEM and adiabatic mapping methods in obtaining routes for a pucker angle transition in α-D-xylulofuranose. The dotted lines mark the adiabatic result whereas the solid lines show the PEM path. The moving configurations of the PEM result are marked by circles. (A) shows the results versus pucker angle and (B) the results in terms of the running r.m.s. displacement. This is defined as the running total of the rms atomic displacement between adjacent positions along the path [18].

angle of 393.5°, shown in Figure 1. Two transition states were also found, at 324°, with a potential energy of 4.96 kcal/mol relative to the start, and another with a higher relative potential energy at 124° (this part of the profile is omitted from Figure 2).

The PEM technique was used to find paths which link the two minima (with no further information taken from the adiabatic mapping). As the full results of the trials are to be published elsewhere [2] only a summary is presented here. The technique found a path through the lower energy transition state located by the adiabatic mapping method, starting from a set of three moving configurations taken from the line through Cartesian co-ordinate space linking the two minima (Figure 2). Three intermediate sample configurations were used on each line section of the path ($N_{\text{inter}} = 3$). The objective function was minized using the Polak-Ribere conjugate gradients technique [10–12]. At the end of the optimization procedure a check was made as to whether the sampling density was high enough by taking an additional sampled point at the mid-point of each interval used in the minimization (this is simply achieved by making a function evaluation with N_{inter} set to $2N_{\text{inter}} + 1$). The use of more moving configurations or sampled configurations was found to have no effect on the peak energy of the resulting path.

Results were not found to be critically dependent on the exact value taken for the power (M) in the objective function. A value of 100 was found to be reasonable. Lower values resulted in a decrease of the average of the path at the cost of the peak energy. In contrast, higher powers resulted in an increase in the average energy with minimal decrease in the peak and furthermore entailed an increase in computational cost as the energies of some positions had to be recalculated to avoid numerical overflows.

As mentioned above, because the lower energy parts of the path contribute little to the objective function artefacts appear in these regions (figure 2). We shall now examine four methods for obtaining a smooth overall path.

3 METHODS

3.1 Steepest Descents

In this procedure the steepest descents path is found by taking small steps (1 = 0.0003 Å) along a unit vector in the direction of the force:

$$\mathbf{X}_{i+1} = \mathbf{X}_i + 1\hat{\mathbf{F}}(\mathbf{X}_i) \qquad (3)$$

The minimization is continued until the norm of the force falls below a preset value. The overall path was found in two sections: one run was started from the peak energy position of the final PEM result and in this run the energy fell to a minimum at 393.5°. The other was made from the intermediate configuration adjacent to the peak energy position which falls to the other minimum. The results of the procedures are as figure 3. It is interesting to note that the steepest descents path is distinct from the adiabatic result, with the former profile lying above the adiabatic when plotted against pucker angle but below when running rms displacement (a measure of the distance along the path) is used. The implications of this are discussed below.

3.2 *Restrained Energy Minimization*

In this procedure energy minimization is applied from the peak energy configuration or an adjacent configuration. A series of positions with descending energy is produced by adding a semi-harmonic restraint on the root mean squared displacement (r_{ms}) of the current configuration from the previously output position to the potential energy function.

$$V_{rest} = \begin{matrix} 0 & \text{if} & r_{ms} < r_{target} \\ K_r(r_{ms} - r_{target})^2 & \text{otherwise} & \end{matrix} \tag{4}$$

The function imposes no penalty if the rms displacement from the previous position is less than the target displacement (r_{target}) thus allowing free energy minimization until the barrier is reached. Minimization was performed using the Polak-Ribere conjugate gradients procedure [10–12]. At the end of minimization a new cycle is started by setting the "previous" configuration to the current position. A target rms displacement between positions of 0.02 Å ws used with a restraint constant K_r of 1000 kcal/(mol. Å^2). As in the steepest descents trial two runs were required to obtain the overall route for the transition.

Figure 3 shows that the results of restrained energy minimization and steepest descents are practically identical for this small system. Restrained energy minimization is the superior method for the larger systems as it is based on a more powerful minimization technique: the procedure can converge to the energy minimum whereas the steepest descents method displays very slow progress as the minimum is approached.

3.3 *Misset Angle Optimization*

This procedure differs from the previous two in that it attempts to find the steepest descents path from the transition state in one run rather than following the descent of a single configuration of the molecule using energy minimization. The procedure adapts the PEM method. The fixed end points of the path are the transiton state configuration and one or other of the energy minima. A restraint term is added to the PEM objective function (equation 1):

$$S_{total}(\mathbf{X}_1, \ldots \mathbf{X}_{N_{move}}) = S_{PEM}(\mathbf{X}_1, \ldots \mathbf{X}_{N_{move}}) + S_{rest}(\mathbf{X}_1, \ldots \mathbf{X}_{N_{move}}) \tag{5}$$

The restraint is on the (misset) angle between the force and the vector along the path at each point considered on the path:

$$S_{rest} = \sum K_\gamma \gamma^2 \tag{6}$$

The sum extends across all the configurations sampled in equation (1). K_γ is a constant and γ is the misset angle of each configuration (**R**). The misset angle is found for each position:

$$\gamma = \cos^{-1}\{\hat{\mathbf{e}}(\mathbf{R}) \cdot \hat{\mathbf{F}}(\mathbf{R})\} \tag{7}$$

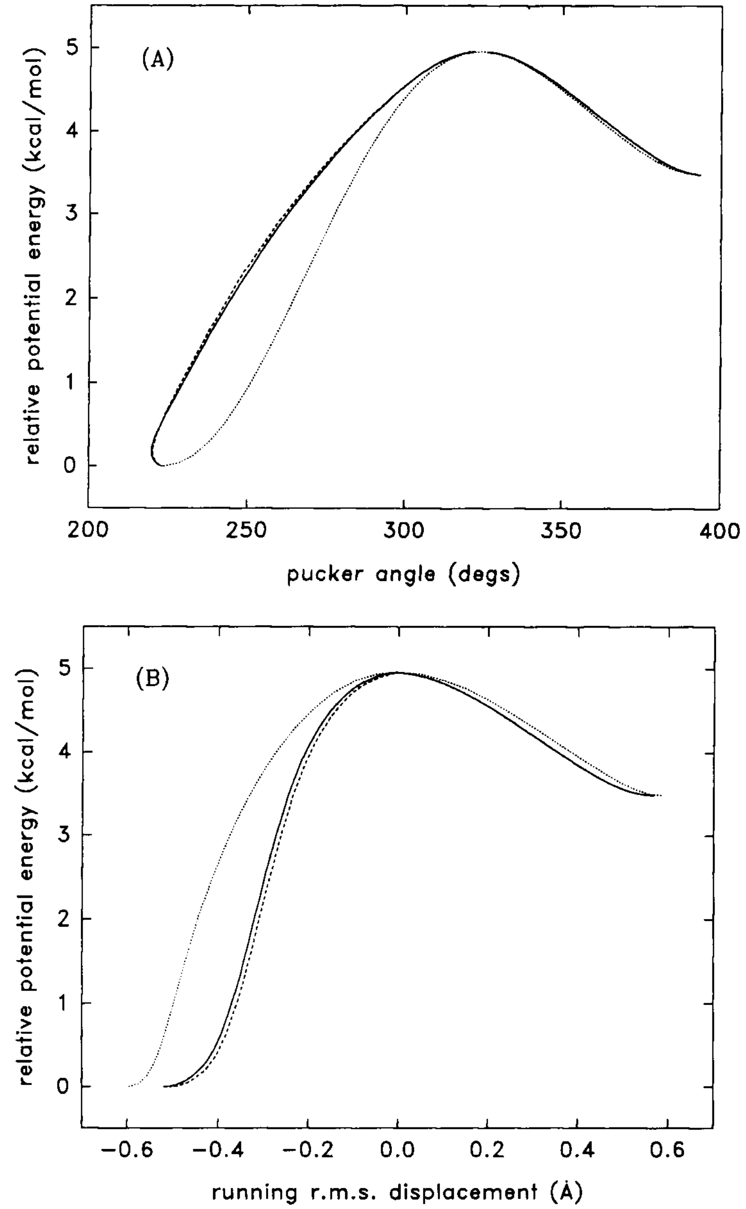

Figure 3 The results of the steepest descent and restrained energy minimization procedures for finding routes between the transition state and energy minima. The solid lines mark the steepest descents result, dashed lines restrained energy minimization and for comparison the adiabatic result is shown with a dotted line. (B) shows that against another reaction co-ordinate the adiabatic result does not have the lowest energy profile.

where **F** (**R**) is the force vector at the configuration, **e** is the vector giving the path direction at the point and the circumflex denotes a unit vector.

A major problem with this procedure is that the restraint involves the force functions (the partial derivatives of the energy). This means that the derivative of the objective function can only be analytically calculated in terms of the second derivatives of the energy. As the program TIC [14] does not include routines to calculate these, the derivative vector of the objective function was found numerically. This proved to extremely computationally costly and only one trial of the procedure was performed.

The trial aimed at finding a route between the lower energy minimum and the transition state. Three moving positions were taken from the "raw" PEM result (figure 2). One intermediate was used in each line section and the restraint constant K_γ was set to 0.5 kcal/(mol. rad^2) (the value results in the initial contribution of the restraint term being comparable to the PEM part). The result of the procedure after partial optimization is shown as Figure 4.

Though the procedure produces a reasonably smooth result it is some way from the steepest descents path (Figure 4). The method was so slow in comparison to restrained energy minimization that further trials were not performed.

The procedure could be made faster using analytical routines for calculate the derivative vector. A better approach may be to adapt the Locally Updated Planes method developed by Elber and co-workers (20, 21]. This procedure finds the steepest descents path for a transition between two fixed positions by minimizing the energies of a series of moving configurations with the constraint that the force vector at each point should be parallel to the path vector. The couching of the requirement as a constraint as opposed to a restraint avoids the need to calculate the second derivatives of the energy function [21]. A weakness of the method is a relatively small radius of convergence. It may be possible to add such constraints to PEM, resulting in a procedure with a high radius of convergence, which will locate the steepest descents path through the transition state in one run.

3.4 PEM Descents

This technique uses the PEM method to find a route downhill either side of the transition state. As noted above, points immediately adjacent to the peak energy position contribute to the PEM objective function. The PEM descents method uses this fact. Only one fixed configuration and one moving configuration are considered (linked by a number of intermediates). The fixed position is initially set to the transition state configuration and the moving position to some point on one or other side of the transition state. The PEM procedure is then used to optimize the line section between the moving and fixed configurations. The highest energy intermediate configuration is "accepted", output, and used to replace the fixed configuration. The process is then restarted. On each cycle the potential energy of the accepted position goes down, in this way PEM is used to optimize a series of positions down to the energy minimum. The results of the procedure on the pucker angle test system are shown as Figure 5. As can be seen the path is close to the adiabatic result and quite distinct from the steepest descents path.

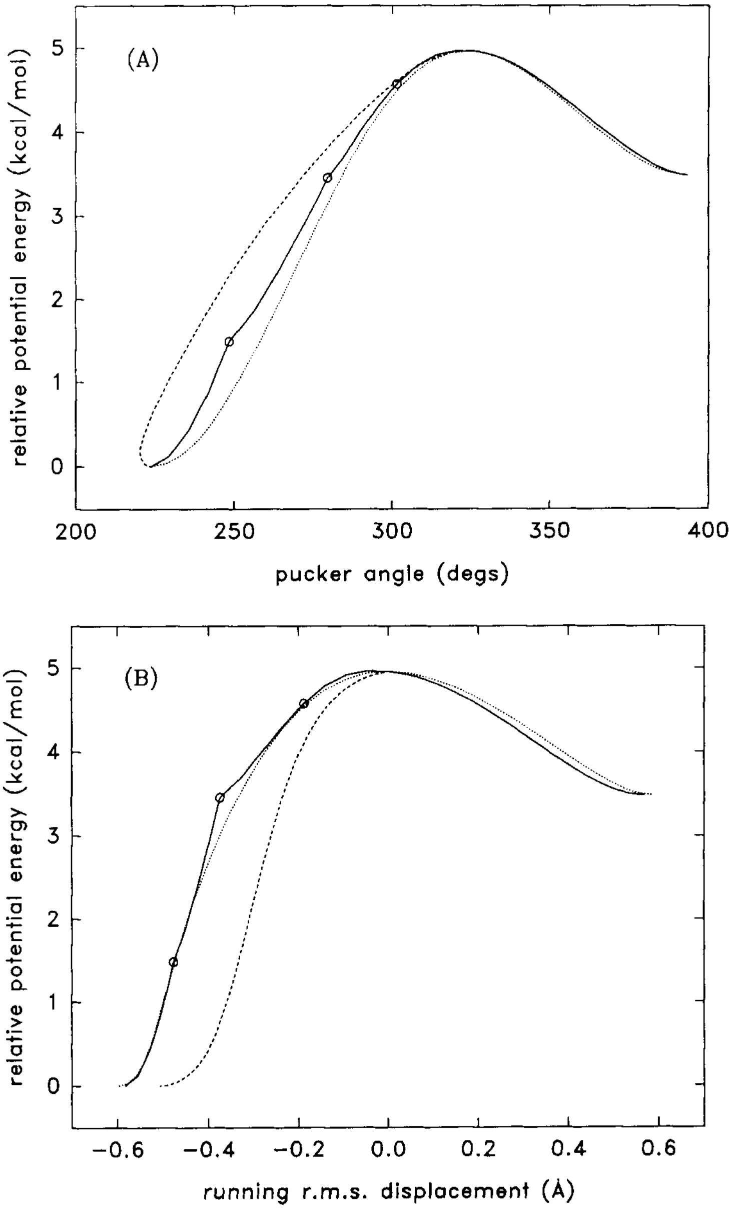

Figure 4 The results of the misset angle optimization procedure are shown with solid lines. For comparison the restrained energy minimization path is marked with a dashed line and the adiabatic result with a dotted line. Moving configurations are marked by circles.

4 DISCUSSION AND CONCLUSIONS

It is interesting to note that even for the small (14 atom) model system used here the steepest descent and adiabatic paths are quite distinct. On reflection, this is not that

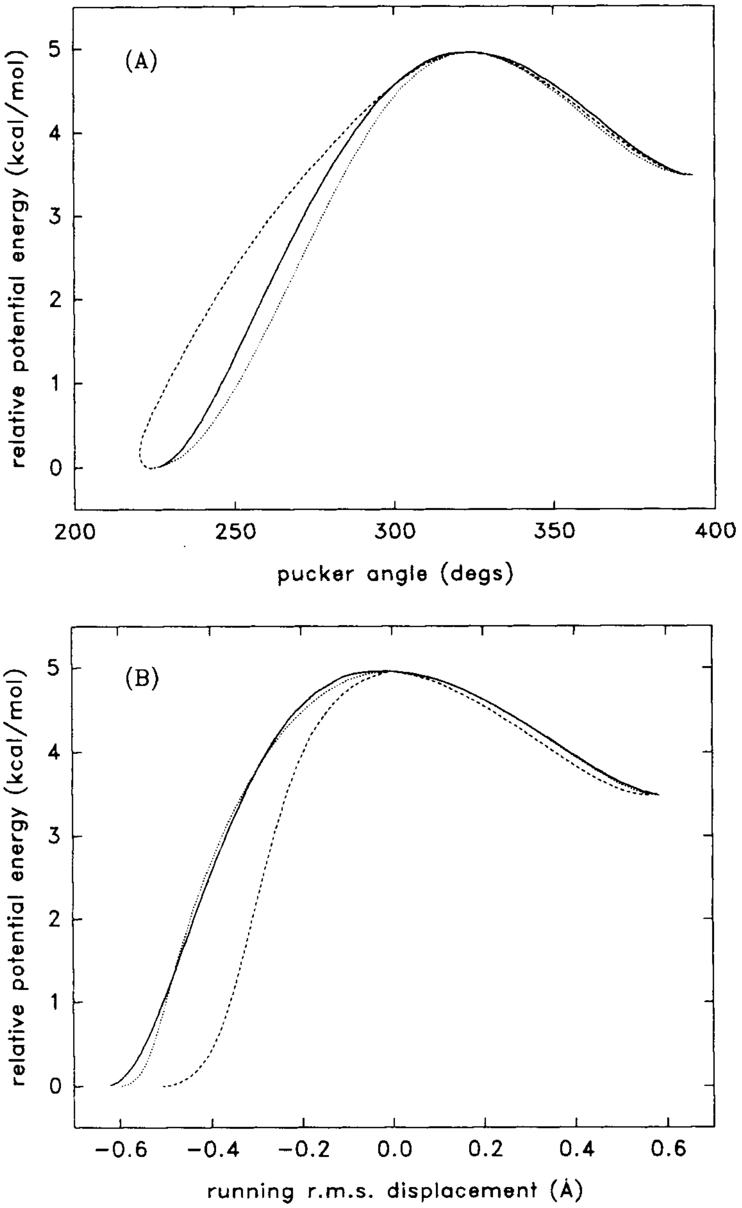

Figure 5 The PEM descents technique. Results are shown with solid lines. For comparison the restrained energy minimization path is marked with a dashed line and the adiabatic result with a dotted line.

surprising: the adiabatic result involves freezing out a single variable (or in this case a single function of the variables) and optimizing the others. This results in the chosen variable being treated in a special manner; the steepest descents path makes no such distinction treating all variables equally. It is important to remember that in the case of two dimensional adiabatic mapping the steepest descents path on the potential energy contour surface may well be distinct from the real steepest descents path.

The question arises as to which the paths is "best". Conventional opinion on this matter suggest that the optimal route is the steepest descents path [19]. We believe that this arises from the definition of the "intrinsic reaction co-ordinate" by Fukui [22]. Quoting directly: "We tentatively define the reaction coordinate as a curve passing through the initial and transition points and orthogonal to the energy equipotential contour surface." The crucial point is that this a tentative definition. Transition State Theory [5–7] states the important features of any path is its peak energy. In this respect every path considered here, even the "raw" PEM result are equally good. Furthermore, as dynamics crucially effect the behaviour of real systems, it is important to note that any of the smooth paths found here could be used to form a suitable reaction co-ordinate for calculations to obtain the potential of mean force [3,4]. This being said, the steepest descents path through the transition state is conceptually simple and easy to find: restrained energy minimization represents the method of choice for calculating it.

Acknowledgements

This work is supported by the UK Science and Engineering Research Council under project grant GR/G49494 and the Molecular Recognition and Computational Science Initiatives.

References

[1] O.S. Smart, "A new method to calculate reaction paths for conformational transitions of large molecules", *Chem. Phys. Letts.* **222**, 405 (1994).

[2] O.S. Smart, "Path Energy Minimization: A new method for the simulation of conformational transitions of large molecules", in *Computer Simulation in Molecular Biology*, J.M. Goodfellow ed., VCH publications, Weinham, Chapter 8, In the press.

[3] R. Elber, "Calculation of the potential of mean force using molecular dynamics with linear constraints: An application to a conformational transition in a solvated dipeptide", *J. Chem. Phys.* **93**, 4312 (1990).

[4] G. Verkhivker, R. Elber and Q.H. Gibson, "Microscopic modeling of ligand diffusion through the protein leghemoglobin-computer-simulations and experiments", *J. Am. Chem. Soc.*, **114**, 7866 (1992).

[5] H. Erying, "The activated complex in chemical reactions", *J. Chem. Phys.* **3**, 107 (1935).

[6] M.G. Evans and M. Polanyi, "Some applications of the transition state method to the calculation of velocity rates especially in solution", *Trans. Faraday Soc.* **31**, 875 (1935).

[7] K.J. Laidler, *Chemical Kinetics*, Third edition, Harper & Row, New York, 1987.

[8] D. Cremer and J.A. Pople, "A general definition of ring puckering coordinates", *J. Am. Chem. Soc.*, **97**, 1354 (1975).

[9] A.D. French and V. Tran, "Analysis of fructofuranose conformations by molecular mechanics", *Biopolymers*, **29**, 1599 (1990).

[10] R. Fletcher and C.M. Reeves, "Functional minimization by conjugate gradients", *Computer J.* **7**, 149 (1964).

[11] E. Polak, *Computational Methods in Optimization*, Academic Press, New York, 1971.

[12] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T Vetterling, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1986, ch. 10.

[13] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghilo, G. Alagona, S. Jr. Profeta, and P.Weiner, "A new force field for molecular mechanical simulation of nucleic acids and proteins", *J. Am. Chem. Soc.*, **106**, 765 (1984).

[14] O.S. Smart, *Simulation of a conformational rearrangement of the substrate in* D*-xylose isomerase*, Ph.D. thesis, University of London, London, 1991.

[15] J.A. McCammon and S.C. Harvey, "Dynamics of Proteins and Nucleic Acids", Cambridge University Press, Cambridge, 1987, section 4.5.

[16] B.R. Gelin and M. Karplus, "Sidechain torsional potential and motion of amino acids in proteins: Bovine Pancreatic Trypsin Inhibitor", *Proc. Nat. Acad. Sci. U.S.A.*, **72**, 2002 (1975).

[17] S.N. Ha, L.J. Madsen and J.W. Brady, "Conformational analysis and molecular dynamics simulations of maltose", *Biopolymers*, **27**, 1927 (1988).

[18] O.S. Smart, J. Akins and D.M. Blow, "Molecular mechanics simulations of a conformational reaarangement of D-xylose in the active site of D-xylose isomerase", *Proteins* **13**, 100 (1992).

[19] K. Müller, "Reaction paths on multidimensional energy hypersurfaces", *Ang. Chem.* **19**, 1 (1980).

[20] A. Utilsky and R. Elber, "A new technique to calculate steepest descents paths in flexible polyatomic systems", *J. Chem. Phys.*, **92**, 1500 (1990).

[21] C. Choi and R. Elber, "Reaction path study of helix formation in tetrapeptides: Effect of side chains", *J. Chem. Phys.*, **94**, 751 (1991).

[22] K. Fukui, "A formulation of the reaction coordinate", *J. Phys. Chem.*, **74**, 4161 (1970).