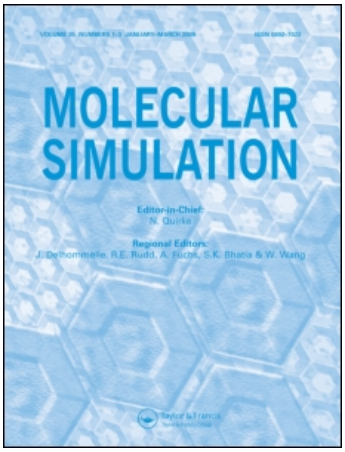

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

A pragmatic method for location of transition states and calculation of reaction paths

Ian Dance[a]
[a] School of Chemistry, University of New South Wales, Sydney, Australia

To cite this Article Dance, Ian(2008) 'A pragmatic method for location of transition states and calculation of reaction paths', Molecular Simulation, 34: 10, 923 — 929
To link to this Article: DOI: 10.1080/08927020802175258
URL: http://dx.doi.org/10.1080/08927020802175258

Molecular Simulation
Vol. 34, Nos. 10–15, September–December 2008, 923–929

Taylor & Francis
Taylor & Francis Group

A pragmatic method for location of transition states and calculation of reaction paths

Ian Dance*

School of Chemistry, University of New South Wales, Sydney, Australia

(*Received 12 February 2008; final version received 1 May 2008*)

An alternative method for the location of transition states using density functional theory is described. The user controlled and managed method, dubbed 'pragmatic', is based on iterations of small-step small-extent energy reductions from trial structures that necessarily oscillate across the potential energy barrier. This automatically optimises variables conjugate to the reaction path, and allows the user to converge trials to the transition state geometry. The method does not require prior knowledge of reactant or product structures, and is valuable for the exploration of complex potential energy surfaces and of unknown mechanisms, and the discovery of intermediates. The illustrations provided are from the complex mechanism in which the conversion of N_2 to NH_3 is catalysed at the FeMo-cofactor of the enzyme nitrogenase. The pragmatic method is successful for reaction paths where geometrical changes are asynchronous, and where the automated linear synchronous transit and quadratic synchronous transit methods are unreliable.

Keywords: transition state; reaction path; density functional; LST/QST

1. Introduction

This paper describes a density functional method developed to locate transition states (TS) for reaction steps in a proposed mechanism for the catalysed conversion of N_2 to NH_3, as effected by the metal sulphide cluster at the active site of the enzyme nitrogenase. This catalyst system and its reactions are neither simple nor straightforward, and conventional procedures sometimes fail (or yield dubious results), and therefore I have deployed a pragmatic method for investigation of reaction profiles.

The catalytic site of nitrogenase is the FeMo-cofactor (FeMo-co, Figure 1), a $Fe_7MoS_9N^c$ metal sulphide cluster, coordinated at Mo by homocitrate, and linked to protein via two amino acids, histidine (coordinated to Mo) and cysteine (coordinated to Fe1) [1]. Aspects of the coordination chemistry of FeMo-co [2] relevant to the discussion here are: (i) the Fe—N^c distances are variable (2.0 to >3 Å); (ii) coordination at these Fe atoms can occur in the *exo* position (*trans* to the Fe—N^c vector) or the *endo* position (*cis* to the Fe—N^c vector) or both; (iii) the equatorial doubly-bridging S atoms (e.g. S2B) can flap around their Fe—Fe vector; (iv) triply-bridging S atoms (e.g. S3B) can lengthen a S—Fe bond to become doubly-bridging, particularly when bearing an H atom. The catalysis occurs at the Fe2, Fe3, Fe6, Fe7 face of FeMo-co [3], and is believed to involve coordination at Fe6 and Fe2. The mechanism considered here involves H atoms bound to S or Fe transferring to bound N_2 and N_2H_x intermediates; the reaction steps are intramolecular [4].

The electronic structure of FeMo-co is complex, as expected for a cluster containing eight transition metals. In addition, the rich vibrational structure of FeMo-co and its ligated intermediates is complicated by the ability of some Fe—N^c and Fe—S bonds to elongate with small energy cost. The method of eigenvalue-following to determine a transition state, which is valuable for simple and straightforward reactions, is expected to be difficult, ambiguous, time-consuming and ineffective for many of the FeMo-co reactions.

The linear synchronous transit (LST) and quadratic synchronous transit (QST) methods for determination of transition state and reaction path are applicable to FeMo-co catalytic steps. LST uses linear interpolation between reactant and product, and conjugate gradient optimisation to obtain a first approximation to the TS, which is then refined using QST which alternates searches for an energy maximum with constrained conjugate gradient minimisations [5,6]. These methods have been investigated for FeMo-co reactions, but have sometimes failed, with the procedure sliding off the saddle towards reactant or product [6].

The background for the chemical investigation was an unknown mechanism for catalysis, necessarily with multiple individual steps (at least one N_2 bond broken and six N—H bonds formed, together with making and breaking bonds with the catalyst), with unknown intermediates and structures, and with multiple hypotheses. In addition there are several probable electronic states for FeMo-co and its derivatives, and a number of electronically

*Email: i.dance@unsw.edu.au

ISSN 0892-7022 print/ISSN 1029-0435 online

DOI: 10.1080/08927020802175258
http://www.informaworld.com

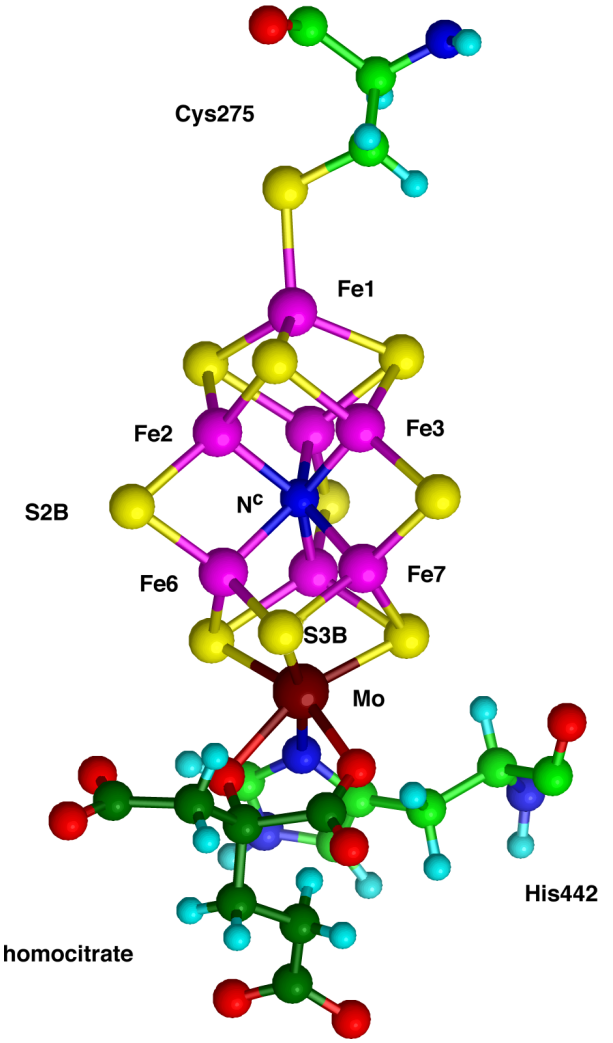

Figure 1. The FeMo-cofactor of the enzyme nitrogenase (atom labels for *Azotobacter vinelandii*, Protein Database crystal structure 1M1N). N^c is the atom at the centre of the Fe_7MoS_9 cluster: homocitrate C atoms are dark green. Catalysis occurs at the Fe2, Fe3, Fe6 and Fe7 face.

different potential energy surfaces to explore, and so reliable density functional calculation of electronic and geometrical structure is required. In this context, in order to prosecute the investigation it was necessary to have a reliable and efficient procedure for determination of TS and exploration of reaction paths. The method developed, dubbed 'pragmatic', is described here.

2. Methodology

The density functional engine is DMol[3] [7], (v4.2) using the standard BLYP functional and the standard double numerical basis sets including polarisation. Each calculation used four processors of an SGI Altix 3700 Bx2 cluster. Calculations are all electron, and spin unrestricted, and use a fine grid with no symmetry. During the SCF calculations the charge mixing between iterations is 0.05, and the spin mixing is normally 0.5. Validation calculations, with reference to experimental geometry and energy data for FeMo-co and related systems, indicate that errors are <0.05 Å in geometry, and <2 kcal mol^{-1} in energy [8]. The model used to calculate FeMo-co is shown in Figure 2, in which 442^{His} is truncated to imidazole, 275^{Cys} is truncated to SCH_3, and homocitrate is truncated to glycolate, $^{-}OCH_2COO^{-}$; this retains the native coordination of all metal atoms. The resting molecular oxidation state is defined by the net charge of -3 [9].

The electronic structure of FeMo-co has *c.* 20 molecular orbitals within 1 eV of the highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbitals: the HOMO–LUMO gap is usually *c.* 0.4 eV. The ground and low-lying electronic states of FeMo-co have been investigated, and are most readily characterised and monitored by the combinations of spins on the individual metal atoms. In resting FeMo-co the electronic state with the lowest energy (and the experimentally required smallest dispersion of Fe—N^c distances) has Fe spin densities ranging in magnitude from 2.7 to 3.2*e* for the seven Fe atoms, and Mo spin 0.3*e*. Additional ligation of any Fe atom in reaction intermediates usually reduces the magnitude of its spin [4]. Control of the electronic state of FeMo-co and its derivatives is effected in DMol by prescribing spin densities (typically $\pm 2.8e$ at Fe without additional ligation) at the beginning of the SCF calculation (keyword Start_Spin_Populations), allowing their optimisation, and checking the complete set of spins on completion. This control is maintained for all calculations

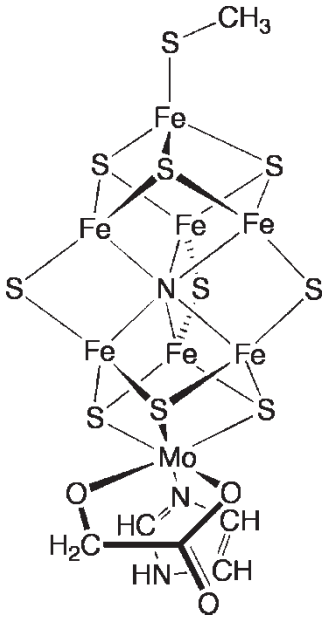

Figure 2. The calculational model for FeMo-co.

of transition state and reaction path, to ensure that a single potential energy surface is being followed. In the following the term 'energy reduction' is used to describe a partial incomplete energy minimisation.

The method for finding the lowest energy saddle point with zero gradient between two local minima (called reactant and product) is the following. A trial structure (1) for the transition state is generated, incorporating any knowledge based on chemical principle and/or precedent. Using energy reduction, limited by small geometrical steps (< 0.05 Å), the downhill direction is observed and followed until any large initial gradients settle. Then, the structure at this point is manually modified to a point (2) expected to be just on the opposite side of the barrier, and small-step energy reduction is again undertaken. If the postulated geometry changes towards the opposite limit, the small-step energy reduction is continued until any large initial gradients settle, but if the geometry change is in the same direction as that previous, the trial geometry (2) is adjusted until the change is towards the opposite limit. Then, this sequence of a small-step energy reduction on one side of the barrier followed by a small-step energy reduction on the opposite side is continued recurrently, as illustrated in Figure 3, in which the solid arrows represent the energy reductions and the broken arrows are the user's geometry adjustments.

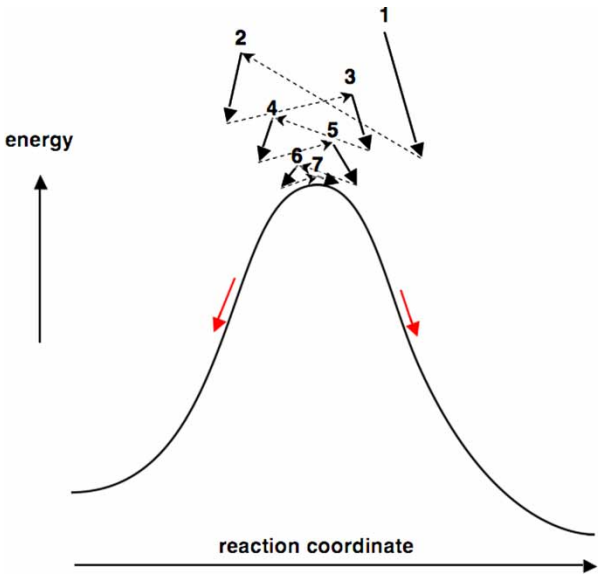

Figure 3. Generalised 2D diagram of the procedure for locating a reaction transition state. The solid arrows represent the course of limited small-step energy reductions from the numbered positions, which are generated (broken arrows) from the previous results on geometry, energy and gradient. The postulated trials necessarily oscillate across the barrier, and therefore optimise all other variables and decrease the energy while approaching the top of the barrier. The reactant and product on the same electronic surface as the transition state are obtained by complete small-step optimisations (red arrows).

Each pair of points in this cycle brackets the TS, and, with the user monitoring progress, the successive displacements across the barrier can be diminished and the TS approached. The progress of the iterations is reflected also in diminishing energy gradients. The significant point is that all other geometrical variables are being optimised towards their values at the TS during the iterations. This occurs because energy reductions alternate across the barrier, which prevents any variable moving towards values in the limiting energy minima, and each energy reduction uses the improvements in those preceding. It is essential that the iterative trials oscillate across the barrier. As the cycles progress (1, 2, 3, 4 . . . 7 in Figure 3) the gradient converges towards zero, the geometric excursions diminish, and the energy converges: the convergence of these properties reveals the validity and precision of the resulting TS. On completion of the calculation the connection of the TS with reactant and product minima is confirmed by extended small-step energy minimisations (red arrows, Figure 3) when the geometry is nudged slightly, yielding the reaction path.

Because the vibrational structure of FeMo-co and its derivatives is complex, with some very low breathing and distortion frequencies, numerical calculation of frequencies is sometimes erroneous (as well as being time-consuming for FeMo-co systems). In practice it has been more expeditious to confirm a TS saddle point through its changes when nudged rather than through frequency calculations.

3. Results

The simple concept is that provided the trials oscillate across the saddle and are not allowed to progress towards local minima, the small-step small-extent energy reduction calculations will automatically optimise the geometrical variables of the saddle. Through user control of the iterating trial geometries the lowest energy point on the saddle can be approached.

In practice the method works well, but needs management by the user. At the lowest level of user control, once geometries on either side of the barrier are obtained, each new trial can be determined as an average, possibly weighted towards the opposite side, of the two previous trials, but always subject to the necessary requirement of oscillating trials. However the course of each small-step energy reduction provides information about the shape of the barrier and the significant geometrical variables, increasing chemical understanding of the system and leading to improved generation of trial geometries and faster convergence. In the energy reductions I usually limit each step by a geometry change of 0.025 Å. In each energy reduction sequence I monitor the energy and geometry changes, and usually terminate after *c.* 10 steps in the

earlier iterations of a TS search, and fewer steps as the process converges. Should there be uncertainty about the relevant reactant or product minima linked to the saddle, complete minimisation can be used as a check at any stage. Refinement of a TS was normally terminated when changes in successive iterations were $<0.5\,\text{kcal mol}^{-1}$ and <0.05 Å.

The pragmatic method needs only marginally more computer time than the LST/QST method (calculated by DMol at the same level), but additional user time is required for monitoring and assessing interim results.

Results from this method have been published [8,10]. Here I illustrate the method with details of three steps from a complete (21 step) chemical mechanism for the complete hydrogenation of N_2 catalysed by FeMo-co [4]. The first example involves the transfer of an H atom from S2B to one of the N atoms of the intermediate containing bound N_2H_2. The reactant, transition state and product structures are shown in part on Figure 4, which plots the energy change through the reaction path as key distances change. In this and subsequent examples the geometrical variables do not change concurrently and there is no straightforward reaction coordinate, and so the abscissa scale is arbitrary, following the points in the small-step energy minimisations that generate the reaction path. Note that in the early preparatory part of this process the S—H bond is not elongated while the N—H distance is shortening by rotating the S—H function and shortening the S—N distance. N—H bond formation is well-progressed (1.4 Å) before the S—H bond elongates, and the N—H bond is largely formed (1.2 Å) at the TS. The small discontinuity (<0.04 Å) in the plot at the TS reflects the precision of the calculation and the geometrical nudging required to follow the trajectories to reactant and product. Application of the LST/QST procedure to this step yielded a transition state with virtually the same geometry (within 0.02 Å) and $0.5\,\text{kcal mol}^{-1}$ lower energy, again reflecting the precision of the pragmatic method.

The second example is an H transfer from Fe2 to an NH group bound to Fe2, that is H transfer across a bond. The results are presented in Figure 5, showing that the preparatory phase involves shortening of N—H (blue) by bending but not lengthening of Fe—H (red), which elongates only on the product side of the TS. The post-TS phase of the process involves changes in three geometrical properties: in particular as the N—H bond is completed the Fe6—N bond breaks, consistent with the retention of four coordinate N. The chemical significance of this reaction step is explained elsewhere [4].

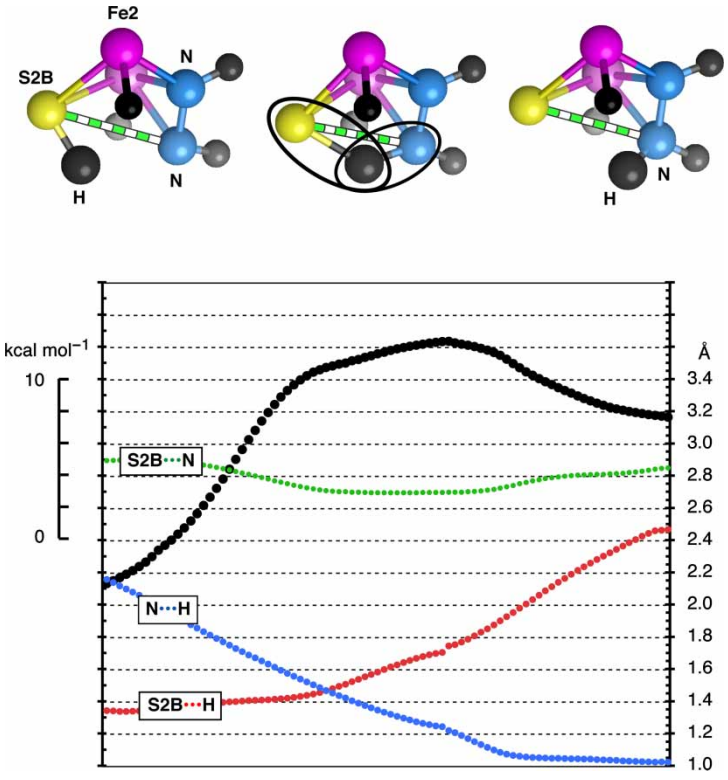

Figure 4. Details of the energy (black circles, left ordinate, kcal mol^{-1}) and relevant geometry (right ordinate, Å) during the reaction step in which an H atom transfers from S2B to N of N_2H_2 bound to the surface of FeMo-co, depicted only as relevant subsections of the complete structure. The moving H atom is drawn larger than the other H atoms (black), the changing S2B—H and H—N bonds are enclosed in the structure of the TS (centre), and the S2B—N distance is marked as green and white stripes.

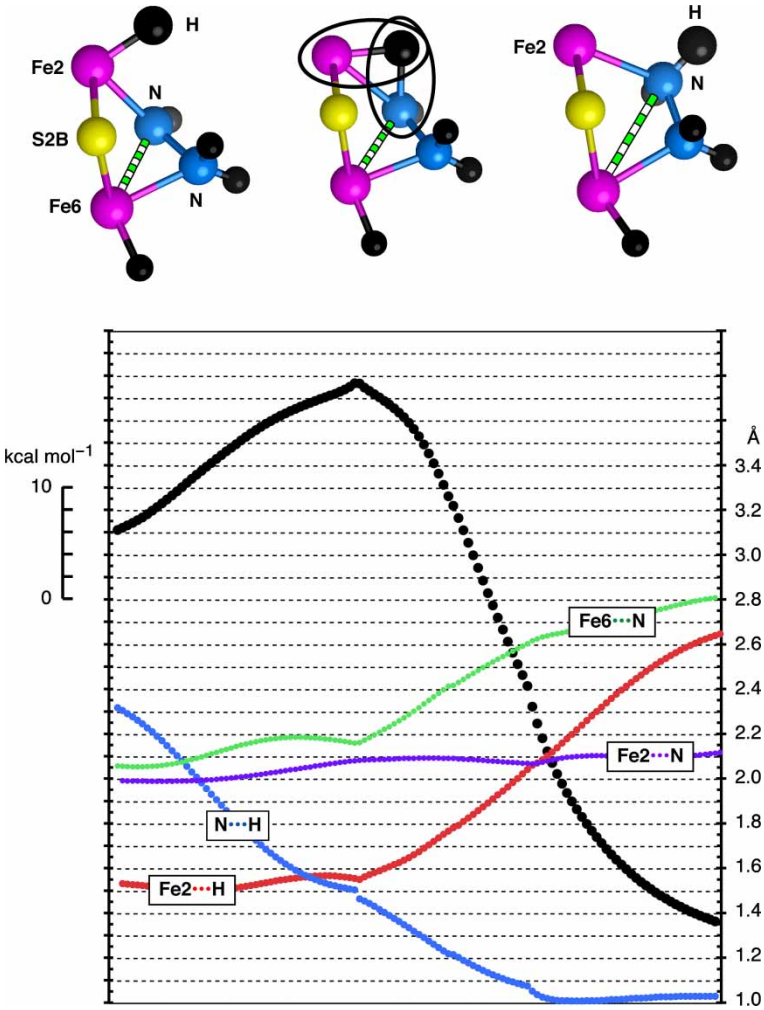

Figure 5. The path of the reaction step in which an H atom transfers from Fe2 to the NH of N_2H_3 bound to the surface of FeMo-co, depicted only as relevant subsections of the complete structure. Axes and colours as for Figure 4: the moving H atom is drawn larger than the other H atoms (black); the changing Fe2—H and H—N bonds are enclosed in the structure of the TS (centre); and the Fe6—N distance which elongates in the completion phase is marked as green and white stripes.

The discontinuities in Figure 5 again show a precision of *c.* 0.5 kcal mol^{-1}, 0.04 Å. The LST/QST methodology failed for this reaction. An LST calculation using the reactant and final product structures had SCF convergence problems, probably because the TS is not on the linear transit between reactant and product. A QST calculation, using the reactant, product and TS structures of Figure 5, lost control and 'fell of the ridge'. Modification of the QST method, using instead as product the first structure where the N—H bond is completed, generated an impossible interpolated geometry.

The third example involves the transfer of an H atom from S3B to the inner N atom of N_2 bound in η^2 geometry to Fe6 (Figure 6). This is a significant process because: (i) it is a crucial first and difficult step in the activation of unreactive N_2, (ii) atom S3B moves considerably, and the Fe6–S3B interaction transforms from non-bond (>3 Å) to bond (2.36 Å) during the H transfer, (iii) the apparently uninvolved Fe6—N^c distance changes during the reaction, as the coordination of Fe6 changes. In the preparatory phase (see Figure 6) of the reaction the S3B-H unit is folding towards N, shortening H—N, shortening S3B—N and shortening S3B—Fe6, without weakening the S—H bond. Then in the transfer phase over the TS the S—H bond is broken and the N—H bond fully formed. In the completion phase the Fe6—S3B bond is formed and the Fe6—N_2H moiety approaches its final conformation. Note that the S3B—N distance (green) shortens during the transfer phase.

During this reaction the Fe6—N^c distance varies between 2.7 and 3 Å. In the coordination chemistry of FeMo-co the Fe—N^c interactions are significant [2], and in the present reaction the Fe6—N^c interaction is essentially non-bonding. A LST/QST calculation on this process yielded a TS with the same S3B—H, N—H, S3B—N and

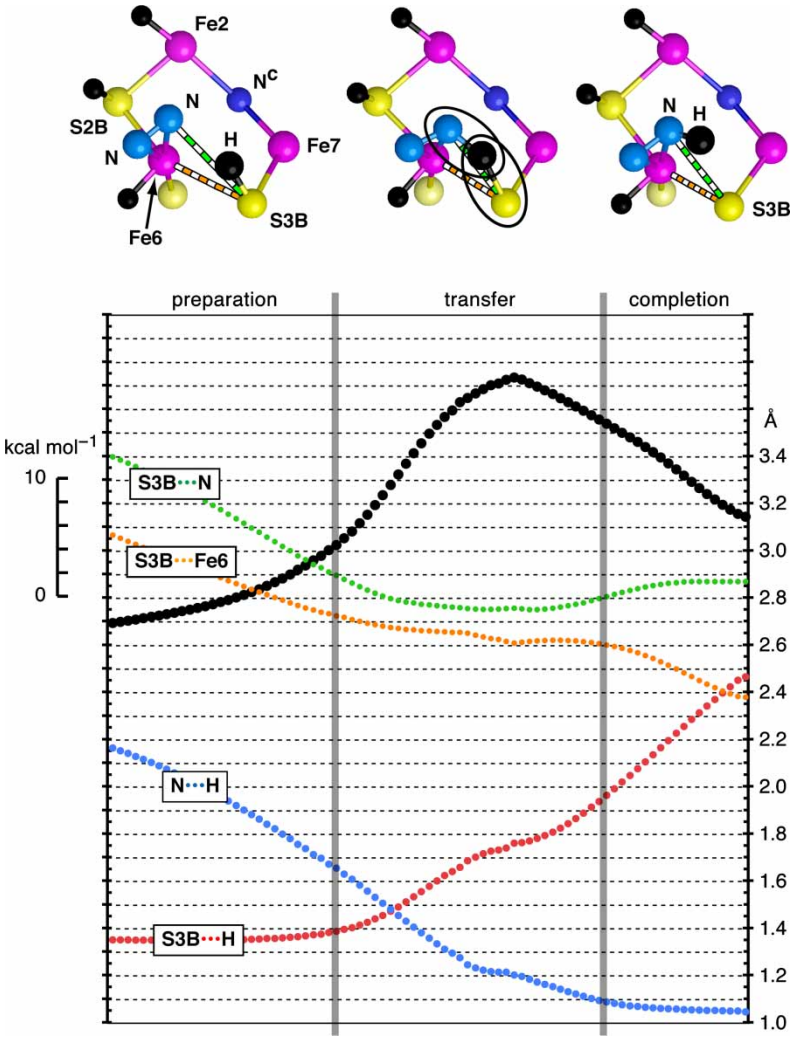

Figure 6. The trajectory for the first H transfer from S3B to N of η^2-coordinated N_2, depicted only as relevant subsections of the complete structure. Axes and colours as for Figure 4: the moving H atom is drawn larger than the other H atoms (black); the changing N—H and H—S3B bonds are enclosed in the structure of the TS (centre); the varying Fe6—S3B distance is marked as green and white stripes; and the varying S3B—N distance is marked as orange and white stripes.

S3B—Fe6 distances as Figure 6, but a longer Fe6—N^c distance (2.97 vs. 2.69 Å), a TS energy lower by 3.5 kcal mol^{-1}, but with very similar energy course and reaction barrier. The small discrepancy between the pragmatic calculation and the LST/QST calculation, occurring only in the Fe6—N^c interaction and the energy involved in its distance variation, is a consequence of the fact that the pragmatic calculation was made with cognisance of the geometries in the preceding and succeeding steps in the mechanism (shorter Fe6—N^c) [4], while the automatic optimisation did not.

In order to assess the performance of the QST method, 27 other TS for possible steps in the FeMo-co reaction mechanism, as determined by the pragmatic method, were subsequently used (with reactant and product structures) for QST calculations. Of these, seven TS were successfully confirmed, nine QST calculations 'fell off the ridge' and towards product, and the other eleven QST calculations were unsuccessful for other reasons such as SCF failure or excessive geometrical distortion.

4. Discussion

So far the pragmatic method has failed to locate a transition state, and has characterised reaction barriers that are low, high, steep, asymmetric, or part of an asynchronous reaction path. To some extent the success of the method relies on the chemical acuity of the user. The method is valuable because it starts in the vicinity of the transition state, and does not require prior knowledge of reactant and product structures or other local minima in the potential energy surface. By focusing on the

saddle-points of the surface, unforeseen local minima and reaction intermediates can be discovered.

The method is useful for the exploration of relatively complex and unconventional systems such as the reactions of FeMo-co. In the second and third examples above, the geometrical variables are not approximately linearly correlated: the various changes occur in different phases in the reaction path. In these cases the linear interpolation of geometry in the LST method is a poor initial approximation of the reaction path, which is asynchronous rather than linear synchronous. In such cases it could be preferable to use the pragmatic method, or to use a chemically knowledgeable putative TS in a QST optimisation.

The application to reactions of FeMo-co is facilitated by the efficiency of the numerical basis sets used in DMol.

An alternative mechanism for the conversion of N_2 to NH_3 catalysed by FeMo-co has been presented by Kastner and Blochl [11]. The main mechanistic difference is the use of NH_4^+ to model proton donation by surrounding amino acids. These authors used damped Car-Parinello molecular dynamics to determine TS, employing a 1D constraint on the atomic positions and fixing the constraint to discrete values around the transition state to maximise the energy.

The pragmatic method has been described here in the context of a molecular system, but it should be applicable to non-molecular (periodic) systems, and be relevant to the materials science focus of this journal issue. Certainly the substrate binding and the H atom transfers catalysed on the surface of FeMo-co are analogous to reactions catalysed on non-molecular surfaces.

Acknowledgements

This research is supported by the Australian Partnership for Advanced Computing, and by the University of New South Wales.

References

[1] (a) O. Einsle, F.A. Tezcan, S.L.A. Andrade, B. Schmid, M. Yoshida, J.B. Howard, and D.C. Rees, *Nitrogenase MoFe-protein at 1.16Å resolution: A central ligand in the FeMo-cofactor*, Science 297 (2002), pp. 1696–1700; (b) D.C. Rees, F.A. Tezcan, C.A. Haynes, M.Y. Walton, S. Andrade, O. Einsle, and J.A. Howard, *Structural basis of biological nitrogen fixation*, Philos. Trans. R. Soc. A 363 (2005), pp. 971–984; (c) J.B. Howard and D.C. Rees, *How many metals does it take to fix N_2? A mechanistic overview of biological nitrogen fixation*, Proc. Natl Acad. Sci. USA 103 (2006), pp. 17119–17124.

[2] I. Dance, *Elucidating the coordination chemistry and mechanism of biological nitrogen fixation*, Chem. Asian J. 2 (2007), pp. 936–946.

[3] (a) L.C. Seefeldt, I.G. Dance, and D.R. Dean, *Substrate interactions with nitrogenase: Fe versus Mo*, Biochemistry 43 (2004), pp. 1401–1409; (b) B.M. Barney, R.Y. Igarashi, P.C. Dos Santos, D.R. Dean, and L.C. Seefeldt, *Substrate interaction at an iron-sulfur face of the FeMo-cofactor during nitrogenase catalysis*, J. Biol. Chem. 279 (2004), pp. 53621–53624; (c) P.C. Dos Santos, R. Igarashi, H.-I. Lee, B.M. Hoffman, L.C. Seefeldt, and D.R. Dean, *Substrate interactions with the nitrogenase active site*, Acc. Chem. Res. 38 (2005), pp. 208–214.

[4] I. Dance, *The chemical mechanism of nitrogenase: Calculated details of the steps in which η^2-N_2 on FeMo-co is converted to NH_3*, submitted (2008).

[5] T.A. Halgren and W.N. Lipscomb, *The synchronous-transit method for determining reaction pathways and locating molecular transition states*, Chem. Phys. Lett. 49 (1977), pp. 225–232.

[6] N. Govind, M. Petersen, G. Fitzgerald, D. King-Smith, and J. Andzelm, *A generalized synchronous transit method for transition state location*, Comput. Mater. Sci. 28 (2003), pp. 250–258.

[7] (a) B. Delley, in *DMol, A Standard Tool for Density Functional Calculations: Review and Advances*, J.M. Seminario and P. Politzer, eds., Vol. 2, Elsevier, Amsterdam, 1995, pp. 221–254; (b) B. Delley, *From molecules to solids with the DMol^3 approach*, J. Chem. Phys. 113 (2000), pp. 7756–7764; DMol3, www reference for DMol^3, www.accelrys.com/mstudio/ms_modeling/dmol3.html (2005).

[8] I. Dance, *The hydrogen chemistry of the FeMo-co active site of nitrogenase*, J. Am. Chem. Soc. 127 (2005), pp. 10925–10942.

[9] I. Dance, *The correlation of redox potential, HOMO energy, and oxidation state in metal sulfide clusters and its application to determine the redox level of the FeMo-co active-site cluster of nitrogenase*, Inorg. Chem. 45 (2006), pp. 5084–5091.

[10] I. Dance, *The mechanistically significant coordination chemistry of dinitrogen at FeMo-co, the catalytic site of nitrogenase*, J. Am. Chem. Soc. 129 (2007), pp. 1076–1088.

[11] J. Kastner and P.E. Blochl, *Ammonia production at the FeMo cofactor of nitrogenase: Results from density functional theory*, J. Am. Chem. Soc. 129 (2007), pp. 2998–3006.