Flexible QM/MM modelling embraces alternative mechanisms for lactate dehydrogenase

Vicent Moliner[a] and Ian H. Williams*[b]

[a] *Departament de Ciències Experimentals, Universitat Jaume I, Box 224, 12080 Castellón, Spain*
[b] *Department of Chemistry, University of Bath, Bath, UK BA2 7AY. E-mail: i.h.williams@bath.ac.uk*

Received (in Cambridge, UK) 10th July 2000, Accepted 21st August 2000
First published as an Advance Article on the web 11th September 2000

A flexible AM1/CHARMM treatment finds two distinct mechanistic pathways across the megadimensional energy hypersurface computed for lactate dehydrogenase catalysed reduction of pyruvate to lactate: these differ in the timing of the hydride transfer and proton transfer components of the reaction.

Current applications of hybrid quantum-mechanical/molecular-mechanical (QM/MM) approaches to modelling of enzyme reactivity commonly neglect flexibility in very large systems. Two recent independent studies of lactate dehydrogenase (LDH) catalysed interconversion of pyruvate and lactate, using similar QM/MM methods, yielded very different transition structures (TSs) and mechanisms.[1,2] We now report the simultaneous existence of two distinct mechanistic pathways across the megadimensional energy hypersurface for this system, and suggest that both earlier studies failed to explore its topography to a sufficient extent. To determine which mechanism is preferred will require use not only of a theoretical method giving reliable energies but also of statistical averaging over many configurations. Our finding is significant because it shows that a rigid approach to modelling of any complex system may easily miss an important feature of reaction mechanism.

The chemical step of the LDH cataysed reaction involves hydride transfer (HT) from the dihydronicotinamide ring of NADH to the carbonyl C atom of pyruvate and proton transfer (PT) to the carbonyl O atom of pyruvate from a protonated histidine residue (Fig. 1). The relative timing of the HT and PT is a matter of some interest in mechanistic enzymology. Ranganathan and Gready[2] (RG) found a mechanism (HT,PT) in which HT preceded PT in a stepwise manner, contrasting with the usual chemical and enzymatic arguments for HT processes and their own results from supermolecule calculations,[3] but in accord both with the assumption of an earlier empirical valence-bond study[4] and with the results of a later QM/MM study of the analogous malate dehydrogenase.[5] On the other hand, we (MTW) found a family of TSs (with differing relative dispositions of active site residues) corresponding to a concerted mechanism (PT/HT) with PT considerably more advanced than HT.[1]

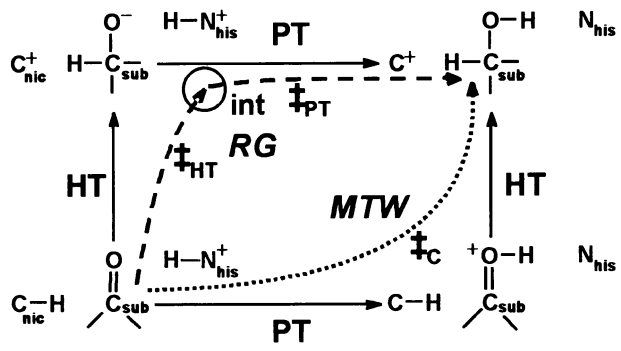

Fig. 1 Reaction map for LDH catalyzed reduction of pyruvate reactant (bottom left) to lactate product (top right); 'his' is histidine-195, 'nic' is nicotinamide, and 'sub' is substrate.

The RG and MTW studies both used the AM1 method[6] for the QM region. RG used the AMBER[7] MM method for the entire dogfish LDH sub-unit ternary complex, but with a relatively small number of atoms allowed to move. In contrast MTW used the CHARMM[8] force field for a truncated (1900 atom) *B. stearothermophilus* LDH but with a very large number of mobile atoms. The first important result of our recent studies is that these and other differences of detail (*e.g.* nature and location of QM/MM link atoms and QM/MM electronic coupling) are not primarily responsible for the different mechanisms found previously. We have now used a large-QM/full-MM model which includes all atoms (~5600) of the monomeric *B. stearothermophilus* LDH subunit,[9] plus NADH, pyruvate, and water molecules within a 20 Å radius ball of water centred on His-195.[10] All atoms were free to move in the optimisations and TS searches. A larger QM region (52 AM1atoms) containing also the ribose moiety was employed, as shown in Fig. 2. Schmidt and Gready have recently asserted that AM1 is likely to be adequate for modelling conformational preferences in QM/MM calculations for the LDH catalyzed reaction.[11]

Approximate saddle points were located by grid searches using constrained ABNR minimization of the CHARMM24b2[12] QM/MM energy.[13] TS refinement in GRACE[1b] employed an explicit hessian for the QM 'core' (Fig. 2) while the MM enzyme 'environment' was continually relaxed to a r.m.s. gradient < 10^{-3} kcal mol^{-1} Å^{-1}. TS optimisation in the core was continued until no element of the gradient vector of the entire system (core plus environment) was larger than 10^{-2} kcal mol^{-1} Å^{-1}. To characterise each saddle point, its hessian was recomputed for a total of 147 atoms (QM core + six amino acid residues in the active-site region) by central finite-differencing of the gradient vector: the single negative eigenvalue of the resultant hessian corresponded to the transition vector. The intrinsic reaction coordinate (IRC) was computed in both directions from each saddle point to confirm each as being the expected TS; energy minimisations from a point along each IRC path yielded structures for the reactant,

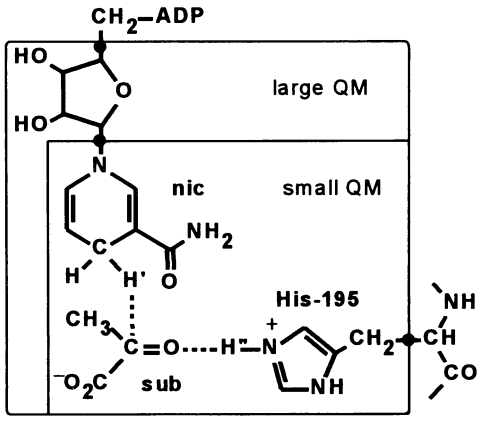

Fig. 2 Large and small QM regions in QM/MM models for LDH; link atoms are indicated as '●'.

DOI: 10.1039/b005549m

Table 1 Selected bond lengths (Å) and energies (kcal mol^{-1}) for AM1/CHARMM large-QM/full-MM optimized structures on the reaction map for LDH catalyzed reduction of pyruvate

	$C_{nic}\cdots H'$	$H'\cdots C_{pyr}$	$O_{pyr}\cdots H''$	$H''\cdots N_{his}$	E_{total}	$E_{QM} + E_{QM/MM}$
reactant	1.130	2.550	2.043	0.998	−12225.1	−514.0
$\ddagger_C$	1.286	1.458	0.996	1.892	−12173.5	−462.7
$\ddagger_{HT}$	1.730	1.234	1.981	1.017	−12155.9	−481.1
int	2.766	1.144	1.700	1.054	−12191.0	−511.4
$\ddagger_{PT}$	2.773	1.140	1.565	1.091	−12183.7	−500.0
product	2.561	1.110	0.978	2.517	−12243.7	−577.5

intermediate, and product. Table 1 contains optimised bond lengths for the HT and PT components of LDH catalysed pyruvate reduction and energies for minima and TSs.

The second important result of this study is that we find reaction paths and TSs for *both* the stepwise HT,PT and the concerted PT/HT mechanisms for both the present large-QM/full-MM model and (results not shown here) small-QM/truncated-MM and small-QM/full-MM models. There is no fundamental disagreement between the MTW and RG results: the apparent discrepancy arose simply because different parts of the hypersurface had been explored in each of the earlier studies. Subtle but significant differences in many coordinates other than just HT and PT dictate whether a particular geometry for the QM region lies within that part of the overall hypersurface belonging to the HT,PT mechanism or else in that different part corresponding to the PT/HT mechanism. It is as if the hypersurface contains a watershed separating two distinct drainage basins: depending upon how a gentle breeze may gust at the crucial moment, a raindrop falling on the Andes watershed may flow to either the Atlantic or the Pacific. So it is with geometry optimisation in a very large and flexible system: two computational runs may start at very similar structures in the core region but be subtly influenced by small differences in their environments to follow separate courses to very different final structures. However, owing to the particular bias provided by the environment, optimisations initiated within the same basin may all tend to converge to the same mechanistic result, despite scanning of the key parameters within the core over adequate ranges of values. A difference in some detail of the QM/MM treatment may of itself generate only a trivial difference in energy for an initial structure but, as a consequence of small differences in the gradient and second derivatives, may cause an optimisation (or TS search) to proceed towards a significantly different stationary point.

The obvious question is which mechanism is preferred: HT,PT or PT/HT? Unfortunately, and perhaps surprisingly, there is no easy answer.[14] The order of the total energies ($E_{QM} + E_{QM/MM} + E_{MM}$) for the three TSs is $\ddagger_{HT} > \ddagger_C > \ddagger_{PT}$. This would suggest that the concerted PT/HT mechanism is preferred over the stepwise HT,PT mechanism, for which HT is the rate-determining step. If, however, only the energy of the QM core within the electrostatic environment of the MM atoms is considered ($E_{QM} + E_{QM/MM}$), then for both models the order of energies of the three TSs is $\ddagger_C > \ddagger_{HT} > \ddagger_{PT}$, favouring the HT,PT mechanism. If no MM atoms are allowed to move in the TS search, then $E_{MM} = 0$ and $E_{total} = E_{QM} + E_{QM/MM}$. However, as flexibility in the MM region is introduced, and the number of atoms contributing to E_{MM} is increased from zero to several thousand, it seems that the mechanistic preference changes. We are unable yet to make a definitive statement as to which mechanism is preferred, but are currently performing *ab initio* QM/MM calculations for this system with a view to obtaining more reliable energetics. We would also like to be able to describe the specific physical interactions within the MM region that favour each of the two mechanisms. Visual inspection of the two structures reveals no obvious significant differences, and as yet the origin of the change in E_{MM} responsible for the mechanistic shift eludes characterisation; we hope to report upon this important point in a full paper. It should be noted that, even for our smallest model, there are several thousand degrees of freedom in the MM region and the critical energetic difference is likely to be the sum of a large number of very small contributions.

Finally, our refined stationary points are not unique. The multiple minima issue has its counterpart for TSs: there exist multiple local saddle points differing in the conformation of the environment. Accurate determination of transition state properties will require statistical averaging over many configurations, each one individually being a transition structure.[1b] Free-energy calculations must not only sample representative configurations along an assumed *in vacuo* reaction path but also must consider the possibility of alternative mechanistic pathways.

We are grateful to the British Council/Acciones Integradas for a collaborative award and HEFCE/EPSRC JREI and Universitat Jaume I for provision of computer time.

Notes and references

1 (*a*) V. Moliner, A. J. Turner and I. H. Williams, *Chem. Commun.*, 1997, 1271; (*b*) A. J. Turner, V. Moliner and I. H. Williams, *Phys. Chem. Chem. Phys.*, 1999, **1**, 1323.
2 S. Ranganathan and J. E. Gready, *J. Phys. Chem. B*, 1997, **101**, 5614.
3 S. Ranganathan and J. E. Gready, *J. Chem. Soc., Faraday Trans.*, 1994, **90**, 2047.
4 A. Yadav, R. M. Jackson, J. J. Holbrook and A. Warshel, *J. Am. Chem. Soc.*, 1991, **113**, 4800.
5 M. A. Cunningham, L. L. Ho, D. T. Nguyen, R. E. Gillilan and P. A. Bash, *Biochemistry*, 1997, **36**, 4800.
6 M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902.
7 S. J. Weiner, P. A. Kollman, D. T. Nguyen and D. A. Case, *J. Comput. Chem.*, 1986, **7**, 230.
8 J. J. Pavelites, J. Gao, P. A. Bash and A. D. Mackerell, *J. Comput. Chem.*, 1997, **18**, 221; A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586.
9 D. B. Wigley, S. J. Gamblin, S. P. Turkenburg, E. J. Dodson, K. Piontek, H. Muirhead and J. J. Holbrook, *J. Mol. Biol.*, 1992, **223**, 317.
10 T. R. Dafforn, I. G. Badcoe, R. B. Sessions, A. S. El Hawrani and J. J. Holbrook, *Proteins: Struct. Funct. Genet.*, 1997, **29**, 228.
11 R. K. Schmidt and J. E. Gready, *Theochem*, 2000, **498**, 101.
12 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187; CHARMM24b2, M. Karplus, Harvard University, 1996.
13 M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.*, 1990, **11**, 700.
14 S. J. Titmuss, P. L. Cummins, A. A. Bliznyuk, A. P. Rendell and J. E. Gready, *Chem. Phys. Lett.*, 2000, **320**, 169.