J|A|C|S
ARTICLES
Published on Web 12/09/2006

A Theoretical Analysis of Rate Constants and Kinetic Isotope Effects Corresponding to Different Reactant Valleys in Lactate Dehydrogenase

Silvia Ferrer,[†] Iñaki Tuñón,*[,†] Sergio Martí,[‡] Vicente Moliner,*[,‡] Mireia Garcia-Viloca,[§] Àngels González-Lafont,[§,⊥] and José M. Lluch[§,⊥]

Contribution from the Departament de Química Física, Universitat de València, València, Spain, Departament de Ciències Experimentals, Universitat Jaume I, Castelló, Spain, Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Barcelona, Spain, and Departament de Química, Universitat Autònoma de Barcelona, Barcelona, Spain

Received July 27, 2006; E-mail: tunon@uv.es; moliner@uji.es

Abstract: In some enzymatic systems large conformational changes are coupled to the chemical step, in such a way that dispersion of rate constants can be observed in single-molecule experiments, each corresponding to the reaction from a different reactant valley. Under this perspective here we present a computational study of pyruvate to lactate transformation catalyzed by lactate dehydrogenase. The reaction consists of a hydride transfer and a proton transfer that seem to take place concertedly. The degree of asynchronicity and the energy barrier depend on the particular starting reactant valley. In order to estimate rate constants we used a free energy perturbation technique adapted to follow the intrinsic reaction coordinate for several possible reaction paths. Tunneling effects are also obtained with a slightly modified version of the ensemble-averaged variational transition state theory with multidimensional tunneling contributions. According to our results the closure of the active site by means of a flexible loop can lead to the formation of different reactant complexes displaying different features in the disposition of some key residues (such as Arg109), interactions with the substrate and number of water molecules in the active site. The chemical step of the reaction takes place with a different reaction rate from each of these complexes. Finally, primary kinetic isotope effects for replacement of the transferring hydrogen of the cofactor with a deuteride are in good agreement with experimental observations, thus validating our methodology.

Introduction

Enzymes are essential for living organisms, making chemical reactions to occur in a time scale compatible with life.[1] The ability of these proteins to speed up chemical processes lowering the activation free energy has been the subject of an increasing number of theoretical approaches during the last few years.[2−15] While the different computational strategies have been successful in reproducing some experimental values, the deep understanding of the molecular machinery involved in enzymatic reactions still requires the solution of many conceptual and practical challenges.

Theoretical studies of proteins involve the exploration of highly dimensional potential energy surfaces on which a hierarchical structure of valleys can be described.[16] As it has been shown in the case of myoglobin,[16] different protein substates or valleys may perform the same function but with different rates. In the case of enzymes this means that different reactant valleys (enzyme−substrate complexes displaying important conformational differences) may contribute to the total reaction flux with different reaction constants.[17−19] A reactant valley contains an ensemble of nuclear configurations that can be explored by means of a long-time molecular dynamics trajectory. The landscape within each valley is extremely rugged,

[†] Universitat de Valencia.
[‡] Universitat Jaume I.
[§] Institut de Biotecnologia i de Biomedicina.
[⊥] Departament de Química, Universitat Autònoma de Barcelona.

(1) Wolfenden, R.; Snider, M. J. *Acc. Chem. Res.* **2001**, *34*, 938.
(2) Gao, J. L.; Truhlar, D. G. *Ann. Rev. Phys. Chem.* **2002**, *53*, 467.
(3) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186.
(4) Martí, S.; Roca, M.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertran, J. *Chem. Soc. Rev.* **2004**, *33*, 98.
(5) Shurki, A.; Villa, J.; Warshel, A. *Abstr. Pap. Am. Chem. Soc.* **2002**, *223*, C59.
(6) Villà, J.; Strajbl, M.; Glennon, T. M.; Sham, Y. Y.; Chu, Z. T.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11899.
(7) Warshel, A. *Abstr. Pap. Am. Chem. Soc.* **1989**, *197*, 101.
(8) Warshel, A.; Villa, J.; Strajbl, M.; Florian, J. *Abstr. Pap. Am. Chem. Soc.* **2001**, *221*, U399.
(9) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9032.
(10) Ma, J. P.; Zheng, X. F.; Schnappauf, G.; Braus, G.; Karplus, M.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14640.
(11) Almarsson, O.; Bruice, T. C. *J. Am. Chem. Soc.* **1993**, *115*, 2125.
(12) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 10540.
(13) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 1472.
(14) Prat-Resina, X.; Gonzalez-Lafont, A.; Lluch, J. M.*THEOCHEM* **2003**, *632*, 297.
(15) Mulholland, A. J. *Drug Discovery Today* **2005**, *10*, 1393.
(16) Ansari, A.; Berendzen, J.; Bowne, S. F.; Frauenfelder, H.; Iben, I. E. T.; Sauke, T. B.; Shyamsunder, E.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 5000.
(17) Min, W.; English, B. P.; Luo, G. B.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. *Acc. Chem. Res.* **2005**, *38*, 923.
(18) Antikainen, N. M.; Smiley, R. D.; Benkovic, S. J.; Hammes, G. G. *Biochemistry* **2005**, *44*, 16835.
(19) Thorpe, I. F.; Brooks, C. L. *J. Am. Chem. Soc.* **2005**, *127*, 12997.

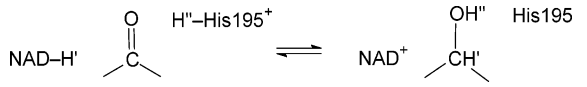
with a huge number of stationary points. The interconversions among all the low-energy configurations within the same valley usually consist of relatively easy local conformational changes and are quite fast at physiological temperatures. In these conditions, a local equilibrium assumption within a given reactant valley can be adopted. Sometimes, a rare event occurs, and the reactive system evolves toward high-energy regions of the same valley in such a way that the dividing hypersurface (transition state) between the reactant region and the product region is eventually reached and the system may cross to its corresponding product valley. Starting from initial configurations obtained by equilibrium molecular dynamics simulations, it is possible to generate multiple minimum-energy paths. Each of them joins a different potential energy minimum inside the reactant valley with a potential energy minimum in the associated product valley, passing through the corresponding transition state structure. This way, a range of different potential energy barriers and free energy barriers can be found.[19−22] From the existence of local equilibrium inside the reactant valley, it is possible to get an ensemble average over paths, thus leading to a single, free energy barrier which determines the catalytic rate constant within the framework of the canonical variational transition state theory (CVT). The existence of different paths and their averaging in enzyme reactions have been addressed in the previous works of Truhlar and co-workers.[23,24]

In addition to local changes, more global conformational changes also exist in proteins in such a way that the interconversion rates among the corresponding conformational substates (reactant valleys) can be much slower than the catalytic rate constants. All the nuclear configurations (and therefore, the stationary points) of a given reactant valley share a distinct global conformational feature which is a fingerprint of the ensemble of structures of the valley. In these conditions even very lengthy normal molecular dynamics simulations initiated at the bottom of a given reactant valley are, in practice, unable to jump to another valley because the free energy barrier for the conformational change is too high. As a result, the ergodicity condition does not hold in practice over the complete ensemble of configurations of the potential energy hypersurface corresponding to the whole reactive system. In other words, while a local equilibrium exists inside each reactant valley, a global equilibrium including all the reactant valleys is very unlikely when the catalytic rates are much greater than the interconversion rates among the different reactant valleys. In this case, the population of each valley is not under thermodynamic control, but under kinetic control. The existence of these relatively stable reactant valleys (or long-lived conformational substates) in comparison with the time scale of the enzymatic reactions is the required molecular basis to understand the dynamic disorder observed in single-molecule studies of some enzyme mechanisms.[17,18,25−28] At single-molecule level, an enzymatic reaction is a stochastic event, and a single-molecule experiment measures the waiting times for the completion of the enzymatic reaction. Multiexponentiality in the probability density of these waiting times is a manifestation of dynamic disorder, that is, fluctuations with time of the catalytic rate constant of a single enzyme molecule, caused by slow transitions among different conformational substates (reactant valleys) of the enzyme having different characteristic values of the rate constant. These fluctuations can occur on a time scale comparable to or longer than that of the enzymatic reaction. The slow interconversion among conformers results in the memory effect associated with the correlations between successive enzymatic turnover times. The overall rate of product formation is no longer governed by a single rate constant, but effectively by a distribution of rate constants. Consequently, catalysis essentially occurs through independent, different catalytic pathways. In other words, each reactant valley will act as an independent enzyme with a different rate constant. Experiments on β-galactosidase enzyme have shown that rate constants for single molecules can span over a broad range of time scales, from 10 to 10^{-3} s.[29] If, conversely, just a reactant valley exists (or in the case of fast interconversion among the valleys), there is not dynamic disorder, and a monoexponential decay is found in the waiting times distribution, indicating that a single rate constant can describe the process. One of the enzymes whose single-molecule studies indicate the existence of dynamic disorder (and, as a consequence, long-lived reactant valleys or conformational substates) is lactate dehydrogenase (LDH). Single-molecule experiments carried out by Xue and Yeung found differences in reactivities among LDH-1 molecules (one of the five human isozymes present in the heart tissue) as large as a factor of 4. They suggest that the origin of these differences may lie in the presence of several stable conformational arrangements of the enzyme.[30,31]

LDHs simultaneously and stereoespecifically interconvert pyruvate to L-lactate and the nicotinamide adenine dinucleotide cofactor from the NADH form to NAD^+.[32−35] In these enzymes binding of cofactor is followed by the substrate in an ordered process that includes the enclosure of the active site by a mobile loop of the protein. This last step is the rate-limiting one in the wild-type enzyme.[36] In this movement of the active-site loop several key residues are brought to close contact with the substrate, including an invariant arginine that is believed to stabilize the transition state (TS) structure. The chemical step of the enzymatic process involves the transfer of a proton from a protonated histidine present in the active site to the carbonyl oxygen atom of pyruvate as well as a hydride transfer from the dihydronicotinamide ring of the cofactor to the carbonyl carbon atom of the substrate (see Scheme 1).[37]

(20) Zhang, Y. K.; Kua, J.; McCammon, J. A. *J. Phys. Chem. B* **2003**, *107*, 4459.
(21) Thorpe, I. F.; Brooks, C. L. *J. Phys. Chem. B* **2003**, *107*, 14042.
(22) Hu, P.; Zhang, Y. K. *J. Am. Chem. Soc.* **2006**, *128*, 1272.
(23) Alhambra, C.; Corchado, J.; Sanchez, M. L.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2001**, *105*, 11326.
(24) Truhlar, D. G.; Gao, J. L.; Garcia-Viloca, M.; Alhambra, C.; Corchado, J.; Sanchez, M. L.; Poulsen, T. D. *Int. J. Quantum Chem.* **2004**, *100*, 1136.
(25) Kou, S. C.; Cherayil, B. J.; Min, W.; English, B. P.; Xie, X. S. *J. Phys. Chem. B* **2005**, *109*, 19068.
(26) Smiley, R. D.; Hammes, G. G. *Chem. Rev.* **2006**, *106*, 3080.
(27) Lu, H. P.; Xun, L. Y.; Xie, X. S. *Science* **1998**, *282*, 1877.
(28) Yang, H.; Luo, G. B.; Karnchanaphanurach, P.; Louie, T. M.; Rech, I.; Cova, S.; Xun, L. Y.; Xie, X. S. *Science* **2003**, *302*, 262.
(29) English, B. P.; Min, W.; van Oijen, A. M.; Lee, K. T.; Luo, G. B.; Sun, H. Y.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. *Nat. Chem. Biol.* **2006**, *2*, 87.
(30) Xue, Q. F.; Yeung, E. S. *Nature* **1995**, *373*, 681.
(31) Tan, W. H.; Yeung, E. S. *Anal. Chem.* **1997**, *69*, 4242.
(32) Clarke, A. R.; Wigley, D. B.; Barstow, D. A.; Chia, W. N.; Waldman, A. D. B.; Hart, K. W.; Atkinson, T.; Holbrook, J. J. *Biochem. Soc. Trans.* **1987**, *15*, 152.
(33) Badcoe, I. G.; Smith, C. J.; Wood, S.; Halsall, D. J.; Holbrook, J. J.; Lund, P.; Clarke, A. R. *Biochemistry* **1991**, *30*, 9195.
(34) Clarke, A. R.; Wilks, H. M.; Barstow, D. A.; Atkinson, T.; Chia, W. N.; Holbrook, J. J. *Biochemistry* **1988**, *27*, 1617.
(35) Deng, H.; Zheng, J.; Clarke, A.; Holbrook, J. J.; Callender, R.; Burgner, J. W. *Biochemistry* **1994**, *33*, 2297.
(36) Clarke, A. R.; Wigley, D. B.; Chia, W. N.; Barstow, D.; Atkinson, T.; Holbrook, J. J. *Nature* **1986**, *324*, 699.
(37) Clarke, A. R.; Waldman, A. D. B.; Hart, K. W.; Holbrook, J. J. *Biochim. Biophys. Acta* **1985**, *829*, 397.

Scheme 1



One of the most studied LDH enzymes is BsLDH (the lactate dehydrogenase of *Bacillus stearothermophilus*). The catalytic rate constant of this enzyme at 25 °C is 250 s^{-1}, and it is assigned to the active-site loop (residues 99−110) closure. In this movement the invariant arginine, Arg109, moves 8 Å from a position in the solvent to one in the active site.[32] A combination of site-directed mutagenesis and kinetic isotope effects (KIEs) allows to conclude that the rate constant of the chemical step should be at least 750 s^{-1} at 25 °C,[32] which means that the phenomenological activation free energy should be lower than 13.5 kcal·mol^{-1}. Primary KIEs were measured by replacing the transferring hydride of the cofactor NADH with a deuteride on the Arg109Gln mutant enzyme, resulting in a 3-fold decrease of the rate constant at 25 °C.[32,36] Schowen and co-workers determined an intrinsic primary KIE of 5.5 at 55 °C for the wild-type BsLDH.[38] Primary KIEs of 3−6 are typical for dehydrogenases.[32] In particular, an experimental value of 2.70 was determined for mutant LDH in which Arg109 was replaced by lysine.

Theoretical calculations have revealed the difficulties associated with the study of this enzyme, in particular for determining the relative timing of the hydride and the proton transfers. Ranganathan and Gready, using hybrid quantum mechanics/molecular mechanics (QM/MM) methods, found a mechanism in which the hydride transfer preceded the proton transfer in a stepwise manner,[39] in agreement with an earlier empirical valence bond study and with QM/MM results of an analogous malate dehydrogenase.[11,40−42] One of us found a family of TS structures, obtained from different starting configurations, corresponding to a concerted mechanism where the proton transfer was considerably more advanced than the hydride transfer.[43] The analysis of these TS structures revealed differences in the interactions established between the substrate and the invariant Arg109. In particular it was found that the TS-stabilizing role of this residue could also be alternatively played by Asn140.[43] It was also described that flexible QM/MM modeling, allowing the enzymatic environment to change while the system advances along the reaction coordinate, could embrace the alternative mechanisms previously described.[44] While all these QM/MM calculations were made using a semiempirical AM1 Hamiltonian, we recently tackled the problem of improving the quality of the potential energy surface including corrections to the quantum description.[45] When these corrections were included, together with a systematic assignment of the protonation state of the different residues according to electrostatic calculations, a concerted mechanism was obtained where the hydride transfer was more advanced than the proton transfer.[45,46] Anyway, different degrees of asynchronicity between the hydride and the proton transfers could be plausible. In a recent transition path sampling study on the human heart isozyme, h-H_4LDH, several reaction trajectories were described that differ in the time lag between the hydride and the proton transfers.[47]

In this paper we present a theoretical study on the pyruvate-to-lactate transformation catalyzed by BsLDH. We specifically address the problem of the proper sampling of the relevant reactant valleys of the system, taking into account that the global process involves not only two chemical events (the proton and the hydride transfers) but also conformational changes of the enzyme that can determine the differences in the reactivities of different single molecules. In our work, we have developed a new strategy to access to a number of reaction valleys by performing a sampling in the high-energy region of the transition states, where the transition among them is easier. From there, the low-energy regions of the reactant valleys are reached going down. We have identified the main features of the reaction valleys, and we have shown that the differences among them can be related to the disposition of the flexible loop closing the active site (a slow conformational change that is the rate-limiting step in the wild-type enzyme). In addition, we have calculated the rate constants associated with each reactant valley through an adaptation of the ensemble averaged-variational transition state theory with multidimensional tunneling (EA-VTST/MT) contributions.[23,24,48] Our theoretical results agree with the single molecular studies that indicate dynamic disorder due to the existence of distinct, well-characterized conformational substates.

2. Computational Details

2.1. Model of the Enzyme−Substrate−Coenzyme Complex. The X-ray structure of BsLDH comes from the PDB code 1LDN [49] which is an octamer, although in this work only the tetramer is used; X-ray studies have proposed that the tetramer is the functional form of BsLDH.[49] In each monomer the crystal structure contains a total of 316 amino acid residues, the cofactor NADH, and the inhibitor oxamate (OXM), which is replaced by pyruvate in our study.

The coordinates of the hydrogen atoms of the protein and the coenzyme were determined using the HBUILD facility of the CHARMM package.[50] The pK_a of titratable aminoacid groups within the protein were calculated using the cluster method by Gilson and co-workers[51,52] as implemented by Field and co-workers.[53] According to this method, each titratable residue in the protein is perturbed by the electrostatic effect of the protein environment. As demonstrated in our previous study on LDH,[45] an inaccurate protonation state of only a few residues of the protein can render erroneous potential energy surfaces (PESs). The most significant change with respect to the use of aqueous solution pK_a values is the fact that Glu199 is found in its protonated form at

(38) Seravalli, J.; Huskey, W. P.; Schowen, K. B.; Schowen, R. L. *Pure Appl. Chem.* **1994**, *66*, 695.
(39) Ranganathan, S.; Gready, J. E. *J. Phys. Chem. B* **1997**, *101*, 5614.
(40) Jackson, R. M.; Gelpi, J. L.; Cortes, A.; Emery, D. C.; Wilks, H. M.; Moreton, K. M.; Halsall, D. J.; Sleigh, R. N.; Behanmartin, M.; Jones, G. R.; Clarke, A. R.; Holbrook, J. J. *Biochemistry* **1992**, *31*, 8307.
(41) Madern, D.; Ebel, C.; Mevarech, M.; Richard, S. B.; Pfister, C.; Zaccai, G. *Biochemistry* **2000**, *39*, 1001.
(42) Read, J. A.; Winter, V. J.; Eszes, C. M.; Sessions, R. B.; Brady, R. L. *Proteins: Struct., Funct., Genet.* **2001**, *43*, 175.
(43) Turner, A. J.; Moliner, V.; Williams, I. H. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323.
(44) Moliner, V.; Turner, A. J.; Williams, I. H. *Chem. Commun.* **1997**, 1271.
(45) Ferrer, S.; Ruiz-Pernía, J. J.; Tuñón, I.; Moliner, V.; Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Chem. Theor. Comput.* **2005**, *1*, 750.
(46) Ferrer, S.; Silla, E.; Tuñón, I.; Oliva, M.; Moliner, V.; Williams, I. H. *Chem. Commun.* **2005**, 5873.
(47) Basner, J. E.; Schwartz, S. D. *J. Am. Chem. Soc.* **2005**, *127*, 13822.
(48) Truhlar, D. G.; Gao, J. L.; Alhambra, C.; Corchado, J.; Garcia-Viloca, M.; Sanchez, M. L.; Kesavan, L. D.; Poulsen, T.; Villa, J. *Abstr. Pap. Am. Chem. Soc.* **2002**, *224*, U480.
(49) Wigley, D. B.; Gamblin, S. J.; Turkenburg, J. P.; Dodson, E. J.; Piontek, K.; Muirhead, H.; Holbrook, J. J. *J. Mol. Biol.* **1992**, *223*, 317.
(50) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(51) Gilson, M. K. *Proteins: Struct., Funct., Genet.* **1993**, *15*, 266.
(52) Antosiewicz, J.; Mccammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415.
(53) Field, M. J.; Amara, P.; David, L.; Rinaldo, D. Personal communication.

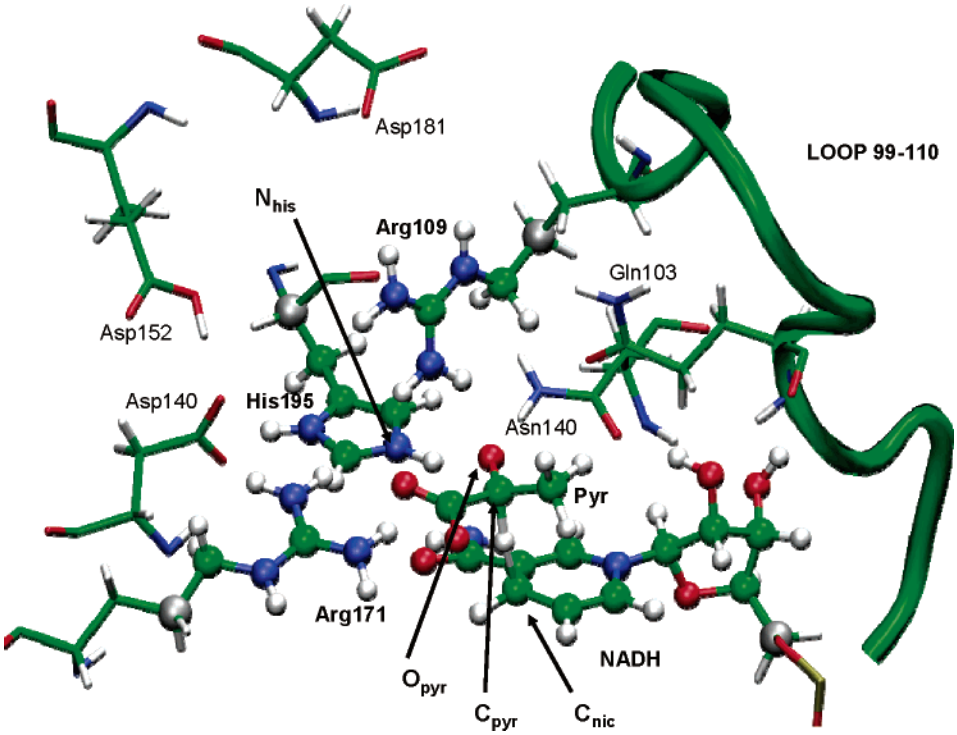


Figure 1. Transition state structure for the pyruvate to lactate transformation in the LDH active center with the cofactor (NADH): The QM region is represented using balls and sticks. The position of the GHO atoms used to define the boundary with the MM region is also indicated as gray balls.

pH = 7. Glu199 is close to Asp168, a residue that interacts directly with His195 and is responsible for maintaining its protonated state in preparation for substrate binding and turnover.[45]

Once all hydrogen atoms were properly added to the heavy atoms of the X-ray PDB structure, the system was partitioned into a QM region consisting of 78 atoms and a MM region containing the rest of the system. The QM subsystem includes the 9 atoms of the pyruvate, 13 atoms of the His195 residue (including the C_α as a boundary atom), 30 atoms of the NADH (which include the dihydronicotinamide and ribose rings and the C5′ ribose atom as a boundary atom), and 13 atoms from each of the two arginines −Arg109 and Arg171− present in the active site (including the C_γ as boundary atoms). In the preliminary calculations used to build the model of the enzymatic system, the QM subsystem was described with the use of the AM1 Hamiltonian.[54] The MM region is represented by the CHARMM27 all-atom force field[55] to represent the MM atoms of the protein, and the three-point-charge TIP3P model[56] is used for water. The boundary atoms are represented by the generalized hybrid orbital (GHO) method.[57] A picture of the active site showing the partition into QM and MM subsystems is presented in Figure 1. We followed a protocol that has been described in a previous paper to solvate the active site and relax the energetically unfavorable contacts.[45] The QM/MM van der Waals interactions were recalibrated. The CHARMM27 van der Waals parameters are expected to work between MM atoms belonging to this force field, but they may not describe correctly a QM(AM1)-MM interaction. The new parameters were obtained following the procedure described by Freindorf and Gao[58] and applied recently by our group.[45]

2.2. Potential Energy Surface and Molecular Dynamics Simulations. The chemical reaction catalyzed by LDH involves two elementary processes: a hydride transfer from a carbon atom of dihydronicotinamide, C_{nic}, to the carbonyl carbon atom of pyruvate, C_{pyr}, and a proton transfer, from the N atom of His195, N_{his}, to the carbonyl oxygen atom of pyruvate, O_{pyr}. Both processes can be described on a single reduced PES obtained as a function of two reaction coordinates. The reaction coordinate R_1 for the hydride transfer is defined in this work as the difference in the distances of the bonds between the transferring hydride and the donor and the acceptor atoms (see Scheme 1 and Figure 1):

$$R_1 = r_{C_{nic}H'} - r_{C_{pyr}H'} \quad (1)$$

For the proton transfer a reaction coordinate R_2 is defined as the difference in the distance of the bonds between the transferring proton and the donor (N_{his}) and the acceptor (O_{pyr}) atoms:

$$R_2 = r_{N_{his}H''} - r_{O_{pyr}H''} \quad (2)$$

The exploration of the PES was then carried out using the R_1 and R_2 coordinates defined before. CHARMM program was employed to carry out this exploration by means of the use of the RESDISTANCE keyword to define the reaction coordinates (R_i).

In order to represent as accurately as possible the energetics of the reaction and taking into account the limitations of the AM1 method[45] we decided to use a modified QM/MM energy function in our simulations:

$$E = E_{QM}^{0,AM1} + E_{QM/MM}^{AM1} + E_{MM} + 2D\_spline\{E_{QM}^{0,MP2}(R_1,R_2) - E_{QM}^{0,AM1}(R_1,R_2)\} \quad (3)$$

where the superscript 0 indicates the gas-phase energy. In this expression we have included a correction term to the gas-phase quantum energy as a function of the R_1 and R_2 coordinates. This correction, calculated as the difference between MP2/6-31G(d,p) and AM1 calculations for a subset of the QM system on structures picked up from the AM1/MM PES,[45] is fitted through the use of a two-dimensional

(54) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(55) MacKerell, A. D.et al. *J. Phys. Chem. B* **1998**, *102*, 3586.
(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(57) Gao, J. L.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714.
(58) Freindorf, M.; Gao, J. L. *J. Comput. Chem.* **1996**, *17*, 386.

cubic spline function,[59] ensuring then the continuity of the energy and the first and second derivatives.

During the simulations, those 14996 atoms >24 Å away from the active site were kept frozen in order to reduce the computational cost, while 7143 atoms were allowed to move. In all the cases, the total energy of the system was minimized with the ABNR[60] method until the norm of the gradient was less than 0.001 kcal·mol^{-1}·Å^{-1}. The GRACEFUL algorithm[61] was used to optimize and characterize an initial transition state structure using a Hessian matrix containing all the coordinates of the QM subsystem,[61] while the gradient norm of the remaining movable atoms was maintained less than 0.01 kcal·mol^{-1}·Å^{-1}.

In order to obtain additional TS structures, we carried out stochastic boundary molecular dynamics (SBMD) simulations in the transition state region, harmonically restraining the R_1 and R_2 coordinates and using Langevin dynamics with the atoms placed more than 24 Å away from the active site kept fixed, and those placed between 20 and 24 Å defined the buffer region. The simulation protocol is described in detail elsewhere,[45,62] and the time step was 1 fs. Some of the configurations generated during the dynamics were used as the starting points for the localization and characterization of new transition state structures. From these structures located with the GRACEFUL algorithm, the intrinsic reaction coordinates (IRCs)[63] were traced by going down to the valleys of the reactants and products in mass-weighted Cartesian coordinates. With this purpose new structures were generated following the gradient vector, except in the vicinity of the transition state structure where the transition vector was used instead. To obtain the free energy profiles associated with some of the reaction paths using a free energy perturbation scheme (see section 3.3.), SBMD simulations were carried out for selected structures appearing between the reactant and the product structures, keeping fixed the coordinates of the QM subsystem. In order to calculate the primary kinetic isotope effect (KIE) for the substitution of the transferred hydride by a deuteride (NAD^{2}H) we introduced quantum effects on the calculation of the rate constants. Thus, the quantum vibrational energy of the quantum part of the system has been evaluated and added to each free energy profile.[64]

In addition, long simulations (500 ps) were carried out for some of the reactant complexes without fixing any coordinate of the QM region in order to characterize the averaged properties of these complexes. In these later simulations, structures were saved every 500 steps for further structural analysis. Unless otherwise noted, a temperature of 298 K has been considered throughout this paper.

2.3. Calculation of the Semi-classical Transmission Coefficient. The contribution of dynamical quantum effects on the rate constant was introduced through the semi-classical transmission coefficient, κ, which has been evaluated for each reaction path by the frozen bath approximation. In this way we tried to account for quantum mechanical tunneling and nonclassical reflections. Contributions of recrossing were not considered here. Insofar as we followed the reaction paths using a quite complete reaction coordinate (defined by all the atoms of the QM system), the corresponding transmission factor should be maximized, and then it should be close to unity. The calculations described in this section were carried out with the CHARMMRATE-2.0 module[65] of CHARMM,[50] which is based on an interface of CHARMM and POLYRATE.[66]

The N-atom system was divided into N_1 primary zone atoms and N_2 secondary zone atoms, where $N_2 = N - N_1$, and N is the total number of atoms. In the frozen bath approximation, only the dynamical effects of the *primary* zone atoms are *explicitly* included in the calculation. In the present work, we tested three model systems that included, respectively, $N_1 = 37$, 52, and 78. For a given TS structure and a specific isotope substitution, all three models yielded similar potential energy barriers, imaginary vibrational frequencies, and semi-classical transmission coefficients, and we decided to use the more efficient 37 atoms primary zone model for further calculations. For each transition state structure, we froze the secondary zone atoms and calculated an isoinertial IRC through the 3×37 dimensional primary zone by treating it as a cluster embedded in the potential field of the fixed secondary zone atoms. In each case, the same initial transition state structure was used for both hydride and deuteride shifts to trace two different IRCs, one for each isotope substitution. From the individual IRCs, individual semi-classical transmission coefficients could be calculated by employing methods fully described in previous works.[23] Geometry optimizations of the primary zone reactant and product species were performed using the BFGS method, whereas the saddle point geometry was relocated using the Newton−Raphson method with Brent line minimization[67] with the gradient components converged to 5×10^{-8} au. The IRCs were traced using the Euler steepest descent method[68] in mass-scaled coordinates with a reduced mass of 1 amu and a step size of 0.005 bohr. The individual semi-classical transmission coefficients $\kappa(T)$ were microcanonically optimized between the small-curvature tunneling (SCT) and the large-curvature tunneling (LCT) values as has been done in previous works.[69−71]

3. Results

3.1. Exploring the Transition State Region. Our computational study begins with the exploration of the transition state region. For this purpose we selected as initial configuration of the system a transition state structure located and characterized after exploration of a reduced PES defined through the use of the R_1 and R_2 coordinates.[45] This transition state structure corresponds to concerted hydride and proton transfers where the former is considerably more advanced than the latter. In this structure the pyruvate is fixed in the active site through interactions of the carboxylate group with Arg171. The carbonyl oxygen atom of pyruvate establishes hydrogen-bond interactions with Arg109, although this role can also be played by Asn140. Other key residues, the proton donor His195 and the NADH cofactor, are also shown in Figure 1. In order to explore the TS region we ran SBMD simulations of the system harmonically restraining the values of the R_1 and R_2 coordinates. After heating at 298 K we performed two different simulations at slightly different temperatures (298 and 300 K) in order to enhance the exploration over a wider range of configurations of the system. These two simulations in fact explored regions of the configurational space with important differences from the point of view of catalysis. Figure 2 shows the evolution of the CA−CB−CD−CG and CA−CB−CD−N dihedral angles of Arg109 and Asn140 in both SBMD simulations. The different orientation reached in the 300 K SBMD after some picoseconds corresponds

(59) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. *J. Phys. Chem. B* **2004**, *108*, 8427.
(60) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J. *J. Comput. Chem* **1983**, *4*, 187.
(61) Martí, S.; Moliner, V.; Tuñón, I. *J. Chem. Theor. Comput.* **2005**, *1*, 1008.
(62) Poulsen, T. D.; Garcia-Viloca, M.; Gao, J. T., D. G. *J. Phys. Chem. B* **2003**, *107*, 9567.
(63) Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363.
(64) Garcia-Viloca, M. A., C.; Truhlar, D. G.; Gao, G. *J. Chem. Phys.* **2001**, *114*, 9953.
(65) Garcia-Viloca, M.; Alhambra, C.; Corchado, J. C.; Sánchez, M. L.; Villà, J.; Gao, J.; Truhlar, D. G. CHARMMRATE-version 2.0; University of Minnesota,: Minneapolis, 2002.
(66) Corchado, C. J. et al. D. G. POLYRATE-version 9.0; University of Minnesota: Minneapolis.
(67) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes*; Cambridge University Press: New York, 1986; Vol 412.
(68) Garrett, B. C.; Redmon, M. J.; Steckler, R.; Truhlar, D. G.; Baldridge, K. K.; Bartol, D.; Schmidt, M. W.; Gordon, M. S. *J. Phys. Chem.* **1988**, *92*, 1476.
(69) Liu, Y. P.; Lynch, G. C.; Truong, T. N.; Lu, D. H.; Truhlar, D. G.; Garrett, B. C. *J. Am. Chem. Soc.* **1993**, *115*, 2408.
(70) Fernandez-Ramos, A.; Truhlar, D. G. *J. Chem. Phys.* **2001**, *114*, 1491.
(71) Liu, Y. P.; Lu, D. H.; Gonzalez-Lafont, A.; Truhlar, D. G.; Garrett, B. C. *J. Am. Chem. Soc.* **1993**, *115*, 7806.

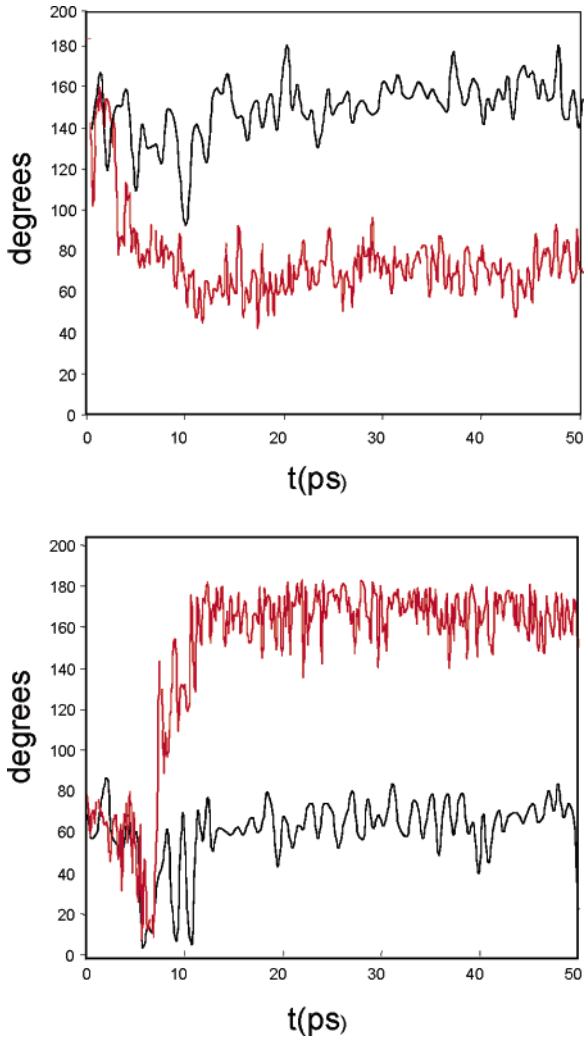


Figure 2. Time evolution of the CA−CB−CD−CG angle of Arg109 (top) and CA−CB−CD−N angle of Asn140 (bottom) during SBMD simulations in the transition state region at 298 K (black) and 300 K (red).

to a different pattern of hydrogen bonds with the carbonyl oxygen atom of pyruvate. In the 298 K SBMD Arg109 establishes a hydrogen bond with the carbonyl oxygen atom of the substrate (O_{pyr}), but in the 300 K SBMD this residue is displaced by Asn140 after some picoseconds. This competition between these two residues has already been reported in our previous computational studies,[43] and it was proposed as a redundant stabilization mechanism that could explain the fact that mutation of Arg109 does not completely destroy the catalytic activity of the enzyme.[36]

Different snapshots of the two SBMD simulations were selected as starting points to locate and characterize transition state structures of the process. The transition state structure search was carried out by means of a combination of graceful and CHARMM programs, defining as active space the coordinates of all the QM atoms, while the rest of coordinates of the mobile atoms form the complementary space. The stationary structures obtained in this way (with zero or nearly zero gradient norms in both the active and complementary spaces) present one and only one imaginary frequency in the active space. We obtained up to 24 transition state structures that are gathered in Table 1. These transition state structures are characterized by the bond-breaking and bond-forming distances corresponding to the proton and hydride transfer, the hydrogen-bond distances to either Arg109 or Asn140, the total potential energy, the value of the imaginary frequency, as well as by the potential energy barrier. The structures located display an important dispersion in these properties. In all cases the transition state structure obtained corresponds to a concerted (but asynchronous) process where the local curvature seems extremely sensitive to fluctuations of the system: the imaginary frequency changes by a factor of 10, from 115i to 1186i cm^{-1}. The most probable value of the imaginary frequencies appears to be about 300i cm^{-1}, but some values are found to be as high as 1200i cm^{-1}. With respect to the hydride transfer the bond-breaking distance ($C_{nic}H'$) ranges between 1.39 and 1.69 Å, whereas the bond-forming distance ($C_{pyr}H'$) is found in a narrower interval, between 1.24 and 1.36 Å. In the case of the proton transfer, the $N_{his}H''$ distance varies between 1.08 and 1.20 Å, and the pyruvate carbonyl oxygen−proton ($O_{pyr}H''$) distance, between 1.35 and 1.66 Å. In terms of the hydride-transfer and proton-transfer coordinates, the dispersion of the transition state structures presented in Table 1 is graphically depicted in Figure 3 where the located structures are represented as a function of R_1 and R_2 transfer coordinates. For all of these structures we traced the IRC going down to the products and reactants valleys. It is evident that the coordinates defining the position of the shifting hydride and the shifting proton always have a very important participation in all the reaction coordinates. However, we have seen in our calculations that the particular composition of the reaction coordinate in terms of those coordinates and the rest of mobile atoms along the 24 distinct reaction paths we have built up depends on each reaction path. The energy barriers obtained as the potential energy difference between the transition state and the reactant structures ranges between 17 and 83 kcal·mol^{-1}. Such a wide distribution of energy barriers is not completely surprising. Very similar results were obtained for the energy barrier of the hydride transfer in dihydrofolate reductase (DHFR), an enzyme that also has a flexible loop in the active site.[19,21]

3.2. Reactant Valleys Analysis. The basic idea exploited in this paper is that reactant enzyme−substrate complexes displaying important differences can be obtained starting from different transition state structures. Effectively, downhill IRCs started at different starting points (the transition state structures) and following different directions (the transition vectors) can lead to quite different reactant valleys. This procedure may be assimilated to a special kind of "chemical simulated annealing" to find different system minima relevant for the chemical process under study. To check this point we selected four reactant structures obtained from the IRCs of the transition state structures appearing in Table 1. In particular we selected the four first transition structures appearing in Table 1 (hereafter denoted as TS1, TS2, TS3, and TS4). We believe that with this selection we cover a significant interval of possible reaction barriers (from 17 to 36 kcal·mol^{-1}; transition state structures involving higher energy barriers are irrelevant to the kinetics). Anyway, it should be here pointed out that this procedure is designed to locate different reactant valleys, but it is not expected to provide an adequate sampling of the full configurational space of the reactant state. Molecular dynamics simulations can be easily carried out within each valley, and then these can be adequately explored, but the relative weighting for each valley remains undetermined (see below).

Table 1. Selected Distances[a] (in Å), Potential Energies (in kcal·mol^{-1}), Imaginary Frequencies (in cm^{-1}), and Potential Energy Barriers (in kcal·mol^{-1}) of the Located Transition State Structures for the Chemical Reaction Catalyzed by LDH

$d(C_{nic}H')$	$d(C_{pyr}H')$	$d(N_{his}H'')$	$d(O_{pyr}H'')$	$d(O_{pyr}HH_{Arg109})$	$d(O_{pyr}HD_{Asn140})$	energy	freq	$\Delta E^{\ddagger}$
1.43	1.31	1.09	1.57	3.04	2.01	−17907	776i	17
1.39	1.36	1.10	1.59	2.20	1.88	−18213	1186i	23
1.58	1.24	1.20	1.35	4.49	1.88	−18382	630i	36
1.39	1.35	1.08	1.66	2.14	1.95	−18166	1095i	34
1.63	1.24	1.16	1.43	4.39	1.82	−18306	463i	32
1.69	1.23	1.15	1.43	4.03	1.88	−18330	356i	83
1.61	1.25	1.14	1.45	4.39	1.86	−18355	557i	37
1.58	1.25	1.17	1.41	4.60	1.88	−18384	352i	34
1.57	1.25	1.15	1.43	4.85	1.84	−18410	620i	22
1.57	1.24	1.18	1.38	3.85	1.89	−18419	222i	45
1.54	1.26	1.12	1.52	4.94	2.00	−18419	194i	23
1.52	1.27	1.11	1.55	4.99	1.85	−18419	257i	23
1.55	1.25	1.12	1.52	5.40	1.87	−18471	215i	40
1.53	1.26	1.12	1.53	4.81	1.85	−18446	285i	18
1.55	1.25	1.14	1.45	4.92	1.86	−18421	352i	22
1.57	1.25	1.15	1.43	5.14	1.85	−18426	502i	38
1.56	1.24	1.18	1.37	4.85	1.84	−18410	542i	22
1.54	1.25	1.18	1.37	4.75	1.84	−18429	285i	41
1.55	1.25	1.14	1.45	4.90	1.87	−18380	317i	38
1.53	1.26	1.13	1.48	5.19	1.85	−18426	477i	56
1.55	1.26	1.16	1.41	4.63	1.85	−18402	721i	21
1.40	1.34	1.08	1.65	2.10	1.97	−18183	1059i	19
1.41	1.32	1.08	1.62	2.29	1.94	−18188	850i	45
1.55	1.24	1.11	1.52	2.30	2.09	−18327	115i	62

[a] If the atom number is not provided, the distance refers to the closest atom among all of the same type.

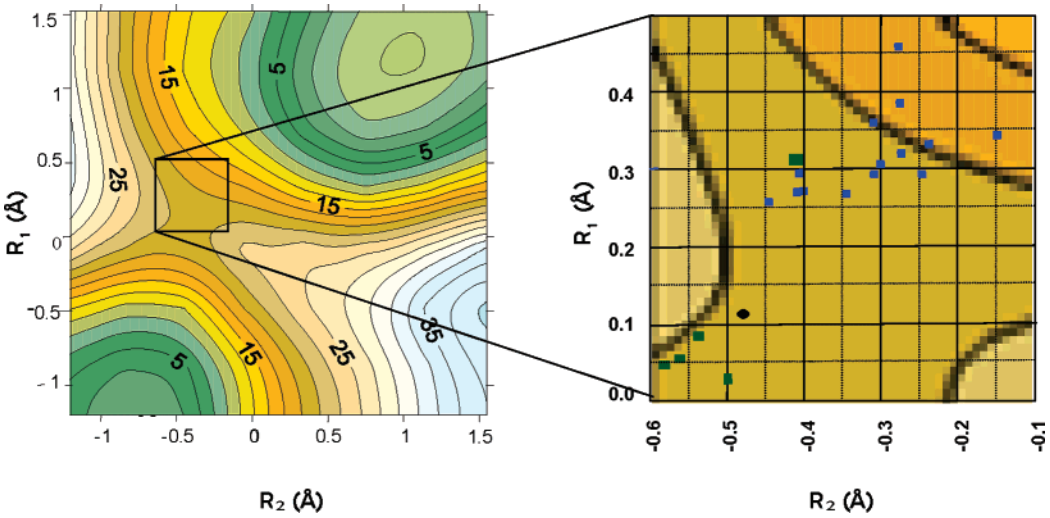


Figure 3. Representation of the located transition state structures on the reduced PES (in kcal·mol^{-1}). The initial transition state structure is represented as a black dot, those coming from the 298 K simulation as green dots, and blue dots represent those coming from the 300 K simulation. The isoenergetical lines represent the PES from which we start to locate the initial transition state structure.

Relatively long SBMD simulations (500 ps) were run for each of those reactant valleys. The analysis of these trajectories effectively showed significant differences during the simulated time. Of course, one may wonder if longer simulation times will show significant overlap of the four trajectories in the configurational space. However, from a practical point of view, and considering the typical simulation times used to obtain free energy profiles, we can consider these complexes as different valleys with each of them contributing in a distinct way to the macroscopic reaction flux. Table 2 provides the averaged values and fluctuations (measured as standard deviations of the averaged property) of some geometrical parameters related to pyruvate obtained during the dynamics of these four reactant complexes.

With respect to the hydrogen-transfer parameters, the donor−hydrogen distances ($C_{nic}H'$ and $N_{his}H''$) are nearly identical in the four simulations. The fluctuations and the differences are much larger for the acceptor−hydrogen distances, obviously because of the significantly lower force constants associated with these interactions. In particular, the reactant obtained from

Table 2. Averaged Geometries[a] of the Four Reactant Complexes Analyzed

	REAC1	REAC2	REAC3	REAC4
$d(C_{nic}H')$	1.13 (0.03)	1.13 (0.02)	1.13 (0.02)	1.13 (0.03)
$d(C_{pyr}H')$	4.3 (0.7)	3.3 (0.3)	3.5 (0.4)	3.2 (0.2)
$d(N_{his}H'')$	1.00 (0.02)	1.01 (0.02)	1.01 (0.02)	1.01 (0.02)
$d(O_{pyr}H'')$	3.0 (0.2)	3.4 (0.4)	3.1 (0.2)	3.0 (0.2)
$d(O_{pyr}H_{nic})$	3.9 (1.0)	2.5 (0.3)	3.2 (0.7)	2.7 (0.4)
$d(O_{pyr}HH_{Arg}109)$	7.1 (0.4)	6.9 (0.6)	6.0 (0.6)	6.7 (0.6)
$d(O_{pyr}HD_{Asn}140)$	6.3 (1.4)	2.6 (0.3)	4.8 (0.7)	6.1 (0.9)
$d(O_{pyr}HH_{Arg}171)$	5.4 (0.2)	5.5 (0.5)	4.0 (0.2)	5.2 (0.5)

[a] Distances, with their standard deviations in parenthesis, are given in Å.

TS1 (REAC1) displays a $C_{pyr}H'$ distance larger than those in the other reactant complexes. The standard deviation is also significantly larger, reflecting the fact that wider fluctuations are observed for REAC1. Part of this difference could be attributed to a compression effect in the other reactants due to two hydrophobic residues (Ile253 and Val33) placed at the other side of the nicotinamide ring of NADH with respect to pyruvate. In a recent study on the human heart isozyme, h-H_4LDH, it was suggested that a residue occupying a similar position (Val31) could play an active role favoring the hydride transfer.[47,72] Analysis of the averaged minimum distance between C_{nic} and carbon atoms of these two residues (Ile253 and Val33) showed that REAC2 and REAC3 display the shorter distances to Ile253 (3.7 and 3.8 Å, respectively), whereas REAC4 presents the shortest distance to Val33 (4.9 Å). It then seems that in REAC1 there is somewhat more room for the nicotinamide ring, and this could contribute to the larger value observed in the standard deviation of the $C_{pyr}H'$ distance. More significant differences that could also explain this characteristic found in REAC1 simulation are found in the analysis of pyruvate−enzyme interactions. In fact, REAC1 trajectory presents the largest value for the averaged rms displacement of pyruvate (1.04, 1.00, 0.84, and 0.83 Å for REAC1−REAC4, respectively). In the four reactant trajectories analyzed here the substrate is anchored by means of hydrogen-bond interactions between the pyruvate carboxylate group and Arg171. This interaction is well conserved during the four simulations. The differences appear in the interactions of the carbonyl oxygen (O_{pyr}). The pattern of interactions established by this atom in each of the simulations is quite different. In all cases there is a strong hydrogen-bond interaction with the protonated His195, the residue responsible for the proton transfer to the substrate. The function and positioning of this residue is modulated by a negatively charged Asp168 which, in turns, establishes strong contacts with Arg171 and a protonated Glu199. The pK_a of this last residue is considerably shifted from its standard value in aqueous solution due precisely to the proximity of Asp168, an essential feature for the energetics of the reaction and that was addressed in a previous study.[46] Coming back to the description of the pyruvate carbonyl oxygen interactions, it is interesting to note that in none of the reactant simulations does Arg109 establish hydrogen-bond interactions with this atom. Considering that this interaction appears in the transition state structures (at least in TS2 and TS4, as discussed in the previous section) one may conclude that the coordinates of this residue should be included in the correct definition of the reaction coordinate, at least for some of the possible reaction paths. The O_{pyr}−Arg109 hydrogen bond observed in some of the transition state structures must be established during the reaction progress. Considering that the hydride transfer precedes the proton transfer in the concerted asynchronous transfer, it could be argued that an additional negative charge (or fraction of charge) on the pyruvate could be the driving force needed to approach this residue. Otherwise, Asn140, a residue candidate to explain the conserved activity of Arg109 mutants,[36] has hydrogen-bond interactions with O_{pyr} during REAC2 simulation. For this reactant valley, this hydrogen bond is kept along the reaction path and then it is also observed in the corresponding transition state structure. Conversely, this hydrogen bond does not exist in any of the other reactant complexes, although it does appear at the corresponding transition state structures (see Table 1). As discussed in the analysis of transition states trajectories, the orientation of the side chain of these two residues is highly correlated and they seem to alternate in the transition state stabilization. In the case of REAC2 and REAC4 we also observed a weak hydrogen bond between the carbonyl oxygen of pyruvate and one of the hydrogen atoms of the amide group of the nicotinamide ring ($O_{pyr}H_{nic}$, distance in Table 2). This hydrogen bond may play a role not only polarizing the carbonyl bond but also assisting the positioning of NADH. In particular, REAC2 and REAC4 are the reactant complexes displaying shorter average distances between the carbon atom of pyruvate and the hydride ($C_{pyr}H'$). Finally, it is interesting to comment on a particular feature observed in REAC3 simulation. In this case the substrate adopts a slightly different orientation in the active site; the pyruvate molecule is slightly rotated around a perpendicular axis to the molecular plane in such a way that a NH_2 moiety of Arg171 is placed in between a carboxylate oxygen and the carbonyl oxygen. In this case we can then occasionally observe a hydrogen-bond interaction between O_{pyr} and Arg171, as reflected in the reduced averaged distance provided in Table 2.

In addition to the reported enzyme−substrate interactions, we have found differences in the interactions with water molecules. Using the CONTACT keyword of CHARMM[50] we computed the average number of water molecules hydrogen bonded to the QM atoms of our system. This averaged number was close to 8 for the simulations corresponding to REAC1, REAC2, and REAC4, but it was 10.5 for REAC3. The standard deviation in all cases was around unity. One of the additional water molecules in REAC3 is hydrogen bonded to the O_{pyr} atom of the substrate. This interaction is found in 32% of the configurations analyzed, and for this subset the $O_{pyr}H_w$ averaged distance was 2.1 ±0.2 Å. An interaction of this type between the substrate and water molecules was found for only a 3% of structures corresponding to the simulation of REAC1 and much less frequently in REAC2 and REAC4 simulations. The other additional water molecule found in REAC3 simulation is placed in the solvation shell of the ribose ring of NADH.

We also found significant differences in the simulations of the four reactant complexes when analyzing the conformations presented by some residues, in particular, Arg109. As explained before, this residue is part of the flexible loop closing the active site. Figure 4 shows the time evolution of the CA−CB−CD−CG dihedral angle of Arg109 during the 500 ps simulations of the four reactant complexes. While in REAC2 and REAC4 this dihedral angle fluctuates around 90°s, in the case of REAC1 and REAC3 it is close to 180°. This orientation of the Arg109 side chain is kept during all the simulation. Accordingly, this residue is able to establish different interactions in each case. Table 3 provides the averaged hydrogen-bond distances and standard deviations between Arg109 and some residues and water molecules. Both in REAC2 and REAC4 we have observed a water molecule hydrogen bonded to Arg109 during the simulation. This feature was not observed in REAC1 and REAC3 simulations, which can be related to the different disposition of the side chains. In addition, Arg109 also forms hydrogen bonds with the carboxylate group of Asp181 in REAC 1 and REAC2, and with the carboxylate group of Gln103 in REAC2.

(72) Basner, J. E.; Schwartz, S. D. *J. Phys. Chem. B* **2004**, *108*, 444.

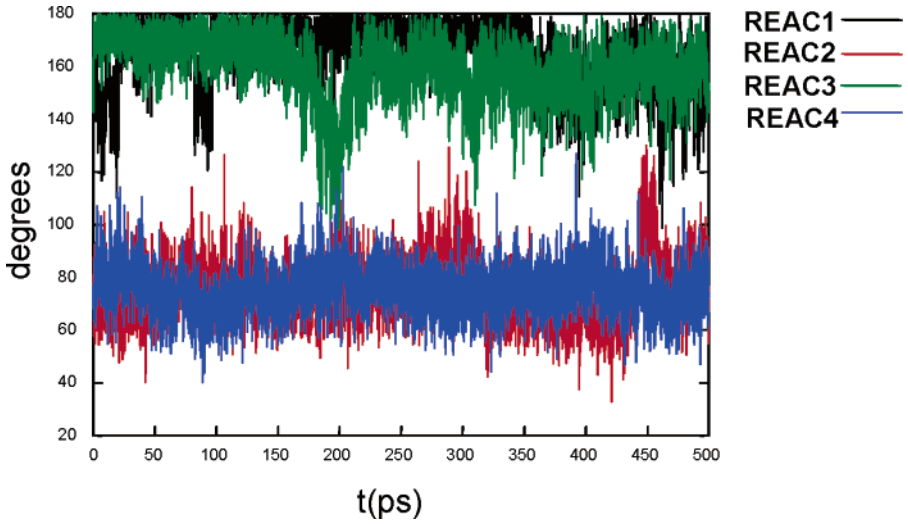


Figure 4. Time evolution of the CA−CB−CD−CG dihedral angle of Arg109 during the 500 ps simulations of the four reactant complexes

Table 3. Averaged Distances (in Å) between Arg109 and other Residues and Water Molecules, and their Standard Deviations in Parentheses

	REAC1	REAC2	REAC3	REAC4
d(HH11$_{Arg109}$−OD1$_{Asp181}$)	5.2 (0.2)	3.5 (1.3)	7.1 (0.8)	8.2 (0.8)
d(HH12$_{Arg109}$−OD1 $_{Asp181}$)	5.4 (0.3)	3.2 (1.6)	6.5 (1.5)	8.1 (1.1)
d(HH21$_{Arg109}$−OD1 $_{Asp181}$)	2.6 (0.7)	5.5 (0.6)	4.2 (0.6)	6.0 (0.5)
d(HH22$_{Arg109}$−OD1 $_{Asp181}$)	4.2 (0.6)	4.8 (1.0)	4.8 (1.3)	6.9 (0.6)
d(HE$_{Arg109}$−OE1$_{Gln103}$)	5.1 (0.3)	3.7 (1.7)	5.4 (0.3)	6.2 (0.9)
d(HH$_{Arg109}$−OH$_{Tyr239}$)	3.2 (0.4)	3.3 (0.8)	2.8 (0.8)	3.0 (0.4)
d(HH$_{Arg109}$−O$_w$)	6.5 (0.5)	2.5 (0.7)	4.7 (0.5)	2.7 (0.5)

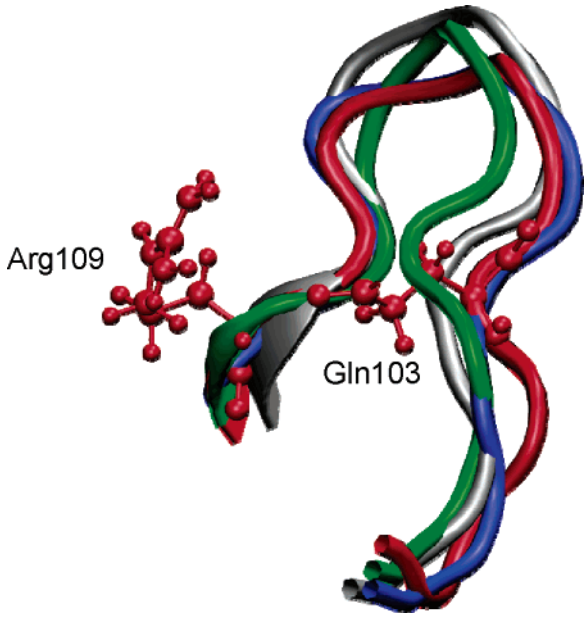


Figure 5. Overlapping of the flexible loop at the end of the four reactant complexes simulations. The color code is the same as in Figure 4. The position of Arg109 and Gln103 in REAC2 is also shown.

In all the reactant complexes a hydrogen-bond interaction between the side chains of Arg109 and Tyr239 is found.

The positioning of the flexible loop closing the active site at the end of the SBMD simulations of the four reactant complexes is represented in Figure 5. A different disposition of the loop for each reactant complex can be observed as a consequence of the different values adopted by the corresponding dihedral angles of the backbone, as seen in Tables 4 and 5, especially in what

Table 4. Average Values of the Backbone Dihedral Angle N_i−CA_i−C_i−N_{i+1} (deg) for Residues 99−110 Belonging to the Active Site Loop, and Standard Deviations in Parenthes

residue	REAC1	REAC2	REAC3	REAC4
99	139.7 (19.4)	159.1 (12.5)	159.2 (8.4)	156.8 (13.9)
100	160.3 (11.1)	121.5 (23.1)	147.8 (13.3)	151.5 (11.2)
101	160.8 (9.6)	161.9 (7.2)	170.8 (5.9)	170.1 (5.9)
102	165.3 (7.8)	142.7 (11.4)	163.8 (10.5)	161.6 (8.0)
103	150.1 (8.5)	136.1 (24.8)	126.3 (12.0)	149.2 (14.5)
104	163.8 (8.5)	160.0 (6.0)	167.0 (6.2)	162.2 (8.1)
105	151.5 (11.0)	79.7 (63.6)	124.1 (40.2)	151.4 (17.0)
106	128.4 (15.0)	156.6 (14.7)	168.2 (7.1)	147.5 (15.3)
107	75.5 (10.6)	67.7 (22.5)	112.5 (14.9)	35.2 (21.3)
108	70.9 (15.4)	109.3 (11.2)	21.4 (7.2)	108.9 (10.0)
109	22.8(8.1)	48.4 (6.6)	28.8 (6.8)	51.0 (6.7)
110	33.6 (8.7)	32.8 (7.7)	39.3 (6.7)	38.4 (8.6)

Table 5. Average Values of the Backbone Dihedral Angle C_{i+1}−N_i−CA_i−C_i (deg) for Residues 99−110 Belonging to the Active Site Loop and Standard Deviations in Parentheses

residue	REAC1	REAC2	REAC3	REAC4
99	123.3 (11.7)	88.2 (12.7)	107.5 (13.1)	124.0 (10.7)
100	148.1 (21.0)	166.8 (8.6)	121.6 (17.9)	143.9 (17.6)
101	97.9 (11.9)	86.9 (11.8)	92.4 (10.0)	106.0 (11.6)
102	116.8 (14.3)	80.5 (10.0)	88.2 (12.3)	93.3 (11.3)
103	90.4 (12.5)	95.9 (10.0)	88.4 (9.8)	81.8 (8.6)
104	80.7 (8.8)	103.0 (27.5)	92.5 (14.3)	104.1 (22.8)
105	68.0 (6.4)	71.3 (8.3)	68.5 (7.2)	65.4 (5.9)
106	116.7 (14.5)	105.1 (20.6)	112.8 (19.8)	119.4 (19.6)
107	133.6 (13.2)	93.4 (13.3)	81.2 (9.2)	107.8 (16.4)
108	135.0 (10.9)	108.8 (22.6)	140.7 (12.5)	71.2 (24.1)
109	96.3 (10.0)	146.6 (11.5)	60.1 (7.0)	143.0 (12.1)
110	64.0 (8.0)	69.6 (8.3)	69.5 (6.5)	70.2 (9.1)

refers to the residues preceding Arg109. The differences in the averaged dihedral angles are significantly larger than the fluctuations observed for some of these residues. The Arg109-Gln103 interaction, observed in the simulation corresponding to the REAC2 complex, deserves some especial attention. The existence of this interaction leads to a slightly different disposition of the loop. In general, conformational changes in this loop can have dramatic consequences on the reaction energetics as will be seen below.

3.3. Free Energy Calculations. The fact that different reactant valleys may contribute to the observed reaction flux could be incorporated, in principle, by means of sufficiently

long sampling trajectories. Considering that the conversion rate between two different reactant complexes can be even slower than the reaction rate, this is not a practical way to explore the different reaction channels except in very simple cases. Another possibility is to identify some of the reactant valleys contributing to the reaction and then to obtain the free energy profiles corresponding to each case by means of much shorter simulations. Finally, the resulting free energy barrier should be conveniently combined, depending on the relative values among the different reactant complexes interconversions and the chemical reaction rates.

In order to obtain free energy profiles corresponding to the pyruvate-to-lactate transformation for each of the reactant valleys analyzed above we first tried to use the umbrella sampling technique. This technique allows, in principle, for a correct sampling of all the degrees of freedom of the system except, obviously, the reaction coordinate. We tried to trace the free energy profiles using umbrella sampling with different definitions of the reaction coordinate of the type:

$$\mathrm{RC} = \alpha_1 r_{\mathrm{C_{nic}H'}} - \alpha_2 r_{\mathrm{C_{pyr}H'}} + \alpha_3 r_{\mathrm{N_{his}H''}} - \alpha_4 r_{\mathrm{O_{pyr}H''}} \quad (4)$$

However, the results were quite discouraging. While we were able to control the behavior of the hydride transfer (monitoring the corresponding transfer coordinate, we observed a soft and continuous variation as the reaction coordinate was changed), the proton transfer invariably took place suddenly, being transferred from His195 to pyruvate in a short time interval of a single window. Attempts to control the variation of the proton-transfer coordinate changing the values of the α_1 weights appearing in the definition of the global reaction coordinate (eq 4) always failed. The free energy profiles obtained in this way displayed unexpectedly high reaction free energy barriers (corresponding to exploration of high-energy regions of the PES) and discontinuous variations along the reaction coordinate. We then decided to employ a different strategy based on the use of the free energy perturbation (FEP) technique. The configurations for which the free energy difference is estimated correspond to those structures obtained along the IRC calculation and are thus characterized by a single coordinate:

$$s = [\sum_{i \in \mathrm{QM}} m_i((x_i - x_{0i})^2 + (y_i - y_{0i})^2 + (z_i - z_{0i})^2)]^{1/2} \quad (5)$$

where x_{0i}, y_{0i}, z_{0i} are the Cartesian coordinates of the QM atoms in the transition state structure,whereas x_i, y_i, and z_i are the coordinates of a structure belonging to the IRC traced from this transition state structure, and m_i are the masses of the atoms. Within this treatment the free energy relative to the reactant can be expressed as a function of the s coordinate as:

$$\Delta G_{\mathrm{FEP}}(s^R \rightarrow s^j) = \Delta E_{\mathrm{QM}}^{0,R\rightarrow j} + \Delta G_{\mathrm{QM/MM}}^{R\rightarrow j} = (E_{\mathrm{QM}}^0(s^j) - E_{QM}^0(s^R)) - k_{\mathrm{B}}T \sum_{i=R}^{i=j-1} \ln\langle \exp \beta(E_{\mathrm{QM/MM}}(s^{i+1}) - E_{\mathrm{QM/MM}}(s^i))\rangle_{\mathrm{MM},i} \quad (6)$$

where E_{QM}^0 is the gas-phase energy of the QM subsystem, including the MP2 two-dimensional correction depending on the R_1 and R_2 coordinates, k_{B} is the Boltzmann's constant, T is temperature, and $\beta = 1/k_{\mathrm{B}}T$. The QM/MM interaction contribution to the free energy difference between two different values of s is obtained by averaging the QM/MM interaction energy (including the polarization energy) overall the MM coordinates of the system obtained for a particular value of the s coordinate. Then we added a term to include the contribution of the vibration of all the QM coordinates (except s) to the free energy profile. This contribution is estimated assuming a quantum harmonic treatment for the vibrational modes of the QM subsystem in the field created by the rest of the enzyme. Then, the quasi-classical free energy profile can be obtained adding the contribution of the QM vibrational degrees of freedom as:

$$G(s^j) - G(s^R) = \Delta E_{\mathrm{QM}}^{0,R\rightarrow j} + \Delta G_{\mathrm{QM/MM}}^{R\rightarrow j} + \Delta G^{j}_{\mathrm{vib,QM}} \quad (7)$$

where

$$\Delta G^j_{\mathrm{vib,QM}} = \sum_{x=1}^{3N-7}\left[\frac{1}{2}h\nu^j_x + k_{\mathrm{B}}T \ln\left(1 - \exp\left(-\frac{h\nu^j_x}{k_{\mathrm{B}}T}\right)\right)\right] - \sum_{x=1}^{3N-6}\left[\frac{1}{2}h\nu^R_x + k_{\mathrm{B}}T \ln\left(1 - \exp\left(-\frac{h\nu^R_x}{k_{\mathrm{B}}T}\right)\right)\right] \quad (8)$$

The frequencies are obtained projecting out the contribution of the reaction coordinate in the Hessian for all the structures but the reactant. Thus, the vibrational term is obtained after summation over the $3N - 6$ highest-frequency vibrational modes (N being the number of QM atoms) in the reactant structure and $3N - 7$ for the rest.[73] The six lowest frequency vibrational modes are assumed to correspond to the relative motion of the MM system with respect to the QM system, which is already included in the second term of the second member of eq 7. The CVT activation free energy is then obtained as:

$$\Delta G^{\mathrm{CVT}} = \max_j(G(s^j) - G(s^R)) \quad (9)$$

In practice, to obtain the free energy profile we run SBMD simulations for the structures appearing along each reaction path. Each simulation window consisted in 2 ps of equilibration and 10 ps of production. We used up to 50 windows for each reaction path and employed the double-wide sampling technique. The free energy profiles obtained in this way for the four reaction paths analyzed here are presented in Figure 6, whereas the activation free energies are given in Table 6. The CVT activation free energies obtained in this quasi-classical approximation do not exactly correlate with the potential energies obtained from the reaction path following. The free energy ordering is FEP2 > FEP4 > FEP3 > FEP1. The term related to the vibrations of the QM subsystem does not introduce additional changes in the relative ordering among the reaction paths, but nonetheless, the quantitative importance is undeniable since it can amount to more than 3 kcal·mol^{-1}.

(73) The vibrational contribution of the QM atoms to the free energy profile was calculated for all the structures of each reaction path and afterwards fitted to a polynomial as a function of the s coordinate. In this process we took into account that the first steps in the IRC were traced following the transition vector, and then the structures closer to the transition structure were not really stationary points after projecting out the component of the reaction path coordinate. See also ref 64.

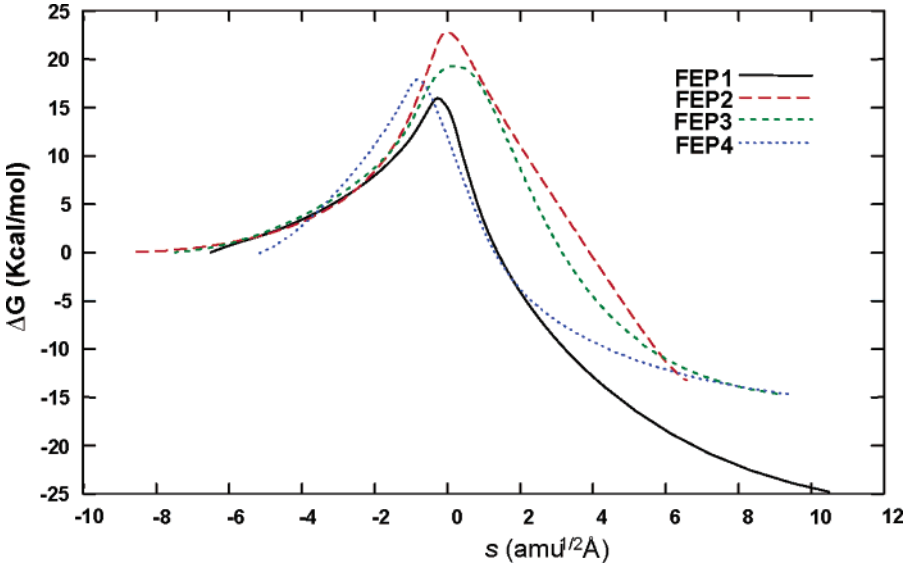


Figure 6. Free energy profiles obtained for the four reaction paths studied.

Table 6. Contributions to the Activation Free Energy (in kcal·mol^{-1}), Semi-classical Transmission Coefficients and Catalytic Rate Constants (in s^{-1}) Obtained for Each of the Four Reaction Paths Analyzed.

	FEP1	FEP2	FEP3	FEP4
ΔG_{FEP}	16.7	23.5	19.4	20.6
$\Delta G_{vib,QM}$	−3.1	−2.5	−2.6	−3.2
ΔG^{CVT}	13.6	21.0	16.8	17.4
κ	4.12	4.05	3.88	3.76
ΔG_{tun}	−0.84	−0.83	−0.80	−0.78
ΔG_{act}	12.8	20.2	16.0	16.6
k_{cat}	2771	0.0095	10.1	4.02

The aforementioned free energy ordering can be clearly correlated with the effect of Arg109 on the potential energy barrier. To illustrate this point we have obtained the difference between the gas-phase energy barrier of the QM subset of atoms calculated with and without the presence of Arg109, at the AM1 level. Calculations were carried out on the optimized transition state structures and the corresponding reactant structures located following the IRC path. In all cases Arg109 plays a stabilizing role in the TS structure relative to the reactants. However, there are significant differences between the four reaction paths. While the first attains the maximum stabilization energy (9.7 kcal·mol^{-1}), in the second (the one corresponding to REAC2) Arg109 plays a minor role (just 1.0 kcal·mol^{-1}). REAC3 and REAC4 display intermediate values (6.7 and 8.2 kcal·mol^{-1}, respectively). It can be seen that the maximum stabilization effect corresponds to the minimum free energy barrier, while the minimum Arg109 effect corresponds to the highest free energy barrier. Thus, the positioning of Arg109 is a key element in determining the reaction rate. As mentioned above, this residue belongs to an enzymatic flexible loop and it seems that different dispositions can be attained, determining the existence of different reactant valleys from which different reaction paths can be followed.

3.4. Semi-classical Transmission Coefficient. We finally evaluated the contribution of tunneling to the free energy barrier as

$$\Delta G_{tun} = -RT \ln \kappa \qquad (10)$$

where R is the gas constant and the phenomenological free energy of activation is obtained as:

$$\Delta G_{act} = \Delta G^{CVT} + \Delta G_{tun} \qquad (11)$$

The values of the semi-classical transmission coefficient and its free energy equivalencies for each of the four reaction paths are given in Table 6. In free energy terms tunneling contributes about 0.8 kcal·mol^{-1} to the free energy barrier reduction, and there are no significant differences between the four reaction paths analyzed. The maximum difference amounts to only 0.06 kcal·mol^{-1}.

3.5. Evaluation of the Reaction Rate Constants Corresponding to the Different Reaction Valleys. The rate constants are obtained through the well-known equation:

$$k_{cat} = \kappa \frac{k_B T}{h} \exp\left(-\frac{\Delta G^{CVT}}{RT}\right) = \frac{k_B T}{h} \exp\left(-\frac{\Delta G_{act}}{RT}\right) \qquad (12)$$

where h is the Planck's constant.

A simple arithmetic average of the four reaction path rate constants presented in Table 6 gives 696 s^{-1}, a value very close to the experimental lower limit estimation. Although this average could be considered as a rough indication of the order of magnitude of the macroscopic reaction rate constant, it is not the right way to proceed. As a matter of fact, each reactant valley acts as an independent enzyme, leading to the formation of the product through separated reaction paths, each one involving its own individual reaction rate constant. These individual enzymatic turnovers can presently be observed in real time by means of single-molecule enzymatic studies.[17,18,30,31] Thus, the macroscopic reaction rate of formation of the product is no longer governed by a single reaction rate constant, but by a distribution of individual reaction rate constants. The goal here is how the overall reaction rate of the ensemble of reactant valleys can be calculated. The individual reaction rate constants should be weighted according to the relative population of each reactant valley. If the interconversion rates among them are much faster than the reaction rate, an equilibrium may be assumed, and the individual rate constants must be weighted using the free energy differences among the substates. Other-

Table 7. Averaged Potential Energies and Interaction Energies between the Enzyme and NADH Plus Pyruvate, and their Respective Standard Deviation in Parentheses (in kcal·mol^{-1}) for the Four Reactant Complexes Analyzed

	REAC1	REAC2	REAC3	REAC4
E	−12181 (34)	−12241 (13)	−12265 (37)	−12261 (44)
$E_{QM/MM}$	−211 (51)	−329 (52)	−411 (44)	−322 (48)

wise, a kinetic control governs the population of each reactant valley, and the probability for the enzyme to be in one of the possible reactant valleys should be calculated after following a lot of dynamical trajectories of the substrate along the entrance channel leading to the active site.

Evaluation of the relative free energy differences among the four reactant valleys studied here is a difficult task because of the nonlocal differences appearing among them. We have not been able to find a single geometrical parameter (or combination of some of them) to follow the transformations among these reactant complexes, and the direct evaluation of the free energy differences is, for the moment, out of our possibilities. To indicate their relative stabilities we have computed the averaged potential energies as well as the interaction energies of the enzyme with NADH plus pyruvate, that is the interaction energy of the enzyme with the cofactor and the substrate. For this purpose we selected 5000 equally distributed snapshots from the 500 ps SBMD simulations of the reactant complexes. For these structures we redefined NADH and pyruvate as the QM subsystem, and then we computed the potential energy of the full QM/MM system and the energy of noninteracting QM and MM subsystems. The interaction energy was then calculated as the difference between these two quantities. Averaged values and the standard deviations are given in Table 7. It is a well-known fact that computation of internal energies and entropies is accompanied by much larger uncertainties than free energies, as reflected by the standard deviations given in Table 7.

As a further evidence of the different characteristics of the reactant complexes analyzed we have obtained quite large differences in the averaged interaction energies between the enzyme and the cofactor plus the substrate. These differences span some 200 kcal·mol^{-1} and are partly compensated by the protein energy, as reflected in the much smaller differences in the total potential energy (up to 80 kcal·mol^{-1}). Thus, it seems that in order to gain interaction energy with the cofactor and the substrate the enzyme must be deformed to increase the enzymatic energy term.

According to these results, the most reactive reactant complex (REAC1) is also the less stable (in potential energy terms). It is tempting to conclude that LDH does not work under the equilibrium regime. Obviously, relative free energies (and also interconversion activation free energies) would be needed in order to conclude this question definitively, but in addition to the potential energy analysis, there are other facts supporting this idea. We have shown that the differences among the four reactant complexes can be related to the disposition of the flexible loop closing the active site. The way in which this loop is folded may determine the interactions established between the substrate and some key residues (such as Arg109 or Asn140) or the number of water molecules found in the active site. Thus, we can imagine a picture of LDH catalysis where the closure of the flexible loop over the active site can lead to different reactant valleys, each of them presenting different inherent reactivities. In any case, there has probably not been any evolutive pressure to reduce the rate constant of the chemical step (or the macroscopic average) under the value of the rate-limiting step (the loop closure over the active site). The stability of these complexes can also be very different, depending on the way in which the loop has been folded. In any case interconversion among them could not be an easy task. One way to go from one reactant valley to another is obviously through the unfolding and refolding of this loop; however, we must take into account that this movement is, in fact, the rate-determining step in the catalytic process, and then the process would be slower than the chemical step.

Table 8. Contributions to the Activation Free Energy (in kcal·mol^{-1}), Semi-classical Transmission Coefficients, Catalytic Rate Constants (in s^{-1}), and KIEs Obtained for Each of the Four Reaction Paths Analyzed in the NAD^{2}H Case

	FEP1	FEP2	FEP3	FEP4
ΔG_{FEP}	16.7	23.5	19.4	20.6
$\Delta G^{D}_{vib,QM}$	−2.5	−1.7	−1.8	−2.6
$\Delta G^{CVT,D}$	14.2	21.8	17.6	18.0
κ^{D}	2.71	2.96	2.76	2.88
k^{D}_{cat}	662	0.0019	2.04	1.16
KIE	4.19	5.05	4.98	3.47

3.6. Kinetic Isotope Effects. We have calculated the kinetic isotope effects as the ratio between the rate constants obtained using NAD^{2}H and NAD^{1}H:

$$\mathrm{KIE} = \frac{k^{H}_{cat}}{k^{D}_{cat}} \tag{13}$$

To obtain the rate constant corresponding to NAD^{2}H we just needed to recalculate the contribution of the QM subsystem vibrations to the free energy profile and the semi-classical transmission coefficient. The classical free energy profile is independent of the nuclear masses, and so it takes the same value when using either NAD^{1}H or NAD^{2}H. The calculations were carried out for the four reaction paths, and the results are summarized in Table 8.

In all the cases, substitution of NAD^{1}H by NAD^{2}H provokes an increase of the CVT activation free energy and a reduction of the semi-classical transmission coefficient. The final result is a reduction of the catalytic rate constant. The resulting kinetic isotope effects range between 3.47 and 5.05, in good agreement with the experimental estimations for this enzyme,[38] LDH mutants, and other enzymes of the dehydrogenase family.[32,36]

4. Conclusions

LDH catalyzes the reversible transformation of pyruvate into lactate. The chemical step consists in a hydride and a proton transfer from the cofactor (NADH) and a protonated histidine (His195), respectively. Our results indicate that the two transfers take place in a single concerted process although the degree of asynchronicity may depend on the particular reaction path followed. We have been able to locate different transition state structures which have been connected by following the corresponding reaction paths to their respective reactant Michaelis complexes. Four of these reactant valleys have been explored in detail by means of relatively long stochastic boundary molecular dynamics simulations. We have found some differences among these complexes related to the positioning of the

flexible loop closing the active site. This is translated to a different enzyme−substrate pattern of interactions, a different number of water molecules in the active site and other differences related to the key residue, Arg109.

Our finding of several different, well-characterized reactant valleys, each one acting as an independent enzyme, with its own individual reaction rate constant, agrees with the dynamical disorder of the enzymes exhibiting fluctuations of catalytic rate constants, as shown by single-molecule enzymatic studies. To reach and explore those reactant valleys we have developed a new strategy which is based on the following points:

(1) Even a long molecular dynamics simulation inside a given reactant valley only generates a set of configurations which belong to a unique ensemble that evolves to the products through the same reaction channel, with the same individual reaction rate constant. Different reactant valleys have to be sampled to account for the contribution of different reaction channels. However, in practice, a finite sampling is presently unable to cover several well-defined reactant valleys with a reasonable computational effort. In this paper we have accessed to a number of reactant valleys by performing molecular dynamics simulations in the high-energy region of the transition states, where the jump between them is easier. From there, the low-energy regions of the reactant valleys are straightforwardly reached by going down, following the gradient. Although this strategy does not provide an adequate sampling of the full configurational space of the reactant state, it at least allows us to locate and characterize different reactant valleys.

(2) Long molecular dynamics simulations in each reactant valley have been done in order to characterize each of them and to ensure that, within the time scale of the simulation, no overlap between the identified reactant valleys occurs. This way, we can consider that each reactant valley acts as an independent enzyme.

(3) Because the concerted hydride and proton transfer involves a complicated geometrical coordinate, we have employed a free energy perturbation technique adapted to follow intrinsic reaction coordinates in order to get significant free energy profiles for each reactant valley. From them, the corresponding individual reaction rate constants have been calculated. Vibrations of the quantum subsystem are incorporated through a harmonic normal mode approximation, whereas tunneling contributions are included using an adaptation of the EA-VTST/MT[23,24,48] method.

The four reaction valleys analyzed display significant differences in the calculated rate constants, thus reproducing the experimental dispersion of reaction rate constants in single-molecule enzyme studies. A simple arithmetic average of the four individual reaction rate constants gives 696 s^{-1}, a value very close to the experimental macroscopic lower-limit estimation. However, averaging among different reactant complexes to obtain the macroscopic rate constant is a question without an obvious answer. To obtain a definitive answer one would need to trace the free energy profiles connecting them. Our estimation through the potential energies displays significant differences but cannot be taken as conclusive. Anyway, taking into account that the movement of the loop is in fact the rate-limiting step in the global reaction, we can imagine a scenario where the reactant complexes interconversion is slower that the chemical reaction, and then global equilibrium cannot be assumed. We have also calculated kinetic isotope effects for each of the four reaction valleys. Our calculated values range between 3.47 and 5.05 in good agreement with experimental determinations of enzyme reactions where hydrogen transfer is rate limiting.

Acknowledgment. We thank the DGI for Projects DGI BQU-2002-00301, BQU2003-04168-C03, and CTQ2005-07115/BQU, BANCAIXA for Project P1A99-03, Generalitat Valenciana for Projects GV01-324, GV04B-21, GV04B-131, and GRUPOS04/28, and Generalitat de Catalunya (2005SGR00400). S.F. acknowledges a doctoral fellowship of the Spanish Ministerio de Ciencia y Tecnología. S.F. also acknowledges the Ministerio for financial support during her visit to the UAB. M.G.-V. thanks the "Ramón y Cajal" research program for financial support.

Supporting Information Available: Complete references 55 and 66. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0653977