

Pyruvate Formate Lyase: A New Perspective

Maria de Fátima Lucas,[†] Pedro Alexandrino Fernandes,[†] Leif A. Eriksson,[‡] and Maria João Ramos^{*,†}

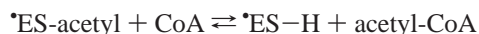
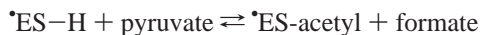
REQUIMTE, Departamento de Química, Faculdade de Ciências, Universidade do Porto, Rua Campo Alegre 687, 4169-007 Porto, Portugal, and Department of Biochemistry, Uppsala University, Box 576, Uppsala S-751 23, Sweden

Received: October 25, 2002; In Final Form: April 1, 2003

The enzymatic conversion of pyruvate into formate by the radical enzyme pyruvate formate lyase is studied using density functional theory with the hybrid B3LYP functional. The residues glycine 734, cysteine 418, and cysteine 419 are known to be fundamental for the reaction to occur. We here represent G734• by CHO–NH–CH•–CO–NH₂, which allows for an efficient π delocalization, whereas the two neighboring cysteines are modeled using methylthiol or as a larger complex that includes part of the protein backbone. The latter representation permits a differentiation of the two cysteines and provides an explanation of the catalytic role of the enzyme. Previous theoretical studies of the mechanism have employed a neutral model to represent pyruvate. In the present work, we have used negatively charged model, which leads to modifications of the behavior of the enzyme. Based on the obtained results, a four-step mechanism is proposed, in which the reaction is initiated by H transfer from C418 via C419 to the glycyl radical. The C418 thiyl radical adds pyruvate from which the formyl radical is detached via a quasi-planar transition state. The previously postulated tetrahedral intermediate does not exist according to the present models. The formyl radical then abstracts a hydrogen atom from C419, forming the formate anion.

1. Introduction

Pyruvate formate lyase (PFL; EC 2.3.1.54) catalyzes the homolytic conversion of pyruvate and coenzyme A into formate and acetyl-CoA¹.



The reaction is fundamental in the anaerobic glucose fermentation route of *Escherichia coli* and other bacteria¹ and resembles functionally the β -ketothiolase reaction of the fatty acid degradation cycle.² In its active form, the enzyme contains a stable glycyl radical (G734) required for catalysis.³ Two cysteines, C418 and C419, have also been established as essential.¹

Previous studies² have shown that mutation of the active site cysteines 418 and 419 (C418A and C419A, respectively) does not affect the activation of the enzyme at glycine 734. However both mutant enzymes are catalytically inactive suggesting that both residues are essential for overall catalysis. It has also been observed that the activated radical form of G734 exchanges hydrogen with the solvent.⁴ The process is faster than expected, and mutation of C419, but not C418, affects the equilibrium. This led to the conclusion⁴ that the hydrogen exchange at the glycyl radical is catalyzed by C419 and that the two residues thus should be in close proximity, as later verified by the X-ray crystallographic structure.⁵ Mutagenesis experiments² have also revealed that the mutation of cysteine 419 into serine does not affect the acetyl exchange with CoA (the second part of the

mechanism), whereas the activity is suppressed when C418 is mutated. This indicates that C418 should be acetylated after the first part of the reaction. It is also observed⁴ that once the enzyme is acetylated the rate of hydrogen exchange of the glycyl radical with the solvent is maintained; that is, C419 should be free. This information is consistent with C418 being the preferred site of acetylation in PFL.

Through the years, several mechanisms have been proposed; the one presented by Kozarich and co-workers⁴ in 1995 was, until recently, the most widely accepted. The mechanism begins with a first hydrogen transfer from C419 to the neighboring G734, after which the cysteine adds to the substrate. However, it did not consider the possibility for the addition of C418, which was thought to be too buried in the protein structure. Only upon observation of the crystallographic structure of the nonactive form of PFL with the substrate analogue oxamate recently determined by X-ray crystallography⁵ (Figure 1), new insight was provided into the roles of the two adjacent cysteines, and it was noted that the crystallographic structure suggests a direct attack from C418 to carbon 2 of pyruvate.⁵

Furthermore, both cysteines participate, in their individual reactions, as thiyl radicals, but the role of C419 is merely to function as a relay for H transfer.⁵ On the basis of previous studies,^{1,2} this recent information, and the observation of the chemical behavior of the substrate analogue methacrylic acid, Knappe and co-workers proposed a new mechanism⁶ for pyruvate conversion, as illustrated in Scheme 1.

It was proposed that the presence of the substrate in the active site triggers the generation of the C418 thiyl radical by H transfer through C419 to •G734 (steps 1 and 2).⁵ The radical cysteine 418 adds to C2 of pyruvate producing a tetrahedral oxyradical intermediate (step 3), which then collapses into a formyl radical anion and acetylated C418 (step 4). It is then expected that the

[†] Universidade do Porto.

[‡] Uppsala University.

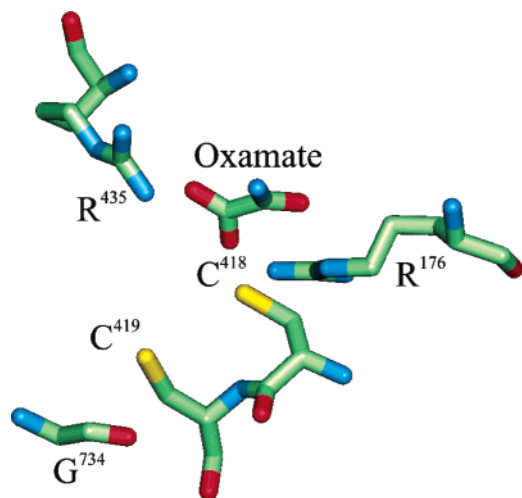
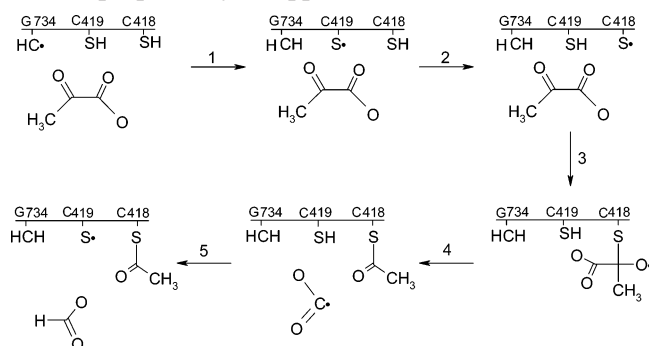


Figure 1. The PFL active site with oxamate substrate analog,⁵ PDB entry 3pfl.

SCHEME 1: Reaction mechanism for the acetylation of PFL as proposed by Knappe and co-workers⁵



formylate radical removes a hydrogen atom on C419 (step 5), regenerating the radical on the protein backbone. The knowledge that C418 acts as the precursor for the second part of the mechanism has always constituted a drawback for other mechanistic proposals that predicted the acetylation to occur on C419,^{3,7} a drawback that generally has been overcome by introducing an additional transesterification step between the two cysteines.^{2,4,8,9} In the present mechanism, no such process is needed because the radical addition takes place directly on C418.

Previous theoretical studies of the catalytic mechanism,⁹ based on information prior to the determination of the X-ray structure, were able to discriminate between two earlier proposals but do not fully comply with the more recent findings. The oxidative degradation of PFL by molecular oxygen has also been explored by computational chemistry methods.¹⁰

In the present work, we focus on the revised mechanism of Scheme 1 and try to establish its adequacy using theoretical chemistry tools. Because the two cysteines are fundamental for the mechanism, two distinct sets of models were used: a smaller model simply composed of CH_3SH and a larger model that includes both the neighboring cysteines side chains and the connecting peptide backbone. The reason for the use of a larger system is related to the fact that although C419 appears to have only a secondary function it is clearly essential for the overall process.

The substrate is a charged species under physiological conditions (pK_a of pyruvic acid is 2.5) and was thus throughout this work modeled as anionic. This results in significant differences compared with earlier work on PFL.⁹ When model-

ing enzymatic reactions in the gas phase, neutral models are normally used (cf. studies of the mechanism of ribonucleotide reductase (RNR)¹¹ or the nucleophilic addition reaction in the Wacker process¹²). It is often stated that neutral models lead to better agreement with experimental results, although many of the reactions considered were actually ionic reactions observed in aqueous solution.¹² It is well established that the use of charged models in gas phase leads to computational difficulties, in terms of both convergence and reliability of the obtained structures and mechanisms. However, this is more likely to occur if the charged model does not include the moieties that directly interact with and stabilize the charged species. This is not the case here because the models include all of the residues directly interacting with pyruvate. In the present work, we have used charged models and were able to obtain converged results throughout. The use of realistically charged models has also enabled us to modify the earlier proposed¹¹ (neutral) mechanism of RNR.¹³

2. Computational Details

All geometries and energies were computed using density functional theory (DFT) and the hybrid functional B3LYP,^{14–16} as implemented in the Gaussian 98¹⁷ program package.

All geometries were determined with the double- ξ basis set 6-31G(d). Frequency calculations were performed to determine the nature of each stationary point. Single-point energies were calculated with the larger basis set 6-311++G(3df,3pd) on the previously optimized geometries. All energies discussed below include zero-point energy (ZPE) corrections obtained at the B3LYP/6-31G(d) level. The part of the protein that was not explicitly included in the model was treated as a homogeneous medium with a dielectric constant of 4, which was shown to give good agreement with experimental results on the active site of proteins and accounts for both the protein and the buried water molecules.^{18,19} Corrections for solvent effects are calculated using the C-PCM polarizable conductor model (Cosmo) developed by Barone and Cossi.^{20–23}

The contribution from the continuum to the total energy was evaluated using the 6-31+G(d) basis set. It was checked that the inclusion of diffuse functions on the geometry optimizations did not account for significant differences. The variations in geometry were not significant, and the final energies calculated with the larger 6-311++G(3df,3pd) basis set were almost equivalent ($\Delta\Delta H_{\text{act}}[6-311++G(3df,3pd)/6-31G(d) - 6-311++G(3df,3pd)/6-31+G(d)] = 0.090$ kcal/mol and $\Delta\Delta H_{\text{reaction}}[6-311++G(3df,3pd)/6-31G(d) - 6-311++G(3df,3pd)/6-31+G(d)] = 0.6$ kcal/mol, with obvious notations). Therefore, for the sake of comparing mechanisms, it was considered unnecessary to use the time-consuming diffuse functions on the geometry optimizations. This reflects the fact that energy is quite insensitive to the quality of the geometry.^{24–26}

Intrinsic reaction coordinates (IRC) were followed from the transition states toward reactants and products.

Given the radical nature of all systems under study care must be taken with spin contamination as a possible source of error. Although the DFT wave function is strictly speaking not an eigenfunction of the operator S^2 , where S is the electron spin operator, the evaluation of $\langle S^2 \rangle$ still provides a useful control of the quality of the data obtained. In the present work, minor spin contamination occurs for the transition states; however, the value of $\langle S^2 \rangle$ never exceeded 0.766 (ideal value of a pure doublet spin eigenstate is 0.750).

Mulliken population analyses were performed to obtain charge and spin densities.

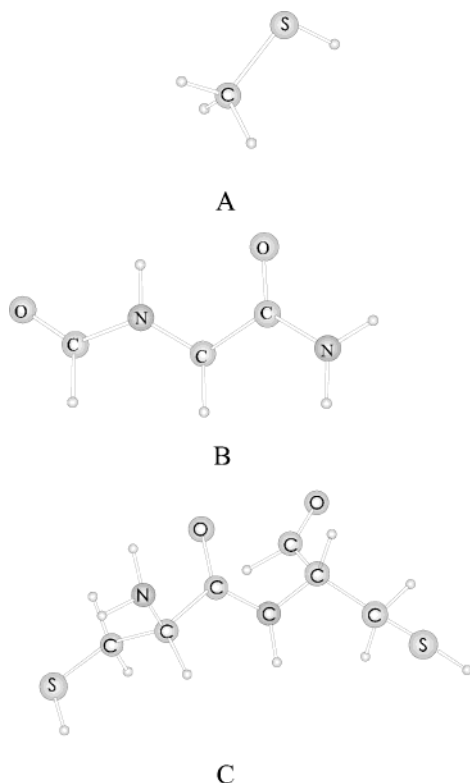


Figure 2. Models used to represent PFL's active center in the present work: (A) methylthiol; (B) glycy radical; (C) C418 + C419.

All complexes illustrated in this work exhibit a large number of minima. The structures discussed in text are not necessarily the absolute minima in a vacuum but rather the structures that more resemble the geometry of the active site.

3. Results and Discussion

Because of computational limitations, it is important to make an adequate choice of models such that it is small enough to allow for accurate calculations yet sufficiently large to incorporate all of the essential features of the system under study. The cysteines investigated in this work are modeled in two different ways. A smaller model, CH₃SH, which has previously been shown to provide a good representation of a cysteine side chain,⁹ and a larger model incorporating both the neighboring cysteines based on the structural data relative to the substrate analogue, oxamate.⁵ For G734*, we used the extended model previously proposed on the basis of the spin distribution of the corresponding backbone-centered radical.⁹ The optimized structures of the models employed are shown in Figure 2.

From the observation that the pyruvate attack at C418 appears to be favored over that at C419,⁵ it is also of interest to explore whether C419 has a stabilizing effect on the carboxylate group of pyruvate, that is, to “hold” the substrate in a position optimal for attack by C418.

As mentioned previously, it is also important to note that in the present work, neutral models have been set aside. Throughout, the anionic forms of the substrate pyruvate and the reaction product formate have been used.

The following sections describe the reactions shown in Scheme 1.

3.1. Hydrogen Transfer between C419 and G734. PFL in its active form carries a stable radical produced by the PFL-activating enzyme with adenosylmethionine as the H-atom abstracting substrate.¹ EPR spectroscopic studies of selectively ¹³C-labeled enzyme have assigned the radical site to Cα of the

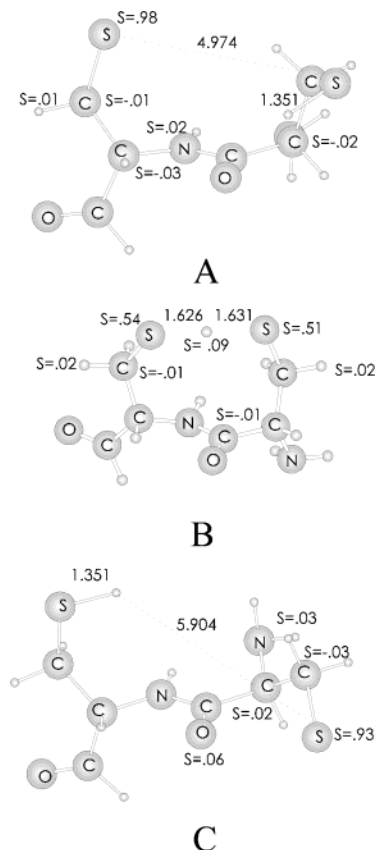


Figure 3. Optimized structures and calculated Mulliken spin for the hydrogen transfer between the two cysteines.

glycine residue 734.³ It has also been established, by means of inhibition experiments, that a sulfur-centered radical is present during the catalytic cycle.² In light of this and given the proximity between G734 and C419, most studies published on this matter^{2,8,9} infer that the first step in PFL catalytic mechanism involves a homolytic hydrogen transfer between these residues.

This process has previously been studied,⁹ and the present calculations—even though they are performed at a slightly higher level—do not present any further improvements. The transition is endothermic by 6.5 kcal mol⁻¹ and has an activation energy of 13.9 kcal mol⁻¹.

3.2. Hydrogen Transfer between Cysteine 418 and Cysteine 419. The presence of the two cysteines in the active site of the enzyme makes functional assignment difficult. Earlier reports on the PFL mechanism have predicted C419 as the site of thiol attack to the substrate.^{7,8} However, the X-ray crystallographic structure of the inhibited enzyme suggests that pyruvate is positioned in the active site such that it favors covalent addition to C418.⁵ The adjacent C419 instead operates as a shuttle for H-atom transfer.

Hydrogen transfer with the cysteines modeled as CH_3SH has been reported with an energy barrier of $2.4 \text{ kcal mol}^{-1}$.⁹ In this work, the H-transfer is modeled also with the larger system C of Figure 2.

In Figure 3, we display the optimized stationary structures and spin distributions along the H-transfer pathway. The transition occurs by means of a transition state similar to the one observed previously.⁹

The process is exothermic ($-4.9 \text{ kcal mol}^{-1}$) according to this model. The use of the smaller methylthiol models does not allow for the reaction to be asymmetric because both reactants and products are identical. An additional difference compared to the small model is concerned with the activation energy. Here,

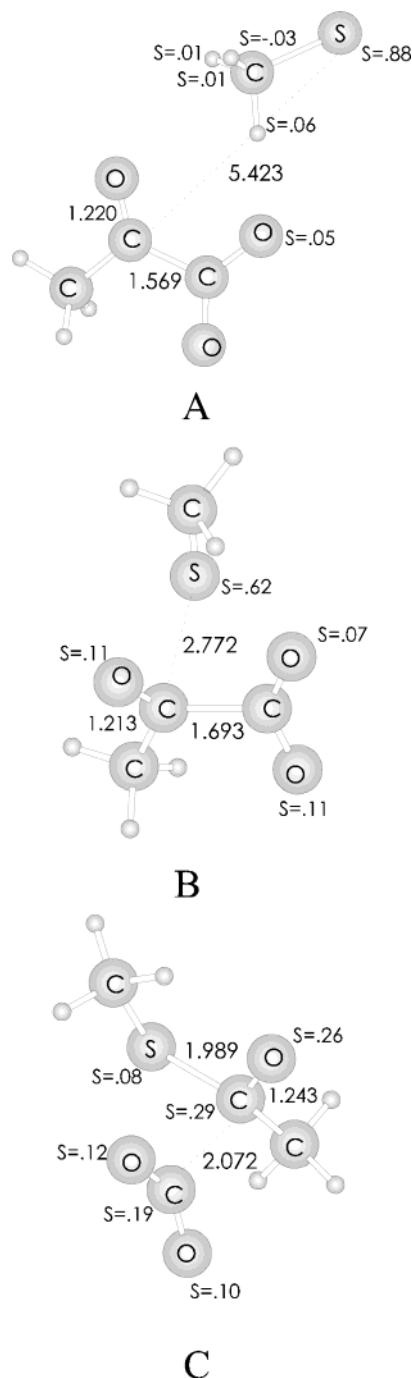


Figure 4. Optimized structures and calculated Mulliken spin for thiol addition to pyruvate.

using methylthiol, we find a barrier of only 3.5 kcal/mol, whereas the inclusion of the connecting protein backbone raises this to 4.9 kcal/mol. When the model CH_3SH is used, the components can orient freely to adopt the most favorable conformation for the transition to occur. The presence of the protein backbone makes the transition more difficult; the C419 sulfur radical is almost 5 Å away from the transferring hydrogen atom of C418.

3.3. Thiol Addition to the Carbonyl Carbon of Pyruvate.

According to Knappe, the next step involves the direct attack of the thiol radical located on the cysteine 418 on the C2 of pyruvate.⁵ The structure determined with oxamate in the active center revealed that the neighboring cysteine 419 could have a stabilizing effect on the substrate. Calculations undertaken with the CH_3SH model of cysteine show that thiol addition occurs

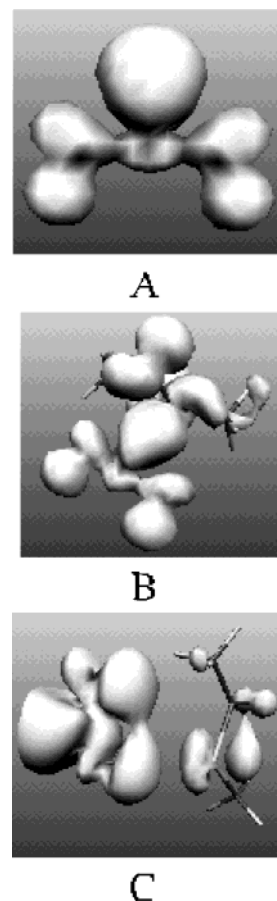


Figure 5. Spin densities for (A) formyl radical in gas phase, (B) formyl radical product from IRC charged transition state, and (C) formyl radical product from IRC neutral transition state.

in one single step and that the transition state has a partially elongated C1–C2 bond in pyruvate (Figure 4). Using a neutral pyruvate model led to a two-step mechanism in which the S–C bond was first formed, after which the complex resides in a shallow minimum as a tetrahedral intermediate. From this, a second small barrier was then surmounted leading to release of (neutral) formic acid radical.⁹ If we examine the structure of the product obtained with the present, charged, model (Figure 4C), it can be seen that the distance separating the leaving group from the original molecule is only 2.072 Å, indicative of maintained interaction between the acetylthiol moiety and the highly reactive formate radical. It is also visible in the figure that there is still a high degree of unpaired spin on the acetylated cysteine.

We can see in Figure 4C that the spin of the formate radical is delocalized over the molecule. It is also interesting to compare the spin density of the formate radical in gas phase to that of the product from the charged model and the product of the neutral model. Figure 5B shows that the spin of the formate moiety of the charged product complex clearly resembles that of the free formate radical. On the contrary, the reaction product from the neutral model, the formic acid radical (Figure 5C), has a spin density less similar to that of the free formate (Figure 5A) and is not delocalized through the protein backbone.

The replacement of pyruvic acid by its base has a direct effect on the mechanistic path in that the tetrahedral intermediate observed with the neutral model is converted into a quasi-planar transition state. In light of this result, it seemed important to improve the model. Given that cysteine 419 was suggested as a possible stabilizer, this led us to consider also this reaction

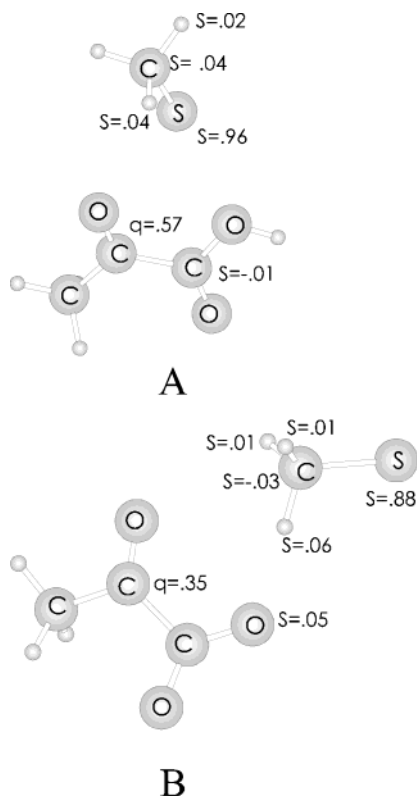


Figure 7. Calculated Mulliken spin and charge densities for (A) pyruvic acid and (B) pyruvate.

TABLE 1: Summary of Reaction Energies (kcal mol⁻¹) and Barriers Calculated in This Work for the Individual Steps in the PFL Mechanism

| reaction | barrier | energy |
|--|---------|--------|
| Gly [•] + Cys → Gly + Cys [•] | 13.9 | 6.5 |
| Cys [•] + Cys → Cys + Cys [•] (lm) ^a | 4.9 | -4.9 |
| Cys [•] + pyruvic acid → tetrahedral intermediate ⁹ | 11.6 | 2.5 |
| Cys [•] + pyruvate → acetyl-Cys + formyl [•] (sm) ^a | 7.7 | 1.2 |
| Cys [•] + pyruvate → acetyl-Cys + formyl [•] (lm) ^a | 11.8 | 7.8 |
| formyl [•] + Cys → formate + Cys [•] | | -8.6 |
| formyl [•] + Gly → formate + Gly [•] | 0.4 | -15.7 |

^a (lm) large model; (sm) small model.

more favorable on the less electrodeficient carbon. By performing a Mulliken population analysis on both pyruvic acid and its conjugate base structures in the presence of methylthiol (Figure 7), we can see that in the protonated form the C2 is more electrodeficient than in the corresponding anionic structure. This leads to an easier attack on pyruvate than on the corresponding acid, as is also seen from the energetics (Table 1).

Additionally, looking at the spin density of the two reactants, we observe a higher degree of spin delocalization on the pyruvate than that in the acid. The absence of the acidic hydrogen permits a redistribution of spin on the pyruvate molecule, especially on the acidic oxygens. These factors combined may explain why the process occurs in one step instead of two.

All pertinent energies determined for PFL's mechanism are summarized in Table 1.

4. Conclusions

Using negatively charged substrate/product models and including part of the protein backbone, we have reevaluated the reaction mechanism of the anaerobic enzyme pyruvate formate lyase (PFL) using the DFT functional B3LYP.

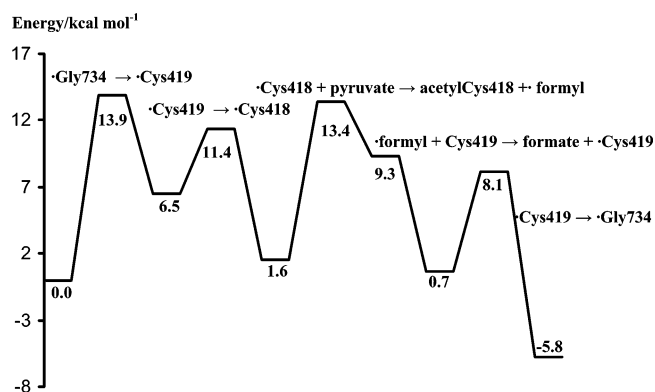


Figure 8. Calculated overall energies for pyruvate conversion in the catalytic mechanism of PFL.

On the basis of the computed data, we propose a mechanism that is consistent with all currently known experimental data. It is proposed that the mechanism is initiated by H transfer from C418 to G734 via C419. C418 adds to pyruvate on C2, and the formylate radical is released. The formylate radical then abstracts a H atom from the H-bonded C419 followed by another H transfer between the G734 and C419. This last step regenerates the radical on the stable G734 site. A fundamental difference in this mechanism over that previously proposed is that the thiol addition occurs by means of a quasi-planar transition state. The postulated tetrahedral intermediate could not be found. This result is consistent with the fact that there is no experimental evidence that supports the presence of such a species. It is also demonstrated that the formylate radical abstracts the H atom from the C419 with little or no barrier and this reaction is stereochemically favored over abstraction from G734.

The potential energy surface including zero-point energy corrections is drawn in Figure 8. As can be seen, the overall process is exothermic by 5.8 kcal mol⁻¹.

All studies reported to date predicted C–C cleavage as the rate-limiting process. According to the results obtained in the present work, the H transfer between the cysteine 419 and glycine 734 is the process with the highest activation energy. This result is consistent with the electron paramagnetic resonance (EPR) studies known for this enzyme, which assigned the spin to the Cα on G734. We also note that the presence of the protein backbone and the catalytically active side chains affect the overall energies but that the overall activation energy needed to drive the entire reaction, 16.1 kcal mol⁻¹, remains the same as in the case of the neutral mechanism.⁹ In the future, further studies should be performed using a larger fragment to clarify the stereochemical role of the enzyme.

Acknowledgment. The Fundação para a Ciência e Tecnologia (Project POCTI/35736/99, Portugal) and the Swedish Science Research Council (VR) are gratefully acknowledged for financial support.

References and Notes

- (1) Knappe, J.; Blaschkowski, H. P.; Gröbner, P.; Schmitt, T. *Eur. J. Biochem.* **1974**, *50*, 253.
- (2) Knappe, J.; Elbert, S.; Frey, M.; Wagner, F. V. *Biochem. Soc. Trans.* **1993**, *21*, 731.
- (3) Wagner, F. V.; Frey, M.; Neugebauer, F. A.; Schäfer, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 996.
- (4) Parast, C.; Wong, K.; Lewisch, S. A.; Kozarich, J. W. *Biochemistry* **1995**, *34*, 2393.
- (5) Becker, A.; Wolf, K.; Kabsch, W.; Knappe, J.; Schultz, S.; Wagner, A. *Nat. Struct. Biol.* **1999**, *6*, 969.
- (6) Plaga, W.; Vielhaber, G.; Wallach, J.; Knappe, J. *FEBS Lett.* **2000**, *466*, 45.

- (7) Unkrig, V.; Neugebauer, F. A.; Knappe, J. *Eur. J. Biochem.* **1989**, 28, 723.
- (8) Leppänen, V.; Merckel, M.; Ollis, D.; Wong, K.; Kozarich, J. W.; Goldman, A. *Structure* **1999**, 7, 733.
- (9) Himo, F.; Eriksson, L. A. *J. Am. Chem. Soc.* **1998**, 120, 11449.
- (10) Gauld, J. W.; Eriksson, L. A. *J. Am. Chem. Soc.* **2000**, 122, 2035.
- (11) Siegbahn, P. E. M. *J. Am. Chem. Soc.* **1998**, 120, 8417.
- (12) Siegbahn, P. E. M. *J. Phys. Chem.* **1996**, 100, 14672.
- (13) Fernandes, P. A.; Eriksson, L. A.; Ramos, M. J. *Theor. Chem. Acc.*, in press.
- (14) Becke, A. D. *J. Chem. Phys.* **1993**, 98, 1372.
- (15) Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648.
- (16) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, 37, 785.
- (17) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.3; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (18) Siegbahn, P. M. E.; Eriksson, L. A.; Pavlov, M. *J. Phys. Chem. B* **1998**, 102, 10622.
- (19) Blomberg, M. R. A.; Siegbahn, P. M. E.; Babcock, G. T. *J. Am. Chem. Soc.* **1998**, 120, 8812.
- (20) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, 107 (8), 3210–3221.
- (21) Miertus, S.; Scrocco, E.; Tomasi, J. *J. Chem. Phys.* **1981**, 55, 117.
- (22) Tomasi, J.; Bonaccorsi, R.; Cammi, R.; Valle, F. J. O. *J. Mol. Struct. (THEOCHEM)* **1991**, 234, 401.
- (23) Tomasi, J.; Bonaccorsi, R. *Croat. Chem. Acta* **1992**, 65, 29.
- (24) Fernandes, P. A.; Ramos, M. J. *J. Am. Chem. Soc.*, in press.
- (25) Stubbe, J.; van der Donk, W. A. *Chem. Biol.* **1995**, 2, 793.
- (26) Zipse, H. *J. Am. Chem. Soc.* **1994**, 116, 576.