High-level QM/MM modelling predicts an arginine as the acid in the condensation reaction catalysed by citrate synthase†

Marc W. van der Kamp,[a] Francesca Perruccio[ab] and Adrian J. Mulholland*[a]

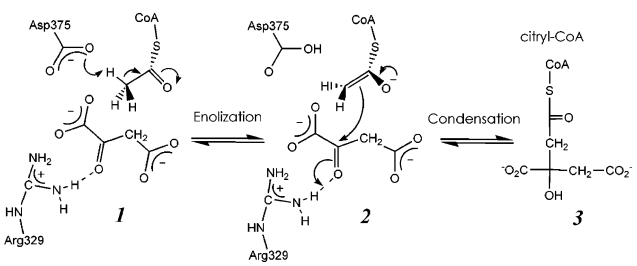
Received (in Cambridge, UK) 21st January 2008, Accepted 25th February 2008
First published as an Advance Article on the web 11th March 2008
DOI: 10.1039/b800496j

High-level *ab initio* quantum mechanical/molecular mechanical (QM/MM) modelling of citryl-CoA formation in citrate synthase reveals that an arginine residue acts as the proton donor; this proposed new mechanism helps to explain how chemical and large scale conformational changes are coupled in this paradigmatic enzyme.

Citrate synthase (CS) catalyses the conversion of oxaloacetate (OAA) to citrate using acetyl-Coenzyme A (acetyl-CoA), the first step of the citric acid cycle. This enzyme is an important model for mechanisms of biological catalysis (in particular for carbon–carbon bond formation). It is further an exemplar of large conformational changes occurring in proteins and has been extensively studied both experimentally and computationally.[1–7] It combines ligase (condensation) and hydrolase activity and achieves a remarkable rate enhancement, without the assistance of metal ions or other cofactors. The reaction starts with proton abstraction from acetyl-CoA by an aspartate residue (Asp375 in pig CS numbering, used throughout). The resulting (enolate[1,2]) intermediate subsequently performs a nucleophilic attack on the carbonyl carbon of OAA. This produces citryl-CoA,[3] which breaks down to citrate and CoA upon hydrolysis. Computational studies have been instrumental in establishing the mechanism of proton abstraction and identifying the enolate intermediate formed by this reaction.[1,2] The mechanism of the crucial condensation step, however, remains unknown. Formation of citryl-CoA requires protonation of what was the carbonyl oxygen of OAA. One of several histidines in the active site was proposed to act as proton donor in this step,[4,5] but previous modelling indicates that they are all neutral.[1] Alternatively, it has been proposed that protonation is associated with hydrolysis,[6] which likely requires a conformational change of the enzyme.[7] Another possibility, however, is that Arg329, which forms a hydrogen bond to the OAA carbonyl oxygen,[5] donates the proton (Scheme 1). Although rare and still somewhat controversial, acid/base catalysis by arginine residues has been proposed in several enzymes.[8,9]

Here, we report the first modelling of the condensation reaction in CS, using high level *ab initio* QM/MM methods. The calculated potential energy profile shows that Arg329 can act as the proton donor, leading to a stable citryl-CoA intermediate. After proton abstraction from acetyl-CoA, carbon–carbon bond formation starts first but is concerted with proton transfer from Arg329 to the former OAA carbonyl oxygen. The energy barriers in the profile are consistent with the experimentally determined reaction rate and indicate that the enolization and condensation steps are closely linked and have similar barriers.

The crystal structure of chicken CS co-crystallized with acetyl-CoA and *R*-malate (PDB entry code 4CSC) was prepared for treatment with QM/MM methods as described previously.[10] Initially, proton abstraction from acetyl-CoA followed by condensation was modelled at the AM1/CHARMM27[11] QM/MM level.[12] Several alternative mechanisms for condensation were tested (see ESI†), but only the mechanism with Arg329 as the proton donor gave structurally and energetically reasonable results. Different approximate reaction coordinates were also considered. Using three different reaction coordinate definitions consecutively gave the most reasonable profile. For enolization, $r1 = d(C_{acetyl\text{-}CoA}H) - d(O_{Asp375}H)$ was used, known to describe this reaction accurately.[1,10] For the first stage of condensation, $r2 = d(C_{acetyl\text{-}CoA}C_{OAA})$ was used, to be followed by $r3 = d(N_{Arg329}H) - d(O_{OAA}H)$ for proton transfer from Arg329. Starting from the AM1-CHARMM27 optimized structure of the approximate transition state for enolization ($r1$ = 0.3 Å), the energy profile for the whole reaction was optimized at the higher B3LYP/6-31+G(*d*)/CHARMM27 QM/MM level. These calculations were performed with QoMMMa, which couples Jaguar and Tinker for QM/MM calculations.[13] The QM region consisted of the thioester part of acetyl-CoA, OAA, the Arg329 sidechain (from Cγ) and the Asp375 and



Scheme 1 Mechanism of citryl-CoA formation in citrate synthase (*via* enolization and condensation) as identified by calculations here.

[a] *Centre for Computational Chemistry, School of Chemistry, University of Bristol, Cantocks Close, Bristol, UK. E-mail: Adrian.Mulholland@bristol.ac.uk*
[b] *Medicinal Informatics, Structure and Design, Pfizer Global R&D, Sandwich, Kent, UK*

† Electronic supplementary information (ESI) available: Details of computational procedures and AM1/CHARMM test profiles. See DOI: 10.1039/b800496j

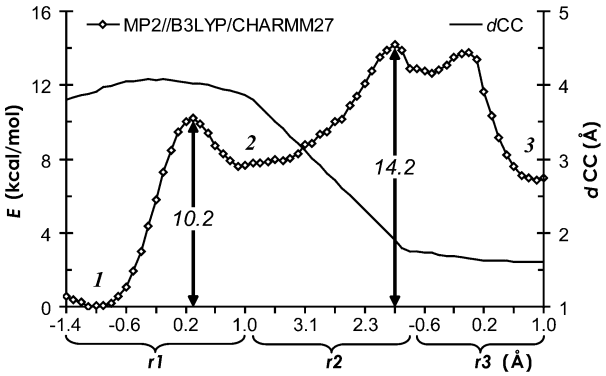


Fig. 1 *Ab initio* QM/MM total energy profile of the CS reaction. Energies are plotted relative to the substrate minimum. The carbon–carbon bond distance (*d*CC) of the forming bond is indicated. The locations of the species labelled in Scheme 1 are indicated; *r*1, *r*2 and *r*3 are the reaction coordinates as defined in the text.

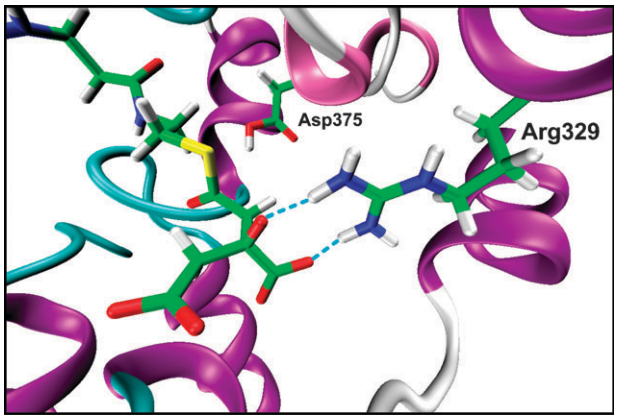


Fig. 2 QM/MM (B3LYP/6-31 + G(*d*)/CHARMM27) optimized conformation of the tetrahedral species in the CS active site. Arg329 is found here to protonate this species to form citryl-CoA. Asp375 is protonated as a result of it having performed the initial proton abstraction from acetyl-CoA (see Scheme 1).

Asp327 sidechains from Cβ (covalent bonds across the QM/MM barrier were modelled using link atoms[13]). For reliable reaction energetics, we used high-level *ab initio* QM/MM single point calculations with Molpro.[14] The resulting MP2/aug-cc-pVDZ//B3LYP/6-31 + G(*d*)/CHARMM27 profile thus obtained is shown in Fig. 1.

The reaction starts with deprotonation of acetyl-CoA, to form an enolate intermediate. The barrier found for this reaction step is 10.2 kcal mol^{-1}. The enolate is calculated to lie 7.6 kcal mol^{-1} above the substrate, with a barrier to the back reaction of 2.6 kcal mol^{-1}, indicative of a short-lived intermediate. This enolate subsequently attacks the carbonyl carbon of OAA. The highest energy conformation (approximate transition state) during this attack has a carbon–carbon distance of 1.9 Å. Its energy relative to the ground state is 14.2 kcal mol^{-1}, which is the highest energy along the reaction. This value agrees well with the activation energy (14.7 kcal mol^{-1}) derived from the experimental rate for the overall reaction.[15] After this point, a very shallow minimum is reached (12.7 kcal mol^{-1} above the ground state), which is essentially a tetrahedral species (see Fig. 2). At this point, proton transfer from Arg329 to the (former) OAA carbonyl oxygen occurs. This proton transfer has a similar, but slightly lower barrier relative to the ground state: 13.8 kcal mol^{-1}. Effectively, carbon–carbon bond formation and proton transfer from Arg329 can be considered concerted. Only after the proton is transferred to the former OAA carbonyl is a citryl-CoA minimum reached with a fully formed carbon–carbon bond. Arginine is an unusual proton donor, as the pK_a value of its sidechain in solution is about 12.5. In the buried active site, however, it is possible that its pK_a will be significantly perturbed, as seen in other cases.[9] Furthermore, the tetrahedral species (formed after the nucleophilic attack of the enolate intermediate on oxaloacetate) is expected to be a very strong base (likely pK_a > 16[16]), making proton transfer from Arg329 possible. Finally, efficient catalysis requires a weak acid, to avoid an over-stable intermediate.

The citryl-CoA state formed in this way is likely to be unstable: Asp375 is protonated and Arg329 deprotonated, whereas from the likely pK_a values of their sidechains the opposite situation will be preferred. The positioning and environment of these two residues in the active site do not permit direct proton transfer between them, nor is there a hydrogen bonding network through which a proton could be transferred. This unstable state may therefore be of kinetic and functional importance in driving the change from ligase to hydrolase activity of the enzyme. When citryl-CoA is used as the substrate for CS, this unstable state is unlikely to be present (Asp375 and Arg329 will be charged). Rearrangement at the active site may therefore be necessary for the formation of acetyl-CoA and OAA from citryl-CoA, which may explain the complex kinetics observed experimentally when it is used as a substrate.[17]

Mutation studies also support the catalytic role of Arg329 proposed here. Man *et al.*[7] reported that the Arg314Gln mutant of *E. coli* CS (Arg314 is equivalent to Arg329 in pig CS) has a greatly impaired overall reaction rate ($\sim 3 \times 10^{-3}$% of wild type at pH = 8). The rate for lyase activity with citryl-CoA as substrate, however, was only moderately impaired (3.5% of wild type). These results can be explained by the findings here: for OAA and acetyl-CoA to form citryl-CoA, Arg329 can donate a proton whereas glutamine cannot. When citryl-CoA is offered as the substrate, Arg329 will be in its normal protonated state, unable to abstract a proton, similar to glutamine and hence displaying a similar reaction rate.

Arg329 plays a pivotal role in the large-scale conformational change of CS from 'open' to 'closed'.[18] It is at the base of an α-helix that needs to shift along its own axis to go from the open to the closed conformation. There is a considerable energy barrier for this shift to occur in the unliganded enzyme,[18] but upon binding of OAA the enzyme closes.[19] Crystals of the open form cannot be grown in presence of OAA, and crystals of the open form crack when OAA is added.[5] With citrate, however, crystals can be grown in both the open and closed forms. The dissociation constants for OAA and citrate show that OAA binds more strongly to citrate synthase than citrate (K_d (OAA) $\approx 5 \times 10^{-3}$ mM and K_d (citrate) $\approx$ 1 mM).[20] This suggests that the strong interaction between OAA and Arg329 holds the enzyme in the

closed form, whereas the weaker interaction with citrate allows reopening. Apparently, a salt-bridge interaction between Arg329 and OAA (in its keto form) is crucial. (In *E. coli* CS, the K_m of OAA increases 13-fold upon mutation of the equivalent arginine.[7]) When Arg329 transfers a proton to form citryl-CoA, the salt-bridge interaction is broken: the positive charge of Arg329 (interacting with the OAA carboxylate) is lost. We can speculate that the change in this interaction may trigger a conformational change of CS, that apparently is necessary for hydrolysis.[7] Our findings give an explanation of the observed lower affinity of the enzyme for citryl-CoA than for OAA. The reduced affinity for citryl-CoA is likely to lead to greater dynamic fluctuations, and may lead to formation of the different 'minor' closed form of the enzyme that has been observed crystallographically[5] and may be associated with hydrolysis. Hydrolysis may be linked to recovering the favoured protonation states of Arg329 and Asp375 (the latter is likely to play a role in hydrolysis[21]). Arg329 is therefore not only crucial for the opening and closing of the enzyme, but, through its involvement in catalysis, it also provides a mechanism for coupling condensation and hydrolysis, and for coupling the chemical and conformational changes during the catalytic cycle.

Clearly, it would be of interest to study the nature of the conformational change in more detail, and (*e.g.* molecular dynamics) simulation would be a good approach. Investigation of the possible large conformational changes following the formation of citryl-CoA cannot be performed at the level of QM/MM theory used here, however, because these methods are too computationally extensive. Extensive multi-nanosecond molecular dynamics simulations would be required to investigate large-scale enzyme conformational changes, and such simulations are currently only feasible with MM methods (see *e.g.* ref. 22).

An Arg-carboxylate motif is found in other enzymes that have been proposed to use an arginine as acid or base.[9] In CS, Arg329 interacts with Asp327 (included in the QM region here, see also Fig. 2) and a carboxylate group of OAA. This is similar to the active sites of fumarate reductase (which also uses an arginine to protonate the substrate in the forward reaction) and L-aspartate oxidase.[23]

High level *ab initio* QM/MM calculations reveal an unexpected mechanism for this key model enzyme. The results are consistent with, and shed new light on, experimental findings. Arginine may function as an acid or base in enzyme reactions more widely than previously anticipated. The mechanism we present here for CS gives detailed insight into the enzymic carbon–carbon bond formation. Beyond this, it also provides understanding of how the enzyme avoids overstabilization of the citryl-CoA intermediate, and suggests how chemical and large-scale conformational changes can be coupled in enzyme catalysis.

MWvdK is funded by the BRISENZ Marie Curie Early Stage Training Centre: he and AJM thank the European Commission for this support. AJM also thanks BBSRC and EPSRC. We thank Prof. M. Danson for useful discussions.

Notes and references

1 A. J. Mulholland and W. G. Richards, *Proteins: Struct., Funct., Genet.*, 1997, **27**, 9.
2 A. J. Mulholland, P. D. Lyne and M. Karplus, *J. Am. Chem. Soc.*, 2000, **122**, 534; W. Yang and D. G. Drueckhammer, *J. Phys. Chem. B*, 2003, **107**, 5986; M. W. Van der Kamp, F. Perruccio and A. J. Mulholland, *J. Mol. Graphics Modell.*, 2007, **26**, 596.
3 G. Löhlein-Werhahn, E. Bayer, B. Bauer and H. Eggerer, *Eur. J. Biochem.*, 1983, **133**, 665.
4 M. Karpusas, B. Branchaud and S. J. Remington, *Biochemistry*, 1990, **29**, 2213.
5 S. J. Remington, *Curr. Top. Cell. Regul.*, 1992, **33**, 209.
6 A. J. Mulholland and W. G. Richards, *THEOCHEM*, 1998, **427**, 175.
7 W. J. Man, Y. Li, C. D. O'Connor and D. C. Wilton, *Biochem. J.*, 1994, **300**, 765.
8 C. G. Mowat, R. Moysey, C. S. Miles, D. Leys, M. K. Doherty, P. Taylor, M. D. Walkinshaw, G. A. Reid and S. K. Chapman, *Biochemistry*, 2001, **40**, 12292.
9 Y. V. Guillén-Schlippe and L. Hedstrom, *Arch. Biochem. Biophys.*, 2005, **433**, 266.
10 M. W. Van der Kamp, F. Perruccio and A. J. Mulholland, *Proteins: Struct., Funct., Bioinf.*, 2007, **69**, 521.
11 A. D. MacKerell, Jr, D. Bashford, M. Bellott, R. L. Dunbrack, Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586; J. J. Pavelites, P. A. Bash, J. Gao and A. D. MacKerell, Jr, *J. Comput. Chem.*, 1997, **18**, 221.
12 M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.*, 1990, **11**, 700.
13 J. N. Harvey, *Faraday Discuss.*, 2004, **127**, 165; *Jaguar 5.0*, Schrodinger, LLC, Portland, Oregon, 2002; J. W. Ponder, *TINKER: Software Tools for Molecular Design, v4.0*, Washington University School of Medicine, Saint Louis, MO, 2003, http://dasher.wustl.edu/tinker/.
14 *Molpro 2002.10*, University College Cardiff Consultants Limited, Cardiff, 2004, www.molpro.net.
15 G. M. Alter, J. P. Casazza, Z. Wang, P. Nemeth, P. A. Srere and C. T. Evans, *Biochemistry*, 1990, **29**, 7557; A. Fersht, *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, W. H. Freeman and Co., New York, 1999, pp. 349–355.
16 S. Takahashi, L. A. Cohen, H. K. Miller and E. G. Peake, *J. Org. Chem.*, 1971, **36**, 1205.
17 G. Pettersson, U. Lill and H. Eggerer, *Eur. J. Biochem.*, 1989, **182**, 119.
18 I. Daidone, D. Roccatano and S. Hayward, *J. Mol. Biol.*, 2004, **339**, 515.
19 A. Kollmann-Koch and H. Eggerer, *Eur. J. Biochem.*, 1989, **185**, 441.
20 C. J. Johansson and G. Pettersson, *Eur. J. Biochem.*, 1974, **42**, 388.
21 L. C. Kurz, T. Nakra, R. Stein, W. Plungkhen, M. Riley, F. Hsu and G. R. Drysdale, *Biochemistry*, 1998, **37**, 9724.
22 U. Pentikäinen, O. T. Pentikäinen and A. J. Mulholland, *Proteins: Struct., Funct., Bioinf.*, 2008, **70**, 498.
23 D. Leys, A. S. Tsapin, K. H. Nealson, T. E. Meyer, M. A. Cusanovich and J. J. Van Beeumen, *Nat. Struct. Biol.*, 1999, **6**, 1113; R. T. Bossi, A. Negri, G. Tedeschi and A. Mattevi, *Biochemistry*, 2002, **41**, 3018.

