

Publishers' page

Publishers' page

Publishers' page

Publishers' page

Dedication Page
(optional)

June 3, 2014 15:32

PSP Book - 9in x 6in

000-main

Contents

1. Working title: Computational Enzyme Engineering and Activity Screening Using Quantum Chemical Methods.	1
1.1 Abstract	1
1.2 Motivation	2
1.3 Introduction	3
1.4 Methods	5
1.4.1 Modeling Approaches	5
1.4.2 Quantum Chemical Methods	7
1.4.3 Software	7
1.5 Activity Screening Applications	8
1.5.1 Overview	8
1.5.2 <i>Candida antarctica</i> Lipase B	9
1.5.3 <i>Bacillus circulans</i> xylanase	11
1.5.4 Further Screening Applications	17
1.6 Conclusions	17
1.6.1 Advantages/Disadvantages	17
1.6.2 Required Future Development Actions	17

Chapter 1

Working title: Computational Enzyme Engineering and Activity Screening Using Quantum Chemical Methods.

Martin R. Hediger^a, Harm Otten^b

^a *Zürich*

^b *København*

m—@—.ch

harm.otten@maxlab.lu.se

1.1 Abstract

Main messages:

- QM Screening of CALB and BCX
- Screening 317 single and 111 double mutants with av. 29 h per barrier and 1 CPU per interpolation point
- Correct mechanisms using PM6
- Barriers in agreement (calc. 18.5, exp. 17.0 kcal/mol)
- 15/22 correctly predicted (using many approximations and not full model)

Book Title

Editor Name

Copyright © Year Pan Stanford Publishing Pte. Ltd.

ISBN

www.panstanford.com

- No bias, more active double mutants identified
- largely automated
- For CALB lowest barriers close to WT, BCX lowest barriers lower than WT

1.2 Motivation

Let us imagine two companies a and b . Both companies use very similar technical equipment to carry out the biotechnological process $X \xrightarrow{\text{Biocatalyst}} Y$ using two different versions of a biocatalyst. Company a uses an enzyme with a rate constant $k_a = 1000s^{-1}$ while company b uses an enzyme with $k_b = 2000s^{-1}$. Letting all other things be equal, the process of company b will therefore only require half the time to produce one Mole of product compared to the time required for company a . Company b therefore can save energy required to keep the reaction volume at a certain temperature. The need for efficient catalysts arises from such an outline and the commercial implications of these considerations are immediate.¹

Increasing the performance of enzymes however is still far from trivial and forms a growing body of research. What is clear though is that the development of such catalysts is costly, in terms of manpower, material and energy – if carried out in the laboratory. A number of companies have in fact formed around this demand: Novozymes (DK), Genzyme (US) or DSM (NL) to name but a few[Meyer *et al.* (2013); Kirk *et al.* (2002); Beilen and Li (2002); Schmid *et al.* (2002)].

Laboratory costs can however be saved to a large part if the development is carried out *in silico*. The proof that computational results are as reliable as experimental results has been provided not too long ago[Claeysens *et al.* (2006)]. The foundation for successful application of computational enzyme engineering has thus been laid out. It is therefore natural to ask: Why hasn't this become more important yet? Why are there not more job openings at pharmaceutical and biotechnology companies looking

¹In this work, we use the terms *enzyme* and *bio-/catalyst* interchangeably.

for computational enzyme engineers? The field, after all, did receive the Nobel price in Chemistry in 2013. As a side note, computational modeling of aero- and fluid dynamics, semiconductor properties, pharmacokinetics/pharmacodynamics or civil engineering is well beyond the point having to justify itself as of strategic relevance. By giving an overview and some details in this article, we hope to, at least in part, answer these questions.

1.3 Introduction

In this review, we aim to provide an introduction into the topic of modeling of enzyme catalysis for people from both within or from outside the field, with emphasis on aspects of activity screening and some emphasis on commercial applicability. We outline the latest achievements, methods in use and required developments to further establish computational enzyme engineering and screening also in commercial practice.

While the process engineering team in company *a* from above is working on improving reaction conditions, consider the enzyme engineer required to engineer the currently used catalyst to match (and possibly outperform) the activity of the catalyst used by company *b*. Following the hypothesis that the function of an enzyme depends on its structure, in order to change the function, some kind of change to the structure has to be introduced and this in general requires the introduction of one or more mutations of amino acids in the primary structure of the enzyme. However, as is well known in the community, it is almost impossible to predict, how a mutated variant of an enzyme will behave[REFS?, Bornscheuer?]. Will it at all be possible to express it and will it even fold and if so, will it be more or less active and specific relative to the wild type? Are its pH properties still the same and is it of similar temperature stability? Even the most reasonably suggested mutations (except for elimination of the catalytic residues) will not allow to predict *qualitatively* with confidence if the mutated enzyme will be more or less active. Therefore it appears as if the only reasonable strategy were to carry out systematic screening

studies.² As was pointed out initially, such studies are expensive in the laboratory and therefore an alternative is to carry them out on the computer. The aim of such initial computational screening is not to provide high-accuracy kinetic data but much rather to provide a reliable, qualitative activity estimate of a preferably large mutant library. It is important to note that while the initial computational screening might be of limited accuracy, its major value is the provision of a systematically developed library of mutants. Based on the screening, promising candidates can then be pipelined into a more demanding computational method for further characterization.

[Physical property of interest:]

The method has to be adapted to the physical property of interest

- Thermal and structural integrity
- pH dependent activity[Ludwiczek *et al.* (2013)]
- Active site hydration
- Binding, *c.f.* drug design where the target protein is constant
- Catalysis: *C.f.* biofuel production or synthesis applications where the substrate is mostly non-variable.

[Different modeling objectives:]

- 1) One Enzyme – Multiple Substrates: These tasks are usually addressed by docking approaches and studies involving millions of substrates have been published[Zhou and Caflisch (2010)].
- 2) Multiple Enzymes – One Substrate: The approaches to this task, as it appears, are much less established. Why is that? A number of factors can be of influence among which are that modeling of different variants of an enzyme will also mean to differentiate between multiple (possibly coupled) protonation states which is non-trivial. Furthermore, modeling of reactivity always involves modeling of a transition state, which is non-trivial and for which there is no standard fail safe approach.

[Application context:]

- Academic: screening can assist in generation of new hypothesis about structure/function relationships
- Commercial: commercial application is restricted by time-to-

²By *screening* we mean the construction of a large library of mutants which are inspected and compared for a specific property like activity.

delivery of predictions, accuracy/reliability, screening spectrum and convenience of use. Notably, in a commercial context, accuracy can be of less value than reliable qualitative prediction of relative activities.

[Review structure and audience:]

This review is written for a more general audience. For specialized details, the reader is referred to the primary literature. Since molecular modeling of enzymes has been reviewed extensively over the recent years, we do not attempt to provide all fundamentals of the subject. Instead, we focus on screening aspects of computational enzyme engineering. The intention is to provide an accessible introduction into the field of quantum chemical modeling of enzymatic reactions, highlight advantages of using quantum chemistry and also point out situations where special care needs to be taken.

By 2014, quantum chemical methods have been in use to study enzymatic reactions for around three decades. What is new however is that quantum chemical methods can now be used also in *in silico* screening assays of enzymatic activity and this is also part of the focus of this review.

1.4 Methods

1.4.1 Modeling Approaches

In all molecular studies, it should be the physical property of interest that guides the selection of the appropriate computational method and not the other way around. At the most fundamental level, computational descriptions of a molecular system is done by (parametric) *force fields* or *ab initio* electronic structure methods.³ Force field methods are useful in studying structural behaviors of enzymes over a significant time period (up to nanoseconds) and can provide details about rearrangement of loop motives. The description of chemical reactivity however requires the description of the electronic structure of the system (molecular mechanics

³ *Molecular mechanics* or *classical mechanics* are other expressions for force field methods.

models do not allow the description of bond formation and breaking processes).

Another categorization can be made among the different modeling approaches among which these days one can name cluster models (kinetic discrimination of mechanisms, transition state testing), hybrid quantum mechanical/molecular mechanical methods (high-accuracy barrier calculations) and full enzyme structure methods (mutant screening studies). Each of these in turn are differently suited to model different properties.

Full enzyme structure methods – General outline of screening method. The presented activity screening method would fall under the last category of the above mentioned. We have used this method in two studies where new functionality has been introduced into an enzyme active site using quantum chemical methods[Hediger *et al.* (2012, 2013a,b)].

In both studies, the basic procedure consists of preparing structures for the enzyme substrate complex (“ES”) and the first intermediate (CALB: “TI”⁴, BCX: “GE”⁵) and calculating the reaction barrier between these two pairs of structures, respectively. The reaction barrier is calculated by interpolating between the ES_{CALB} and TI or ES_{BCX} GE structures, *i.e.* by preparing a set of 10 intermediate structures which approximate the structure of the system along the reaction coordinate.⁶ By evaluating the energies for each interpolation frame for the two reactions ES_{CALB} → TI or ES_{BCX} → GE, it becomes possible to obtain an approximate potential energy surface of the reaction, Fig. 1.1. The quality of the results of this approach are strongly dependent on how careful the modeling of each of these structures is carried out. A frequent reason for meaningless results is significant structural rearrangement during the geometry optimization, requiring to interpolate between very different conformations of the two stationary points on each side of the reaction barrier.

Finally, the activities of different variants of an enzyme are

⁴Tetrahedral intermediate

⁵Glycosyl enzyme

⁶In the literature, the terms *interpolation*, *linear transit scan*, *reaction coordinate calculation* or *adiabatic mapping* are frequently used to indicate the same kind of procedure

ordered by comparing the activation energy required for the rate determining step of the overall reaction – low activation energies correspond to high anticipated activity of the variant. Implicitly this assumes high substrate concentration such that k_{cat} becomes rate limiting, a reasonable assumption under industrial conditions.

Requirements. The most basic requirements are that a structure representing a bound intermediate along the reaction pathway is available. Alternatively, a homology model could be used. Furthermore, ideally the mechanism for the rate determining step is established in the literature. In case no generally accepted mechanism is reported, a preceeding study could analyse kinetically competing mechanism alternatives.

1.4.2 Quantum Chemical Methods

Every method has it's chemistry. The results of one quantum chemical method need to be verified, ideally against high-level standard reference methods. If a method can reproduce the canonical mechanism then there is a good chance that at least qualitatively reliable conclusions can be made using this method.

- AM1/RM1/PM3/PM6/PM7/MOZYME
- DFT/B3LYP/Dispersion corrections/M06/...
- Modeling procedures: Define constraints- how does it affect the calculation (examples in literature?)
- Explain available software: GAMESS, MOPAC
- Explain parameters of PM6: MOZYME, gradient convergence and NDDO cutoff

1.4.3 Software

Methods are of limited use if they can not be applied, therefore available modeling software is a very important factor.

Modeling software is available, PyMOL, Schrödinger, Chem?? (Walter Thiel, commercially) but everything is still very much custom work (especially the data analysis, programs output format is like from last century), this is probably the major reason preventing the use of QM as standard method in industry.

Compare docking and MD methods have software which is designed towards user friendliness and has become established in industry. GTKDynamo is an exceptionally good example for what it should look like, unfortunately only small community.

Many studies claim that “*future developments are directed at further automating the presented approach*”. Obviously these remarks are directed at making the method industrially applicable by avoiding the requirement to carry out complicated setup and configuration steps[Rathore *et al.* (2013)]. In reality, such future developments, to the great unfortune of the whole field, are rarely realized – speculation on the reasons is beyond the scope of this article. Nevertheless, almost no publication fails to mention the potential applications in drug design. Computational considerations are only rarely taken into account but greatly help at putting the whole approach into perspective[Buch *et al.* (2011)].

1.5 Activity Screening Applications

1.5.1 Overview

Independent of the method, a screening assay should be

- Fast
- easy to setup
- qualitative accuracy

Practical applicability of QM based enzyme screening has a couple of problems:

- The structures of the enzyme in the relevant state of the reaction need to be available, ie. an enzyme-substrate complex (ES) and a transition state (TS) analog (usually an inhibited species)
- If the mechanism is not established, the reaction can usually follow various pathways – these would need to be checked in order to find a best guess at the rate determining step
- Identification of the transition state is difficult and depends intimately on the structure, optimizing two

structures differing by 0.1 Å in a particular bond length can result in identification or not identification of the TS

1.5.2 *Candida antarctica* Lipase B

In the studies related to CALB[Hediger *et al.* (2012, 2013a)], we answer the question if quantum chemical methods can be used for enzyme activity screening, develop an efficient modeling protocol and verify the calculated predictions with experimental results. The model system used in this study was CALB and the reaction studied is the hydrolysis of *N*-benzyl-2-chloroacetamide.

UPCOMING

The questions answered in this study are:

- Is it possible to screen for mutant activity using quantum chemical methods?
- Which quantum chemical model can reproduce the accepted mechanism?
- What are the time requirements depending on the model size?
- How should the program be configured?
- What modeling procedure/sequence is needed? Which tools are useful?

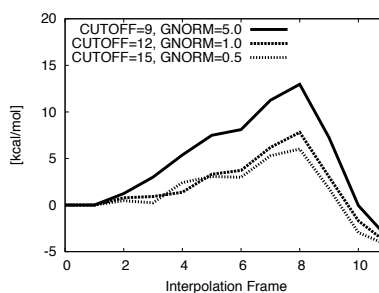
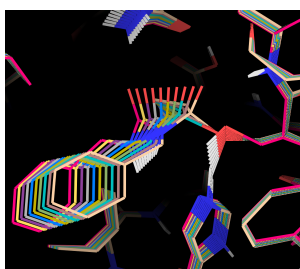


Figure 1.1 Intermediate structures along reaction coordinate of CALB reaction.

CALB is used as a versatile catalyst in a variety of applications among which enzymatic kinetic resolution of enantiomeric

mixtures is of high pharmaceutical importance[Gotor-Fernández *et al.* (2006)]. While CALB is also known for its reactive promiscuity[Svedendahl *et al.* (2008)], if the catalyst can be engineered such as to increase its performance for certain of these applications, possible bottle-necks in production could be widened. About the cutoff: this is only relevant in studies where experimental data for many mutants is available. Then it can be used to assess the lower barrier for the accuracy. Given all approximations, x out of y mutants are correctly predicted. So any method improving on the approximations should provide at least at many correct hits. About the cutoff: this is only relevant in studies where experimental data for many mutants is available. Then it can be used to assess the lower barrier for the accuracy. Given all approximations, x out of y mutants are correctly predicted. So any method improving on the approximations should provide at least at many correct hits. A number of weak points remain:

- Selection of model ambiguous/not full structure
- Not systematically screening all mutants
- No other protein system/enzymatic reaction tested
- Binding effects not considered
- Large rearrangements can not be modeled
- Semi-empirical methods might not be adequate for metals
- Subsequent steps of reaction not modeled (eg. product leaving of active site)
- Different ionization states of residues

Real world applications of BCX: production of biofuels and papers, CALB: pharmaceutical and fine chemicals[REFS?].

1.5.3 *Bacillus circulans* xylanase

In order to address the points summarized above and to show general applicability of the approach outlined for CALB, the xylanase of *Bacillus circulans* (BCX) is engineered to be catalytically active towards an artificial substrate (*ortho*-nitrophenol- β -xylobioside), Fig. 1.2. In this study, the mechanism

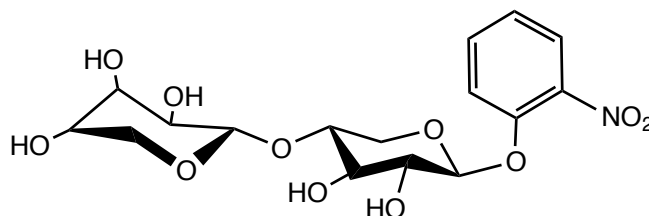


Figure 1.2 *ortho*-nitrophenol- β -xylobioside substrate used in BCX study consisting of two xylose units and ONP.

was taken from the literature as reported previously [Joshi *et al.* (2000, 2001)]. In the following, we outline the work reported in the study and illustrate the sequence of simulation and modeling steps.

WT reaction barrier modeling. In a first step, the wild-type (WT) reference reaction barrier is established. This is a crucial step because all subsequent conclusions about mutants are depending on how good the WT reaction barrier is estimated. As was done for CALB, only the rate determining step is modeled, Fig. 1.3. As outlined in the methods section, the reaction barrier

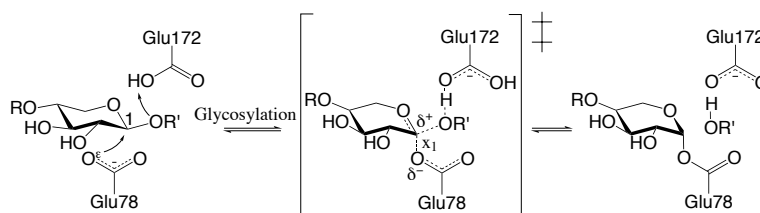


Figure 1.3 Rate determining step in conventional glycosylation. x_1 : constrained reaction coordinate; R : xylose; R' : *ortho*-nitrophenol (ONP). C_1 indicating nucleophilic carbon of first xylose unit [Hediger *et al.* (2013b)].

is mapped out by fixing the reaction coordinate (x_1 in Fig. 1.3) at discrete, decrementing values while optimizing the rest of the protein structure. Since the reaction consists of two concerted events (nucleophilic attack by E78 and proton transfer from E172 to ONP), in principle a two-dimensional potential energy surface could be mapped out to determine the reaction barrier. However, since proton transfer is associated with a rearrangement of the electronic structure, this reaction coordinate is left unconstrained and is handled solely by the quantum chemical method, *i.e.* the proton is free to transfer from or remain on E172. Forcing the quantum chemical model to do something it would not if left unconstrained most likely results in an increased reaction barrier and therefore it is advisable to keep the number of constraints as low as possible.

A PDB structure comprising a covalently linked inhibitor is available (1BVV) and is used as a starting point for the modeling procedure because it gives a very clear indication of how the substrate is located in the active site. As shown in Fig. 2 in Hediger *et al.* (2013b), two possible modeling pathways were worked out. In the first one, the ES complex is prepared from the GE structure *before* the GE structure is optimized (see footnote 5). Alternatively, the GE structure is optimized first and then used as a template for the ES structure. Within this alternative approach, the preparation of the ES structure then just requires minor modifications of substrate conformation before the structure is optimized.

Because in this second approach the ES structure very closely resembles the optimized GE structure, it is possible to concentrate the optimization of the ES structure only on the active site, while the remote part of the enzyme remains fixed at the optimized geometry of the GE structure. This results in very large gains in computational efficiency, which depend on the layer of residues being optimized around the active site.⁷ Fixing, *i.e.* not reoptimizing an already optimized part of the enzyme however also results in slightly increased energy of the ES structure relative to

⁷The structural constraints are parameters to the calculation program and are prepared programmatically using a PYMOL script. The way the parameters are submitted to the calculation may vary depending on the software used for calculations.

the fully optimized GE structure and so the reaction barriers will be decreased. This increase (of ES energy relative to GE) results from multiple minor structural strain forming at the interface between the fixed and optimized layers of the structure. The observations are summarized in Fig. 1.4. Based on this second

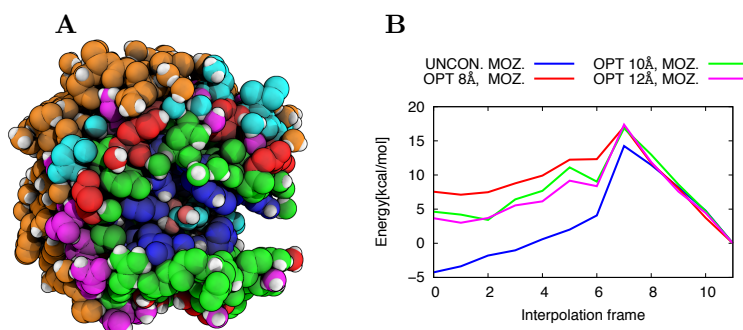


Figure 1.4 **A:** Illustration of optimization layers of BCX with bound ONP substrate. Color coded residue layers have at least one atom within the indicated distance to any atom of the substrate: *green* 8 Å, *red* 10 Å, *purple* 12 Å, *cyan* 14 Å. *Brown* residues remain at their optimized GE geometry, *dark blue* are residues which are screened for active mutations. The substrate is visible in the center of the dark blue residues. **B:** Calculated reaction barriers for different layers of optimized residues (based on Hediger *et al.*(2013b)).

approach, *i.e.* using the optimized GE structure as the template for the ES structure, and when not applying any constraints (apart from the reaction coordinate x_1), the reaction barrier is found to be 18.5 kcal/mol which is in very close agreement to the experimentally reported value of 17.0 kcal/mol obtained from transition state theory[Joshi *et al.* (2000)]. This strongly supports that the outlined modeling procedure, the configuration of the quantum chemical method and the included/omitted physical effects provide a meaningful model of the enzyme kinetics. As pointed out above, when applying constraints on remote parts of the structure, this can decrease the reaction barrier because the ES structure is increased in energy relative to the GE reference point (which is set to zero). However, it is reasonable to assume that this relative increase will be the same for all studied structures and

so cancels out once mutants are compared relative to each other. Overall, application of constraints to remote parts of the enzyme is necessary to reach the desired calculation efficiency of one reaction barrier per day when running each interpolation step in parallel. Furthermore, application of constraints has been widely adopted by the community[Himo (2006); Siegbahn and Himo (2009); Liao and Thiel (2012); Lonsdale *et al.* (2013)]. The time required for optimizing the structures with or without constraints are found to differ by a factor of six (Fig. 4B in Hediger *et al.* (2013b)).

Mutant activity screening. By a similar approach as the one outline for the modeling of the WT reaction, it is found that different ways of modeling the set of mutants results in significantly different levels of efficiency, both in terms of required manual efforts as well as in computational time requirements. We refer to the original publication for details[Hediger *et al.* (2013b)] but point out that when, similarly as for the WT, the ES structures of the mutants are based on templates of optimized mutant GE structures, again large gains in computational efficiency are possible (see Fig. 7 in Hediger *et al.* (2013b)).

Using PYMOL[Schrödinger, LLC (2010)], models of all possible single mutants within a given distance of the substrate are prepared (excluding the catalytically active E78 and E172).⁸ It is noted that in the current implementation of the presented approach, all mutated side chains are in their default ionization state at physiological pH. Furthermore, in a number of cases it is not possible to introduce a large side chain in a spatially restricted environment, such cases are discarded. Lastly, instead of using a side chain rotamer library, each mutated side chain is locally optimized (using an empirical builtin PYMOL function) in the environment of the enzyme before being submitted to the quantum calculation. This approach allowed to screen a set of 317 single mutants, where on average 29 hours are required to complete the calculations of a full reaction barrier using one CPU for each interpolation point.⁹ From this set of single mutants, for

⁸The PYMOL script used for the preparation of the mutants is available at <https://github.com/mzhKU/Enzyme-Screening/blob/master/vsc-bcx.py>.

⁹All obtained barriers are provided at <http://www.scribd.com/doc/133445214/Supp-Mat-Paper-4>.

every position i, j, \dots in the active site, the mutation resulting in the mutant with the lowest barrier X_i, Y_j, \dots is identified. Then, from this set of single mutants $\{K_i\}$, all possible double mutants (X_i, Y_j) are formed. Analysis of the reaction barriers of the obtained candidates reveals that the set of double mutants on average has lower activation barriers than the set of single mutants. Furthermore, the lowest barriers are found among the double mutants.

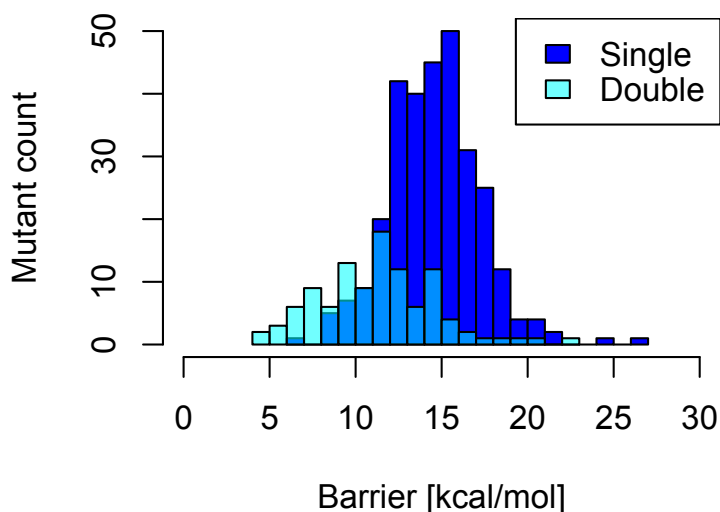


Figure 1.5 Barrier distribution of single and double BCX mutants[Hediger *et al.* (2013b)]. The plot is a histogram of reaction barriers, i.e. how many mutants are identified with a given reaction barrier.

Rationalization of mutants. Analysis of the set of mutants indicates that position 127 has strong influence on the activity (9 of 20 best single mutants carry a mutation at that position). Inspection of the structural arrangement of the Q127W, W9D/E and N35E mutations points to a number of possible explanations of the observed lower barriers, Fig. 1.6. As discussed above, the mechanism is postulated to involve a catalytically active, negatively charged E78 residue as the nucleophile. If Q127 stabilizes the negative charge by hydrogen bonding, the

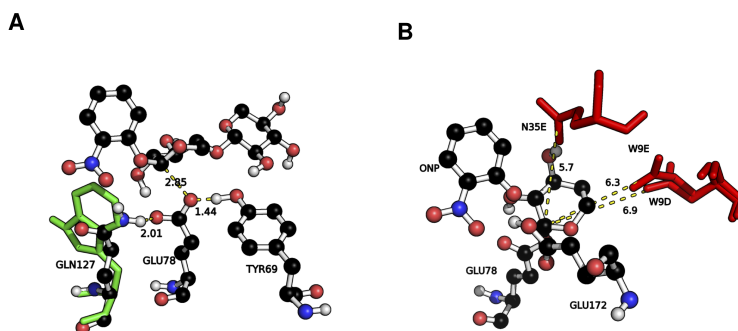


Figure 1.6 Rationalization of reaction barriers. **A** Overlay of WT (black carbon spheres) and Q127W side-chain (green). **B** Stabilizing, coulombic interactions between negative W9D/E or N35E (red) with the nucleophilic, partially positively charged C_1 of the substrate. Distances in Å[Hediger *et al.* (2013b)].

nucleophilic character of E78 is reduced. Replacing the potential hydrogen bond donor Q127 with a non-hydrogen bonding substitute that can preserve overall structural integrity with a side chain of similar size, such as Q127W, could result in increased nucleophilicity of E78, Fig. 1.6 A. Furthermore, as illustrated in Fig. 1.3, the transition state involves a partially positive charge on C_1 of the substrate. Negative charges on side chains (such as in W9D/E) could act as stabilizers of this negative charge and increase catalytic activity in doing so, *i.e.* by stabilization of the transition state. Similarly, when ionized, N35E could act in stabilizing C_1 but when in neutral protonation state, it could be acting as a hydrogen bond donor to the catalytically active E172. This rationalization of N35 mutations would be related to a recently proposed *reverse protonation* mechanism[Joshi *et al.* (2000)], which still is the subject of ongoing research.

1.5.4 Further Screening Applications

The cocaine hydrolase guys

Make a cocaine hydrolase better[Gao *et al.* (2006)].

Biofuel production

The bottle neck in biofuel preparation from biobased materials is hydrolysis of cellulosic materials[REFS?]. Development of improved catalysts clearly could boost this field and has immediate impact on company performance.

Organic synthesis

Gotor *et al.*

1.6 Conclusions

1.6.1 Advantages/Disadvantages

- Bond formation/breaking can be modeled
- Large systems might require constraints to reach convergence
- Calculation times depend strongly on model chemistry used
- Not yet as established and convenient to use like e.g. docking software

1.6.2 Required Future Development Actions

Experimental catalytic data (k_{cat}) of many mutants to validate against.

June 3, 2014 15:32

PSP Book - 9in x 6in

000-main

Bibliography

- Beilen, J. and Li, Z. (2002). Enzyme technology: an overview, *Current Opinion in Biotechnology* **13**, 4, pp. 338–344.
- Buch, I., Giorgino, T. and De Fabritiis, G. (2011). Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations, *Proceedings of the National Academy of Sciences* **108**, 25, pp. 10184–10189.
- Claeyssens, F., Harvey, J., Manby, F., Mata, R., Mulholland, A., Ranaghan, K., Schütz, M., Thiel, S., Thiel, W. and Werner, H. (2006). High-Accuracy Computation of Reaction Barriers in Enzymes, *Angewandte Chemie* **118**, 41, pp. 7010–7013.
- Gao, D., Cho, H., Yang, W., Pan, Y., Yang, G., Tai, H.-H. and Zhan, C.-G. (2006). Computational design of a human butyrylcholinesterase mutant for accelerating cocaine hydrolysis based on the transition-state simulation, *Angewandte Chemie International Edition* **45**, 4, pp. 653–657.
- Gotor-Fernández, V., Busto, E. and Gotor, V. (2006). Candida antarctica lipase b: an ideal biocatalyst for the preparation of nitrogenated organic compounds, *Advanced synthesis & catalysis* **348**, 7-8, pp. 797–812.
- Hediger, M. R., De Vico, L., Rannes, J. B., Jäckel, C., Besenmatter, W., Svendsen, A. and Jensen, J. H. (2013a). In silico screening of 393 mutants facilitates enzyme engineering of amidase activity in calb, *PeerJ* **1**, p. e145.
- Hediger, M. R., De Vico, L., Svendsen, A., Besenmatter, W. and Jensen, J. H. (2012). A computational methodology to screen activities of enzyme variants, *PLoS ONE* **7**, 12, p. e49849, doi:10.1371/journal.pone.0049849, URL <http://dx.doi.org/10.1371/journal.pone.0049849>.
- Hediger, M. R., Steinmann, C., De Vico, L. and Jensen, J. H. (2013b). A computational method for the systematic screening of reaction barriers in enzymes: searching for bacillus circulans xylanase mutants with greater activity towards a synthetic substrate, *PeerJ* **1**, p. e111.

20 | Bibliography

- Himo, F. (2006). Quantum chemical modeling of enzyme active sites and reaction mechanisms, *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **116**, 1, pp. 232–240.
- Joshi, M., Sidhu, G., Nielsen, J., Brayer, G., Withers, S. and McIntosh, L. (2001). Dissecting the electrostatic interactions and pH-dependent activity of a family 11 glycosidase, *Biochemistry* **40**, 34, pp. 10115–10139.
- Joshi, M., Sidhu, G., Pot, I., Brayer, G., Withers, S. and McIntosh, L. (2000). Hydrogen bonding and catalysis: a novel explanation for how a single amino acid substitution can change the ph optimum of a glycosidase1, *Journal of molecular biology* **299**, 1, pp. 255–279.
- Kirk, O., Borchert, T. V. and Fuglsang, C. C. (2002). Industrial enzyme applications, *Current opinion in biotechnology* **13**, 4, pp. 345–351.
- Liao, R.-Z. and Thiel, W. (2012). Comparison of qm-only and qm/mm models for the mechanism of tungsten-dependent acetylene hydratase, *Journal of Chemical Theory and Computation* **8**, 10, pp. 3793–3803.
- Lonsdale, R., Houghton, K. T., Zurek, J., Bathelt, C. M., Foloppe, N., de Groot, M. J., Harvey, J. N. and Mulholland, A. J. (2013). Quantum mechanics/molecular mechanics modeling of regioselectivity of drug metabolism in cytochrome p450 2c9, *Journal of the American Chemical Society* **135**, 21, pp. 8001–8015.
- Ludwiczek, M. L., D'Angelo, I., Yalloway, G. N., Brockerman, J. A., Okon, M., Nielsen, J. E., Strynadka, N. C., Withers, S. G. and McIntosh, L. P. (2013). Strategies for modulating the ph-dependent activity of a family 11 glycoside hydrolase, *Biochemistry* **52**, 18, pp. 3138–3156.
- Meyer, H.-P., Eichhorn, E., Hanlon, S., Lütz, S., Schürmann, M., Wohlgemuth, R. and Coppolecchia, R. (2013). The use of enzymes in organic synthesis and the life sciences: perspectives from the swiss industrial biocatalysis consortium (sibc), *Catalysis Science & Technology* **3**, 1, pp. 29–40.
- Rathore, R., Sumakanth, M., Reddy, M. S., Reddanna, P., Rao, A. A., Erion, M. D. and Reddy, M. (2013). Advances in binding free energies calculations: Qm/mm-based free energy perturbation method for drug design, *Current pharmaceutical design* **19**, 26, pp. 4674–4686.
- Schmid, A., Hollmann, F., Park, J. B. and Bühler, B. (2002). The use of enzymes in the chemical industry in europe, *Current Opinion in Biotechnology* **13**, 4, pp. 359–366.
- Schrödinger, LLC (2010).

- Siegbahn, P. E. and Himo, F. (2009). Recent developments of the quantum chemical cluster approach for modeling enzyme reactions, *JBIC Journal of Biological Inorganic Chemistry* **14**, 5, pp. 643–651.
- Svedendahl, M., Carlqvist, P., Branneby, C., Allnr, O., Frise, A., Hult, K., Berglund, P. and Brinck, T. (2008). Direct epoxidation in candida antarctica lipase b studied by experiment and theory, *ChemBioChem* **9**, 15, pp. 2443–2451, doi:10.1002/cbic.200800318, URL <http://dx.doi.org/10.1002/cbic.200800318>.
- Zhou, T. and Caffisch, A. (2010). High-Throughput Virtual Screening Using Quantum Mechanical Probes: Discovery of Selective Kinase Inhibitors, *ChemMedChem* **5**, 7, pp. 1007–1014.