



**HAL**  
open science

## Computational design of a lipase for catalysis of the Diels-Alder reaction

Mats Linder, Anders Hermansson, John Liebeschuetz, Tore Brinck

► **To cite this version:**

Mats Linder, Anders Hermansson, John Liebeschuetz, Tore Brinck. Computational design of a lipase for catalysis of the Diels-Alder reaction. *Journal of Molecular Modeling*, 2010, 17 (4), pp.833-849. 10.1007/s00894-010-0775-8 . hal-00602989

**HAL Id: hal-00602989**

**<https://hal.science/hal-00602989>**

Submitted on 24 Jun 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Editorial Manager(tm) for Journal of Molecular Modeling  
Manuscript Draft

Manuscript Number: JMM01313R1

Title: Computational design of a lipase for catalysis of the Diels-Alder reaction

Article Type: Original paper

Keywords: Candida Antarctica lipase B, near attack conformer, molecular dynamics, density functional theory, in silico, promiscuous reactivity, rational engineering

Corresponding Author: Prof. Tore Brinck, Ph.D.

Corresponding Author's Institution: Royal Institute of Technology (KTH)

First Author: Mats Linder, MS

Order of Authors: Mats Linder, MS; Anders Hermansson, BS; John Liebschuetz, PhD; Tore Brinck, Ph.D.

Abstract: Combined molecular docking, molecular dynamics (MD) and density functional theory (DFT) studies have been employed to study catalysis of the Diels-Alder reaction by a modified lipase. Six variants of the versatile enzyme *Candida Antarctica* lipase B (CALB) have been rationally engineered *in silico* based on the specific characteristics of the pericyclic addition. A kinetic analysis reveals that hydrogen bond stabilization of the transition state and substrate binding are key components of the catalytic process. In the case of substrate binding, which has the greater potential for optimization, both binding strength and positioning of the substrates are important for catalytic efficiency. The binding strength is determined by hydrophobic interactions and can be tuned by careful selection of solvent and substrates. The MD simulations show that substrate positioning is sensitive to cavity shape and size, and can be controlled by a few rational mutations. The well-documented S105A mutation is essential to enable sufficient space in the vicinity of the oxyanion hole. Moreover, bulky residues on the edge of the active site hinders the formation of a sandwich-like near-attack conformer (NAC), and the I189A mutation is needed to obtain enough space above the face of the  $\alpha,\beta$ -double bond on the dienophile. The double mutant S105A/I189A performs quite well for two of three dienophiles. Based on binding constants and NAC energies obtained from MD simulations combined with activation energies from DFT computations, relative catalytic rates ( $k_{cat}/k_{uncat}$ ) of up to  $10^3$  are predicted.

Response to Reviewers: We are happy with the favorable comments by the reviewers. We have corrected the manuscript as suggested by reviewer 1

1. A sentence explaining catalytic promiscuity has been added.
2. A paragraph discussing the relationship between the NAC concept and the Reaction Force has been added.

We have shortened the manuscript as suggested by reviewer 2. In particular, we have moved information regarding the structural relaxation of protein structure in MD simulations to the supplementary material.



Noname manuscript No.  
(will be inserted by the editor)

# Computational Design of a Lipase for Catalysis of the Diels-Alder reaction

Mats Linder · Anders  
Hermansson · John  
Liebschuetz · Tore  
Brinck

Received: date / Accepted: date

**Abstract** Combined molecular docking, molecular dynamics (MD) and density functional theory (DFT) studies have been employed to study catalysis of the Diels-Alder reaction by a modified lipase. Six variants of the versatile enzyme *Candida Antarctica* lipase B (CALB) have been rationally engineered *in silico* based on the specific characteristics of the pericyclic addition. A kinetic analysis reveals that hydrogen bond stabilization of the transition state and substrate binding are key components of the catalytic process. In the case of substrate binding, which has the greater potential for optimization, both binding strength and positioning of the substrates are important for catalytic efficiency. The binding strength is determined by hydrophobic interactions and can be tuned by careful selection of solvent and substrates. The MD simulations show that substrate positioning is sensitive to cavity shape and size, and can be controlled by a few rational mutations. The well-documented S105A mutation is essential to enable sufficient space in the vicinity of the oxyanion hole. Moreover, bulky residues on the edge of the active site hinders the formation of a sandwich-like near-attack conformer (NAC), and the I189A mutation is needed to obtain enough space above the face of the  $\alpha,\beta$ -double bond on the dienophile. The double mutant S105A/I189A performs quite well for two of three

dienophiles. Based on binding constants and NAC energies obtained from MD simulations combined with activation energies from DFT computations, relative catalytic rates ( $v_{cat}/v_{uncat}$ ) of up to  $10^3$  are predicted.

**Keywords** Diels-Alder · Rational design · CALB · DFT · Molecular Dynamics

## 1 Introduction

The Diels-Alder reaction is remarkably rare in nature, [1] which may be attributed to a large activation energy due to the highly ordered transition state. To date, only a few naturally occurring enzymes coined "Diels-Alderase" have been characterized, [2] and these are very substrate-specific. The mechanistic studies performed on Macrophomate Synthase (MPS), the only enzyme performing bimolecular Diels-Alder, [3] also show that the catalytic mechanism is likely to be a two-step, Michael-Aldol pathway, rather than a concerted [4 + 2] addition. [4] Attempts to employ artificial 'enzyme mimics', for example cyclodextrin, catalytic antibodies and other macrocycles [5] that utilize substrate recognition and transition state (TS) binding, have yielded only modest results due to product inhibition. [6]

To find a versatile Diels-Alderase is highly desirable, owing to the synthetic use of creating up to four stereocenters in one concerted step. The concept of non-covalent, hydrogen bond (H-bond) catalysis of Diels-Alder using carbonyl-activated dienophiles has undergone rapid development in the last decade, [7, 8, 9] and since activation of carbonyl-containing substrates to nucleophilic attack is very common in enzymes, in particular the large hydrolase family (EC 3.x), the idea to utilize a promiscuous member of such an enzyme family for Diels-Alderase activity is not far-fetched, and has been suggested in the literature. [10]

Among the hydrolase family, lipases are by far the most commonly utilized for *in vitro* catalysis of organic reactions, such as (dynamic) kinetic resolution of secondary alcohols, transesterifications and acylations. [11, 12, 13, 14] Many lipases have been found to be stable under a broad range of temperature and pH, and active in organic solvents, [15, 16] which greatly expands their utility in synthetic applications. More recently, evidence has emerged of catalytic promiscuity [17, 18] (*i.e.* the ability of an enzyme to exhibit catalysis of other than its native reaction) towards *e.g.* aldol condensations, [19, 20] epoxidations [21] and Michael additions, [22, 23, 24, 25, 26] not seldom enhanced by point mutations. [27] An often-seen feature in promiscuous catalysis is that only parts of the catalytic machinery are used, so that removal of one or more of

Mats Linder, Anders Hermansson and Tore Brinck  
Physical Chemistry, Royal Institute of Technology (KTH)  
Tel.: +46-(0)8-7908210  
E-mail: linder@physchem.kth.se, tore@physchem.kth.se

John Liebeschuetz  
Cambridge Crystallography Data Centre  
Tel.: +44-(0)1223-762532  
Fax.: +44-(0)1223-33603  
E-mail: john@ccdc.cam.ac.uk

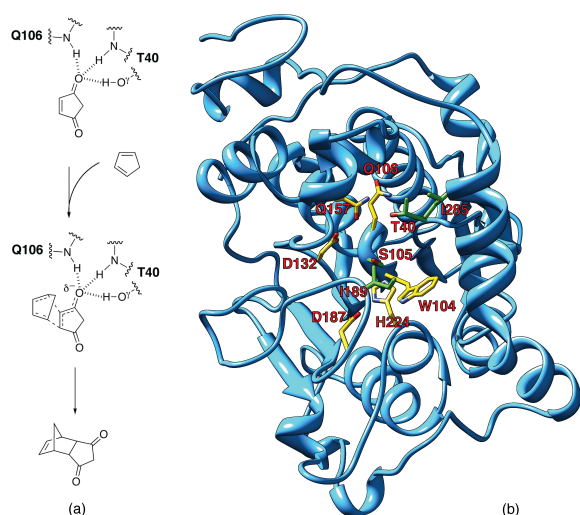


Fig. 1: (a) The hypothetic CALB-catalyzed Diels-Alder mechanism, (b) CALB crystal structure (PDB entry 1LBT), with residues important for the present study explicitly shown. Residues subject to mutations are shown in green.

these residues might actually enhance the desired activity. [28] Following these concepts, the hypothetical reaction mechanism for hydrolase-catalyzed Diels-Alder is shown in Fig. 1a. Note that the only part of the catalytic machinery thought to influence Diels-Alder catalysis is the oxyanion hole, rendering the catalytic triad obsolete. This effectively removes the dependence on dynamic binding in the terminology of Kirby, [6] and means that there is a large degree of freedom in screening and (re)design of candidate enzymes.

Building on previous studies of *Candida Antarctica* Lipase B (CALB), performed by us and collaborators, [19, 20, 22, 21] we set out to test this versatile enzyme for Diels-Alder activity. CALB was chosen because of its high documented catalytic promiscuity and tolerance for a broad range of solvents and reaction conditions. Fig. 1b shows the structure of CALB, which belongs to the large  $\alpha/\beta$  hydrolase superfamily. [29]

Initial experimental efforts with CALB wild type (wt) and S105A mutant (shown to have enhanced catalytic activity for carbon-carbon bond formation [20, 22, 23, 24]) did not indicate any Diels-Alder activity. [30] Therefore, we turned our attention to modeling the enzyme-substrate interactions, in order to *i*) identify the obstacles for CALB Diels-Alder activity, and *ii*) explore how rational design might improve the system. We have employed both molecular dynamics (MD) and quantum chemical (QC) methods, as well as molecular docking, in order to obtain all facets of the complex enzyme-substrate(s) system. To our knowledge, *in silico* studies of enzymes with non-natural substrates are

scarce, and we therefore seek to find a theoretical framework for prediction of reactivity, catalytic power, and productive combinations of enzyme species and substrates. Unlike most other computational design studies which focus on optimum transition state stabilization, including the groundbreaking *de novo* designed enzymes by Baker and colleagues, [31, 32] we attempt to monitor the full reaction pathway, including substrate binding and substrate-substrate interactions. This approach is supported by a very recent study by Simón and Goodman, [33] who show that oxyanion holes are not optimized for maximum TS stabilization, but rather for minimizing the reaction barrier.

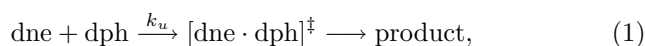
As seen in our previous work, H-bond catalysts can provide an enthalpic TS stabilization of  $\sim 1-5$  kcal $\cdot$ mol $^{-1}$ , depending on reactants and polarizability of the medium. [34] These values are expected to be similar using lipase oxyanion holes, so the main concern for the viability of lipases such as CALB therefore lies in the ligand-receptor dynamics. From data on molecular catalysis reported in the literature, we can conclude that rather high catalyst concentrations are required for a reasonable activation, typically around 1-20 mol-%, [35, 36, 37, 38] which is not possible with enzymes. Thus the binding constant between the enzyme and both substrate must be sufficiently low.

Six CALB variants were simultaneously studied in this work. The S105A mutant was explored alongside the wild-type (wt), due to its previously reported utility. I189 was identified as a steric constraint near the oxyanion hole. Thus, the I189A and S105A/I189A mutants were probed. I189 and I285 constitute the 'channel' dividing the active site in the small 'stereoselectivity pocket' and larger hydrophobic (or acyl) pocket. [39] This channel is the most sterically crowded region near the oxyanion hole, yet the region where the sandwich-like DA TS must be located in order to exploit the catalytic machinery of CALB. Consequently, we also modeled the I285A and S105A/I285A mutants.

## 2 Theory & Methods

### 2.1 Kinetics and Binding Dynamics

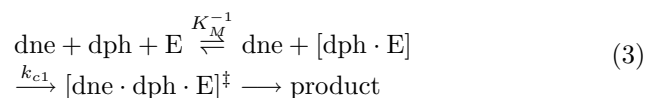
In order to estimate the catalytic potential of the Diels-Alder reaction by a given enzyme, we need expressions relating the different states along the catalytic pathway, as well as a means to calculate their relative energies. The reaction formula for the uncatalyzed reaction is given in eq. 1.



for which the rate expression is

$$v_u = k_u[\text{dne}][\text{dph}] \quad (2)$$

As a first approximation, we can consider the case where we have no specific interaction between the diene and the enzyme active site, i.e. there is no productive binding of the diene (note that the dienophile is assumed to bind to the oxyanion hole with the binding constant  $K_M$ , as in 1a).



This reaction can be described by the well-known Michaelis-Menten kinetics.[40]

$$K_M = \frac{[\text{E}][\text{dph}]}{[\text{dph} \cdot \text{E}]} = \frac{([\text{E}^0] - [\text{dph} \cdot \text{E}])([\text{dph}])}{[\text{dph} \cdot \text{E}]} \quad (4)$$

The rate expression is then given by eq. 5.

$$v_{cat} = k_{c1}[\text{dne}][\text{dph} \cdot \text{E}] = \frac{k_{c1}[\text{E}^0][\text{dne}][\text{dph}]}{K_M + [\text{dph}]} \quad (5)$$

Dividing eq. 5 by eq. 2, we obtain the ratio

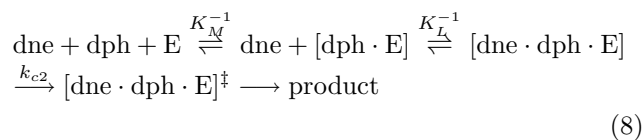
$$\frac{v_{cat}}{v_u} = \frac{k_{c1}}{k_u} \frac{[\text{E}^0]}{K_M + [\text{dph}]} \quad (6)$$

We see that according to eq. 6, the catalytic effect is maximized by using high enzyme and low dienophile concentrations. Moreover, the relative rate is independent of diene concentration, and as long as  $K_M \ll [\text{dph}]$ , the value of  $K_M$  has little influence on the relative rate. It should be noted that this model simplifies the comparison between the catalyzed and the uncatalyzed reaction. The entropic cost of bringing the diene to the dienophile is similar for the two reactions, and the solvation/desolvation effects are also similar. Thus, the relative rate can easily be estimated from two sets of quantum chemical calculations; one on the uncatalyzed reaction and one on the catalyzed reaction with a simple model of the active site.

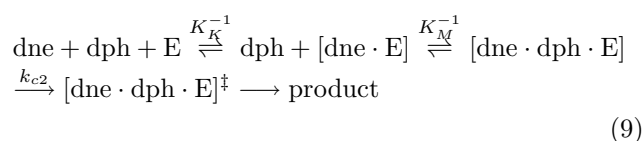
$$\frac{k_{c1}}{k_u} = e^{-(\Delta G_{c1}^\ddagger - \Delta G_u^\ddagger)/RT} \approx e^{-(\Delta E_{c1}^\ddagger - \Delta E_u^\ddagger)/RT} \quad (7)$$

Eq. 6 describes well the kinetics we have for a reaction catalyzed by an organic catalyst without a binding pocket. It is also expected to work well for enzymes when we have a nonpolar diene and a nonpolar solvent, i.e. when there is no binding to the active site due to hydrophobic or electrostatic effects. Clearly for this model, small organic catalyst are more favored since they allow for higher catalyst concentration.

In an enzyme it is also possible to have a specific binding of the diene to the active site. This can lead to increased reaction rates, since the entropic cost of bringing the reactants together is compensated for by the attractive binding of diene. In this case the model have to be refined according to eq. 8.



By neglecting any bias on the order of substrate binding, one can equivalently write



The rate expression is in this case given by eq. 10. [41, 42]

$$v_{cat} = \frac{k_{c2}[\text{E}^0][\text{dne}][\text{dph}]}{K_M K_L + K_M[\text{dne}] + K_L[\text{dph}] + [\text{dne}][\text{dph}]} \quad (10)$$

The rate relative the uncatalyzed reaction is given by

$$\frac{v_{cat}}{v_u} = \frac{k_{c2}}{k_u} \frac{[\text{E}^0]}{K_M K_L + K_M[\text{dne}] + K_L[\text{dph}] + [\text{dne}][\text{dph}]} \quad (11)$$

The difference between  $k_{c1}$  and  $k_{c2}$  is the reference state. By comparing eqs. 3, 8 and 9, one realizes that

$$k_{c2} = k_{c1} K_L \quad (12)$$

Inserting this expression into eq. 11 yields the eq. 13.

$$\frac{v_{cat}}{v_u} = \frac{k_{c1}}{k_u} \frac{[\text{E}^0]}{K_M + [\text{dph}] + \frac{K_M[\text{dne}]}{K_L} + \frac{[\text{dne}][\text{dph}]}{K_L}} \quad (13)$$

We note that for large  $K_L$ , i.e. no specific binding of the diene by the enzyme, eq. 13 reduces to eq. 6. However, it is important to note that the value of  $k_{c1}$  is not independent of  $K_L$ . When  $K_L$  is low there are interactions in the active site that stabilize the diene relative its free solvated state. These interactions are also likely to be present at the transition state, and lower its energy relative the free state. Thus, the approach to calculate  $\frac{k_{c1}}{k_u}$  from quantum chemical calculations using a simple model of the active site is no longer applicable. Fig. 2 illustrates the relation between the various states discussed in the text, as well as the difference between  $k_{c1}$  and  $k_{c2}$ .

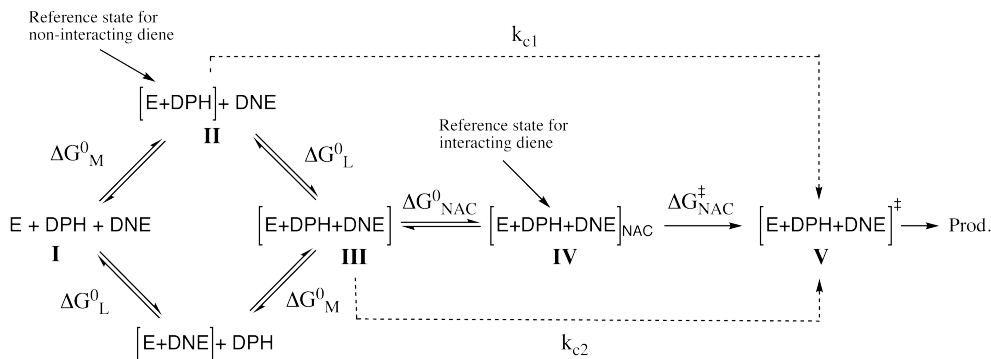


Fig. 2: Reaction pathway showing the difference between the catalyzed and uncatalyzed reactions.  $\Delta G_M^0$  and  $\Delta G_L^0$  refer to the indexing in eqs. 3-11.

The challenge now lies in accurately calculating the (relative) free energies of the relevant states. The free energy of activation for the uncatalyzed reaction can be calculated directly with quantum chemical methods by correcting the energies for solvation effects and thermal effects, using implicit solvent models and independent particle statistical mechanics, respectively. Binding constants between substrate and enzyme can be calculated from MD trajectories, using classical methods based on statistical sampling, *e.g.* MM/PB(GB)SA [43] or Linear Interaction Energy (LIE). [44, 45] The main problem is computing the relative free energy of the enzyme-catalyzed TS. In principle this can be accomplished by free energy integration methods and a QM/MM description of the enzyme system. However, these methods are very computationally demanding due to the need for extensive sampling of the configurational space, and there are still technical problems associated with the representation of the quantum-classical interface. Instead, we prefer to use a simple model of the oxyanion hole as a QC representation of the enzyme, and seek a common reference state along the reaction coordinate for the quantum and classical regimes. One such reference state is the 'near-attack conformer' (NAC). [46, 47]

The NAC can be viewed as a point, or an area, on the potential energy surface, located between the bound enzyme-substrate complex and the TS, where the geometry of the system is preorganized for forming the TS and the atoms of the bonds that are formed in the TS are within van der Waals distance. Since the geometry of the NAC is similar to the TS, the interactions between the active site and the substrate(s) are similar in the NAC and the TS, except for those that directly influence the bond breaking/forming. Thus it can be anticipated that quantum chemical calculations with even a small model of the active site can reproduce the free energy difference between the two states.

Justification for the division of the potential energy surface also comes from the reaction force concept of Toro-Labbe and coworkers. [48, 49, 50] They have shown for a number of different reaction types that there is a narrow transition region on the potential energy surface (PES) where the main changes in electronic properties take place, and where the system goes from activated reactants to activated products. Structural changes not associated with bond breaking/formation and environmental effects, such as solvation, mainly influence the PES before and after the transition region.

The free energy of activation going from the enzyme substrate complex to the TS is then given by eq. 14.

$$\Delta G_{c2}^\ddagger = \Delta G_{NAC} + \Delta G_{NAC}^\ddagger \quad (14)$$

where  $\Delta G_{NAC}$  is the free energy of forming a NAC, and  $\Delta G_{NAC}^\ddagger$  is the free energy of activation from the NAC. MD simulations provide information of the NAC frequency, which is obtained by statistical analysis of the geometries in the trajectory. Analogous to transition state theory, one can regard the relationship between NACs and non-NACs as an equilibrium. Thus, the relative population of NACs from a set of trajectory points gives an estimate of  $\Delta G_{NAC}$ .

If the diene binds tightly to the active site and in a geometry that is close to the NAC geometry,  $\Delta G_{NAC}$  will be low; typically for a natural enzyme binding a natural substrate,  $\Delta G_{NAC}$  is in the order of 0.5-2 kcal·mol<sup>-1</sup>. [46, 47] However, if the substrate binds loosely to the active site and retains substantial freedom of moment,  $\Delta G_{NAC}$  will be higher. Even higher values of  $\Delta G_{NAC}$  are expected if the substrate binds tightly but in a non-productive fashion, *i.e.* in a geometry that is very different from the NAC. In the case that  $\Delta G_{NAC}$  is high, its determination by unconstrained MD simulations also becomes less accurate since only a very small number of configurations will be NACs and the sampling will be

insufficient. This is particularly true if stringent criteria for NAC geometries are enforced.

On the basis of  $\Delta G_{\text{NAC}}$  obtained from the MD simulation and  $\Delta G_{\text{NAC}}^\ddagger$  from the quantum chemical calculation it is possible to estimate  $\Delta G_{c_2}^\ddagger$ . Combined with the free energy estimate of the uncatalyzed reaction,  $k_{c_2}/k_u$  can be calculated (*c.f.* Fig. 2):

$$\frac{k_{c_2}}{k_u} = e^{-(\Delta G_{c_2}^\ddagger - \Delta G_u^\ddagger)/RT} \quad (15)$$

Thus, by a combination of MD simulation and quantum chemical calculations it is possible to estimate the catalytic performance for different substrates and enzyme combinations using eq. 11. There is, unfortunately, substantial uncertainty involved in the determination of the different energy terms, which makes it hard to obtain accurate global energy estimates. However, relative rates are likely to be of better precision. In particular, for ranking the performance of different enzyme variants, e.g. mutations, the  $\Delta G_{\text{NAC}}$  can be of great importance. For this purpose less stringent NAC criteria can be employed, which improves the statistical significance of the results. A first estimate of the quality of a given enzyme variant or substrate is however most easily obtained by visual and statistical analysis. The ability to form NACs has for this purpose been used as an important indicator of the viability of a given enzyme variant.

## 2.2 Enzyme variants

The CALB model was obtained from the PDB structure 1LBT. [39] The entry was chosen due to the presence of a large lipid analog covalently bound to the catalytic S105, which putatively induces a fully active conformation of the enzyme. The all-atom RMSD to two other common structures in modeling is small; 1TCA [51] = 0.29, 1LBS [39] = 0.44. The 1LBT structure was stripped of all water molecules and protonated (default protonation state) in the viewer Hermes, a part of the GOLD Suite. [52] This structure was subsequently used as the basis for mutant models, which were generated in SwissPDB viewer. [53]

## 2.3 Molecular docking

Selection of dienophiles and generation of initial studies were accomplished by docking simulations, using the GOLD Suite. [54, 55, 56, 57] The docking procedure was conducted according to the following protocol: The built-in genetic algorithm (GA) was used to

generate different poses, which were evaluated with the ChemScore scoring function. [58, 59] The space of possible poses was limited to relevant ones by inducing hydrogen bond constraints between the dienophile carbonyl oxygen and T40-NH, T40-O<sup>γ</sup>H and Q106-NH. To somewhat counteract the limitations created by these constraints, the 'diverse solution' flag was switched on, using default parameters.  $10^5$  GA runs were used for each docking pose, and 50 poses were generated for each molecule. Statistics were calculated on the top 10 of these solutions, based on total score. The poses were further analyzed according to a composite scoring function, which is explained in detail in the Electronic Supplementary Material (ESM). After numerical and visual analysis, the best poses of each dienophile were saved for docking of the diene. The same set of parameters were used as for docking the dienophiles, except now, no constraints were used to limit the possible poses.

Docking in the MD-relaxed variants was performed analogously. Since the average structures (Section 3.1) resulted in unphysical positions of a large portion of hydrogens, all hydrogens were stripped off and the enzymes were reprotonated in GOLD. The H-bond constraints were limited to T40-NH and Q106-NH since the T40-O<sup>γ</sup>H was seen to contribute less to binding of the substrates, and turned away somewhat during relaxation.

## 2.4 Molecular dynamics simulations

All MD simulation were performed with the AMBER 9 & 10 packages, [60] and the AMBER 99 force field [61] was employed throughout. All initial structures were taken from molecular docking. Force field parameters for the nonprotein atoms (dienophiles and diene) were generated from the general Amber force field (gaff), [62] using the Antechamber program. The generation of gaff-parameters was based on AM1 geometries, and the AM1-BCC approach was used for partial charges. [63, 64] All amino acid residues were assumed to be in their natural protonation state.

The LEaP program was used to create topology files, neutralize the enzyme complexes with Sodium ions, and add an octahedral box of solvent in a minimum layer of 8 Å. Acetonitrile (ACN) was treated by the available united-atom model, [65] while chloroform (CHL) [66] and water (WAT, described by TIP3P) [67] parameters were taken from all-atom models.

Simulations were performed under periodic boundary conditions and the long-range electrostatics was handled by particle Mesh Ewald summation. Initial minimization was performed in two steps of 1000 and 2500 iterations, respectively, where the protein was held fixed



during the first step. Next, the systems were heated to 300 K over 20 ps with mild restraints on the protein atoms ( $10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}$ ). Unconstrained simulations followed for 2-4 ns at 300 K and 1.0 bar. 2 fs time steps were used, both during warmup and the following simulations, and hydrogen movements were constrained by the SHAKE algorithm. [68]

Trajectories were based on snapshots taken every 0.2 ps. Statistics were obtained using the PTRAJ tool within AMBER. For structural analysis, an average structure was generated based on the last 1000 snapshots (200 ps). In these structures, hydrogen positions were poorly defined, so when performing subsequent docking, all hydrogens were stripped off and the relaxed enzymes were reprotonated in GOLD.

For simulations of substrates **1** and **3** in water (see sections 3.4 and 3.5), the binding free energies of both diene and dienophile were estimated using the LIE method. [44, 45] LIE approximates  $\Delta G_{bind}$  by an electrostatic contribution  $V^{elec}$  (Coulomb interactions) and a non-polar (van der Waals) contribution  $V^{vdw}$ . These contributions are weighted by coefficients  $\alpha$  and  $\beta$ , calibrated to Free Energy Perturbation (FEP) data and empirical data, respectively.

$$\Delta G_{bind} = \alpha[\langle V_{l-s}^{elec} \rangle_{bound} - \langle V_{l-s}^{elec} \rangle_{free}] + \beta[\langle V_{l-s}^{vdw} \rangle_{bound} - \langle V_{l-s}^{vdw} \rangle_{free}], \quad (16)$$

where  $\alpha$  has been determined to 0.43 and  $\beta$  to 0.18 for the neutral molecules used here. [69] 300 samples were taken from the last half of the 2 ns trajectories. The LIE method requires explicit simulations of the free solvated substrates; they were thus simulated for 500 ps, from which 200 samples were used. Both average values and standard deviation were stable with respect to a 5-fold increase in sampling rate.

It should be noted that we also attempted to calculate binding free energies using the MM/PBSA and MM/GBSA approaches, but they lead to more divergent results and irrational magnitudes.

## 2.5 DFT calculations

A model of the oxyanion hole was created from the parent crystal structure 1LBT, using the docked structures as starting guesses for the dienophile and TS. The central peptide atoms of the backbone Q106 and T40 nitrogens were included together with adjacent  $\alpha$ -carbons, in order to correctly model the electrostatics of the hydrogen bonds. The T40 side chain was included as well as methyl group on residue 105, representing both serine and alanine. Finally, the entire Q106 side chain was included because it is seen to H-bond to the back of the

T40 OH-group in CALB crystal structures. A graphical representation of the oxyanion hole model is provided in ESM Fig. 5.

Geometry optimizations were performed at the B3-LYP/6-31G(d) level, [70, 71, 72, 73] for the [E-dph], NAC-[E-dph-dne] (the  $C_{dne}^4-C_{dph}^\alpha$  and  $C_{dne}^1-C_{dph}^\beta$  distances being fixed at 4.0 Å, see the discussion in Section 3.3) and [E-dph-dne]<sup>‡</sup> states. Gaussian 03 was used for all calculations. [74] An *endo* conformation of the TS was chosen, since it is generally accepted to be the kinetically favored geometry. [75] Because of the truncated oxyanion hole model, several atoms were kept frozen throughout the optimization, as illustrated in ESM Fig. 5. As this generally leads to more than one negative force constant in the calculated Hessian, no thermodynamic corrections were made on the oxyanion hole structures. Where stated, estimates of the corrections were instead made by using the Hessian from the corresponding uncatalyzed reactions. This is a reasonable estimate since eigenstates of the enzyme are expected to change little during the course of reaction. The transition state were detected by identifying the most negative imaginary frequency as being the formation of the two incipient C-C bonds.

Single point calculations on the optimized structures were performed using a 6-31G++(d,p) basis set and the C-PCM solvent model. [76, 77] We used a dielectric constant of  $\epsilon = 4$  to simulate the polarity of a typical enzymatic environment, and the shape of the cavity was determined by the United Atom Topological Model. [78]

## 2.6 Graphical representation

All molecular representations were created using the UCSF Chimera package. [79, 80] Diagrams were created using XMGrace [81] and Plot. [82]

# 3 Results and discussion

## 3.1 Enzyme variants

To test for relative stability and look for unexpected distortions of the enzyme variants, 4 ns MD simulations were carried out on each variant in the three solvents. In all simulations, the backbone RMS trajectories level out after 1 ns, indicating that 4 ns is enough simulation time for this study.

For structural comparison, average structures of the last 200 ps of simulations were abstracted. Their backbone RMS deviations (RMSDs) with respect to the

crystal structure are shown in ESM Table 1. No major structural change occurs in any solvent, compared to that of the wild type, which suggests that all mutants are potentially stable. As expected, the water simulations display the lowest RMSDs with respect to the crystal structure, while the CHL-relaxed variants show the highest RMSD values.

An extensive analysis of solvent and mutant influences on the structure is given in the ESM. Here, we summarize by noting that the essential catalytic machinery is well-preserved, but that the T40 side-chain moves away from its crystal position where the OH-group stabilizes the oxyanion. A question thus arises concerning its role in catalysis, which will be addressed below. The S105 residue, where present, tends to move into the designated oxyanion hole region, an observation that advocates mutation. Finally, the I189A and I285A mutations did not generate any major structural distortions, and should thus be viable.

ACN was chosen for further MD studies of enzyme-substrate interactions, since it does not significantly distort the cavity, while it is a moderately polar solvent, which is believed to improve the affinity between the substrates and the hydrophobic active site. Moreover, using an organic solvent is of interest in synthesis, and (immobilized) CALB is known to operate well in dry environments.

### 3.2 Molecular docking

Among a larger set of molecules, benzoquinone **1**, maleic anhydride **2** and cyclohexenone **3** (Fig. 3), all three common reactants for Diels-Alder, were selected for further study. One strong motivation for this choice is that all dienophiles are cyclic, thus locked in an *s-trans* conformation. Although *s-cis* is widely considered to be a more activated conformation, [75] a *trans* dienophile has the  $\beta$ -carbon further away from the oxyanion hole when bonded, which might be necessary in the crowded active site. In addition, rings are more facile to model since the conformational space is limited.

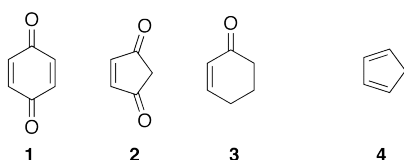


Fig. 3: Structure of Diels-Alder substrates used in this study.

From each dienophile docking, one pose was selected onto which the diene (cyclopentadiene **4**) was subse-

quently docked, using the same set of parameters as for the dienophile runs. **4** was chosen due to its frequent use in Diels Alder in the literature, and again the fact that the cyclic structure limits its mobility and facilitates simulation. The best [E·dph·dne] complex for each variant and dienophile was then subjected to MD simulations, together with the best [E·dph] complexes.

### 3.3 MD simulations of substrates

Both the [E·dph] and [E·dph·dne] complexes were simulated for 2 ns each, since the relaxation of the uncomplexed variants indicated that the RMSD leveled off after 1 ns. From the simulations, trajectories of backbone RMSDs, H-bond distances between the oxyanion hole and carbonyl oxygen, and B-factors were calculated. We note in passing that the atomic fluctuations are similar to those of the uncomplexed variants in ACN, but that there is no clear trend regarding how the flexibility is affected by different point mutations, similar to the uncomplexed simulations (ESM Fig. 7). Nor is there any significant difference between the flexibility pattern for complexes of different substrates. RMS deviations over the last 0.2 ns of simulation are summarized in Table 1. Distance trajectories for oxyanion hole hydrogens and dienophile oxygen are presented in Fig. 4.

The occurrence of H-bonding between dienophile and oxyanion hole was evaluated using a threshold distance of 2.75 Å between the hydrogen atom and the acceptor, and calculating the fraction of saved geometries fulfilling this threshold for each H-bond donor (Q106-NH, T40-NH and T40-O $\gamma$ H). To avoid bias originating from the unrelaxed starting guess, only the last 1 ns (5000 frames) of each simulation was used. The fractions for each variant and dienophile are presented in Fig. 5. One can see, both from Fig. 4 and Fig. 5, that in some cases, H-bonding of the dienophile is slightly induced by the introduction of a diene, which effectively blocks the escape route. But overall, the pattern from the [E·dph] simulations is consistent.

From the H-bond trajectories in Fig. 4, one clearly sees that the two most important donors are the backbone nitrogens of T40 and Q106. T40-O $\gamma$ H participates only sparsely. The available crystal structures of CAL-B [51, 39] reveal that T40 accepts H-bonds from the amide side chains of Q106 and Q157 (Fig. 1), and at least one is maintained in the simulations, accompanied by a slight distortion of the backbone, also observed in the uncomplexed variants (ESM Fig. 2). Thus, it is questionable that T40-OH participates in the oxyanion hole during substrate recognition. This supports the previous observation that only two donors are important in the oxyanion hole at any one time, even in enzymes

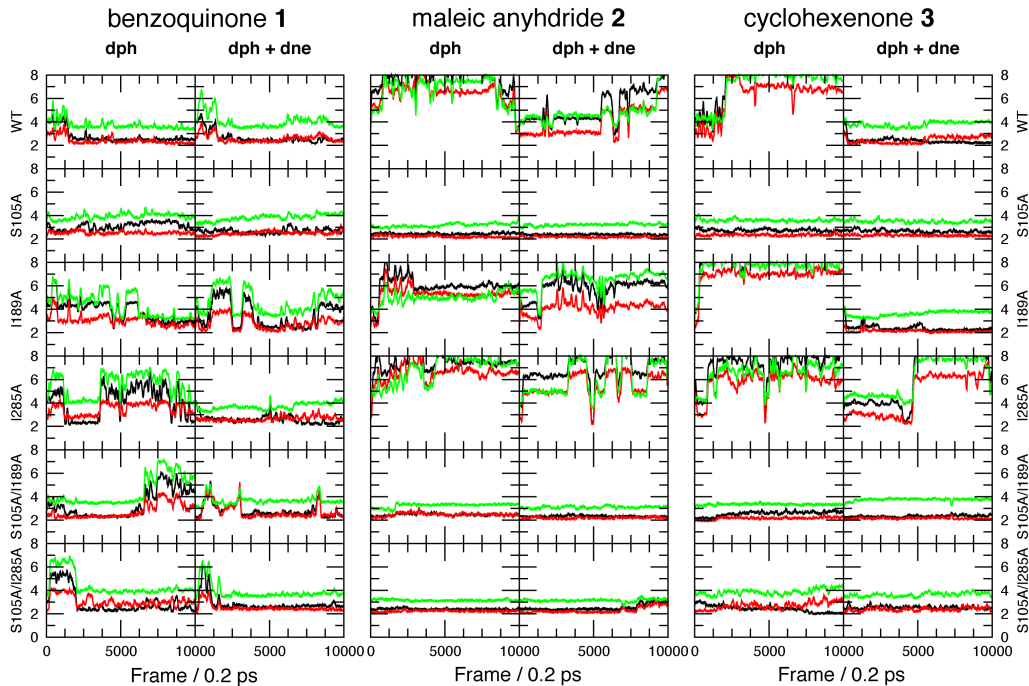


Fig. 4: H-bond trajectories from the 2 ns simulations of all substrates in the 6 variants. Distances are between Q106-NH (black), T40-NH (red) and T40-O<sup>7</sup>H (green), and acceptor oxygen without and with cyclopentadiene **4** present. Distances are in Angstroms and each trajectory is presented as a 100-frame running average, for clarity, meaning that extreme values are cut off.

Table 1: Backbone RMS deviations ( $\text{\AA}$ ) for substrate complexes compared to corresponding uncomplexed variant in ACN.<sup>a</sup>

| Variant     | 1           | 1+4         | 2           | 2+4         | 3           | 3+4         |
|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| WT          | 0.83 (0.93) | 0.94 (0.75) | 0.76 (0.85) | 0.80 (1.13) | 0.90 (0.80) | 0.99 (0.75) |
| S105A       | 1.00 (0.87) | 1.07 (0.88) | 0.94 (1.01) | 1.06 (1.01) | 0.98 (0.86) | 1.06 (1.04) |
| I189A       | 0.85 (1.04) | 0.75 (1.03) | 0.92 (0.97) | 1.01 (0.92) | 0.92 (1.06) | 1.07 (0.97) |
| I285A       | 0.89 (0.92) | 0.76 (0.76) | 1.34 (1.41) | 0.71 (0.99) | 0.84 (1.06) | 1.13 (1.31) |
| S105A/I189A | 1.05 (1.15) | 0.98 (1.09) | 1.16 (1.09) | 1.04 (0.86) | 1.41 (1.39) | 1.05 (1.04) |
| S105A/I285A | 1.18 (1.04) | 1.16 (0.92) | 1.24 (0.90) | 1.16 (1.19) | 1.24 (0.77) | 1.13 (0.88) |

<sup>a</sup> Based on average structures from the last 0.2 ns of simulation (1000 frames). Numbers in parentheses refer to RMS deviations with the 1LBT crystal structure.

having donors in the sidechains. [20, 21] However, as seen in the 1LBS crystal structure, [39] the covalent intermediate (and possibly the TS) binds to all three H-bond donors, perhaps by some geometric distortion induced by the tetrahedral structure.

To evaluate the frequency of NACs between diene and dienophile, we used a similar procedure as given above for H-bonds. We have used a slightly less rigorous definition than the original one stated by Bruce and colleagues, [46, 47] as this allows us to reduce the sampling of data. Most important for our model to be

valid is that the NAC definition is identical in both the MD and QC regime (see section 2.1). The 'loose NAC' is thus defined as a structure where both the dne-C<sup>1</sup>-dph-C <sup>$\alpha$</sup>  and dne-C<sup>4</sup>-dph-C <sup>$\beta$</sup>  distances are  $\leq 4.0 \text{ \AA}$  simultaneously. As with the H-bond evaluation, only the last 5000 frames (1 ns) are taken into account. Fig. 6 shows the fractions of structures fulfilling these criteria, accompanied by the corresponding fractions with only one incipient bond being within 4.0  $\text{\AA}$ . Note that for **1**, which has two identical  $\alpha, \beta$ -double bonds, data is presented for the one with the highest num-

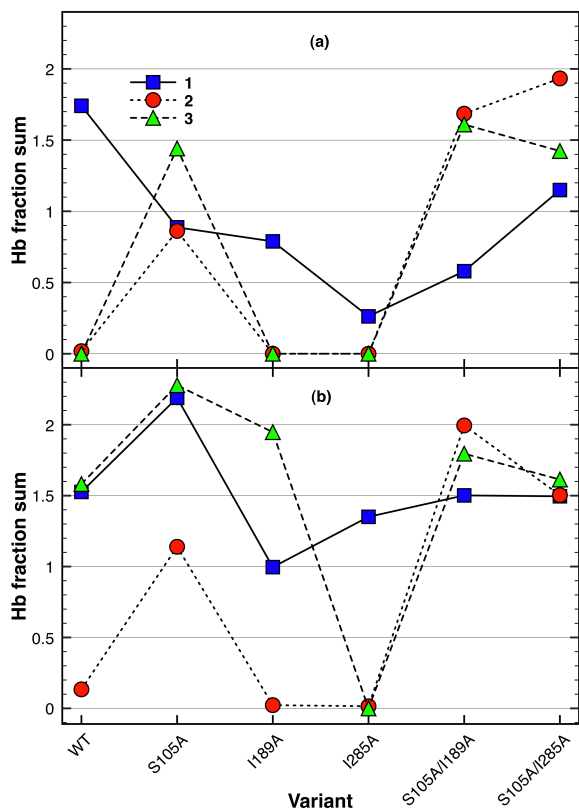


Fig. 5: Sum of fraction of the 3 putative H-bonds between oxyanion hole and substrate acceptor (Q106-NH, T40-NH and T40-O<sup>γ</sup>H, threshold 2.75 Å), calculated on the last half of each simulation. (a) Dienophile complex, (b) dienophile + diene complex.

ber of frames fulfilling the NAC criterion. It should be noted that angular parameters are only evaluated visually, therefore the calculated NAC frequencies are overestimations and should be taken as upper limits. With these loose NAC criteria, we obtain up to 1–4 % NAC-structures with **1**, corresponding to a  $\Delta G_{\text{NAC}}$  of  $\sim 2.8$ – $1.9$  kcal/mol. Even though seldom in close association with the dienophile, the diene is clearly attracted by the hydrophobic cavity, which is expected to compensate for the entropic penalty of bringing the two reactants together. Following the distance trajectories (not shown), it is apparent that the diene occasionally diffuses away before termination. This happens in the [1·S105A], [1·I189A], [1·S105A/I189A], [2·WT], [2·S105A], [2·I285A], [2·S105A/I189A], [3·I285A] and [3·S105A/I285A] complexes. It is difficult to tell whether this is a stochastic event or a direct effect of the particular substrate-enzyme combination. The relatively high fraction of NACs in **1** is partly due to the molecular symmetry; benzoquinone has two identical double bonds that could be subject to attack. A more detailed discussion for each dienophile follows below.

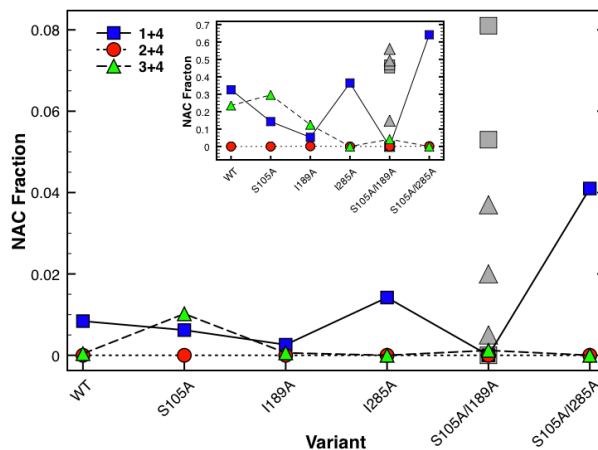


Fig. 6: Simultaneous incipient bond NAC fractions (threshold 4.4 Å). The inset shows corresponding fraction with only one of two distances fulfilled. Results from additional simulations of **1** and **3** in S105A/I189A (see section 3.4) are shown in gray.

*Benzoquinone.* Statistically, benzoquinone **1** is the dienophile that appears to be best suited for Diels-Alder in a CALB variant. However, no simulation of **1** + **4** display any real ‘sandwich’-like conformation between the substrates, which is essential for the appropriate orbital overlap. One reason is that **1** is a long molecule that clashes with residues surrounding the catalytic triad, especially I189, hence it cannot assume a productive conformation, which exposes the flat face to the solvent, as seen in Fig. 7a. We therefore expected the situation to improve when introducing the I189A mutation, but were surprised to see that the the NAC frequencies in I189A and S105A/I189A were in fact lower than in all other variants. From visual analysis, it appears that **1** tends to fall out of the active site in absence of I189, possibly due to lack of confinement. On the other hand, it is clear that I189 hinders a good sandwich-like NAC conformation in the other variants, including S105A/I285A, which at first glance seems to fulfill the loose NAC criteria as defined (see Fig. 7b). But because I285 is on average further away from the oxyanion hole, we deem it less influential on forming activated complexes.

*Maleic anhydride.* Maleic anhydride **2** clearly needs the S105A mutation to bind to the oxyanion hole (Fig. 4). Interactions with the diene is poor in all variants, as illustrated by virtually no detected NACs (Fig. 6). The diene remains reasonably close to **2** in the two double mutants, but the position of both diene and dienophile is less consistent than for the other substrates; in S105A/I285A, the diene resides in the small subpocket, close to W104 (Fig. 8a), while it prefers the large subpocket

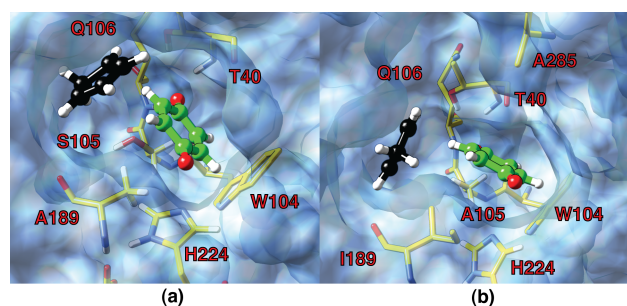


Fig. 7: Selected snapshots from MD simulations of **1** in (a) I189A and (b) S105A/I285A, showing selected residues and hydrogens. Both illustrations show typical T-shaped interactions. The reactants come close, but not typically in the sandwich-like conformation required for TS formation. In (a) it is shown that S105 blocks part of the oxyanion hole, forcing **1** to tilt away and bind less strongly. (b) shows that the I285A mutation does not create as much space in the immediate vicinity of the oxyanion hole. There is no increase in sandwich-like conformations relative to the S105A mutant.

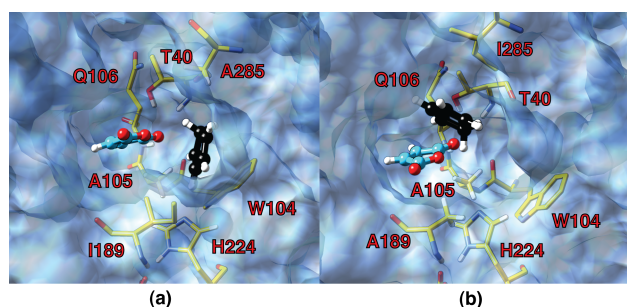


Fig. 8: Selected snapshots from MD simulations of **2** in (a) S105A/I285A and (b) S105A/I189A, showing selected residues and hydrogens. In both illustrations, it is shown how **2** tilts to the left, effectively making the  $\alpha,\beta$ -double bond unavailable for attack. (a) shows an unusual diene position, made available because of the dienophile tilt. A more frequent position is shown in (b), but the buriedness of the dienophile hinders attack by the diene.

in S105A/I189A (which is the common region in most simulations), as displayed in Fig. 8b. **2** also tends to tilt more deeply into the cavity, which effectively blocks the  $\alpha,\beta$  double bond; see Fig. 8a and b. This could be due to undesired interaction with the second, equivalent carbonyl group.

*Cyclohexenone.* Like **2**, cyclohexenone **3** is more tightly bound to the oxyanion hole with the S105A mutation present (Fig. 4). The double bond generally faces the large subpocket, where the diene resides in all simulations except the S105A/I285A variant, where it escapes quickly. From Fig. 6, we see that the NAC frequencies are not very impressive, and that the 1-bond NAC frequency tends to decrease with mutations of the bulky I189 and I285 residues. The latter observa-

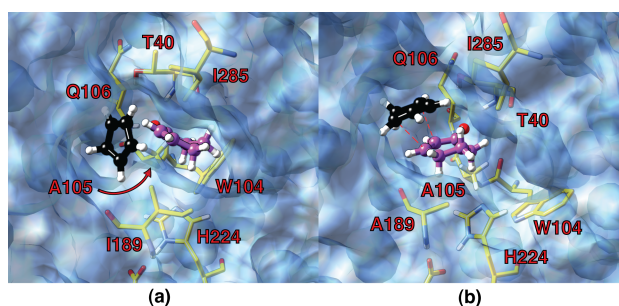


Fig. 9: Selected snapshots from MD simulations of **3** in (a) S105A and (b) S105A/I189A, showing selected residues and hydrogens. In (a) it clearly shows that with both isoleucins present, there is insufficient space above and below the dienophile for TS formation. (b) shows an example of a good NAC conformation. The indicated incipient bond distances are:  $C_{dne}^{\alpha}-C_{dph}^{\alpha} = 3.79 \text{ \AA}$ ,  $C_{dne}^1-C_{dph}^{\beta} = 4.15 \text{ \AA}$ . Closer contact is hindered by T40.

tion can be a confinement effect; visual analysis reveals that S105A/I189A produces the desired sandwich-type conformation of substrates, albeit with a low NAC frequency. The S105A mutation alone generates T-shaped diene-dienophile interactions (Fig. 9a), similar to the few interactions produced with **2**. It is evident that with both I189 and I285 present, little space is available for a Diels-Alder TS. [**3**·**4**·S105A/I189A] is the one simulation where at least one conformation fulfills the more rigorous NAC criterion that the two incipient bonds should be within  $3.5 \text{ \AA}$ , while also maintaining a reasonable NAC geometry (Fig. 9b).

Even when conditions are seemingly optimized for Diels-Alder, such as in S105A/I189A, the  $C^{\alpha}$  of the dienophiles is obscured by the loop containing T40, in particular the T40 side chain. This is a further reason why so few 2-bond NAC conformations are detected, although 1-bond proximity is quite common in many instances.

### 3.4 Further optimization

The two main purposes with each point mutation was to create more space for the dienophile to fit into the oxyanion hole without having to be covalently bound to the enzyme, and make possible the NAC geometry necessary for Diels-Alder. S105 tends to block the oxyanion hole during a large part of all simulations; consequently, as Fig. 5a shows, the best H-bonding is obtained with the variants containing the S105A mutation. This is further evidenced by the H-bond trajectories. The fact that **1** breaks the pattern somewhat could be due to favorable interactions of the conjugated molecule with *e.g.* W104 and some of the hydrophobic residues.

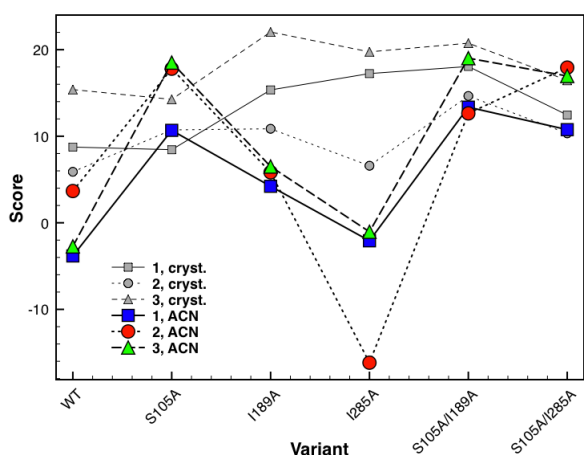


Fig. 10: Docking consensus score compared between the 6 variants and solvent-relaxed structures versus the crystal form. The consensus score is calculated according to the protocol given in the supplementary information, and includes contributions from individual dienophile, diene and TS poses.

The I189A and I285A mutants were investigated since the large isobutyl side chains tend to clash with either the dienophile or the diene during molecular docking. In terms of H-bonding to the oxyanion hole, however, neither mutant provided any major improvement. If anything, the dienophile recognition became worse, which may have to do with unrealistic starting guesses. Since the molecules start within the active site, clashes with I189 or I285 may actually lock the substrate in position, while the smaller alanine replacement allows for escape. In the two double mutants, The H-bond frequency increases, but now without any steric clashes locking the substrates in the active site.

From the above discussion, we draw the conclusions that *i*) the S105A mutation is necessary to obtain acceptable binding between the oxyanion hole and the cyclic enones **1-3**; *ii*) the I189A mutation makes possible an adequate positioning of the diene for a pericyclic reaction; *iii*) T40 obstructs the diene from approaching the dienophile C $\alpha$ , while the alcohol side chain is obsolete in the oxyanion hole; *iv*) **2**, although the most reactive of the dienophiles, fails to assume favorable geometries in the active site; *v*) **1** associates weakly with the oxyanion hole, possibly due to more delocalized charge, but has general good affinity to the diene, although it is too bulky to assume productive NACs in most variants; *vi*) **3** binds strongly to all S105A mutants and produces good NAC-like geometries in some variant, but at low frequencies. It is also a less reactive substrate.

Additional support for the S105A and I189A mutants are found when comparing results of molecular docking in 4 ns relaxed variants (last 1000 frame av-

erage structures) with those in the unrelaxed crystal structure (as was used originally). When docking dienophiles in the unrelaxed variants, the results are quite invariant to mutations. In contrast, in ACN and WAT-relaxed structures, S105A mutants give rise to higher scores, while the remaining three scored poorly. This is mainly due to clashes; the steric constraints in the relaxed variants are much higher, unless S105 is substituted (see figure S1). Furthermore, a 'consensus' approach, utilizing dienophile, diene and TS composite scores,\* shows that when considering the complete reaction pathway in the relaxed structures, S105A/I189A gives the highest score for both **1** and **3** (see Fig. 10).

Based on the above discussion, we therefore conducted the following additional investigations:

(A) *Multiple simulations of 1 and 3 with 4 in S105A/I189A.* One problem with our approach is that it is difficult to sample the conformational space of the diene in the cavity. We wanted to see if more NACs could be found in S105A/I189A by additional MD simulations, since post-relaxation docking so strongly indicated this variant to be favorable. Therefore, using three different starting geometries and velocities, three additional short (1 ns) simulations were performed on these two systems, using the same protocol as previously. The geometries were selected from the previous simulation, selecting ones where NAC-like interactions were present.

(B) *1 and 3 with 4 in S105A/I189A, using water as solvent.* To investigate if water, which is more polar than ACN, increases the affinity between the substrates and active site, one 2 ns simulation was executed for each of the [S105A/I189A-**1**], [S105A/I189A-**1-4**], [S105A/I189A-**3**] and [S105A/I189A-**3-4**] complexes. The generated trajectories were also used to calculate LIE energies.

(C) *Simulation of a T40A/S105A/I189A variant in ACN.* Since a general obstacle for forming good NACs is the T40 sidechain, a triple mutant T40A/S105A/I189A was generated *in silico*, and relaxed in ACN for 4 ns, while 2 ns simulations of [T40A/S105A/I189A-**3**] and [T40A/S105A/I189A-**3-4**] were performed in parallel.

Results from (A) are presented in Table 2, giving H-bond data and NAC statistics. Protein RMSDs and flexibilities were similar to the previous 2 ns simulations. The high NAC fractions, displayed in gray in Fig. 6, give further support to the S105A/I189A variant. Once again, we see evidence that the T40 alcohol

\*See the ESM for a detailed description of scoring and consensus.

Table 2: H-bond and NAC statistics for the three 1 ns simulations of **1** and **3** with **4** in S105A/I189A.<sup>a</sup>

| Entry                         | dph H-bond $\bar{d}$ [ $\text{\AA}$ ] <sup>b</sup> |             |                       | NAC fractions <sup>c</sup> |                    |
|-------------------------------|--|-------------|-----------------------|----------------------------|--------------------|
|                               | Q88-NH   | T40-NH      | T40-O $\gamma$ H      | Loose <sup>d</sup>         | Tight <sup>e</sup> |
| 1 <b>1</b> , I                | 2.87 (0.61)  | 3.01 (0.61) | 3.47 (0.07)           | 0.053                      | 0.0016             |
| 2 <b>1</b> , II               | 2.37 (0.91)  | 2.30 (0.91) | 3.62 ( $\bar{}$ 0.01) | 0.081                      | 0.0022             |
| 3 <b>1</b> , III <sup>f</sup> | 3.05 (0.35)  | 2.47 (0.81) | 4.21 (0.01)           | 0                          | 0                  |
| 4 <b>3</b> , I                | 2.75 (0.52)  | 2.31 (0.91) | 3.76 ( $\bar{}$ 0.01) | 0.037                      | 0.0004             |
| 5 <b>3</b> , II               | 2.36 (0.87)  | 2.14 (0.98) | 3.83 ( $\bar{}$ 0.01) | 0.020                      | 0.0002             |
| 6 <b>3</b> , III              | 2.58 (0.75)  | 2.14 (0.98) | 3.43 ( $\bar{}$ 0.01) | 0.005                      | 0.0002             |

<sup>a</sup> 1 ns simulations, 5000 saved geometries. <sup>b</sup> Distances are averaged over all 5000 frames. Numbers in parentheses show the fraction of distances below a threshold value of 2.75  $\text{\AA}$ . <sup>c</sup>Based on the two-bond criterium described in the text. <sup>d</sup>4.0  $\text{\AA}$  threshold. <sup>e</sup> 3.5  $\text{\AA}$  threshold. <sup>f</sup> The diene left the cavity shortly after the beginning of the simulation, similar to the original 2 ns run.

is no participant in the oxyanion hole in our system. Intriguingly, the loose NAC geometries are now quite frequent in two of three simulations with each substrate. We also detect some 'tight' NAC poses, although they are still very scarce. With this set of simulations, we have shown that **1** can bind to the S105A/I189A and form NACs, and that the null result from the first setup was not generalizable, but that multiple simulations are preferable for sufficient sampling. However, **1** seems to be bonded more weakly to the oxyanion hole than **3**, as implied by the average H-bond distances in Table 2. Visual analysis of the trajectories also reveals that **1** tumbles around in the active site more than **3**, and occasionally falls out of the oxyanion hole for a short time. The best single simulation in this set is entry 2 in Table 2, in which a sandwich-like NAC is maintained almost constantly (see the supplementary information for an animation). These data further multiple trajectories are indeed important for efficient sampling of the conformational space, which is a standard approach of *e.g.* Pleiss and coworkers. [83, 84]

The saturated framework of **3** makes it more difficult to form sandwich-like interactions with **4** than for **1**. There is a smaller volume available for the diene to move in to form NACs with **3**, and this volume is often occupied by surrounding residues, most dominantly I285, T40 and Val140. The prevalent interaction is a T-shaped geometry, with **4** on the double bond side of **3**. Still, **4** displays good affinity with the active site and dienophile.

Results from the water simulations (B) are given in Table 3. We note that while **3** sticks tightly to the oxyanion hole, **1** tumbles around slightly, as observed in ACN. Hence the high average H-bond distances. We see a 14-fold increase in NAC fractions for **3**, compared to the initial simulation in ACN, and the NAC fraction for **1** is even higher. The relationship is more or less in concert with the findings in (A). More importantly, in both simulations, there is an event where the diene temporarily falls out of active site but readily moves

back in. In ACN, this does not occur; if the diene starts to diffuse out, it does not stop. Thus, a highly polar solvent is preferable when working with substrates such as those in the present study. The issue of Diels-Alder catalysis by water[75] is not explored further here, but clearly, there will be competition.

An extensive analysis of the T40A/S105A/I189A triple mutant (C) is given in the ESM. Here it suffices to say that while H-bonding of the dienophile was surprisingly improved, the formation of NACs with the diene was poor. Both phenomena are believed to arise from a relief of steric crowding around the oxyanion hole. This causes allows the dienophile to bind harder and become more buried, which in turn makes it less accessible to the diene. This finding highlights the importance of delicate tuning around the catalytic site in order to align both reactants.

### 3.5 Binding Constants

We have estimated the binding constants ( $K_M$  and  $K_L$ ) of substrates **1**, **3** and **4** to S105A/I189A in water using the LIE method, see Table 4. The results for the dienophiles are in agreement with earlier experimental and theoretical studies on the binding of small ketones and aldehydes to CALB and other lipases; the binding constants in water are typically in the range of 0.1-1 M, and generally decrease with increasing hydrophobicity of the substrates. [21, 85, 86] Our lowest binding constant is obtained for cyclopentadiene **4**, which has no specific binding to the active site. This supports the interpretation that the main driving force for binding is hydrophobic interactions. In earlier study on the epoxidation of butenal in CALB wt and S105A we obtained experimentally a  $K_M$  value of 0.2 M in aqueous buffer, whereas no saturation was found in ACN up to concentrations of 3M. [21] Thus, to obtain sufficiently low binding constants we conclude that nonpolar substrates and water as solvent should be used.

Table 3: RMS, H-bond and NAC statistics for simulations of **1** and **3** with and without **4** in S105A/I189A, with WAT as solvent.<sup>a</sup>

| Entry      | RMSD [Å] <sup>d</sup> | dph H-bond $\bar{d}$ [Å] <sup>b</sup> |             |                      | NAC fractions <sup>c</sup> |                    |
|------------|-----------------------|---------------------------------------|-------------|----------------------|----------------------------|--------------------|
|            |                       | Q88-NH                                | T40-NH      | T40-O <sup>γ</sup> H | Loose <sup>e</sup>         | Tight <sup>f</sup> |
| <b>1</b>   | 1.01 (0.91)           | 2.85 (0.39)                           | 2.15 (0.98) | 3.55 (0.00)          | 0.085                      | 0.0020             |
| <b>1+4</b> | 0.95 (0.80)           | 3.11 (0.32)                           | 3.13 (0.33) | 3.19 (0.12)          |                            |                    |
| <b>3</b>   | 1.05 (1.03)           | 3.64 (0.01)                           | 2.35 (0.88) | 3.52 (0.03)          | 0.017                      | 0.0006             |
| <b>3+4</b> | 0.83 (0.76)           | 2.35 (0.94)                           | 2.16 (0.98) | 3.40 (0.01)          |                            |                    |

<sup>a</sup> 2 ns simulations, 10000 saved geometries. <sup>b</sup> Distances are averaged over last 5000 frames. Numbers in parentheses show the fraction of distances below a threshold value of 2.75 Å. <sup>c</sup>Based on the two-bond criterium described in the text. <sup>d</sup>Backbone atoms, vs. the 4 ns relaxed structure (vs. crystal structure) <sup>e</sup>4.0 Å threshold. <sup>f</sup>3.5 Å threshold.

Table 4: Calculated estimates of binding parameters.<sup>a</sup>

| MD <sup>b</sup>        | Benzoquinone <b>1</b>  |              |                     |
|------------------------|------------------------|--------------|---------------------|
|                        | I→II                   | II→III       | III→IV <sup>c</sup> |
| $\Delta\Delta E_{el}$  | 2.6                    | 0.6          | -                   |
| $\Delta\Delta E_{vdW}$ | -1.4                   | -1.2         | -                   |
| $\Delta G$             | 1.2                    | -0.7         | 1.8                 |
| const.                 | $K_M = 7.3$            | $K_L = 0.33$ | $K_{NAC} = 0.047$   |
| MD <sup>b</sup>        | Cyclohexenone <b>3</b> |              |                     |
|                        | I→II                   | II→III       | III→IV <sup>c</sup> |
| $\Delta\Delta E_{el}$  | 1.3                    | -0.5         | -                   |
| $\Delta\Delta E_{vdW}$ | -1.6                   | -1.0         | -                   |
| $\Delta G$             | -0.3                   | -0.4         | 2.4                 |
| const.                 | $K_M = 0.61$           | $K_L = 0.48$ | $K_{NAC} = 0.017$   |

<sup>a</sup> Roman numbers refer to 2. All energies are in kcal·mol<sup>-1</sup>. <sup>b</sup> Electrostatic and van der Waals energies calculated by the LIE method, and scaled by 0.43 and 0.18, respectively. <sup>c</sup> Calculated from NAC averaged over all 5 [S105A/I189A-dph-dne] simulations in ACN and WAT.

### 3.6 DFT calculations

Solvent-corrected B3LYP/6-31++G(d,p) relative energies are reported in Table 5, together with corresponding values for the uncatalyzed reaction. Energies for [E·dph] are also presented for reference. Fig. 11 shows the [E·dph] and TS optimized geometries, including oxyanion hole H-bond distances.

It should be noted that because of the observation from MD that the T40-OH frequently points away from the oxyanion hole, we were led to investigate if it has any influence on the stabilization of the TS. For this comparison, suitable snapshots were extracted from the S105A/I189A trajectories with **4** and **1**, **2** and **3**, respectively, and oxyanion hole models were generated as described in the methods section. It was found that the crystal structure version of the model consistently stabilized the TS more than the relaxed structure models, while the dienophile stabilization was more similar in both (slightly better in the crystal model for **2** and **3**, and slightly better in the relaxed model for **1**). This finding supports the above proposal that all three H-bond donors present in the CALB oxyanion hole contribute to TS stabilization, while the association with a car-

bonyl substrate does not necessarily involve T40-OH. We concluded that the crystal structure geometry was a better representation of the enzyme-TS complex, and therefore only present results obtained from this model.

We note that in the activated TS,  $d_\beta < d_\alpha$ , which is a general finding for this type of reaction. The  $d_\beta$  bond forms earlier along the reaction coordinate[87], which could somewhat mitigate the steric hindrance from T40. H-bond distances are generally smaller in the TS than in the [E·dph] complex, as seen in Fig. 11, which reflects the oxyanion hole preference for (partial) negative charge. It was somewhat surprising to see that the T40-OH distances are constantly among the shortest in the model, but this illustrates the sensitivity of enzyme-substrate bond dynamics. With the loop in the right position, T40-OH is a strong H-bond donor, while it is often excluded from the oxyanion hole by the enzyme motion. Moreover, neither dienophile nor TS assume the otherwise optimum planar geometry between carbonyl group and proton donors. This is in concert with the finding that oxyanion holes are optimized to give a minimum barrier between the TS and substrate complex, which occurs at a geometry which is not a minimum in either case. [33]

With the [E·dph] complex taken as reference state (simple model), **1** is the dienophile that undergoes the largest TS stabilization ( $\Delta G_{c1}^\ddagger - \Delta G_u^\ddagger = -3.9$  kcal·mol<sup>-1</sup>). The large  $\Delta G_u^\ddagger$  of **3** is surprising considering this is a known Diels-Alder reactant. The relative rate constants ( $k_{c1}/k_u$ ) of all three dienophiles ( $\sim 10^1 - 10^2$ ) correlate well to experimental rates for H-bonding molecular catalysts. [35] Moreover, they confirm the limitation of H-bond stabilization as the only means of rate enhancement; one cannot obtain more than a few kcal·mol<sup>-1</sup> in stabilization, and thus  $\leq 10^3$  in  $k_{c1}/k_u$ .

### 3.7 Kinetics

Using the LIE computed  $K_M$  values and the  $k_{c1}/k_u$  values from the quantum chemical calculations we can estimate relative catalytic rate using the simplified ki-



Table 5: B3LYP/6-31++G(d,p) energies in oxyanion hole model<sup>a</sup>

| Dph      | $d_\beta - d_\alpha$ [Å] <sup>b</sup> | $\Delta G_u^\ddagger$ | $\Delta G_{c1}^\ddagger$ <sup>c</sup> | $\frac{k_{c1}}{k_u}$ | $\Delta G_{NAC}^d$ | $\Delta G_{NAC}^\ddagger$ <sup>e</sup> | $\Delta G_{c2}^\ddagger$ <sup>f</sup> | $\frac{k_{c2}}{k_u}$ |
|----------|---------------------------------------|-----------------------|---------------------------------------|----------------------|--------------------|--|---------------------------------------|----------------------|
| <b>1</b> | 0.52 (0.00)                           | 28.1                  | 24.2                                  | 660                  | 1.8                | 13.2                                   | 15.1                                  | $3.1 \cdot 10^9$     |
| <b>2</b> | 0.20 (0.00)                           | 24.8                  | 23.5                                  | 8.6                  | -                  | 11.4                                   | -                                     | -                    |
| <b>3</b> | 0.21 (0.28)                           | 40.0                  | 36.7                                  | 46                   | 2.4                | 26.3                                   | 28.7                                  | $3.0 \cdot 10^7$     |

<sup>a</sup> All energies are in kcal·mol<sup>-1</sup>. Gibbs free energies are estimated using thermodynamic corrections added from the corresponding uncatalyzed reaction. <sup>b</sup> Incipient bond distances in TS:  $d_\alpha = C_{dne}^4 - C_{dph}^\alpha$ ,  $d_\beta = C_{dne}^1 - C_{dph}^\beta$ . Values in parentheses refer to the corresponding uncatalyzed reaction. <sup>c</sup> Relative to the model [E·dph] complexes depicted in Fig. 11a-c. <sup>d</sup> NAC defined by  $d_\alpha = d_\beta = 4.0$  Å.  $\Delta G_{NAC}$  obtained in Table 4. <sup>e</sup> Relative to QC NAC model. <sup>f</sup> From eq. 14.

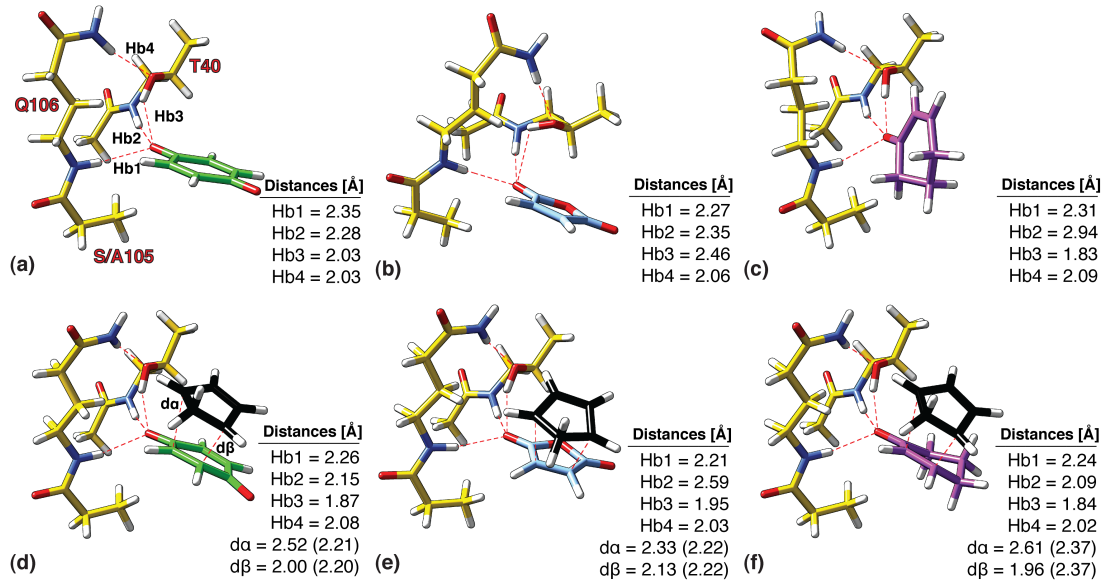


Fig. 11: B3LYP/6-31G(d) optimized geometries of oxyanion hole-bound dienophiles (a-c) and TSs (d-f) model. Distances within parentheses refer to the uncatalyzed TS.

netic model that assumes no binding of the diene ( $K_L$ ). For substrates **1** and **3** the relative rate expressions are (inserting relative rate constants from Table 5):

$$\frac{v_{c1}}{v_u} = \frac{660[E^0]}{7.3 + [dph]}$$

$$\frac{v_{c1}}{v_u} = \frac{46[E^0]}{0.61 + [dph]}$$

Inserting standard enzyme and substrate concentrations ( $10^{-5}$  M and 0.001 – 1 M, respectively) in the above expressions, it is clear that without binding of the diene,  $v_{c1}/v_u < 0.001$ , *i.e.* no catalytic effect can be expected. It is not realistic to raise the  $k_{c1}/k_u$  much above 660 by utilizing hydrogen bonding for catalysis. The corresponding  $v_{c1}/v_u$  values for the most efficient organic catalyst can be estimated to approximately 10-100. [35] The dependence on the catalyst concentration clearly explains the superiority of small organic catalyst over enzymes for catalyzing the Diels-Alder reac-

tion; organic catalysts can be, and typically are, used in concentrations  $> 1$  M.

The picture changes if we also consider productive binding of the diene. Using eq. 11, and our calculated  $K_M$ ,  $K_L$  and  $k_{c2}/k_u$  values, we obtain the following rate expressions for substrates **1** and **3**, respectively.

$$\frac{v_{c2}}{v_u} = \frac{3.1 \cdot 10^9[E^0]}{2.4 + 7.3[dne] + 0.33[dph] + [dne][dph]}$$

$$\frac{v_{c2}}{v_u} = \frac{3.0 \cdot 10^7[E^0]}{0.29 + 0.61[dne] + 0.48[dph] + [dne][dph]}$$

The  $k_{c2}/k_u$  values are of similar magnitude to those we previously have obtained for Michael-type additions and direct epoxidation in wt and S105A CALB. [22, 21] Inserting the standard enzyme and substrate concentrations we obtain rate enhancements of the order  $10^3$  for **1** and  $10^2$  for **3**, respectively. These values are competitive with conventional molecular catalysts, such as

Schreiner-type thioureas, [35] which give modest rate enhancements of about 10, despite that they are used at much higher concentrations. However, it should be emphasized that the rate enhancements are likely to be overestimated. In particular, the description of the NAC is generous and the  $\Delta G_{\text{NAC}}$  values could be considered as lower limits. Moreover, the estimates come with a standard deviation of the same order as the average value. Note that to sample instances of  $K_{\text{NAC}} < 10^{-3}$ , we need an order of magnitude longer simulation time, which is beyond the scope of this study. There is also an uncertainty in the rate constants for the uncatalyzed reaction, where we have not considered hydrogen-bonding catalysis by the solvent explicitly. Still, the results are encouraging, and clearly demonstrate that with productive binding of both dienophile and diene to the active site it is possible to obtain rate-enhancements of similar magnitude or better to those of molecular catalysts.

#### 4 Summary and Conclusion

Using standard steady-state kinetics and Eyring rate-theory, we have derived expressions for calculating the rate-enhancements for catalysis of bimolecular reactions by enzymes or molecular catalysts. The input for the calculations is a combination of data from molecular dynamics simulations and quantum chemical calculations. MD simulations on solvated substrates and solvated enzyme-substrate systems provide means for obtaining substrate binding constants and free energies for forming NAC complexes. Quantum chemical calculations on model systems allows estimates of activation energies for the enzyme and uncatalyzed reactions.

Our analysis shows that the amount of transition state stabilization that can be obtained from hydrogen-bonding is limited to a few kcal/mol, and it is difficult to optimize enzymes or substrates to enhance this effect further. It is not realistic to significantly surpass the hydrogen bonding effect that is obtained with organic catalysts, which have the great advantage of allowing for much higher catalyst concentrations due to their smaller size. Instead the key component to obtain effective catalysis in enzyme catalysts lies in substrate binding. Favorable interactions between the substrates and the active site can compensate for the entropic penalty of bringing the two substrates together. The binding constants for small organic molecules toward lipases are typically in the range of 0.1-1 M in aqueous buffer. Since this is similar to the typical substrate concentrations, even small improvements in binding can have significant effects on reaction rates. Hydrophobic interactions are the main driving force for binding, thus

decreasing substrate polarity and increasing substrate size will improve the catalytic performance. The catalytic effect increases with solvent polarity, and the best effect is expected in aqueous solution. Another key parameter, related to substrate binding, is the free energy required to form near-attack conformers,  $\Delta G_{\text{NAC}}$ . The relatively narrow cavity of CALB hinders the formation of the proper sandwich-like NAC required for the pericyclic reaction pathway. This is also the parameter that is most productive to optimize by point mutations. Extensive MD simulations were performed on a number of rationally-designed enzyme mutants, and it was found that the S105A/I189A variant gives up to 5% 'loose' NAC geometries in ACN for one given simulation, and even higher in WAT (8.5%). The corresponding value for the wild-type in ACN was below 1%.

On the basis of our computed parameters we have analyzed the catalytic performance of the S105A/I189A mutant in water. The computed rate enhancements with standard enzyme and substrate concentrations are in the order of  $10^2$ - $10^3$ . This is comparable to the catalytic effects that we have seen in earlier studies of catalysis of promiscuous reactions in CALB.[21, 22] Whereas we expect the computed rate enhancements to be overestimated, the results clearly point toward the potential for effective Diels-Alder catalysis in redesigned enzymes. We intend to continue optimization of enzyme-catalysts for this important reaction based on the current findings. In addition, we have provided important information for designing experiments to enhance the potential for detecting Diels-Alder catalysis in enzymes.

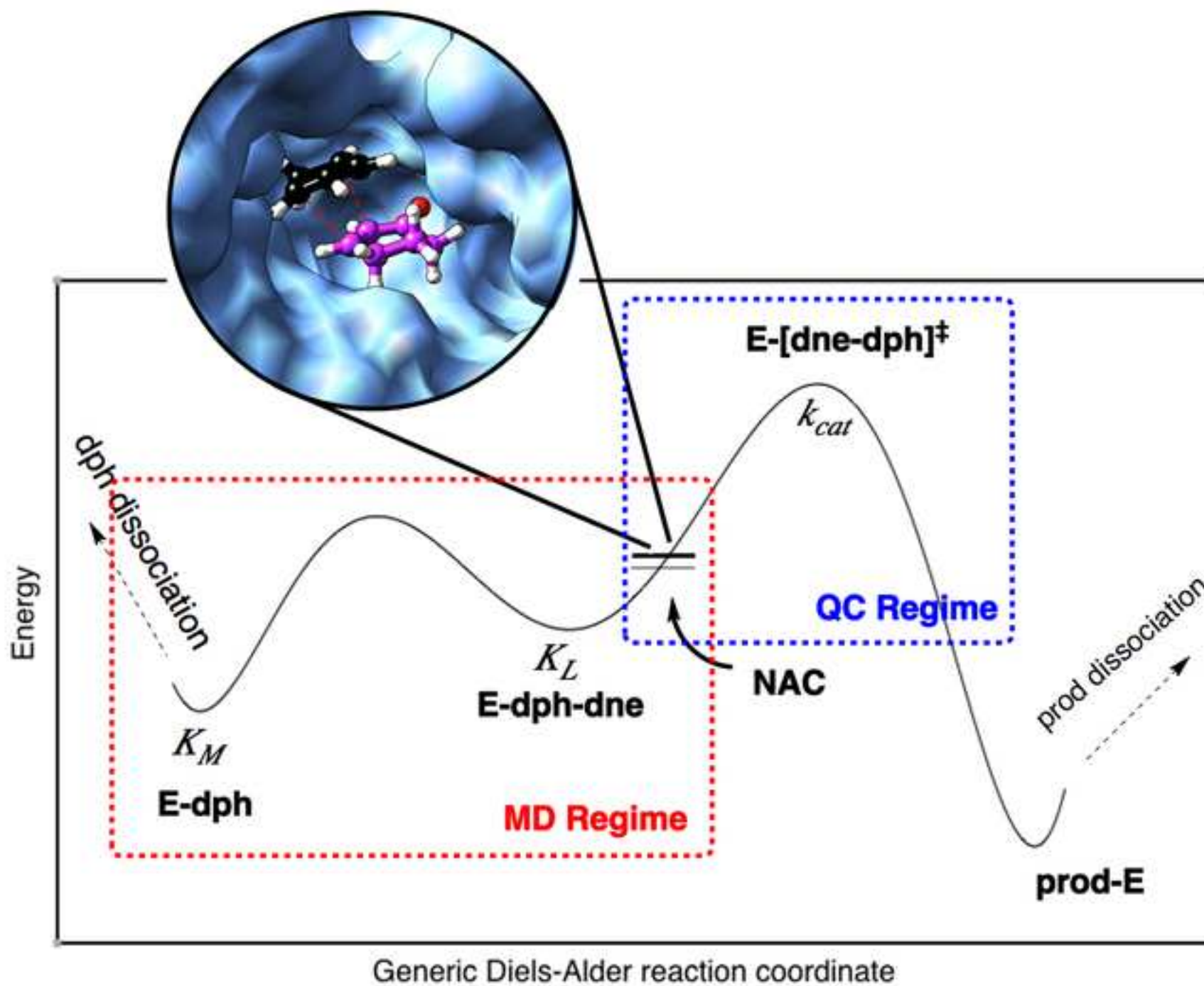
**Acknowledgements** This work has been supported by the Swedish Research Council (VR) and Cambridge Crystallographic Data Centre (CCDC). Dr. Peter Carlqvist is also warmly acknowledged for fruitful discussions.

#### References

1. Kelly, W. L. *Org. Biomol. Chem.* **2008**, *6*, 4483–4493.
2. Pohnert, G. *ChemBioChem* **2001**, *2*, 873–875.
3. Ose, T.; Watanabe, K.; Mie, T.; Honma, M.; Watanabe, H.; Yao, M.; Oikawa, H.; Tanaka, I. *Nature* **2003**, *422*, 185–189.
4. Guimaraes, C.; Udier-Blagovic, M.; Jorgensen, W. *J. Am. Chem. Soc.* **2005**, *127*, 3577–3588.
5. Kim, S. P.; Leach, A. G.; Houk, K. N. *J. of Org. Chem.* **2002**, *67*, 4250–4260.
6. Kirby, A. J. *Angew. Chem. Int. Ed.* **1996**, *35*, 706–724.
7. Pihko, P. M. *Angew. Chem. Int. Ed.* **2004**, *43*, 2062–2064.

8. Schreiner, P. R. *Chem. Soc. Rev.* **2003**, *32*, 289–296.
9. Zhang, Z.; Schreiner, P. R. *Chem. Soc. Rev.* **2009**, *38*, 1187–1198.
10. Zhang, X.; DeChancie, J.; Gunaydin, H.; Chowdry, A. B.; Clemente, F. R.; Smith, A. J. T.; Handel, T. M.; Houk, K. N. *J. Org. Chem.* **2007**, *73*, 889–899.
11. Kazlauskas, R. J.; Weissfloch, A. N. E.; Rappaport, A. T.; Cuccia, L. A. *J. Org. Chem.* **1991**, *56*, 2656–2665.
12. Stecher, H.; Faber, K. *Synthesis* **1997**, 1–16.
13. Kazlauskas, R.; Weber, H. *Curr. Opin. Chem. Biol.* **1998**, *2*, 121–126.
14. Pamies, O.; Bäckvall, J.-E. *Chem. Rev.* **2002**, *103*, 3247.
15. Zaks, A.; Klibanov, A. M. *Proc. Natl. Acad. Sci.* **1985**, *82*, 3192–3196.
16. Klibanov, A. M. *Nature* **2001**, *409*, 241–246.
17. O'Brien, P. J.; Herschlag, D. *Chem. Biol.* **1999**, *6*, R91–R105.
18. Hult, K.; Berglund, P. *Trends Biotech.* **2007**, *25*, 231–238.
19. Branneby, C.; Carlqvist, P.; Magnusson, A.; Hult, K.; Brinck, T.; Berglund, P. *J. Am. Chem. Soc.* **2003**, *125*, 874–875.
20. Branneby, C.; Carlqvist, P.; Hult, K.; Brinck, T.; Berglund, P. *J. Mol. Catal. B* **2004**, *31*, 123–128.
21. Svedendahl, M.; Carlqvist, P.; Branneby, C.; Allnr, O.; Frise, A.; Hult, K.; Berglund, P.; Brinck, T. *ChemBioChem* **2008**, *3*, 2443–2451.
22. Carlqvist, P.; Svedendahl, M.; Branneby, C.; Hult, K.; Brinck, T.; Berglund, P. *ChemBioChem* **2004**, *6*, 331–336.
23. Svedendahl, M.; Hult, K.; Berglund, P. *J. Am. Chem. Soc.* **2005**, *127*, 17988–17989.
24. Svedendahl, M.; Jovanovic, B.; Berglund, P. *Chem-CatChem* **2009**, *1*, 252–258.
25. Strohmeier, G. A.; Sovic, T.; Steinkellner, G.; Hartner, F. S.; Andryushkova, A.; Purkarthofer, T.; Glieder, A.; Gruber, K.; Griengl, H. *Tetrahedron* **2009**, *65*, 5663–5668.
26. Torre, O.; Alfonso, I.; Gotor, V. *Chem. Commun.* **2004**, 1724–1725.
27. Kazlauskas, R. J. *Curr. Opin. Chem. Biol.* **2005**, *9*, 195–201.
28. Peracchi, A. *Trends Biochem. Sci.* **2001**, *26*, 497–503.
29. Ollis, D. L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S. M.; Harel, M.; Remington, S. J.; Silman, I.; Schrag, J.; Sussman, J. L.; Verschueren, K. H.; Goldman, A. *Protein Eng.* **1992**, *5*, 197–211.
30. Svedendahl, M.; Berglund, P. *Personal communication*, 2009.
31. Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453*, 190–195.
32. Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, I., C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387–1381.
33. Simón, L.; Goodman, J. M. *J. Org. Chem.* **2010**, *75*, 1831–1840.
34. Linder, M.; Brinck, T. *Org. Biomol. Chem.* **2009**, *7*, 1304–1311.
35. Wittkopp, A.; Schreiner, P. R. *Chem. Eur. J.* **2003**, *9*, 407–414.
36. N. Thadani, A.; R. Stankovic, A.; H. Rawal, V. *Proc. Natl. Acad. Sci.* **2004**, *101*, 5846–5850.
37. Ishihara, K.; Nakano, K.; Akakura, M. *Org. Lett.* **2008**, *10*, 2893–2896.
38. Kano, T.; Tanaka, Y.; Osawa, K.; Yurino, T.; Maruoka, K. *Chem. Comm.* **2009**, *15*, 1956–1958.
39. Uppenberg, J.; Oehrner, N.; Norin, M.; Hult, K.; Kleywegt, G. J.; Patkar, S.; Waagen, V.; Anthonson, T.; Jones, T. A. *Biochemistry* **1995**, *34*, 16838–16851.
40. Michaelis, L.; Menten, M. L. *Biochem. Z.* **1913**, *49*, 333.
41. Cleland, W. W. *Enzymes* **1970**, *2*, 1–65.
42. Dalziel, K. *Enzymes* **1970**, *11*, 2–60.
43. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889–897.
44. Åqvist, J.; Medina, C.; Samuelsson, J. *Protein Eng.* **1994**, *7*, 385–391.
45. Marelius, J.; Hansson, T.; Åqvist, J. *Int. J. Quantum Chem.* **1998**, *69*, 77–88.
46. Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 2595–2605.
47. Bruice, T. C.; Lightstone, F. C. *Acc. Chem. Res.* **1999**, *32*, 127–136.
48. Toro-Labbe, A.; Gutierrez-Oliva, S.; Murray, J.; Politzer, P. *Mol. Phys.* **2007**, *105*, 2619–2625.
49. Labet, V.; Morell, C.; Grand, A.; Toro-Labbe, A. *J. Phys. Chem. A* **2008**, *112*, 11487–11494.
50. Toro-Labbe, A.; Gutierrez-Oliva, S.; Murray, J.; Politzer, P. *J. Mol. Model.* **2009**, *15*, 707–710.
51. Uppenberg, J.; Hansen, M. T.; Patkar, S.; Jones, T. A. *Structure* **1994**, *2*, 453–454.

52. *GOLD 4.1*, 2008, [http://www.ccdc.cam.ac.uk/products/life\\_sciences/gold/](http://www.ccdc.cam.ac.uk/products/life_sciences/gold/).
53. Guex, N.; Peitsch, M. *Electrophoresis* **1997**, *18*, 2714–2723, <http://www.expasy.org/spdbv/>.
54. Jones, G.; Willett, P.; Glen, R. *J. Mol. Biol.* **1995**, *245*, 43–53.
55. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727–748.
56. Nissink, J. W. M.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. *Proteins* **2002**, *49*, 457–471.
57. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. *Proteins* **2003**, *52*, 609–623.
58. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.
59. Baxter, C. A.; Murray, C. W.; Clark, D. E.; Westhead, D. R.; Eldridge, M. D. *Proteins* **1998**, *33*, 367–382.
60. Case, D. et al. *AMBER 10*, 2008.
61. Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
62. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
63. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132–146.
64. Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623–1641.
65. Guardia, E.; Pinzon, R.; Casulleras, J.; Orozoco, M.; Luque, F. J. *Mol. Simulat.* **2001**, *26*, 287–306.
66. Cieplak, P.; Caldwell, J. W.; Kollman, P. A. *J. Comput. Chem.* **2001**, *22*, 1048–1057.
67. Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926–935.
68. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
69. Almlöf, M.; Brandsdal, B. O.; Åqvist, J. *J. Comput. Chem.* **2004**, *25*, 1242–1254.
70. Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
71. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
72. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
73. Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724–728.
74. Frisch, M. J. et al. *Gaussian 03, Revision C.02*, 2003.
75. Kong, S.; Evanseck, J. *J. Am. Chem. Soc.* **2000**, *122*, 10418–10427.
76. Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
77. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669–681.
78. Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210–3221.
79. UCSF Chimera, 2004, <http://www.cgl.ucsf.edu/chimera>.
80. Pettersen, E.; Goddard, T.; Huang, C.; Couch, G.; Greenblatt, D.; Meng, E.; Ferrin, T. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
81. Turner, P. J. *Grace v.5.1.20*, 2007, <http://plasmagate.weizmann.ac.il/Grace>.
82. Wesemann, M. *Plot, v. 0.997*, 2007, <http://plot.micw.eu/>.
83. Trodler, P.; Pleiss, J. *BMC Struct. Biol.* **2008**, *8*, 9.
84. Trodler, P.; Schmid, R.; Pleiss, J. *BMC Struct. Biol.* **2009**, *9*, 38.
85. Nini, L.; Sarda, L.; Comeau, L.-C.; Boitard, E.; Dubes, J.-P.; Chahinian, H. *Biochim. Biophys. Acta* **2001**, *1534*, 34–44.
86. Chahinian, H.; Nini, L.; Boitard, E.; Dubes, J.-P.; Comeau, L.-C.; Sarda, L. *Lipids* **2002**, *37*, 653–662.
87. Garcia, J.; Mayoral, J.; Salvatella, L. *J. Am. Chem. Soc.* **1996**, *118*, 11680–11681.



|   |
|---|
| Noname manuscript No.<br>(will be inserted by the editor) |
|---|

---

## Computational Studies of a Lipase-mediated Diels-Alder reaction – Electronic Supplementary Material

Mats Linder · Anders Hermansson ·  
Tore Brinck · John Liebschuetz

Received: date / Accepted: date

### 1 Analysis of relaxed enzyme variants

Here follows a thorough analysis of solvent and mutant effects on the structure of the CALB variants. Backbone RMSDs of the average relaxed structures are given in Table 1.

The flexibility of each residue has been taken as the B-factor of the backbone atoms, displayed in Fig. 1. We note that the 'nucleophilic elbow', [1] to which S/A105 belongs, is rigid in all variants and solvents. The loop containing T40 is also rigid, although residues 30-50 have B-factors of up to 25 in some variants. This region appears to be most flexible in CHL. Furthermore, the  $\alpha 5$  and  $\alpha 10$  helices (residues 142-146 and 268-287, respectively) are among the most flexible in all simulations, together with the terminal regions, in agreement with recent studies. [2, 3] The relative flexibility of different regions in different solvents varies, as does the relative flexibility between different variants. On the whole, each variant is most flexible in water, which is generally accepted to be the case for most enzymes. [2, 4] But the difference is smaller than one might expect, perhaps due to the high documented flexibility caused by the AMBER FF99 force field, which seems to offset the need for

---

Mats Linder  
Physical Chemistry, Royal Institute of Technology  
Tel.: +46-(0)8-7908429  
E-mail: [linder@physchem.kth.se](mailto:linder@physchem.kth.se)

Tore Brinck  
Physical Chemistry, Royal Institute of Technology  
Tel.: +46-(0)8-790xxxx  
E-mail: [tore@physchem.kth.se](mailto:tore@physchem.kth.se)

John Liebeschuetz  
Cambridge Crystallography Data Centre  
Tel.: +44-(0)1223-762532  
Fax.: +44-(0)1223-33603  
E-mail: [john@ccdc.cam.ac.uk](mailto:john@ccdc.cam.ac.uk)

Table 1: Backbone RMS deviations ( $\text{\AA}$ ) of relaxed structures<sup>a</sup>

| Variant     | ACN  | CHL  | WAT  |
|-------------|------|------|------|
| WT          | 1.02 | 1.46 | 1.04 |
| S105A       | 1.08 | 1.38 | 0.84 |
| I189A       | 1.01 | 1.33 | 0.76 |
| I285A       | 1.04 | 1.46 | 0.96 |
| S105A/I189A | 1.07 | 1.64 | 1.00 |
| S105A/I285A | 1.32 | 1.44 | 0.92 |

<sup>a</sup> Average structure from last 0.2 out of 4 ns simulations, relative to the 1LBT crystal structure.

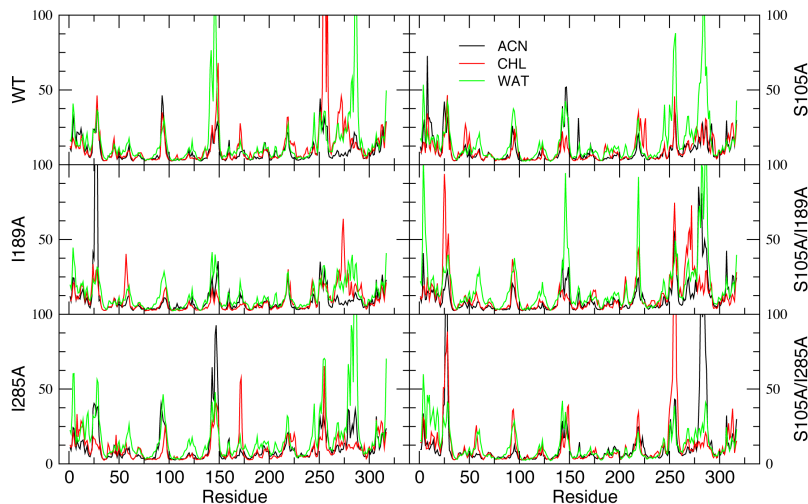


Fig. 1: B-factor per residue, calculated from the 4 ns simulation of the 6 CALB variants, in 3 solvents.

surface-bound water molecules. Interestingly, the  $\alpha$ 10 helix is very flexible in chloroform in the wt, I189A, S105A/I189A and S105A/I189A variants.

Fig. 2 shows the active sites of the superimposed, relaxed structures in each solvent. In agreement with Fig. 1, it is seen that the nucleophilic elbow backbone is hardly distorted. However, we observe significant movement of several side chains, most notably T40, S105 and W104. In species without the S105A mutation (Fig. 2a, c and d), the S105 OH group moves towards the oxyanion hole in all solvents. The T40 OH group moves away from its crystal position in all solvents and variants except I285A (Fig. 2d). W104 tends to move closer to the catalytic triad, especially in WAT, which effectively makes the active site smaller. A similar, albeit less prominent movement is observed in CHL

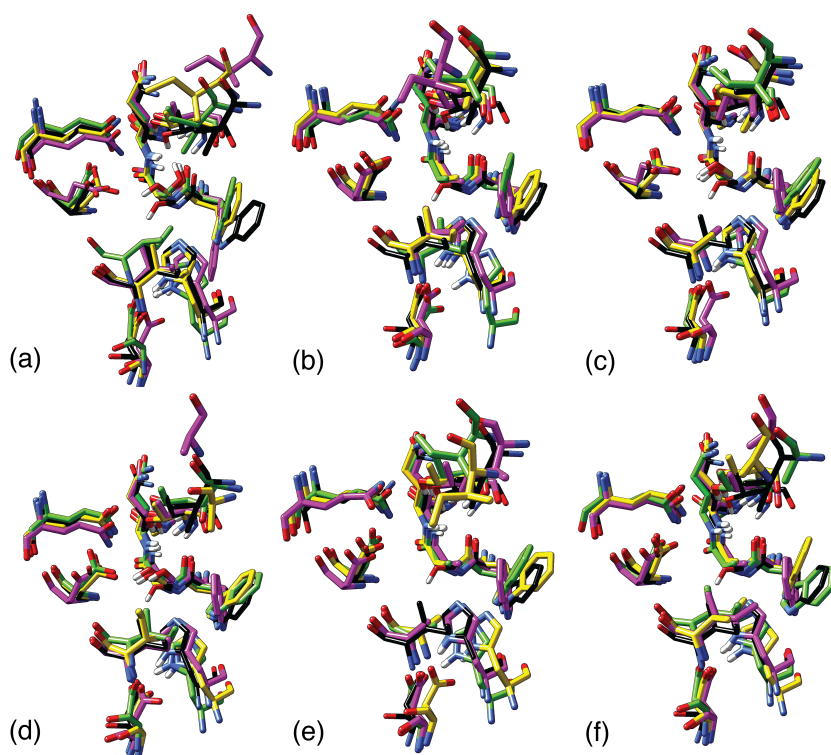


Fig. 2: Superposition of selected residues (40, 104, 105, 106, 134, 157, 187, 189, 224, 285) of the average structure of from simulations in ACN (yellow), CHL (green) and WAT (magenta), relative to the 1LBT crystal structure (black). The superposition was made on the backbone of all 317 residues. For clarity, only selected hydrogens are displayed, belonging to the oxyanion hole or catalytic triad. (a) WT, (b) S105A, (c) I189A, (d) I285A, (e) S105A/I189A, (f) S105A/I285A. Stereoviews of these superpositions are presented in Fig. 6.

(exceptions being I285A and S105A/I285A, Fig. 2d and f). ACN is the solvent in which W104 is the least distorted (except in wt and S105A/I285A, Fig. 2a and f). These observations can be rationalized by the fact that the structure 1LBT [5] contains a co-crystallized covalent inhibitor, inducing conformational changes, and that the relaxed structures correspond to an empty enzyme. But since no nucleophilic attack by S105 is expected to occur in the Diels-Alder mechanism, its role and therefore position is uncertain, and could well be one that hinders access to the oxyanion hole. This argument further motivates a special interest in S105A mutants. Moreover, some interactions are detected between S105 and D134, which also causes the amide of Q157 to flip over.

Fig. 2 also illustrates what happens when the two isoleucine residues are mutated. The I189A mutation (Fig. 2c and e) creates additional space below and outside the oxyanion hole, as expected, and no significant geometric distortion of the variants is detected. The I285A mutation (Fig. 2d and f) creates



more space above the oxyanion hole; this is expected to be the face of attack from the diene, but the net effect is more difficult to predict, since I285 resides slightly further away from the oxyanion hole and tends to be more flexible, as seen in both Fig. 1 and Fig. 2.

## 2 Docking composite and consensus scoring

The ChemScore [6, 7] scoring function in GOLD generates a total score based on an estimate of Gibbs Free Energy of binding,  $\Delta G_{binding}$ , which is modified by a set of penalty terms (dependent of what type of constraints are applied). In this study, the score is given by eq. 1.

$$CS = -\Delta G_{binding} - P_{clash} - cP_{internal}, \quad (1)$$

where  $c$  is a coefficient.  $\Delta G_{binding}$ , in turn, is given by eq. 2.

$$\Delta G_{binding} = v_0 + v_1 P_{Hb} + v_2 P_{metal} + v_3 P_{lip0} + v_2 P_{rot}, \quad (2)$$

where all  $P$ 's are block functions with empirical parameters that are used to describe different types of interactions,[8] and  $v_i$  are regression coefficients. Similarly,  $P_{clash}$  and  $P_{internal}$  are block functions describing energy penalties when the ligand clashes with protein atoms or is distorted.

Since molecular docking software is designed mainly for drug discovery and similar applications, there is a bias in the scoring function to promote large ligands with many contacts. But since we seek to design a catalytic process involving two substrates that are turned over by the active site, our demands on a quite different to drug discovery. Most importantly, there must be solvent-accessible space left in the active site after docking the first ligand (in our case the dienophile). A ligand filling out the entire active site will acquire a high score in GOLD, but is unfit as a Diels-Alder reactant. Furthermore, we wish to promote H-bonding to the oxyanion hole and, since we induce a binding constraint between oxyanion hole donors and dienophile, penalize clashes that may arise because of this. Last, the smaller the ligand, the larger the number of ways in which it may dock into the active site. In order to find ligands that bind unambiguously in one conformation but is still small enough to enable Diels-Alder, we seek to include a term describing the conformational variety.

A simple way to promote these parameters is to create a 'composite' scoring function, explicitly including terms that describe H-bonding, steric clashes and conformational variety. We used eq. 3:

$$S_{comp}(n) = \left( \frac{\langle CS \rangle}{i_{atoms}} \right)^{1/2} + \langle P_{Hb} \rangle - \langle P_{clash} \rangle - \langle RMSD \rangle, \quad (3)$$

which is based on average scores of the top  $n$  docking poses from a set of solutions. The total  $CS$  score is normalized by the number of heavy atoms in the ligand.[9]  $P_{Hb}$  and  $P_{clash}$  are taken directly from the ChemScore output

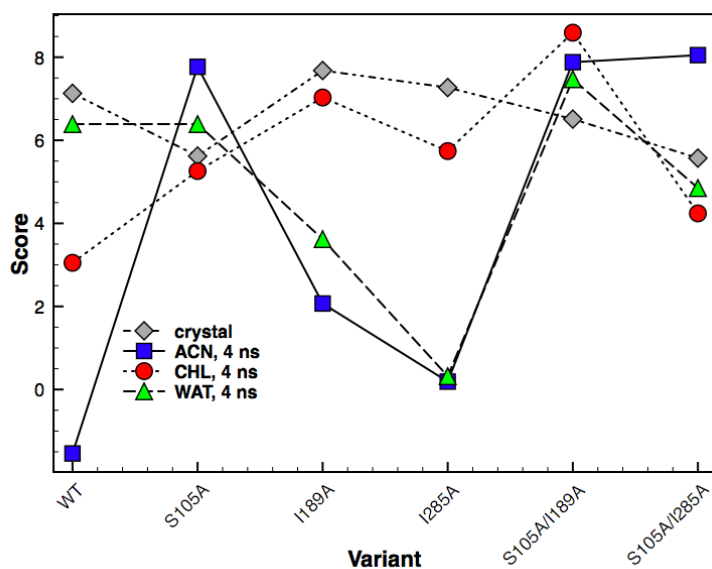


Fig. 3: Sum of composite scores of dienophiles **1-3** in the six enzyme variants, using the crystal structure and solvent-relaxed average structures.

and are thus essentially counted twice. *RMSD* is the root mean square deviation of the top  $n$  solutions to each other, meaning that a low value represents an consistent binding mode. The above expression obviously has no physical meaning; it is simply a means to sum up terms important to describe the system. All terms are of the same order magnitude, with the largest weight being placed on the normalized fitness score.

It should be noted that much analysis must still be done by visual interpretation, since the terms in eq. 3 do not include any measure of *e.g.* solvent accessibility to  $\alpha,\beta$ -double bond of the dienophile. GOLD includes utilities to do such descriptions, but they have not been used here due to the large diversity of the initial set of ligands used. Composite scoring still has a great value in internally ranking selected poses.

Fig. 3 compares the sum of composite scores for the three dienophiles in this study when docked in the enzyme variants. It is shown how the score depends on whether the enzyme was taken as the crystal structure or pre-relaxed in MD. In the crystal structure, there is little variation due to mutations, while the effects can be large when the enzyme is allowed to relax. Analysis of the separate terms of eq. 3 reveals that the largest portion of the difference is due to changes in the clash term. For variants including S105, it is generally very large. Note that in all solvents, the variant with the highest composite score is S105A/I189A.

A 'consensus' scoring procedure has also been used to map several points on the reaction pathway. Docking has been performed of the diene with a

dienophile already bound to the oxyanion hole (the pose being selected from the set of dienophile solutions), as well TS models with fixed incipient bond distances. These were generated based on DFT-optimized structures, and were then treated as a single molecule by GOLD. The consensus score was calculated according to equation 4.

$$S_{consensus} = S_{comp,dph} + S_{comp,dne} + S_{comp,TS}. \quad (4)$$

Calculated consensus score for each dienophile in ACN-relaxed variants and crystal structures are displayed in figure 10 in the main article.

### 3 Investigations on a triple mutant T40A/S105A/I189A

The triple mutant (C) statistics are summarized in Fig. 4a. The RMSD of the relaxed uncomplexed variant to the crystal structure is 1.2. Structurally, the most remarkable feature is that the oxyanion loop (in ACN), now containing A40, is somewhat more distorted than in previous variants, which has the effect that the the Q106-N-A40-N distance increases by 0.8 Å relative to the crystal structure. A similar distortion can be seen in the S105A/I189A and S105A/I285A variants, but is accompanied by a concordant movement of the nucleophilic elbow, resulting in a smaller increase or even reduction of the N-N distance, as illustrated in Fig. 4b. We also note that the RMSD is not completely leveled off in the 4 ns simulation, which is indicative of larger geometric distortions caused by three mutations in close vicinity.

The backbone RMSDs based on the last 1000 frames of both triple mutant complexes with **3** and **3 + 4** were 1.1 Å with respect to the crystal structure, and 1.0 Å with respect to the relaxed uncomplexed structure. Quite surprisingly, the H-bonding ability of the modified oxyanion hole was remarkable, both with and without the presence of a diene, as shown in 4a. Based on the last 5000 recorded frames, 90-97 % distances were within the 2.75 Å threshold in both the dienophile and diene simulations (with average distances between 2.2 and 2.4 Å), but the diene did not form any NAC and fell out of the oxyanion hole after  $\approx$  1 ns. To investigate whether or not this was a stochastic event, we performed two additional 2 ns simulations with **3** and **4**. Backbone RMSDs and H-bond statistics were similar to the first run, but neither produced any NAC. In one case, the diene remained in the active site, but mostly in a T-shaped interaction with **3**. We attribute the finding that the additional T40A mutation did not give better NAC formation to that the dienophile is allowed to become more buried (hence the enhanced H-bonding), and this introduces new hindering residues, effectively blocking attack on the double bond. Most notable is Q157, which has taken over the role of obscuring the dienophile C $^{\alpha}$ . See Fig. 4c for an illustration.

To summarize, the triple mutant provides enhanced binding of the dienophile due to the modification of the oxyanion hole, but structural modifications in the active site leads to poor coordination of the diene. These results are still important, as they suggest that spacious oxyanion holes with two donors, such

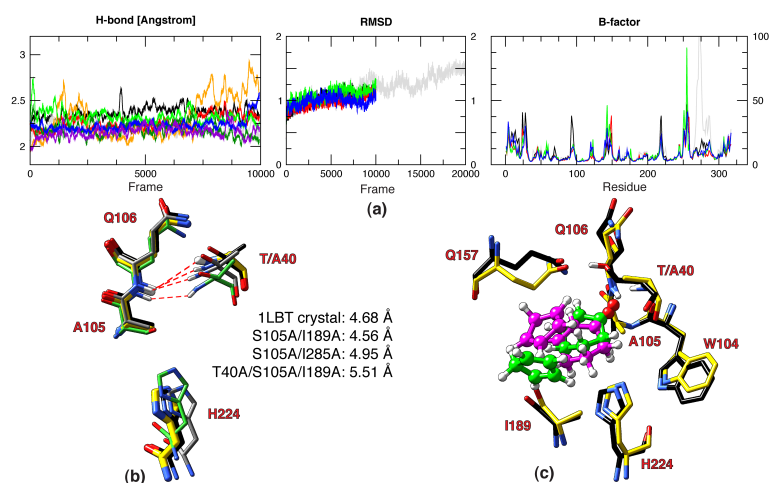


Fig. 4: (a) Simulation statistics of the T40A/S105A/I189A variant. The color legend is as follows: For RMSD and B-factor, gray=naked enzyme; black=dph simulation; red=dph+dne, 1st run; green=dph+dne, 2nd run; blue=dph+dne, 3rd run. For H-bond trajectories (100-frame running averages), Q106 and A40 are represented by black & gray (dph run), red & orange (dph+dne, 1st run), green & dark green (dph+dne, 2nd run), and blue & violet (dph+dne, 3rd run). (b) A comparison of the oxyanion hole geometry between T40A/S105A/I189A (yellow, thick sticks), the 1LBT crystal (green), S105A/I189A (gray) and S105A/I285A (black). All structures refer to the last 200 ps averages of 4 ns relaxation in ACN. His224 is shown for the sake of orientation. The distances displayed are calculated between Q106-N and T/A40-N. (c) Representative snapshots of the [enz·3·4] complexes in T40A/S105A/I189A (yellow, green substrates) and S105A/I189A (black, magenta substrates). Note how the smaller methyl side chain on A40 allows the dienophile to become more buried, and how Q157 moves closer in the triple mutant, both resulting in less accessibility to the  $\alpha,\beta$ -double bond. In contrast, the S105A/I189A variant displays a much more ideal NAC-like conformation of the substrates, although the dienophile C $^{\alpha}$  is still obscured by the T40 side chain.

as those present in *Pseudomonas cepacia* and *Candida rugosa* lipases, might be preferable over the one in CALB for binding the dienophile. Furthermore, these enzymes have comparatively large cavities, [10, 11, 12] which we have shown is an important feature to obtain the necessary Diels-Alder TS conformation.

## References

1. Holmqvist, M. *Curr. Prot. Peptide Sci.* **2000**, *1*, 209–235.
2. Trodler, P.; Pleiss, J. *BMC Struct. Biol.* **2008**, *8*, 9.
3. Skjot, M.; De Maria, L.; Chatterjee, R.; Svendsen, A.; Patkar, S. A.; Ostergaard, P. R.; Brask, J. *ChemBioChem* **2009**, *10*, 520–527.
4. Klivanov, A. M. *Trends Biotechnol.* **1997**, *15*, 97–101.

- 
5. Uppenberg, J.; Oehrner, N.; Norin, M.; Hult, K.; Kleywegt, G. J.; Patkar, S.; Waagen, V.; Anthonsen, T.; Jones, T. A. *Biochemistry* **1995**, *34*, 16838–16851.
  6. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.
  7. Baxter, C. A.; Murray, C. W.; Clark, D. E.; Westhead, D. R.; Eldridge, M. D. *Proteins* **1998**, *33*, 367–382.
  8. 2010, For more details, see the GOLD documentation, <http://www.ccdc.cam.ac.uk/support/documentation/gold/4.1/gold/>.
  9. Pan., Y.; Huang, N.; Cho, S.; MacKerrell, A. J. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 267–272.
  10. Schrag, J. D.; Li, Y.; Cygler, M.; Lang, D.; Burgdorf, T.; Hecht, H.-J.; Schmid, R.; Schomburg, D.; Rydel, T. J.; Oliver, J. D.; Strickland, L. C.; Dunaway, C. M.; Larson, S. B.; Day, J.; McPherson, A. *Structure* **1997**, *5*, 187–202.
  11. Lang, D.; Mannesse, M.; De Haas, G.; Verheij, B., H.M. Dijkstra *Eur. J. Biochem.* **1998**, *254*, 333–340.
  12. Grochulski, P.; Bouthillier, F.; Kazlauskas, R. J.; Serreqi, A. N.; Schrag, J. D.; Ziomek, E.; Cygler, M. *Biochemistry* **1994**, *33*, 3494–3500.

#### 4 Supplementary figures

Supplementary figures as referred to in the main article.

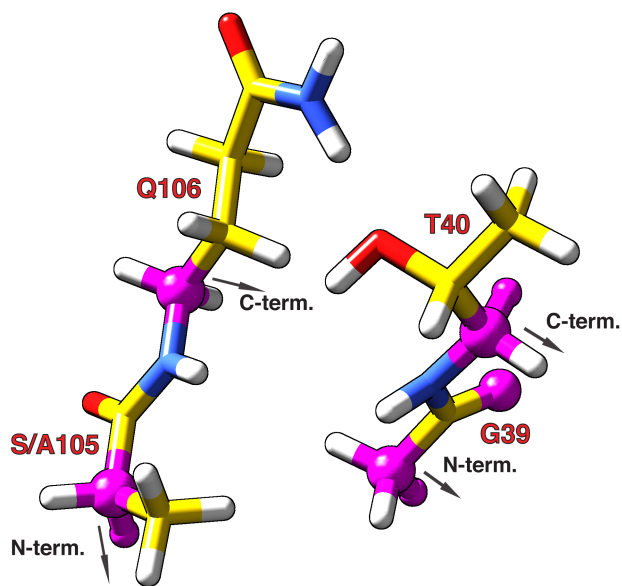


Fig. 5: The oxyanion hole model. Frozen atoms are colored in magenta.

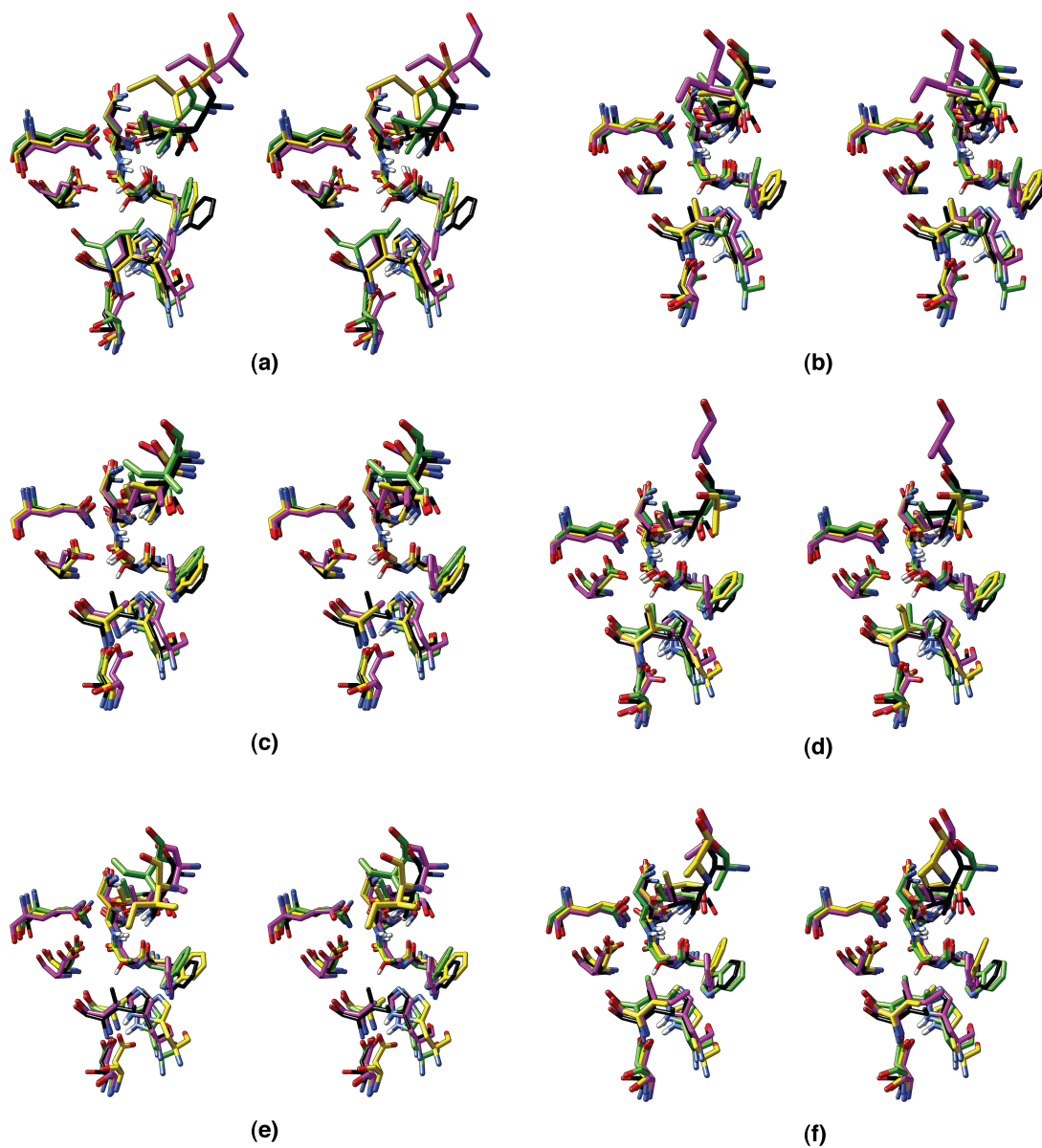
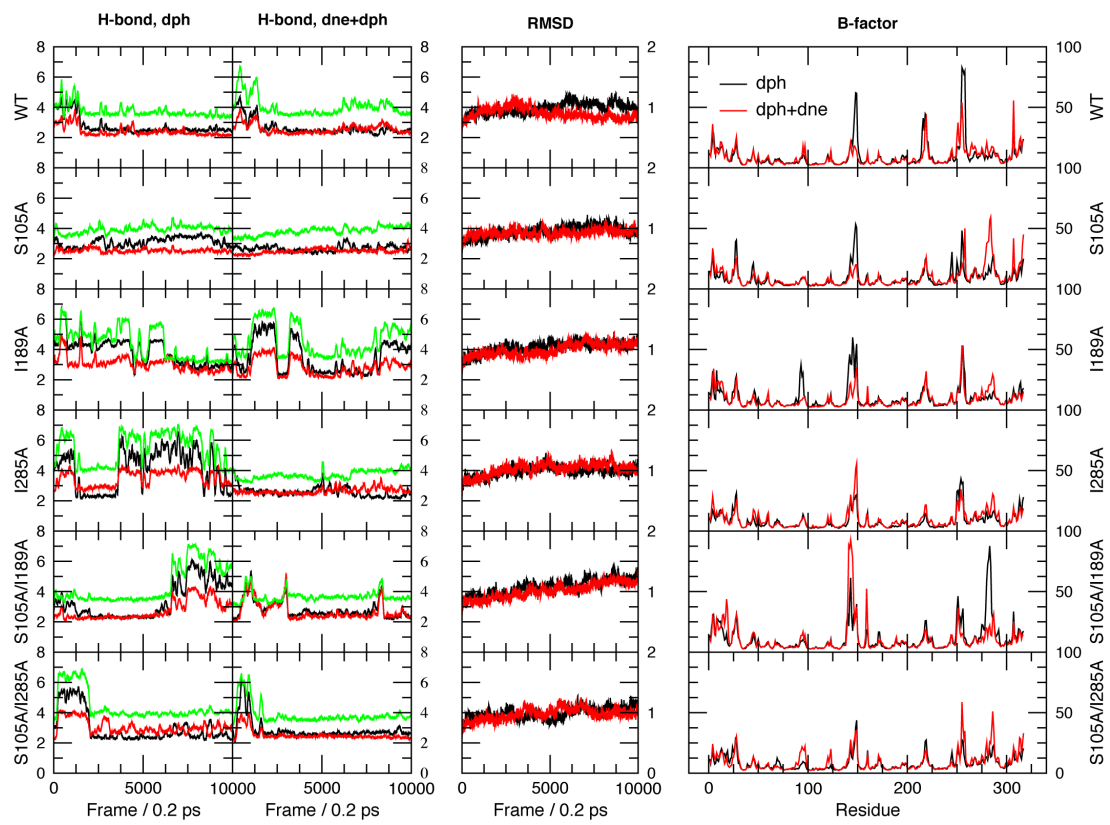


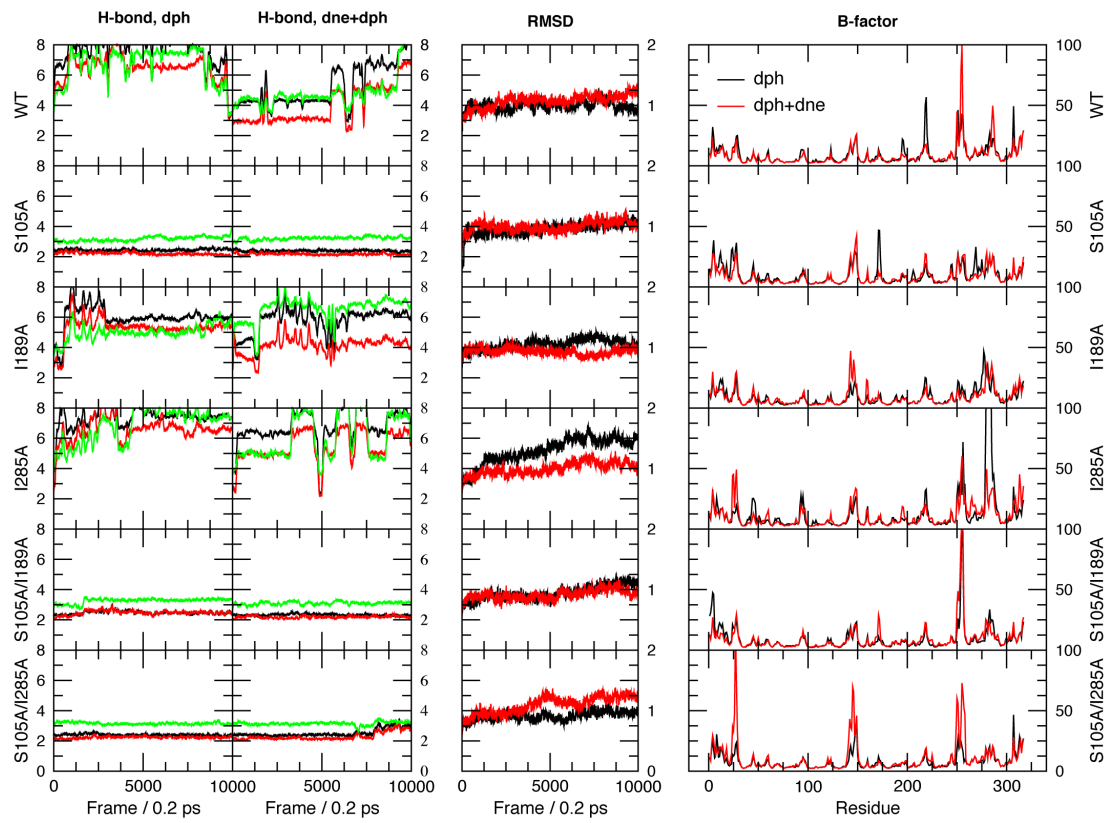
Fig. 6: Stereoviews of selected active site residues of the 6 enzyme variants in the crystal structure 1LBT (black), superimposed with average relaxed structures from ACN (yellow), CHL (green) and WAT (magenta) simulations. For clarity, only selected hydrogens are displayed. (a) WT, (b) S105A, (c) I189A, (d) I285A, (e) S105A/I189A, (f) S105A/I285A.



(a)

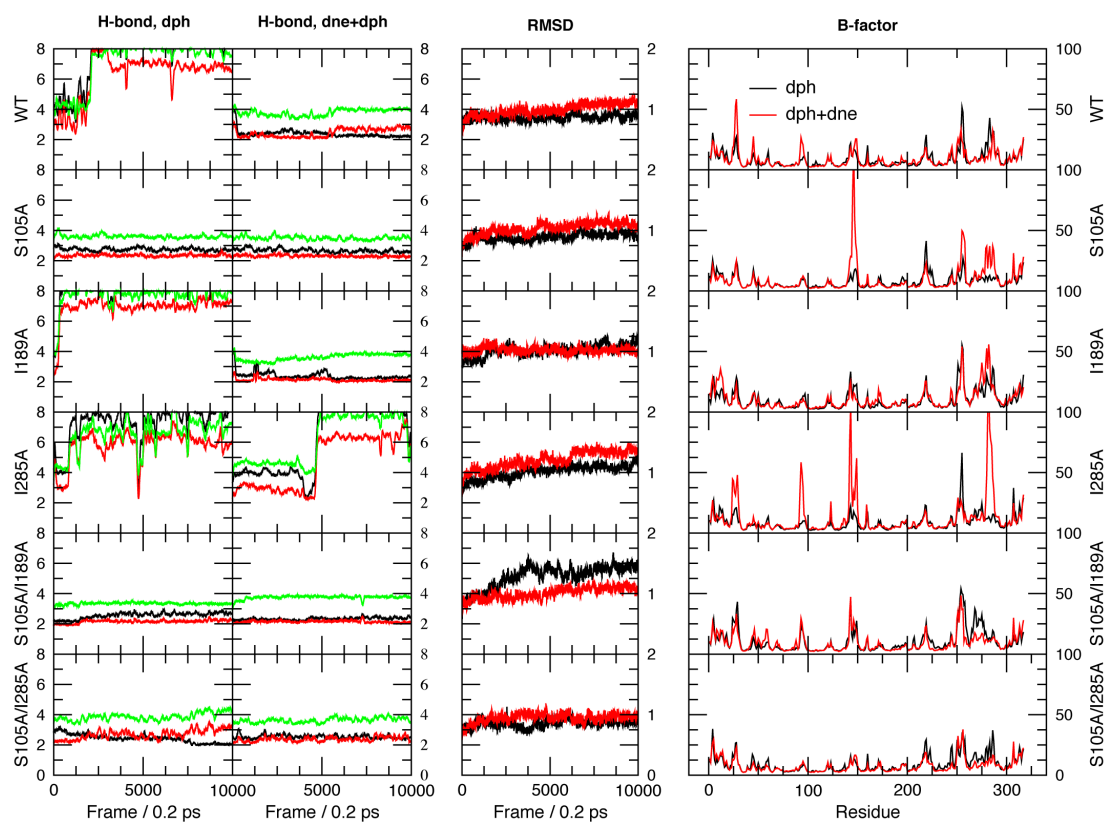
Fig. 7





(b)

Fig. 7: (cont.)



(c)

Fig. 7: (cont.) Statistics from the 2 ns simulations in the 6 variants. (a) benzoquinone **1**, (b) maleic anhydride **2**, (c) cyclohexenone **3**. Two leftmost columns: Distance between Q106-NH (black), T40-NH (red) and T40-O<sup>γ</sup>H, and acceptor oxygen without (left) and with (right) diene present. Distances are in Angstroms and each trajectory is presented as a 100-frame running average, for clarity. Two rightmost columns: RMSD trajectory (left) and B-factor per residue, for dienophile (black) and dienophile+diene (red) simulation, respectively.

This is pdfTeX, Version 3.1415926-1.40.10 (Web2C 2009) (format=latex 2010.4.6) 31 MAY 2010  
12:28

entering extended mode

%&-line parsing enabled.

\*\*brinck\_2010-dielsalder\_revised.tex

(./brinck\_2010-dielsalder\_revised.tex

LaTeX2e <2009/09/24>

Babel <v3.8l> and hyphenation patterns for english, usenglishmax, dumylang, nohyphenation, german-x-2009-06-19, ngerman-x-2009-06-19, ancientgreek, ibycus, arabic, basque, bulgarian, catalan, pinyin, coptic, croatian, czech, danish, dutch, esperanto, estonian, farsi, finnish, french, galician, german, ngerman, monogreek, greek, hungarian, icelandic, assamese, bengali, guajrati, hindi, assamese, malayalam, marathi, oriya, panjabi, tamil, telugu, indonesian, interlingua, irish, italian, kurmanji, latin, latvian, lithuanian, mongolian, mongolian2a, bokmal, nynorsk, polish, portuguese, romanian, russian, sanskrit, serbian, slovak, slovenian, spanish, swedish, turkish, turkmen, ukenglish, ukrainian, uppersorbian, welsh, loaded.

(c:/TexLive2009/texmf/tex/latex/aries/svjour3.cls

Document Class: svjour3 2006/03/15 v3.1

LaTeX document class for Springer journals

(c:/TexLive2009/texmf-dist/tex/latex/base/fleqn.clo

File: fleqn.clo 1998/08/17 v1.1c Standard LaTeX option (flush left equations)

\mathindent=\dimen102

)

Class Springer-SVjour3 Info: extra/valid Springer sub-package (-> \*.clo)

(Springer-SVjour3) not found in option list of \documentclass

(Springer-SVjour3) - autoactivating "global" style.

(c:/TexLive2009/texmf/tex/latex/aries/svglov3.clo

File: svglov3.clo 2006/02/03 v3.1 style option for standardised journals

SVjour Class option: svglov3.clo for standardised journals

)

LaTeX Font Info: Redefining math symbol \Gamma on input line 146.

LaTeX Font Info: Redefining math symbol \Delta on input line 147.

LaTeX Font Info: Redefining math symbol \Theta on input line 148.

LaTeX Font Info: Redefining math symbol \Lambda on input line 149.

LaTeX Font Info: Redefining math symbol \Xi on input line 150.

LaTeX Font Info: Redefining math symbol \Pi on input line 151.

LaTeX Font Info: Redefining math symbol \Sigma on input line 152.

LaTeX Font Info: Redefining math symbol \Upsilon on input line 153.

LaTeX Font Info: Redefining math symbol \Phi on input line 154.

LaTeX Font Info: Redefining math symbol \Psi on input line 155.

LaTeX Font Info: Redefining math symbol \Omega on input line 156.

\logodepth=\dimen103

\headerboxheight=\dimen104

\betweenumberspace=\dimen105

\aftertext=\dimen106

\headlineindent=\dimen107

\c@inst=\count79

\c@auth=\count80

\instindent=\dimen108

\authrun=\box26

\authorrunning=\toks14

```

\titrun=\box27
\titlrunning=\toks15
\combirun=\box28
\c@lastpage=\count81
\rubricwidth=\dimen109
\c@section=\count82
\c@subsection=\count83
\c@subsubsection=\count84
\c@paragraph=\count85
\c@subparagraph=\count86
\spthmsep=\dimen110
\c@theorem=\count87
\c@case=\count88
\c@conjecture=\count89
\c@corollary=\count90
\c@definition=\count91
\c@example=\count92
\c@exercise=\count93
\c@lemma=\count94
\c@note=\count95
\c@problem=\count96
\c@property=\count97
\c@proposition=\count98
\c@question=\count99
\c@solution=\count100
\c@remark=\count101
\c@figure=\count102
\c@table=\count103
\abovecaptionskip=\skip41
\belowcaptionskip=\skip42
\figcapgap=\dimen111
\tabcapgap=\dimen112
\figgap=\dimen113
\bibindent=\dimen114
\@tempcntc=\count104
) (c:/TexLive2009/texmf-dist/tex/generic/babel/babel.sty
Package: babel 2008/07/06 v3.8l The Babel package
(c:/TexLive2009/texmf-dist/tex/generic/babel/english.ldf
Language: english 2005/03/30 v3.3o English support from the babel system
(c:/TexLive2009/texmf-dist/tex/generic/babel/babel.def
File: babel.def 2008/07/06 v3.8l Babel common definitions
\babel@savecnt=\count105
\U@D=\dimen115
)
\l@canadian = a dialect from \language\l@american
\l@australian = a dialect from \language\l@british
\l@newzealand = a dialect from \language\l@british
) (c:/TexLive2009/texmf-dist/tex/generic/babel/swedish.ldf
Language: swedish 2005/03/31 v2.3d Swedish support from the babel system
Package babel Info: Making " an active character on input line 90.
)) (c:/TexLive2009/texmf-dist/tex/latex/graphics/graphicx.sty
Package: graphicx 1999/02/16 v1.0f Enhanced LaTeX Graphics (DPC,SPQR)

```

```

(c:/TexLive2009/texmf-dist/tex/latex/graphics/keyval.sty
Package: keyval 1999/03/16 v1.13 key=value parser (DPC)
\KV@toks@=\toks16
) (c:/TexLive2009/texmf-dist/tex/latex/graphics/graphics.sty
Package: graphics 2009/02/05 v1.0o Standard LaTeX Graphics (DPC,SPQR)
(c:/TexLive2009/texmf-dist/tex/latex/graphics/trig.sty
Package: trig 1999/03/16 v1.09 sin cos tan (DPC)
) (c:/TexLive2009/texmf-dist/tex/latex/latexconfig/graphics.cfg
File: graphics.cfg 2009/08/28 v1.8 graphics configuration of TeX Live
)
Package graphics Info: Driver file: dvips.def on input line 91.
(c:/TexLive2009/texmf-dist/tex/latex/graphics/dvips.def
File: dvips.def 1999/02/16 v3.0i Driver-dependant file (DPC,SPQR)
))
\Gin@req@height=\dimen116
\Gin@req@width=\dimen117
) (./mciteplus.sty
Package: mciteplus 2008/01/15 v1.0 enhanced mcite
\@mcitetmptoksA=\toks17
\@mcitetmpdima=\dimen118
\@mcitetmpdimb=\dimen119
\@mcitetmpcnta=\count106
\@mcitetmpcntb=\count107
\c@mcitebibitemcount=\count108
\c@mcitesubitemcount=\count109
) (c:/TexLive2009/texmf-dist/tex/latex/natbib/natbib.sty
Package: natbib 2009/11/07 8.31a (PWD, AO)
\bibhang=\skip43
\bibsep=\skip44
LaTeX Info: Redefining \cite on input line 694.
\c@NAT@ctr=\count110
) (c:/TexLive2009/texmf-dist/tex/latex/base/makeidx.sty
Package: makeidx 2000/03/29 v1.0m Standard LaTeX package
) (c:/TexLive2009/texmf-dist/tex/latex/jknaptx/mathrsfs.sty
Package: mathrsfs 1996/01/01 Math RSFS package v1.0 (jk)
\symrsfs=\mathgroup4
) (c:/TexLive2009/texmf-dist/tex/latex/amsmath/amsmath.sty
Package: amsmath 2000/07/18 v2.13 AMS math features
\@mathmargin=\skip45
For additional information on amsmath, use the '?' option.
(c:/TexLive2009/texmf-dist/tex/latex/amsmath/amstext.sty
Package: amstext 2000/06/29 v2.01
(c:/TexLive2009/texmf-dist/tex/latex/amsmath/amsgen.sty
File: amsgen.sty 1999/11/30 v2.0
\@emptytoks=\toks18
\ex@=\dimen120
)) (c:/TexLive2009/texmf-dist/tex/latex/amsmath/amsbsy.sty
Package: amsbsy 1999/11/29 v1.2d
\pmbraise@=\dimen121
) (c:/TexLive2009/texmf-dist/tex/latex/amsmath/amsopn.sty
Package: amsopn 1999/12/14 v2.01 operator names
)

```

`\inf@bad=\count111`  
LaTeX Info: Redefining `\frac` on input line 211.  
`\uproot@=\count112`  
`\leftroot@=\count113`  
LaTeX Info: Redefining `\overline` on input line 307.  
`\classnum@=\count114`  
`\DOTSCASE@=\count115`  
LaTeX Info: Redefining `\ldots` on input line 379.  
LaTeX Info: Redefining `\dots` on input line 382.  
LaTeX Info: Redefining `\cdots` on input line 467.  
`\Mathstrutbox@=\box29`  
`\strutbox@=\box30`  
`\big@size=\dimen122`  
LaTeX Font Info: Redefining font encoding OML on input line 567.  
LaTeX Font Info: Redefining font encoding OMS on input line 568.

Package amsmath Warning: Unable to redefine math accent `\vec`.

`\macc@depth=\count116`  
`\c@MaxMatrixCols=\count117`  
`\dotsspace@=\muskip10`  
`\c@parentequation=\count118`  
`\dspbrk@lvl=\count119`  
`\tag@help=\toks19`  
`\row@=\count120`  
`\column@=\count121`  
`\maxfields@=\count122`  
`\andhelp@=\toks20`  
`\eqnshift@=\dimen123`  
`\alignsep@=\dimen124`  
`\tagshift@=\dimen125`  
`\tagwidth@=\dimen126`  
`\totwidth@=\dimen127`  
`\lineht@=\dimen128`  
`\@envbody=\toks21`  
`\multlinegap=\skip46`  
`\multlinetaggap=\skip47`  
`\mathdisplay@stack=\toks22`  
LaTeX Info: Redefining `\[` on input line 2666.  
LaTeX Info: Redefining `\]` on input line 2667.  
) [c:/TexLive2009/texmf-dist/tex/latex/tools/multicol.sty  
Package: multicol 2008/12/05 v1.6h multicol column formatting (FMI)  
`\c@tracingmulticols=\count123`

Package multicol Warning: May not work with the twocolumn option on input line 135.

`\mult@box=\box31`  
`\multicol@leftmargin=\dimen129`  
`\c@unbalance=\count124`  
`\c@collectmore=\count125`  
`\doublecol@number=\count126`

```

\multicoltolerance=\count127
\multicolpretolerance=\count128
\full@width=\dimen130
\page@free=\dimen131
\premulticols=\dimen132
\postmulticols=\dimen133
\multicolsep=\skip48
\multicolbaselineskip=\skip49
\partial@page=\box32
\last@line=\box33
\mult@rightbox=\box34
\mult@grightbox=\box35
\mult@gfirstbox=\box36
\mult@firstbox=\box37
\@tempa=\box38
\@tempa=\box39
\@tempa=\box40
\@tempa=\box41
\@tempa=\box42
\@tempa=\box43
\@tempa=\box44
\@tempa=\box45
\@tempa=\box46
\@tempa=\box47
\@tempa=\box48
\@tempa=\box49
\@tempa=\box50
\@tempa=\box51
\@tempa=\box52
\@tempa=\box53
\@tempa=\box54
\c@columnbadness=\count129
\c@finalcolumnbadness=\count130
\last@try=\dimen134
\multicolovershoot=\dimen135
\multicolundershoot=\dimen136
\mult@nat@firstbox=\box55
\colbreak@box=\box56
) (c:/TexLive2009/texmf-dist/tex/latex/caption/caption.sty
Package: caption 2010/01/09 v3.1m Customizing captions (AR)
(c:/TexLive2009/texmf-dist/tex/latex/caption/caption3.sty
Package: caption3 2010/01/14 v3.1m caption3 kernel (AR)
\captionmargin=\dimen137
\captionmargin@=\dimen138
\captionwidth=\dimen139
\caption@indent=\dimen140
\caption@parindent=\dimen141
\caption@hangindent=\dimen142
Package caption Info: Unknown document class (or package),
(caption)      standard defaults will be used on input line 1084.
)

```

Package caption Warning: Unsupported document class (or package) detected,  
(caption) usage of the caption package is not recommended.  
See the caption package documentation for explanation.

Package caption Info: \@makecaption = \long macro:#1#2->\captionstyle \ifx \@ca  
p type \fig@type \vskip \figcapgap \fi \setbox \@tempboxa \hbox {\floatlegendst  
yle #1\floatcounterend }\capstrut #2}\ifdim \wd \@tempboxa >\hsize {\floatlegen  
dstyle #1\floatcounterend }\capstrut #2\par \else \hbox to\hsize {\leftlegendl  
ue \unhbox \@tempboxa \hfil }\fi \ifx \@captive \fig@type \else \vskip \tabcapg  
ap \fi .

\c@ContinuedFloat=\count131

) (c:/TexLive2009/texmf-dist/tex/latex/footmisc/footmisc.sty

Package: footmisc 2009/09/15 v5.5a a miscellany of footnote facilities

\FN@temptoken=\toks23

\footnotemargin=\dimen143

\c@pp@next@reset=\count132

Package footmisc Info: Declaring symbol style bringhurst on input line 855.

Package footmisc Info: Declaring symbol style chicago on input line 863.

Package footmisc Info: Declaring symbol style wiley on input line 872.

Package footmisc Info: Declaring symbol style lampport-robust on input line 883.

Package footmisc Info: Declaring symbol style lampport\* on input line 903.

Package footmisc Info: Declaring symbol style lampport\*-robust on input line 924

.

) (./brinck\_2010-dielsalder\_revised.aux)

\openout1 = `brinck\_2010-dielsalder\_revised.aux'.

LaTeX Font Info: Checking defaults for OML/cmm/m/it on input line 13.

LaTeX Font Info: ... okay on input line 13.

LaTeX Font Info: Checking defaults for T1/cmr/m/n on input line 13.

LaTeX Font Info: ... okay on input line 13.

LaTeX Font Info: Checking defaults for OT1/cmr/m/n on input line 13.

LaTeX Font Info: ... okay on input line 13.

LaTeX Font Info: Checking defaults for OMS/cmsy/m/n on input line 13.

LaTeX Font Info: ... okay on input line 13.

LaTeX Font Info: Checking defaults for OMX/cmex/m/n on input line 13.

LaTeX Font Info: ... okay on input line 13.

LaTeX Font Info: Checking defaults for U/cmr/m/n on input line 13.

LaTeX Font Info: ... okay on input line 13.

Unknown journal: specify \journalname{<name of your journal>} in preamble

Package caption Info: Begin \AtBeginDocument code.

Package caption Info: Incompatible package detected (regarding \@caption).

(caption) \@caption = \long macro:#1[#2]#3->\par \addcontentsline {  
\csname ext@#1\endcsname }{#1}{\protect \numberline {\csname the#1\endcsname }}  
\ignorespaces #2}}\begingroup \@parboxrestore \@makecaption {\csname fnum@#1\en  
dcsname }{\ignorespaces #3}\par \endgroup .

Package caption Warning: \caption will not be redefined since it's already

(caption) redefined by a document class or package which is

(caption) unknown to the caption package.

See the caption package documentation for explanation.



Package caption Info: End \AtBeginDocument code.  
SVJour3 warning: Missing \journalname

LaTeX Font Warning: Font shape `OT1/cmr/m/n' in size <8.5> not available  
(Font) size <8> substituted on input line 30.

LaTeX Font Warning: Font shape `OT1/cmr/m/n' in size <16> not available  
(Font) size <17.28> substituted on input line 30.

LaTeX Font Warning: Font shape `OT1/cmr/bx/n' in size <16> not available  
(Font) size <17.28> substituted on input line 30.

LaTeX Font Info: Try loading font information for U+rsfs on input line 30.  
(c:/TexLive2009/texmf-dist/tex/latex/jknapltx/ursfs.fd  
File: ursfs.fd 1998/03/24 rsfs font definition file (jk  
)

LaTeX Font Info: Calculating math sizes for size <8.5> on input line 30.

LaTeX Font Warning: Font shape `OT1/cmr/m/n' in size <4.25> not available  
(Font) size <5> substituted on input line 30.

LaTeX Font Warning: Font shape `OML/cmm/m/it' in size <8.5> not available  
(Font) size <8> substituted on input line 30.

LaTeX Font Warning: Font shape `OML/cmm/m/it' in size <4.25> not available  
(Font) size <5> substituted on input line 30.

LaTeX Font Warning: Font shape `OMS/cmsy/m/n' in size <8.5> not available  
(Font) size <8> substituted on input line 30.

LaTeX Font Warning: Font shape `OMS/cmsy/m/n' in size <4.25> not available  
(Font) size <5> substituted on input line 30.

LaTeX Font Warning: Font shape `OMX/cmex/m/n' in size <8.5> not available  
(Font) size <8> substituted on input line 30.

LaTeX Font Warning: Font shape `U/rsfs/m/n' in size <8.5> not available  
(Font) size <8> substituted on input line 30.

LaTeX Font Warning: Font shape `U/rsfs/m/n' in size <4.25> not available  
(Font) size <5> substituted on input line 30.

Overfull \hbox (1.64522pt too wide) in paragraph at lines 33--34  
\OT1/cmr/m/n/10 nam-ics (MD) and den-sity func-tional the-ory (DFT) stud-  
[]

[1

]

Overfull \hbox (5.75746pt too wide) in paragraph at lines 44--45  
\OT1/cmr/m/n/10 Li-pase B (CALB), per-formed by us and collaborators,[],  
[]

! LaTeX Error: Cannot determine size of graphic in calb-scheme.png (no Bounding Box).

See the LaTeX manual or LaTeX Companion for explanation.

Type H <return> for immediate help.

...

l.48 ...x{0.30}{\includegraphics{calb-scheme.png}}

Try typing <return> to proceed.

If that doesn't work, type X <return> to quit.

File: calb-scheme.png Graphic file (type eps)

<calb-scheme.png>

Underfull \hbox (badness 1383) in paragraph at lines 49--49

\OT1/cmr/m/n/8 residues im-por-tant for the present study ex-plic-itly shown.  
[]

Overfull \hbox (4.0769pt too wide) in paragraph at lines 55--56  
\OT1/cmr/m/n/10 pro-vide an en-thalpic TS sta-bi-liza-tion of  $\text{\Oms/cmsy/m/n/10}^{\text{\Oms/cmsy/m/n/10}}$   
 $\text{\Oms/cmsy/m/n/10}^{\text{\Oms/cmsy/m/n/10}}$  1-5 kcal $\text{\Oms/cmsy/m/n/10}^{\text{\Oms/cmsy/m/n/10}}$   $\text{\Oms/cmsy/m/n/10}^{\text{\Oms/cmsy/m/n/10}}$  mol $\text{\Oms/cmsy/m/n/10}^{\text{\Oms/cmsy/m/n/10}}$ ,\$  
[]

Overfull \hbox (5.03635pt too wide) in paragraph at lines 55--56  
\OT1/cmr/m/n/10 de-pend-ing on re-ac-tants and po-lar-iz-abil-ity of the medium  
.  
[]

Overfull \hbox (0.25838pt too wide) in paragraph at lines 70--72  
\OT1/cmr/m/n/10 As a first ap-prox-i-ma-tion, we can con-sider the case where  
[]

Overfull \hbox (2.59172pt too wide) in paragraph at lines 77--79  
\OT1/cmr/m/n/10 This re-ac-tion can be de-scribed by the well-known Micha-  
[]

Underfull \hbox (badness 10000) has occurred while \output is active

□□  
□

[2]

! LaTeX Error: Cannot determine size of graphic in dynamics.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.133 ...phics[width=0.75\textwidth]{dynamics.png}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: dynamics.png Graphic file (type eps)

<dynamics.png> [3]

Underfull \vbox (badness 10000) has occurred while \output is active □

Underfull \vbox (badness 10000) has occurred while \output is active □

Underfull \hbox (badness 10000) has occurred while \output is active

□□  
□

[4]

Overfull \hbox (20.21135pt too wide) in paragraph at lines 149--150

\OT1/cmr/m/n/10 nat-u-ral sub-strate, \$\OML/cmm/m/it/10 ^^AG[]\$ \OT1/cmr/m/n/10  
is in the or-der of 0.5-2 kcal\$\OMS/cmsy/m/n/10 ^^A\$\OT1/cmr/m/n/10 mol\$[]\$.

□

Overfull \hbox (1.2585pt too wide) in paragraph at lines 157--158

\OT1/cmr/m/n/10 zyme com-bi-na-tions us-ing eq. 11[]. There is, un-for-tu-natel  
y,

□

[5]

Overfull \hbox (2.06403pt too wide) in paragraph at lines 170--171

\OT1/cmr/m/n/10 files, neu-tral-ize the en-zyme com-plexes with Sodium ions,

□

Overfull \hbox (8.45302pt too wide) in paragraph at lines 170--171

\OT1/cmr/m/n/10 avail-able united-atom model, [□] while chlo-ro-form (CHL)

[]

Overfull \hbox (2.14742pt too wide) in paragraph at lines 172--173  
\OT1/cmr/m/n/10 handled by particle Mesh Ewald summation. Initial min-  
[]

Overfull \hbox (7.78625pt too wide) in paragraph at lines 174--175  
\OT1/cmr/m/n/10 within AM-BER. For structural analysis, an average struc-  
[]

Overfull \hbox (3.11844pt too wide) in paragraph at lines 176--177  
\OT1/cmr/m/n/10 diene and dienophile were estimated using the LIE meth-  
[]

Underfull \vbox (badness 1253) has occurred while \output is active []

Underfull \hbox (badness 10000) has occurred while \output is active  
[]  
[]

[6]  
Overfull \hbox (1.11954pt too wide) in paragraph at lines 209--210  
\OT1/cmr/m/n/10 Among a larger set of molecules, benzoquinone \OT1/cmr/bx/n/1  
0 1\OT1/cmr/m/n/10 , maleic  
[]

! LaTeX Error: Cannot determine size of graphic in BMC\_reactants.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.213 ...0.7}{\includegraphics{BMC\_reactants.png}}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: BMC\_reactants.png Graphic file (type eps)  
<BMC\_reactants.png>

! LaTeX Error: Cannot determine size of graphic in BMC\_Hb.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.225 ...raphics[width=0.85\textwidth]{BMC\_Hb.png}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: BMC\_Hb.png Graphic file (type eps)  
<BMC\_Hb.png>

LaTeX Font Warning: Font shape `OT1/cmr/bx/n' in size <8.5> not available  
(Font) size <8> substituted on input line 226.

[7]  
Underfull \hbox (badness 10000) has occurred while \output is active  
[]  
[]

[8]

! LaTeX Error: Cannot determine size of graphic in Hb\_dph\_dne.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.252 ...x{0.08}{\includegraphics{Hb\_dph\_dne.png}}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: Hb\_dph\_dne.png Graphic file (type eps)  
<Hb\_dph\_dne.png>

! LaTeX Error: Cannot determine size of graphic in nac2\_4.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.262 ...lebox{0.31}{\includegraphics{nac2\_4.png}}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: nac2\_4.png Graphic file (type eps)  
<nac2\_4.png>

Overfull \hbox (2.2018pt too wide) in paragraph at lines 259--267  
\OT1/cmr/m/n/10 [\OT1/cmr/bx/n/10 3\$\OMS/cmsy/m/n/10 ^^A\$\OT1/cmr/m/n/10 S105A/  
I285A] com-plexes. It is dif-fi-cult to tell whether

[]

! LaTeX Error: Cannot determine size of graphic in ben\_snapshots\_trans.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.272 ...includegraphics{ben\_snapshots\_trans.png}}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: ben\_snapshots\_trans.png Graphic file (type eps)  
<ben\_snapshots\_trans.png>  
Overfull \hbox (1.84682pt too wide) in paragraph at lines 269--276  
\OT1/cmr/m/it/10 Ben-zo-quinone. [] \OT1/cmr/m/n/10 Sta-tis-ti-cally, ben-zo-q  
uinone \OT1/cmr/bx/n/10 1 \OT1/cmr/m/n/10 is the dien-  
[]

! LaTeX Error: Cannot determine size of graphic in mal\_snapshots\_trans.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.281 ...includegraphics{mal\_snapshots\_trans.png}}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: mal\_snapshots\_trans.png Graphic file (type eps)  
<mal\_snapshots\_trans.png>  
Overfull \hbox (2.95287pt too wide) in paragraph at lines 278--285  
\OT1/cmr/m/n/10 is less con-sis-tent than for the other sub-strates; in S105A/  
[]

[9]

! LaTeX Error: Cannot determine size of graphic in cyh\_snapshots\_trans.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.291 ...includegraphics{cyh\_snapshots\_trans.png}}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: cyh\_snapshots\_trans.png Graphic file (type eps)  
<cyh\_snapshots\_trans.png>  
Underfull \hbox (badness 10000) has occurred while \output is active  
[]  
[]

[10]

! LaTeX Error: Cannot determine size of graphic in consensus\_BMC.png (no BoundingBox).

See the LaTeX manual or LaTeX Companion for explanation.  
Type H <return> for immediate help.

...

l.307 ....30}{\includegraphics{consensus\_BMC.png}}

Try typing <return> to proceed.  
If that doesn't work, type X <return> to quit.

File: consensus\_BMC.png Graphic file (type eps)  
<consensus\_BMC.png>  
LaTeX Font Info: Try loading font information for OMS+cmr on input line 312.

(c:/TexLive2009/texmf-dist/tex/latex/base/omscmr.fd  
File: omscmr.fd 1999/05/25 v2.5h Standard LaTeX font definitions  
)

LaTeX Font Info: Font shape `OMS/cmr/m/n' in size <7> not available  
(Font) Font shape `OMS/cmsy/m/n' tried instead on input line 312.  
LaTeX Font Info: Font shape `OMS/cmr/m/n' in size <5.94997> not available  
(Font) Font shape `OMS/cmsy/m/n' tried instead on input line 312.

Overfull \hbox (9.18231pt too wide) in paragraph at lines 320--321  
\OT1/cmr/m/it/10 (C) Sim-u-la-tion of a T40A/S105A/I189A variant in ACN.  
[]

[11]  
Overfull \hbox (1.5352pt too wide) in paragraph at lines 394--395  
\OT1/cmr/m/n/10 the [E\$OMS/cmsy/m/n/10 ^^A\$OT1/cmr/m/n/10 dph] and TS optimized ge-ome-tries, in-clud-ing oxy-  
[]

Underfull \hbox (badness 10000) has occurred while \output is active  
[]  
[]

[12]

LaTeX Warning: File `TS\_diagram2.png' not found on input line 415.

! LaTeX Error: File `TS\_diagram2.png' not found.

See the LaTeX manual or LaTeX Companion for explanation.

Type H <return> for immediate help.

...

l.415 ...cs[width=0.85\textwidth]{TS\_diagram2.png}

Try typing <return> to proceed.

If that doesn't work, type X <return> to quit.

! LaTeX Error: Cannot determine size of graphic in TS\_diagram2.png (no Bounding Box).

See the LaTeX manual or LaTeX Companion for explanation.

Type H <return> for immediate help.

...

l.415 ...cs[width=0.85\textwidth]{TS\_diagram2.png}

Try typing <return> to proceed.

If that doesn't work, type X <return> to quit.

File: TS\_diagram2.png Graphic file (type eps)

<TS\_diagram2.png>

Underfull \vbox (badness 10000) has occurred while \output is active []

Overfull \hbox (9.81189pt too wide) in paragraph at lines 422--423

\OT1/cmr/m/n/10 largest TS sta-bi-liza-tion (\$\OML/cmm/m/it/10 ^^AG[] \OMS/cmsy/m/n/10 ^^@ \OML/cmm/m/it/10 ^^AG[] \OT1/cmr/m/n/10 = \OMS/cmsy/m/n/10 ^^@\OT1/cmr/m/n/10 3\OML/cmm/m/it/10 :\OT1/cmr/m/n/10 9\$ kcal\$\OMS/cmsy/m/n/10 ^^A\$\OT1/cmr/m/n/10 mol\$[]\$).

[]

[13]

Underfull \hbox (badness 10000) has occurred while \output is active

[] []

[]

[14]

Overfull \hbox (0.25847pt too wide) in paragraph at lines 452--453

\OT1/cmr/m/n/10 rate-enhancements for catal-y-sis of bi-molec-u-lar re-ac-tions

[]

Overfull \hbox (1.14732pt too wide) in paragraph at lines 456--457



\OT1/cmr/m/n/10 mu-tant in wa-ter. The com-puted rate en-hance-ments with  
□

Overfull \hbox (12.14326pt too wide) in paragraph at lines 458--459  
□\OT1/cmr/m/n/8 This work has been sup-ported by the Swedish  
□

Underfull \vbox (badness 3108) has occurred while \output is active []

[15] (./brinck\_2010-dielsalder\_revised.bbl  
Underfull \hbox (badness 10000) in paragraph at lines 77--85  
□\OT1/cmr/m/n/10 Zhang, X.; DeChancie, J.; Gu-nay-din, H.;  
□

Underfull \hbox (badness 1460) in paragraph at lines 77--85  
\OT1/cmr/m/n/10 Chowdry, A. B.; Clemente, F. R.; Smith, A.  
□

Underfull \hbox (badness 4217) in paragraph at lines 141--148  
□\OT1/cmr/m/n/10 Branney, C.; Car-lqvist, P.; Mag-nus-son, A.;  
□

Missing character: There is no Å in font cmr10!  
Missing character: There is no ^ in font cmr10!

Underfull \hbox (badness 3158) in paragraph at lines 157--165  
□\OT1/cmr/m/n/10 Svedendahl, M.; Car-lqvist, P.; Bran-neby, C.;  
□

Underfull \hbox (badness 2173) in paragraph at lines 157--165  
\OT1/cmr/m/n/10 Allnr, O.; Frise, A.; Hult, K.; Berglund, P.;  
□

Underfull \hbox (badness 3158) in paragraph at lines 166--173  
□\OT1/cmr/m/n/10 Carlqvist, P.; Sve-den-dahl, M.; Bran-neby, C.;  
□

Underfull \hbox (badness 10000) in paragraph at lines 239--249  
□\OT1/cmr/m/n/10 Rothlisberger, D.; Kher-son-sky, O.; Wol-la-  
□

Underfull \vbox (badness 10000) has occurred while \output is active []

Underfull \hbox (badness 10000) has occurred while \output is active

[]  
[]

[16]

Underfull \hbox (badness 5245) in paragraph at lines 398--404

[]\OT1/cmr/m/n/10 Uppenberg, J.; Hansen, M. T.; Patkar, S.;  
[]

Underfull \hbox (badness 2096) in paragraph at lines 405--411

[]\OT1/cmr/m/it/10 GOLD 4.1\OT1/cmr/m/n/10 , 2008, <http://www.ccdc.cam.ac.uk/>-  
[]

Underfull \hbox (badness 1852) in paragraph at lines 501--507

[]\OT1/cmr/m/n/10 Guardia, E.; Pin-zon, R.; Ca-sulleras, J.; Oro-  
[]

Underfull \vbox (badness 10000) has occurred while \output is active []

Underfull \hbox (badness 10000) in paragraph at lines 610--616

[]\OT1/cmr/m/it/10 UCSF Chimera\OT1/cmr/m/n/10 , 2004,  
[]

Underfull \hbox (badness 10000) in paragraph at lines 631--636

[]\OT1/cmr/m/n/10 Wesemann, M. \OT1/cmr/m/it/10 Plot, v. 0.997\OT1/cmr/m/n/10 ,  
2007,  
[]

) [17] (./brinck\_2010-dielsalder\_revised.aux)

LaTeX Font Warning: Size substitutions with differences  
(Font) up to 1.28pt have occurred.

LaTeX Warning: Label(s) may have changed. Rerun to get cross-references right.

)

Here is how much of TeX's memory you used:

4027 strings out of 93775

62484 string characters out of 1149208

144637 words of memory out of 3000000

7276 multiletter control sequences out of 15000+50000

11436 words of font info for 43 fonts, out of 1200000 for 2000

714 hyphenation exceptions out of 8191

36i,20n,38p,4309b,522s stack positions out of 5000i,500n,10000p,200000b,5000s

Output written on brinck\_2010-dielsalder\_revised.dvi (17 pages, 113932 bytes).

