J|A|C|S
ARTICLES
Published on Web 11/16/2002

Reaction Mechanism of Soluble Epoxide Hydrolase: Insights from Molecular Dynamics Simulations§

Birgit Schiøtt*,† and Thomas C. Bruice*,‡

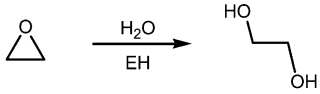
Contribution from the Department of Chemistry, Aarhus University, 8000 Aarhus, Denmark and the Department of Chemistry and Biochemistry, University of California, Santa Barbara, California, 93106

Received July 25, 2002

Abstract: Molecular dynamics simulations have been performed to gain insights into the catalytic mechanism of the hydrolysis of epoxides to vicinal diols by soluble epoxide hydrolase (sEH). The binding of a substrate, 1S,2S-*trans*-methylstyrene oxide, was studied in two conformations in the active site of the enzyme. It was found that only one is likely to be found in the active enzyme. In the preferred conformation the phenyl group of the substrate is π-sandwiched between two aromatic residues, Tyr381 and His523, whereas the other conformation is π-stacked with only one aromatic residue, Trp334. Two simulations were carried out to 1 ns for each conformation to evaluate the protonation state of active site residue His523. It was found that a protonated histidine is essential for keeping the active site from being disrupted. Long time scale, 4 ns, molecular dynamics simulation was done for the structure with the most likely combination of binding conformation and protonation state of His523. Near Attack Conformers (NACs) are present 5.3% of the time and nucleophilic attack on either epoxide carbon atom, ~75% on C^1 and ~25% on C^2, is found. A maximum of one hydrogen bond between the epoxide oxygen and either of the active site tyrosines, Tyr465 and Tyr381, is present, in agreement with experimental mutagenesis results that reveal a slight loss in activity if one tyrosine is mutated and essential loss of all activity upon double mutation of the two tyrosines in question. It was found that a hydrogen bond from Tyr465 to the substrate oxygen is essential for controlling the regioselectivity of the reaction. Furthermore, a relationship between the presence of this hydrogen bond and the separation of reactants was found. Two groups of amino acid segments were identified each as moving collectively. Furthermore, an overall anti-correlation was found between the movements of these two individually collectively moving groups, made up by parts of the cap-region, including the two tyrosines, and the site of the catalytic triad, respectively. This overall anti-correlated collective domain motion is, perhaps, involved in the conversion of E·NAC to E·TS.

Introduction

Oxiranes (epoxides) are formed from an oxidative metabolism of unsaturated endogenous[1] and xenobiotic compounds. Many of the oxiranes derived from xenobiotics have been classified as mutagenic and carcinogenic initiators.[2] It is of vital importance for the living organism to be able to defend itself against such harmful chemicals and their metabolites. Epoxide hydrolases (EH)§ (EC 3.3.2.3) are the first line of defense against oxiranes and they have therefore attracted much attention over the past decades.[3,4,5] Recent evidence, however, suggest that the major role of the enzyme is in the general area of biochemical mediation.[6] EHs catalyze the conversion of oxiranes to more soluble and easy excretable vicinal diols by formal addition of a water molecule (eq 1).



(1)

EHs are ubiquitous enzymes that have been isolated from a wide variety of organisms.[7−9] Enzymes from microbial sources

* Correspondence can be addressed to either author. E-mail: birgit@chem.au.dk. Fax: +45 8619 6199. Phone: +45 8942 3953. E-mail: tcbruice@bioorganic.ucsb.edu. Fax: (805) 893−2229. Phone: (805) 893−2044.

† Department of Chemistry, Aarhus University, Aarhus, Denmark.

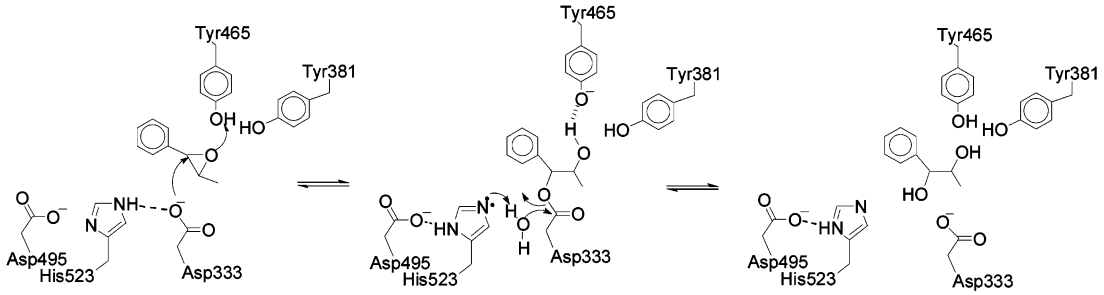
‡ Department of Chemistry and Biochemistry, University of California, Santa Barbara, California.

§ Abbreviations: EH, epoxide hydrolase; sEH, soluble epoxide hydrolase; mEH, microsomal epoxide hydrolase; E·S, enzyme-substrate; NAC, near attack conformation; MD, molecular dynamics; tMSO, 1S,2S-*trans*-methylstyrene oxide; rmsd, root-mean-square deviation; SD, steepest descent; ABNR, adopted basis Newton-Raphson.

(1) Greene, J. F.; Newman, J. W.; Williamson, K. C.; Hammock, B. D. *Chem. Res. Tox.* **2000**, *13*, 217−226.
(2) Fretland, A. J.; Omiecinski, C. J. *Chem.-Biol. Interac.* **2000**, *129*, 41−59.
(3) Oesch, F. *Xenobiotica* **1973**, *3*, 305−340.
(4) Ota, K.; Hammock, B. D. *Science* **1980**, *207*, 1479−1481.
(5) Wixtram, R. N.; Hammock, B. D. *Biochemical Pharmacology and Toxicology*; Zakim, D., Vessey, D. A., Eds.; Wiley: New York, 1985; Vol. 1, 1−93.
(6) (a) Zeldin, D.; Wei, S.; Falck, J. R.; Hammock, B. D.; Snapper, J. R.; Capdevila, J. H. *Arch. Biochem. Biophys.* **1995**, *316*, 443−451. (b) Morriseau, C.; Beetham, J. K.; Pinot, F.; Debernard, S.; Newman, J. W.; Hammock, B. D. *Arch. Biochem. Biophys.* **2000**, *378*, 321−332. (c) Armstrong, R. N. *Drug Metab. Rev.* **1999**, *31*, 71−86.
(7) Archer, I. V. J. *Tetrahedron* **1997**, *53*, 15 617−15 662.
(8) Pedragosa-Moreau, S.; Archelas, A.; Furstoss, R. *J. Org. Chem.* **1993**, *58*, 5533−5536.
(9) Mischitz, M.; Kroutil, W.; Wandel, U.; Faber, K. *Tetrahedron: Asymmetry* **1995**, *6*, 1261−1272.

10.1021/ja021021r CCC: $22.00

Scheme 1



are easily accessible through fermentation techniques, and thus, the potential use of EHs as highly asymmetric biocatalysts is of much interest in organic synthesis.[10,11] Enantiopure epoxides and vicinal diols, both very important building blocks in total synthesis efforts, can be obtained from racemic epoxides by this biotransformation through kinetic resolution experiments on a preparative scale.[12]

Mammalian epoxide hydrolases are present in the liver to detoxify epoxides. Three of the known SEHs are very substrate specific toward various metabolites of endogenous compounds, cholesterol oxide hydrolase, hepoxilin A_3 hydrolase, and leukotriene A_4 hydrolase.[2] Two enzymes, each with a broader range of substrate selectivity, have been found, microsomal (mEH) and soluble epoxide hydrolases (sEH). These two forms of epoxide hydrolase show low sequence homology[13] though both are members of the α/β hydrolase fold family[14] possessing an α/β-sheet core domain and a cap domain.[15] The active sites of all EHs are located between the two domains of the structure and contain an Asp-His-Asp catalytic triad and two neighboring tyrosine residues located in the cap-domain. It has been found, based on several studies with O^{18} labeled water,[16] that the enzyme reaction is initiated through nucleophilic attack of an aspartate residue (Asp333), located at the "nucleophilic elbow," on an oxiran carbon atom to form a glycol−monoester−enzyme intermediate. Furthermore, from mutagenic studies, Tyr381 and Tyr465 in the active site have been identified as essential and have been assigned as general acid(s) catalyst(s) through hydrogen (bond) donation to the intermediate.[17] Very recently high-resolution X-ray crystal structures of EHs have become available; three structures have been solved of apo-enzymes (murine sEH,[18] *Agrobacterium radiobacter* AD1[19] and *Aspergillus Niger*[20]) and three structures of murine sEH with bound urea inhibitors.[18,21] Due to difficulties in crystallizing membrane bound proteins, all solved structures are sEHs, though *Aspergillus Niger* represents a border enzyme being a soluble mEH.[20]

All solved structures have the topological features inevitably of the active site. The three-dimensional structures of the active site of the sEH enzymes are remarkably identical regardless of whether a urea inhibitor is present or not. This can be appreciated by inspection of Figure 1 where an overlay of the five solved sEH structures are depicted using the murine sEH numbering scheme. Note that the charge-relay aspartate is missing for the microbial sEH, rather a water molecule is positioned to hydrogen bond to Nδ1 of the general base, histidine. The structure of the catalytic domain shows great similarity to the active site of haloalkane dehalogenase.[18]

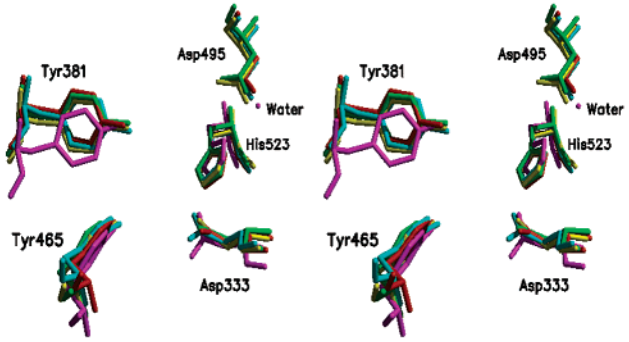


Figure 1. Stereoimage of an overlayof the active site from the five X-ray structures of sEH. Yellow, orange and green: murine sEH with bound urea inhibitors (PDB codes 1CR6, 1EK2, and 1EK1, respectively); cyan, apo-enzyme of murine sEH (PDB code 1CQZ); magenta, apo-enzyme of *Agrobacterium radiobacter* AD1 (PDB code 1EHY). Numbering scheme refers to murine sEH.

The reaction mechanism of Scheme 1 has been proposed for hydrolysis of *trans*-β-methylstyrene oxide by sEH.[18] No direct evidence accounting for the protonation state of the important residues of the active site is available based on the high-resolution X-ray structures. As outlined in the proposed mechanism,[18] the protonation state of His523 is extremely unsettled, because all three possibilities were included in the proposal.[18]

Two important mechanistic questions emerge after inspection of the proposed mechanism by Argiriadi et al.;[18] what is the protonation state of His523 in the active site of the enzyme and how is the formed tyrosinate stabilized? Armstrong and Cassidy have recently questioned the formation of a tyrosinate in the active site.[22] Rather they suggest, that the forming glycol-

(10) Orru, R. V. A.; Faber, K. *Curr. Opion. Chem. Biol*. **1999**, *3*, 16−21.
(11) Chiappe, C.; Rubertis, A. D.; Marioni, F.; Pelagotti, F. *J. Mol. Catal. B* **2001**, *14*, 125−129.
(12) Archelas, A.; Furstoss, R. *Curr. Opion. Chem. Biol*. **2001**, *5*, 112−119; Archelas, A.; Furstoss, R. *Trends Biotech*. **1998**, *16*, 108−116.
(13) Beetham, J. K.; Grant, D.; Arand, M.; Garbarino, J.; Kiyosue, T.; Pinot, F.; Oesch, F.; Belknap, W. R.; Shinozaki, K.; Hammock, B. D. *DNA Cell Biol.* **1995**, *14*, 61−71.
(14) Nardini, M.; Dijkstra, W. *Curr. Opion. Struct. Biol*. **1999**, *9*, 732−737.
(15) (a) Arand, M.; Wagner, H.; Oesch, F. *J. Chem. Biol.* **1996**, *271*, 4223−4229; (b) Rink. R.; Fenneman, M.; Smids, M.; Dehmel, U.; Janssen, D. B. *J. Biol. Chem*. **1997**, *272*, 14 650−14 657.
(16) (a) Lacourciere, G. M.; Armstrong, R. N. *J. Am. Chem. Soc*. **1993**, *115*, 10466−10 467. (b) Borhan, B.; Jones, A. D.; Pinot, F.; Grants, D. F.; Kurth, M. J.; Hammock, B. D. *J. Biol. Chem*. **1995**, *270*, 26 923−26 930.
(17) (a) Rink, R.; Spelberg, J. H. L.; Pieters, R. J.; Kingma, J.; Nardini, M.; Kellogg, R. M.; Dijksatra, B. W.; Janssen, D. B. *J. Am. Chem. Soc*. **1999**, *121*, 7417−7418; (b) Rink, R.; Kingma, J.; Spelberg, J. H. L.; Janssen, D. B. *Biochemistry* **2000**, *39*, 5600−5613; (c) Yamada, T.; Morisseau, C.; Maxwell, J. E. M. A. Argiriadi, Christianson, Hammock, B. D. *J. Biol. Chem*. **2000**, *275*, 23 082−23 088.
(18) Argiriadi, M. A.; Morrisseau, C.; Hammock, B. D.; Christianson, D. W. *Proc. Natl. Acad. Sci*. **1999**, *96*, 10 637−10 642.
(19) Nardini, M.; Ridder, I. S.; Rozeboom, H. J.; Kalk, K. H.; Rink, R.; Janssen, D. B.; Dijkstra, B. W. *J. Biol. Chem*. **1999**, *274*, 14 579−14 586.
(20) Zou, J.; Hallberg, B. M.; Bergfors, T.; Oesch, F.; Arand, M.; Mowbray, S. L.; Jones, T. A. *Structure* **2000**, *8*, 111−122.
(21) Argiriadi, M. A.; Morrisseau, C.; Goodrow, M. H.; Dowdy, D. L.; Hammock, B. D.; Christianson, D. W. *J. Biol. Chem*. **2000**, *275*, 15 265−15 270.
(22) Armstrong, R. N.; Cassidy, C. S. *Drug Metab. Rev*. **2000**, *32*, 327−338.

monoester enzyme intermediate is charged (alkoxy group) and hydrogen bonded to the tyrosine residues. On the basis of fluorescence experiments, Armstrong et al. outlined the kinetic mechanism for the similar mEH class of enzymes, revealing a reversible alkylation reaction to be found, and excluding a mechanism that involves major conformational changes in the active site after alkylation of the enzyme and before the rate-limiting hydrolysis step.[23]

In this paper, we present the results of an investigation of the structure and dynamic behavior of murine sEH with bound substrate, tMSO, based upon molecular dynamics (MD) simulations in order to look deeper into the mechanistically unsolved questions discussed above. Specifically, the conformation of bound substrate, the protonation state of His523, as well as the regioselectivity of the nucleophilic attack at the epoxide ring have been the focus of our analysis. Also, the possible coupling between enzyme dynamics and catalysis is discussed from an analysis of correlated motions within the catalytic domain.

Methods

General Setup. All simulations of murine soluble epoxide hydrolase with bound substrate, *trans*-methylstyrene oxide, in the active site were performed with the program CHARMM (version 25b2)[24] and using the CHARMM22 all-atom force field parameters.[25] The structures reveal that mammalian sEHs are large dimeric enzymes having 544 amino acids in each monomer. Furthermore, it is found that each monomer consists of two compact domains which are joined together by a meandering strand. The C-terminus (Val235-Ala544) contains the catalytic domain whereas the N-terminus (Arg4-Gly218) is the vestigial domain, resembling the active site of haloacid dehalogenase.[18] The initial structures for all simulations are extracted from the catalytic domain of strand B in the crystal structure of murine sEH with bound urea inhibitor (PDB entry 1CR6).[18] The urea inhibitor was deleted and a substrate molecule was manually docked into the active sites of the dimeric enzyme using MidasPlus.[26] The structure of the substrate was taken from our previous work optimized at the HF/6-31+G(d,p) level of theory.[27] The resulting structure was further manipulated before MD, as hydrogens were added to the X-ray coordinates of the enzyme and crystal water molecules using CHARMM. The protonation states of all histidines were carefully evaluated based on analysis of the binding environment of each of the histidines (see Supporting Information for details). A total of four simulations have been carried out to elucidate the binding conformation of 1S,2S-*trans*-methylstyrene oxide as well as the protonation state of His523. The two conformations in the study have the phenyl group of the substrate stacked with either the indole group of Trp334 or with His523, conformation A and B, respectively. This choice is based on an earlier study by us which used molecular mechanics minimization procedures for tMSO docked with the phenyl group of the substrate superimposed on either of the two large groups of the urea inhibitor of the cocrystallized enzyme-ligand complex.[27] For each conformation, two simulations were done, one with a neutral His523, and the hydrogen atom found on Nδ1 and one with a histidinium ion. The resulting four enzyme-ligand complexes (approximate dimensions were 100 Å × 100 Å × 40 Å) were minimized in two sets of minimizations before being subjected to stochastic boundary molecular dynamics. First, the potential energy of the systems were minimized keeping non-hydrogen atoms of the enzyme system fixed by using a combination of steepest descent (SD, 500 steps) and adopted-basis Newton−Raphson (ABNR, 10 000 steps or until convergence criteria has been met). The minimized coordinates were subjected to a second minimization in which all non-hydrogen atoms were constrained by using weak harmonic constraints (force constant of 1 kcal/(mol·Å^2)) − SD 500 steps followed by ABNR until convergence has been met. The minimized structures of the four studied complexes are depicted in Figure 2, with the important active site residues and the substrate shown. Hydrogens have been omitted for clarity, though the hydrogens on the imidazole ring of His523 are shown to specify the protonation state of this important residue. These four complexes are termed confB-pos, confB-neu, confA-pos, and confA-neu, respectively, referring to the binding conformation and the charge of His523.

tMSO Modeling. The substrate is modeled by using the HF/6-31+G(d,p) optimized structure from our previous work.[27] Atomic partial charges for use in the force field calculations are computed by fitting point charges to the electrostatic potential derived from the HF/6-31+g(d,p) electron density using the ChelpG procedure.[28] There are no parameters in the CHARMM22 force field to describe an ether type oxygen. Therefore, Accelrys−CHARMm parameters as included in the program QUANTA 98[29] were used to model the substrate accordingly. The computed partial charges, the added values for force constants and van der Waals terms required to model tMSO are tabulated in the Supporting Information.

Molecular Dynamics Simulations. The same protocol for stochastic boundary molecular dynamics[30] simulations was used for the four minimized structures of Scheme 1. Each minimized structure was placed in a 32 Å in radius sphere of TIP3P[31] water molecules centered at C1 of the substrate (benzylic carbon). If an oxygen atom of a TIP3P water molecule was within 2.8 Å of an enzyme or substrate heavy atom, then it was deleted. To avoid cavities in the solvation sphere, the water sphere was rotated and the existing water−enzyme−substrate system was re-solvated and water molecules closer than 2.8 Å to other atoms were deleted. This re-solvation procedure was done four times sequentially, resulting in a total of approximately 23 500 atoms and overall ∼2300 water molecules solvating the catalytic domain of strand B. The resulting model for confB-pos is included in the Supporting Material. The water layer surrounding the catalytic domain is up to 12 Å thick. A spherical boundary potential was used to prevent the water from "evaporating" from the solvent surface.[32] Atoms outside the 30 Å sphere were frozen

(23) Laughlin, L. T.; Tzeng, H.-F.; Lin, S.; Armstrong, R. N. *Biochemistry* **1998**, *37*, 2897−3904.
(24) Brooks, B. R.; Brucolleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(25) MacKerell Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher III, W. E.; Roux, B.; Schlenkrih, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3585−3616.
(26) MidasPlus Version 2.1, Computer Graphics Lab., University of California, San Francisco, 1997.
(27) Lau, E. Y.; Newby, Z. E.; Bruice, T. C. *J. Am. Chem. Soc*. **2001**, *123*, 3350−3357.
(28) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem*. **1990**, *11*, 361−373.
(29) *QUANTA*: Molecular Simulations Inc.: San Diego, CA, 1998, as of July 2001, Accelrys.
(30) Brooks III, C. L.; Karplus, M. *J. Mol. Biol*. **1989**, *208*, 159−181.
(31) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impley, R. W.; Klein, M-. L. *J. Chem. Phys.* **1983**, *79*, 925−935.

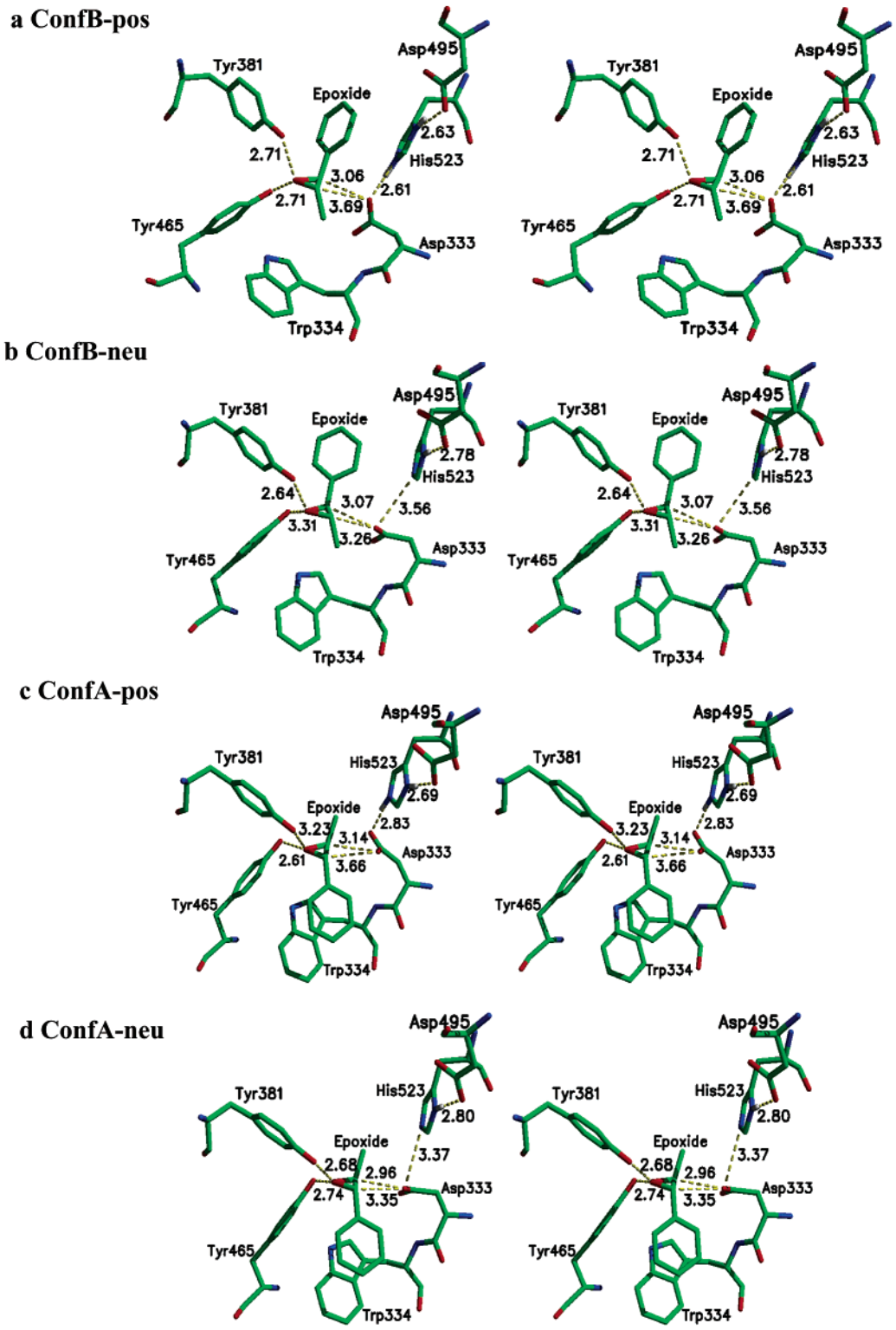


Figure 2. Stereo images of the minimized structures of the four enzyme-ligand complexes studied. The names refer to the binding conformation and the protonation state of His523. Carbon atoms are shown in green, hydrogens in white, oxygens in red and nitrogens in blue. For clarity, only hydrogen atoms involved in hydrogen bonding from His523 are shown.

during dynamics. A buffer region, 2 Å in thickness, was used between 28 and 30 Å from C1 of the substrate. The system was coupled to a heat bath (300 K) using a frictional coefficient of 250 ps^{-1} on heavy-protein atoms.[30] The oxygen atoms of the water molecules had a frictional coefficient of 62 ps^{-1}.[30] Constraining forces for the heavy protein atoms in the buffer region were constructed from their average Debye−Waller *B*-factors[18] and applying the equation $\langle \Delta r^2 \rangle^{1/2} = (3B/8\pi^2)^{1/2}$.[33] All atoms within 28 Å of the origin were treated by the ordinary equations of motions, using Verlet dynamics.[34] Atoms in the

(32) Brünger, A. T.; Brooks, C. L.; Karplus, M. *Proc. Natl. Acad. Sci.* **1985**, *82*, 8458−8462.

buffer zone were treated by Langevin dynamics.[30] A force-shifting function was used to cutoff the Coulombic terms at 13 Å and a switching function was used between 10 and 11 Å for the van der Waals energy.[35] A time step of 1 fs was used and the nonbonded list was updated every 20 time steps. Coordinates were saved every 100 time steps and the SHAKE algorithm[36] was used to constrain bonds containing hydrogen atoms to their equilibrium distance (tolerance 10^{-6}). The four simulations were carried out to approximately 1 ns. One simulation (confB-pos) was continued to 4 ns, because it showed the most promising binding conformation based on the initial ~1 ns of dynamics.

Trajectories of amino acids 235−544 of strand B (the catalytic domain) after the initial 200 ps of heating and equilibration were considered for analysis. Positional fluctuations around the average structure were calculated for the backbone α-carbon atoms and compared to those obtained from the crystal structure *B*-factors. The extent of correlated motions was evaluated by calculating the covariance matrix and cross correlations[37] in CHARMM for the backbone α-carbons in the reaction zone.

Results and Discussion

Minimized Structures. The minimized structures of the four E·S complexes studied are depicted in Figure 2. Names are chosen to describe the binding conformation of the substrate in the active site and the charge of His523. When the substrate is bound in conformation B the phenyl ring is π-sandwiched between His523 and Tyr381, whereas in the A-conformation only one π-interaction is possible for the substrate, with Trp334. An other important stabilizing interaction for substrate binding is the possible formation of hydrogen bonds from the epoxide oxygen atom to the two tyrosine residues, Tyr381 and Tyr465. All minimized structures show good to fair distances between the tyrosine Oγ's, and the epoxide oxygen atom with respect to hydrogen bond formation. Inspection of the minimized structures of ConfB-pos (Figure 2 a) and ConfB-neu (Figure 2 b) reveals that the π-interactions are structurally different. Recent very high-level ab initio calculations explained how stabilizing π-interaction could be found by having either of two different intermolecular geometries, orthogonal (as in ConfB-neu) and parallel-displaced as in ConfB-pos.[38] This difference in the substrate binding is most likely caused by the increased distance between Nϵ2 of His523 and the nearest oxygen atom of Asp333 upon removal of His523(Hϵ2). However, the overall structure of the four minimized E·S complexes, which serve as starting structures for the dynamics, are very similar. The average root-mean-square deviation (RMSD) for each Cα in the protein between the minimized structures and the crystal structure was 0.232 , 0.129, 0.232, and 0.230 Å, respectively, for confB-pos, confB-neu, confA-pos, and confA-neu. The observed differences in dynamics can therefore not be due to discrepancies in the starting structures.

Stability of MD Simulations. All four MD-simulations were stable over the duration of the simulation because the variation in the temperature were less than 3 K; thus the differences observed in the dynamics cannot be ascribed to instabilities in the simulation procedure. The RMSDs for the Cα backbone atoms relative to the crystal structure were calculated for all four simulations and figures are included in the Supporting Material. Heating and equilibration corresponds to the initial 0.2 ns and production dynamics are present after 0.2 ns for all four simulations. An RMSD of ~2.0 Å is found after 4 ns for ConfB-pos indicating a stable complex.

The positional fluctuations of the Cα in the protein backbone in the catalytic domain were calculated from the full 4 ns MD trajectory for confB-pos and compared to those derived from the Debye−Waller *B*-factors of the crystal structure (in Supporting Material) by use of the equation $\langle \Delta r^2 \rangle^{1/2} = (3B/8\pi^2)^{1/2}$.[33] The positional fluctuations from the confB-pos simulation is in fine qualitative agreement with those obtained from crystallography because areas with low flexibility in the crystal are reproduced nicely in the simulation. However, the magnitude of the fluctuations in the simulation is lower than those obtained from the crystal structure. This discrepancy can be due either to insufficient sampling of all possible conformations of sEH within the time scale of the simulation[39] or to static (lattice) disorder within the crystal. The crystal structure was solved at 2.8 Å and could contain lattice defects and other systematical errors (e.g., absorption, extinction, disorder[40]) which would lead to increased B-factors.[41] This increased B-factor when converted to positional fluctuations will have an enhanced value in which only a portion of the value is actually from the thermal motions. Very few residues gave higher fluctuations in the simulation than found in the crystal, and it was found that they all are located at the surface of the catalytic domain of the protein. In the simulation, these residues are solvent exposed and therefore expected to show more motion than in the crystal where their motions can be attenuated due to packing effects.

Overall, based on the calculated RMSDs as a function of simulation time for the four sampled structures and on the positional fluctuation for confB-pos, we believe that the local motions sampled by the tMSO-sEH complex in the MD trajectories are reasonable and thus the long 4 ns simulation for confB-pos represents a realistic picture of the dynamics of this enzyme when complexed to the substrate. In the following, we present a comparative discussion of the dynamics of the four studied structures of Figure 2 based on ~1 ns simulations before we turn to a detailed analysis of the 4 ns simulation of the preferred confB-pos structure.

1 ns Simulations of the Four Complexes. The calculated RMSD-plots for the four studied binding conformations of tMSO to sEH look very similar—see the Supporting Information for figures. When the substrate is bound in conformation B the geometry with a neutral His523 leads to an initially steeper increase in RMSD reaching 1.2 Å after approximately 50 ps, whereas with the histidinium ion, it takes about 230 ps to reach an RMSD of 1.2 Å. In the production dynamics phase, after 200 ps, the two curves are very identical, both only showing very slow increase to ~1.5 Å at 1 ns. Whether amino acid His

(33) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987.
(34) Verlet, L. *Phys. Rev.* **1967**, *159*, 98−113.
(35) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667−683.
(36) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Chem.* **1977**, *23*, 327−341.
(37) Ichiye, T.; Karplus, M. *Proteins* **1991**, *11*, 205−217.
(38) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10 887−10 893.
(39) Clarage, J. B.; Romo, T.; Andrews, B. K.; Pettitt, B. M.; Phillips, G. N. *Proc. Natl. Acad. Sci.* **1995**, *92*, 3288−3292.
(40) Iversen, B. B.; Larsen, F. K.; Figgis, B. N.; Reynolds, P. A.; Schultz, A. J. *Acta Cryst*, **1996**, *B52*, 932−931.
(41) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Prespective of Dynamics, Structure and Thermodynamics*; John Wiley and Sons: New York, 1987.

523 is neutral or protonated does apparently not affect the stability of the enzymatic structure much. From the calculated RMSDs, it can be speculated that the presence of the hydrogen bond between nucleophilic Asp333 and His523 in ConfB-pos does not influence the overall protein structure, but may have a more local effect on the active site structure. When the substrate is bound in conformation A, not much difference in the RMSDs whether His523 is neutral or protonated is found. A similar lack in protein structure dependence upon the protonation state of a structural Histidine is also found in inosine−uridine nucleoside hydrolase with respect to His241.[42]

The catalytic efficacy of the enzyme is dependent on how often the nucleophile and electrophile are present in near attack conformations (NACs).[43] A NAC is a ground state structure that geometrically must be formed prior to reaching the transition state of the studied reaction. It has been suggested, based on a detailed study of intramolecular ester formation, that the more NACs formed in an E·S complex the higher efficacy of the enzyme.[43,44] This idea of NAC formation as a measure of enzyme efficacy has recently attracted much attention in the literature.[45,46] To further evaluate and compare the four simulations, we analyzed the conformations sampled during production dynamics for NAC formation. The geometric criteria for NAC formation in the alkylation step of Scheme 1, being an S_N2 reaction, was defined based on structural data from our former ab initio study[27] and on van der Waal radii. A separation of less than 3.2 Å between the nucleophilic oxygen of Asp333 and the electrophilic epoxide carbon atom must be present. This has previously been shown to be a reasonable distance for discriminating reactive conformers in reactions where a oxygen nucleophile attacks a carbon center.[43] The attacking angle from the Asp333 oxygen via the attacked epoxide carbon to the leaving group (the epoxide oxygen), must be 145 ± 15°. This number is based on the results from the ab initio study where all structures of reactants, transition states and products had this angle between 140° and 148°. However, these two criteria are not enough to ensure a correct attacking geometry, because the cone-shell spanned in space around epoxide(C1) by these two limitations includes room in space occupied by the hydrogen and phenyl group bound at the epoxide(C1)—see the illustration in Figure 3. To ensure the correct orientation of the nucleophile in the NAC structures, it was necessary to limit the angle spanned between epoxide(C2), epoxide(C1) and the attacking oxygen of Asp333 to 90 ± 15° (termed directing angle) and also ensuring that the attacking group is in an anti orientation relative to the epoxide oxygen, see Figure 3 for a schematic illustration for these criteria for NAC-formation.

The structures collected during the molecular dynamics simulations of the four complexes were analyzed for NAC formation (Table 1); attacks by either carboxyl oxygen of Asp333 at either carbon atom of the epoxide ring were considered, giving a total of four possible reaction pathways for each binding conformation. The obtained numbers for NAC formation are listed in Table 1 for the four pathways in each 1 ns simulation. From an inspection of Table 1, it is obvious that

Table 1. Number of NACs Formed during Production Dynamics (0.2 to ∼1 ns) for the Four Complexes[a]

	OD1-C1	OD2-C1	OD1-C2	OD2-C2
confB-pos	0 (1)	648 (1045)	0 (0)	40 (59)
confB-neu	0 (0)	0 (146)[b]	0 (0)	3 (13)
confA-pos	0 (0)	0 (0)	0 (0)	0 (2)
confA-neu	0 (0)	0 (0)	0 (0)	0 (12)
confB-pos (4ns)	5 (6)	1526 (1923)	1 (1)	496 (515)

[a] The numbers in parentheses refer to the total number of NACs formed from 0 ns to ∼1 ns. The columns refer to all possible attacks by two carboxylate oxygens of Asp333 at two epoxide carbon atoms. The last entry represents the 4 ns simulation for ConfB-pos. The labels refer to Figure 3. [b] The last NAC-formation was found at 182,4 ps.

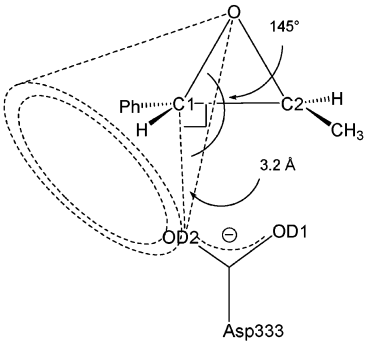


Figure 3. Schematic presentation of the structural criteria for NAC formation is shown. The attacking distance form Asp333(OD2) to Epoxide-(C1) and the attacking angle Asp333(OD2)- - - -Epoxide(C1)−Epoxide-(O) are outlined by dotted lines. The cone on the figure shows all the space that satisfy the NAC criteria when only considering attacking distance and angle.

the most optimal binding conformation of tMSO with regard to nucleophilic attack by Asp333 on the epoxide ring is ConfB-pos because most NACs are formed. The other three conformations simply fail to produce a significant number of NACs during the ∼1 ns of simulation. Several attempts to constrain the system during heat-up and equilibration all produced similar results − *no NACs are forming unless His523 is protonated and the substrate is bound in conformation B.* Closer looks into the dynamics of the three conformations lacking NAC formation revealed that the starting structure for the simulations, the active site with bound substrate, was destroyed/perturbed/changed in various ways during the 1 ns simulations.

The finding of a charged His523 is in agreement with the experimentally determined X-ray structures from which the presence of Hϵ2 in His523 seems likely when regarding the heteroatomic distances measured from the structures. The distances between His523(Nϵ2) and the nearest carboxyl oxygen of Asp333 varies from 2.36 to 2.70 Å for the five structures shown in Figure 1, and all are positioned in space to make almost linear hydrogen bonds. *Aspergillus Niger*, the soluble mEH, on the other hand shows a separation of 3.17 Å and no suitable geometry for making a hydrogen bond which may indicate a neutral histidine for this particular enzyme during the first reaction step. As a consequence of the four 1 ns simulations, the ConfB-pos conformation was extended beyond 1 ns continuing for a total of 4 ns.

4-ns Simulation of ConfB-pos. With a total of close to 25 000 atoms in the solvated enzyme-ligand complex, the ConfB-pos MD simulation represents, according to Daggett, a descent size study with regard to simulation time and the number of atoms.[47] As argued above the system is stable with an RMSD

(42) Mazumder, D.; Bruice, T. C. *J. Am. Chem. Soc*. In Press.
(43) Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc*. **1996**, *118*, 2595. Lightstone, F. C.; Bruice, T. C. *Acc. Chem. Res*. **1999**, *32*, 127−136.
(44) Bruice, T. C. *Acc. Chem. Res* **2002**, *35*, 139−148.
(45) Kollman, P. A.; Kuhn, B.;. Peräkylä, M *J. Phys. Chem. B* **2002**, *106*, 1537−1542.
(46) Field, M. J. *J. Comput. Chem*. **2002**, *23*, 48−58.

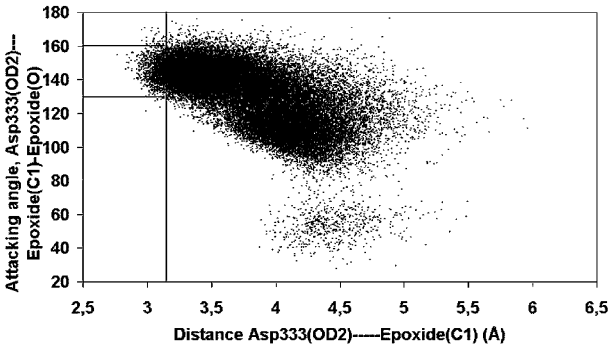


Figure 4. Conformations sampled by the nucleophilic Asp333(OD2) and the epoxide during dynamics. Structures within the small box formed by the lines represent NACs, with a separation of less than 3.2 Å and an attacking angle of 145 ± 15°.

of 2 Å after 4 ns with the substrate bound in the active site (see the Supporting Information for the plot). The trajectory was analyzed with respect to formation of NACs for the four possible reaction pathways, considering both of the oxygens at Asp333 as nucleophile and both electrophilic carbon centers at the epoxide ring. The results are displayed in Table 1 as the last entry. The trends observed during the first 1 ns simulation of this binding geometry are still valid for the full simulation period. Basically, only Asp333(OD2) is responsible for NAC-formation. Furthermore, reactive conformers are produced corresponding to benzylic (C1) as well as homobenzylic (C2) attacks at the epoxide.

Benzylic Attack. Let us first turn to a structural characterization of the NACs leading to attack of Asp333(OD2) at the benzylic position of the epoxide substrate, epoxide(C1). The criterion for formation of a NAC in this *pseudo*-S_N2 displacement reaction is a separation of less than 3.2 Å between the incoming nucleophile and the electrophilic center, Asp333(OD2)···Epoxide(C1), in combination with an attacking angle from nucleophile via electrophile to the leaving group, Asp333(OD2)··· Epoxide(C1)−Epoxide(O), of 145 ± 15°. Of the 38 000 structures sampled during production dynamics, 1731 fell into this category. Figure 4 graphically shows the distribution of attacking distance and angle among all conformations sampled. The calculated average Asp333(OD2)···Epoxide(C1) attack distance and Asp333(OD2)···Epoxide(C1)−Epoxide(O) attack angle from all sampled structures during the MD simulation were 3.80 ± 0.43 Å and 130.2 ± 18.9°, respectively. It is evident that many structures are formed which structurally is just beyond the limits set for a NAC. Of the 1731 structures inside the small box containing NAC-structures in Figure 4, only 1526 are true NACs when the two remaining criteria for NAC-formation are regarded, namely limiting the room in space that leads to a reactive structure (see Figure 3 and the last entry of Table 1). The average properties for a NAC were an Asp333(OD2)···Epoxide(C1) separation of 3.12 Å, an attacking angle of 144.5° and a directing angle of 93.5°. From Figure 4 it is revealed, that it is mostly the upper limit in separation between the nucleophile and the electrophile that defines these NAC-structures. Increasing the maximally allowed separation for NAC-formation by only 0.05 Å to 3.25 Å increases the number of NACs by more than 50%, whereas an increase in attacking angle by e.g., 5° did not make a significant difference, because

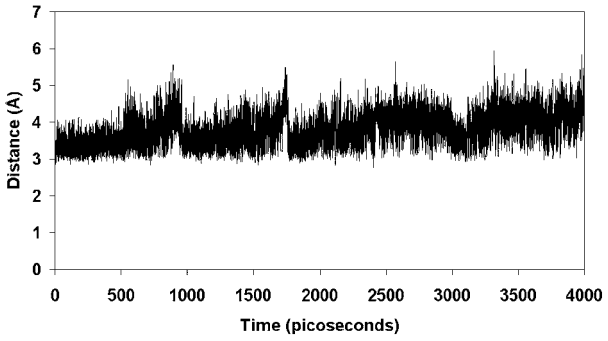


Figure 5. Variation in the separation between Asp333(OD2) and the benzylic carbon atom of the epoxide (C1) measured during dynamics.

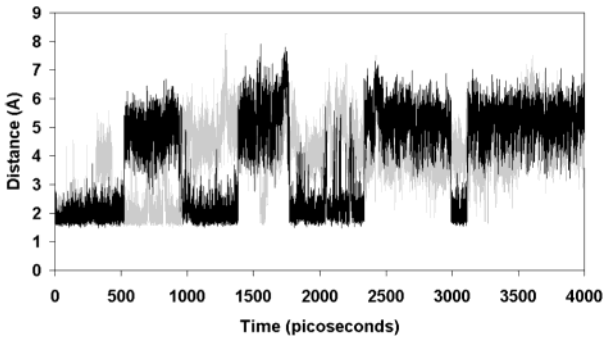


Figure 6. Observed changes in separation between the epoxide oxygen and the hydroxyl hydrogen of the two active site tyrosines, Tyr381 (black) and Tyr465 (grey) during the MD simulation.

almost none of the sampled conformations had attacking angles outside the 145 ± 15° range when keeping the separation between nucleophile and electrophile below 3.5 Å.

In Figure 5, the variation over time of the important O···C attacking separation is shown. It has been found experimentally that the two tyrosines (Tyr381 and Tyr465) in the active site are important for catalysis, and it has been proposed that they in concert function as hydrogen bond donors toward the evolving negatively charged oxygen atom during the S_N2 displacement at the epoxide.[17] The distance between each of the hydroxyl hydrogens of the tyrosines and the epoxide oxygen was recorded over time (Figure 6).

Two Distinct Ground State Conformations for Benzylic Attack? A noticeable observation is made when comparing the variations of the attacking distance, Asp333(OD2)···Epoxide-(C1), over time (Figure 5), with the making and breaking of a hydrogen bond between the epoxide oxygen and the hydroxyl group of Tyr465 (Figure 6). The two structural features seem to be related. When a hydrogen bond from the epoxide oxygen to Tyr465 is present (0−535 ps, 960−1383 ps, 1785−2335 ps, 3003−3112 ps), the O···C attack distance is short and less fluctuating as seen in Figures 5 and 7 (see later), whereas when this hydrogen bond is not present this separation is longer and more fluctuating. A similar observation has been previously made from an MD simulation of Catechol *O*-Methyltransferase,[48] for which the S_N2 displacement reaction of catecholate on S-adenosylmethionine is associated with a close contact between the leaving group (S-adenosylhomocysteine) and the C_β atom of a tyrosine residue. No similar relation between the separation of reactants and the existence of a hydrogen bond between epoxide oxygen and Tyr381 is reflected in the figures.

(47) Daggett, V. *Curr. Opion. Struct. Biol*. **2000**, *10*, 160−164.

(48) Lau, E. Y.; Bruice, T. C. *J. Am. Chem. Soc*. **1998**, *120*, 12 387−12 395.

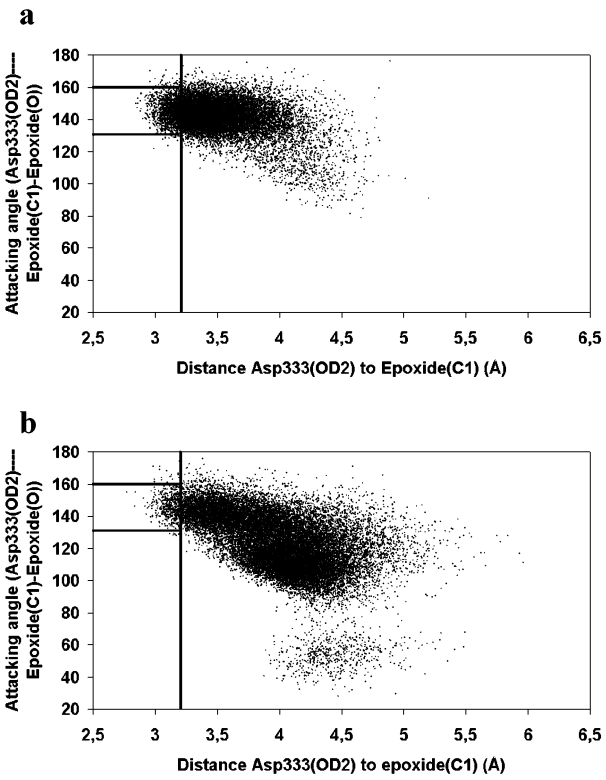


Figure 7. Areas in conformational space sampled by the two distinct populations sampled for ConfB-pos. The conformations sampled in the *tight* group of structures is shown in graph a, whereas the *loose* structures are found in graph b. NAC structures are located inside the small box of the plot.

To further explore whether the relation between attacking distance and hydrogen bonding to Tyr465 is important for NAC formation, the sampled structures were divided into two groups. In the first group, all conformations having the Tyr465 hydrogen bond are present. This group is called *tight* due to the apparent shorter separation between the carboxylate oxygen of Asp333 and epoxide(C1). The second group, termed *loose* to picture the larger and more fluctuating separation between the reactants, contained all sampled structures without the hydrogen bond in question. Plots similar to that of Figure 4 are provided for the two groups in Figure 7.

As expected, the *tight* group showed a more confined sampling of conformations, whereas the *loose* group lead to a much broader region in conformational space being sampled, and therefore, many nonreactive conformations are present in this group. In total, the *tight* group contained 14 170 structures (37% of all structures), of which 1056 were NACs (7.5%), whereas the *loose* group contained 23 830 structures but only 470 NACs (2.0%). This leads to at least two distinct populations of NACs sampled during the MD simulation for Asp333(OD2) attack at epoxide(C1), namely those formed when Tyr465 hydrogen bonds to the epoxide, the *tight* NAC structure, and those formed when this hydrogen bond is not present and the active site apparently is less compressed, the *loose* NAC structure. NAC-formation occurs approximately 4 times more often with the compression by Tyr465 in effect than in the *looser* active site. Overall, *tight* NACs for Asp333(OD2) attack at epoxide(C1) are formed in 4.0% of all structures sampled and *loose* NACs are formed in 1.3% of the structures sampled during production dynamics. From the Boltzmann distribution and using

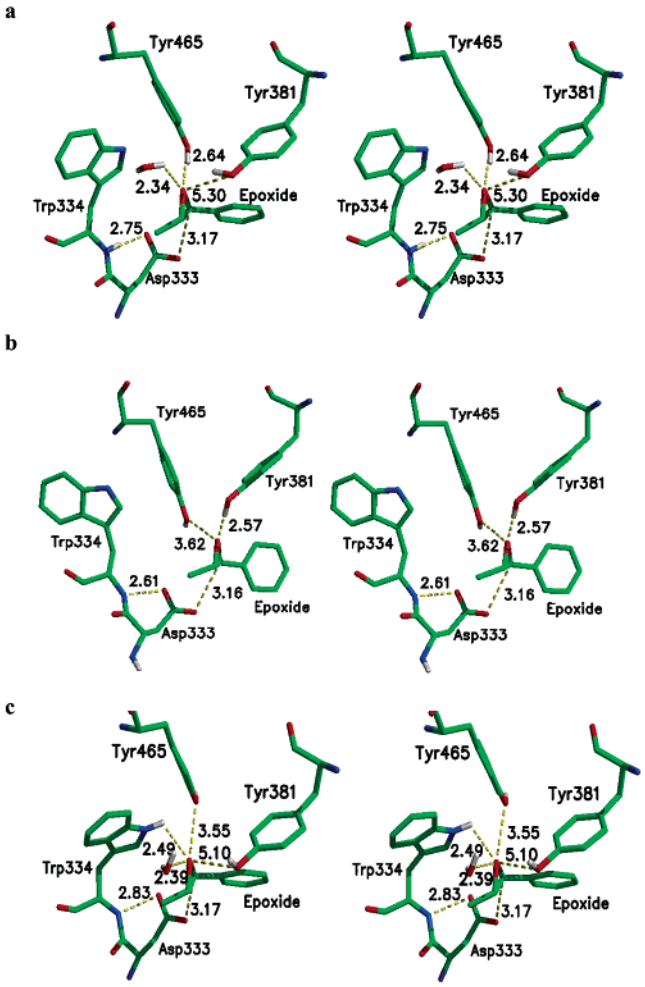


Figure 8. Stereoimages of the near attack conformations (NACs) formed in sEH for the *tight* structure at 1110.6 ps (a), and the *loose* structure, one with a hydrogen bond from the epoxide oxygen to Tyr381 at 809,8 ps (b) and one without at 2599.9 ps (c). The numerical values listed correspond to the distances in Ångstrøm. Oxygen atoms are shown in red, nitrogen in blue, carbon in green, and hydrogen in white. For clarity, only hydrogens involved in hydrogen bonds with the substrate or nucleophile are displayed.

these population probabilities for the two distinct NAC populations, the difference in energy between a *loose* and a *tight* NAC is calculated to be approximately 0.5 kcal/mol. However, this difference in energy between the two distinct NAC populations for attack of Asp333(OD2) at epoxide C1 does not provide any quantitative prediction of relative rates of reaction. Presumably, $\Delta G^{\ddagger}$ should be less positive when Tyr465 is in position to be a general acid catalyst. Accurate calculations, by e.g., QM/MM methods, are needed to evaluate the relative height of reaction barriers as well as the energy of the reactants and product structures in the two different hydrogen bonding environments. An interesting aspect in this regard will be to study how the transition states interacts with the active site tyrosines (work in progress).

Structures of NACs for Benzylic Attack. Essential characteristics of one *tight* and two *loose* NACs are provided in Figure 8. The structure of a *tight* NAC (Figure 8a) at 1110.6 ps, reveals the epoxide oxygen hydrogen bonds to Tyr465 and a water molecule. Tyr381 is hydrogen bonded to another water molecule (not shown) that has entered the active site. In the *tight* NAC structure, the distance between hydroxyl oxygen of Tyr381 and the epoxide oxygen is 5.30 Å. A *loose* NAC-

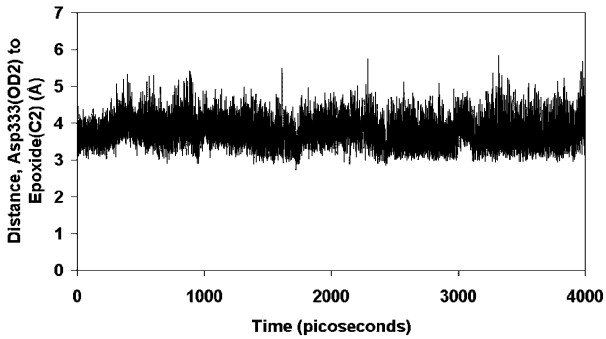


Figure 9. Variation over time in the measured separation between Asp333-(OD2) and epoxide (C2).

structure (Figure 8b) identified at 809.8 ps shows a fine hydrogen bond between Tyr381 and the epoxide oxygen, whereas the hydroxyl group of Tyr465 is pointing away from the hydrogen bond acceptor, the distance between the two oxygen atoms in question being 3.62 Å. A second *loose* NAC is depicted in Figure 8c, reflecting those NACs that are found with no hydrogen bond between any of the Tyrosines and the epoxide. As reflected in the picture, Trp334 is rotated so that Hϵ1 points into the active site and makes an interaction with the epoxide oxygen possible. In all the pictured NACs, Asp333-(OD1) is hydrogen bonded to the backbone NH of Trp334. In our former ab initio study in gas- and solution-phase of this reaction, it was found that the most stable conformation of the reactants had both carboxylic oxygens interacting with the epoxide carbons. It was proposed in that paper that an important task for the enzyme was to destabilize this conformation to favor a structure with only one O···C interaction present. The enzyme accomplishes this task by forming a hydrogen bond between Asp333(OD1) and the peptidic Trp334(NH). This is achieved by rotating the carboxyl plane to an almost perpendicular orientation relative to the epoxide plane.

Homobenzylic Attack. Let us now turn to an analysis of the registered NACs in Table 1 leading to attack by Asp333-(OD2) at the homobenzylic position of the substrate, at C2. The variation over time of the separation between Asp333(OD2) and Epoxide(C2) (O···C) is shown in Figure 9. The average separation between Asp333(OD2) and Epoxide(C2) was calculated to be 3.78 ± 0.34 Å, which is very similar to that calculated for attack at the C1 position. The total of 496 NACs during production dynamics leading to attack at the C2 position is divided as follows; 35 (0.29%) NAC structures are identified among the 14 170 sampled structures belonging to the *tight* active site, whereas 461 (1.93% of the 23 830 sampled structures) were found when having a *loose* active site. When comparing to the numbers reported for attack at epoxide(C1), it is seen, that the presence of a hydrogen bond between Tyr465 and the epoxide oxygen is a stronger regulating factor for C2 attack as compared to C1 attack. The likelihood for NAC formation is almost 8 times greater in the *loose* active site than in the *tight* active site. For attack at C1, it was the other way around, because the likelihood of NAC formation was greatest in the *tight* active site by about a factor of 4.

In Figure 10, two typical NAC structures are shown for attack at C2 in the *loose* active site. In both NACs the epoxide is hydrogen bonded to a water molecule. The structure in Figure 10a (at 1556.4 ps) has a hydrogen bond between Tyr381 and the epoxide oxygen, whereas Figure 10b contains a structure

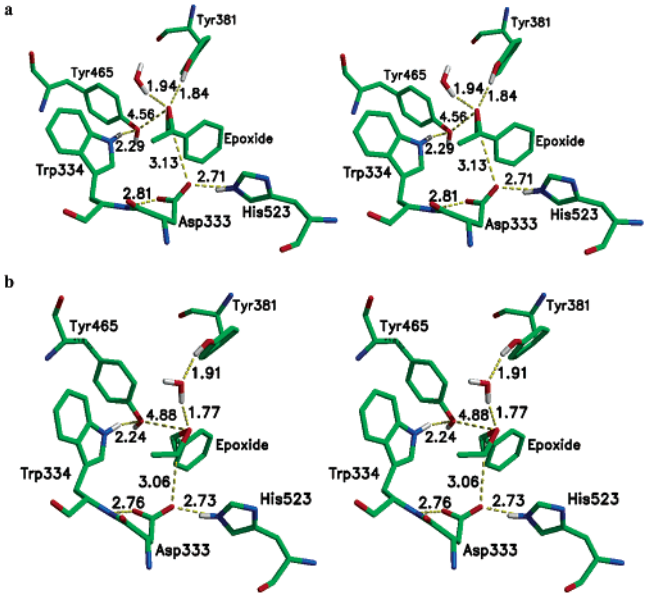


Figure 10. Stereoimages of the near attack conformations formed in sEH for nucleophilic attack by Asp333(OD2) at the homobenzylic C2 of the substrate in the *loose* active site at 1556,4 ps (a) and 2380,5 ps (b), respectively. The numerical values listed correspond to the distances in Ångstrøm. Oxygen atoms are shown in red, nitrogen in blue, carbon in green, and hydrogen in white. For clarity, only hydrogens involved in hydrogen bonds around the substrate or nucleophile are displayed.

(at 2380.5 ps) where only one hydrogen bond was found stabilizing the epoxide oxygen. The average NAC structure for attack at C2 had a Asp333(OD2)···epoxide(C2) separation of 3.13 Å, an attacking angle of 148.3° and a relative directing angle, Asp333(OD2)- - - -epoxide(C2)−epoxide(C1), of 95.4°.

Correlated and Anti-Correlated Motion. Differing views have been published on the plausibility of a linkage between enzyme mechanisms and correlated and anti-correlated motions in enzymes.[49−51] Much work, including analysis of long-term MD data with a variety of enzymes, is required prior to arriving at an understanding of the relationship (if any) between motions and rates of reactions. Recently, Radkiewicz, and Brooks suggested that a link exists between protein dynamics and catalysis in the dihydrofolate reductase enzymatic reactions.[49] From molecular dynamics studies of dihydrofolate reductase they found *only* in the Michaelis complex large areas of the protein exhibiting anti-correlated motion, which were suggested to be linked to the catalytic reaction that is initialized in this complex. In Figure 11 the calculated cross-correlations are displayed for the ConfB-pos simulation during the last half of the trajectory. To ensure that the cross-correlation coefficients have converged, a similar calculation was carried out for the full trajectory. The correlation pattern obtained for the full production dynamics period is essentially identical (see the Supporting Information for a plot) to that of the last half of the MD simulation displayed in Figure 11, thus reflecting the converged nature of the movements within the enzyme substrate complex. The segments of the protein showing correlated motions (yellow and red in Figure 11) reflect the overall secondary and tertiary structures (loops, sheets, helices, etc.) of the enzyme. Areas corresponding to enzyme fragments that

(49) Radkiewicz, J. L.; Brooks III, C. L*: J. Am. Chem. Soc*. **2000**, *122*, 225−231.
(50) Bruice, T. C.; Benkovic, S. J. *Biochemistry* **2000**, *39*, 6267−6274.
(51) Hammes, G. G. *Biochemistry*, **2002**, *41*, 8221−8228.

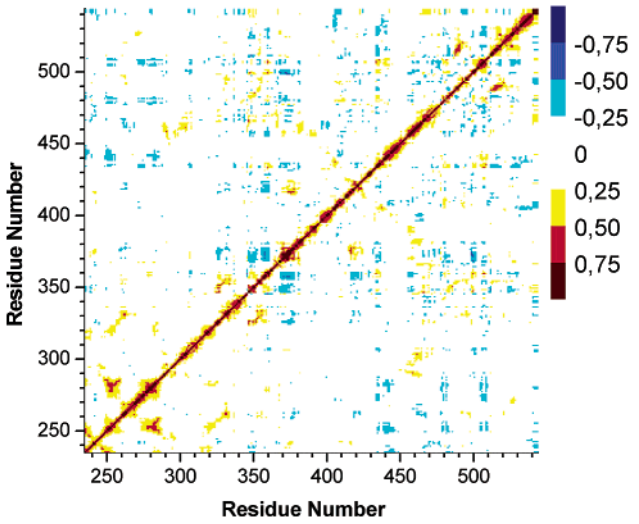


Figure 11. Plots of the cross-correlations coefficients, C(*i*,*j*), between the Cα's in the catalytic domain calculated during the second half of the trajectory. A similar plot of the full production dynamics is found in the Supporting Information. Areas in red and yellow correspond to residues that display correlated motions during dynamics, whereas areas in blue represent residues that are anti-correlated.

Table 2. Calculated Cross-correlation Coefficients between the Catalytic Triad Residues and the Hydrogen Bond Donors for the Last Half (2−4 Ns) of the Trajectory for ConfB-pos

	Tyr381	Tyr465	His523	Asp333	Asp495
Tyr381	1.00				
Tyr465	0.25	1.00			
His523	−0.14	−0.24	1.00		
Asp333	−0.22	−0.17	0.37	1.00	
Asp495	−0.11	−0.26	0.53	0.37	1.00

show anti-correlated movements (colored blue) are present, indicating that the simulated complex is indeed reactive.[49]

The calculated cross-correlation coefficients between the five amino acid residues in the active site are listed in Table 2. The numbers reveal that the amino acids of the catalytic triad move in a concerted manner because significant positive cross-correlation coefficients are computed. The two potential hydrogen bond donors also move concerted. Furthermore, it is seen that the two sides of the active site to a certain extent move in an anti-correlated fashion because all six calculated cross-correlation coefficients between any of the catalytic triad residues and either of the tyrosines are negative.

When looking closer at Figure 11, it becomes evident that bands run through the plot, indicating that the movements of all the residues within a certain string of the enzyme are similar with respect to the rest of the enzyme. One band is, for example, seen for the string containing residues 430−450 and another is identified for residues 270−285. From a closer look at the pattern of cross-correlations in Figure 11, for these two bands, it is seen that they are complementary, by which we mean that if the residues from 430 to 450 shows positive correlation with another string, e.g., residues 500−510, then the sequence from amino acid 270−285 shows anti-correlated motion with this particular string. The bands described all consist of amino acids belonging to the same tertiary structure.

Collective Domain Movement. To get a better idea as to whether there are segments of the catalytic domain in the enzyme that participate in larger collectively moving domains,

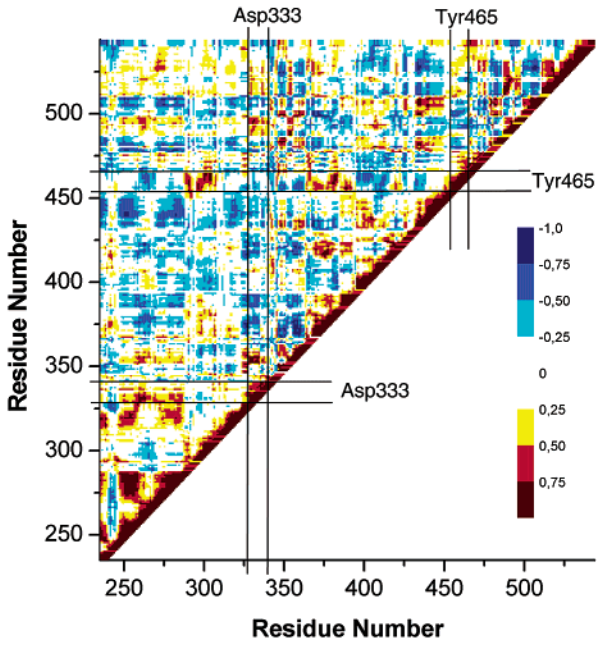


Figure 12. Plot of the super-correlation coefficients SCC(*i*,*j*)'s calculated during the last half of the trajectory. Areas in yellow-red corresponds to amino acid sequences that overall with respect to all other residues move correlated. Areas in blue, similarly, correspond to sequences that, move anti-correlated with regard to all other residues (see the Supporting Information for details).

as proposed originally by Ichiye and Karplus,[37] we analyzed the cross-correlation matrix calculated for the amino acid residues of the reaction zone further. An analysis of the similarity in the cross-correlation coefficients between all pairs of amino acid residues was undertaken to evaluate whether there are amino acid segments that have identical or similar patterns in the plotted cross-correlation coefficients of Figure 11. For this purpose we define a super-correlation coefficient, SCC(*i*,*j*), that quantifies how correlated the overall movement of two residues (*i* and *j*) are during dynamics with respect to all other residues, N (see the Supporting Information for details). Areas of overall oppositely moving amino acid residues are present in the super-correlation plot in Figure 12. Thus, dark blue islands are present, which was originally predicted in the early paper by Ichiye and Karplus.[37] In a 100 ps MD study by Swaminathan et al., such an anti-correlation was identified between parts of the flap and the cantilever regions of HIV-1.[52] The above analysis of the trajectory for the confB-pos structure revealed that the formation of a hydrogen bond between the oxygen atom of the epoxide substrate and Tyr465 seemed related to the attacking distance, which again showed to be a structural parameter used by the enzyme to produce NACs. We now want to explore further whether the two amino acid residues (Tyr465 and Asp333) involved in this relation belongs to domains of the enzyme that participate in collective movements.

By inspection of Figure 12 it is found that Tyr465 belongs to a string of amino acid residues from 453 to 467, because these residues forms a band in the plot. The figure furthermore reveals that this segment moves collectively with other segments because yellow-red islands are found in the plot for Tyr465.

(52) Swaminathan, S.; Harte, Jr., W. E.; Beveridge, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 2717−2721.

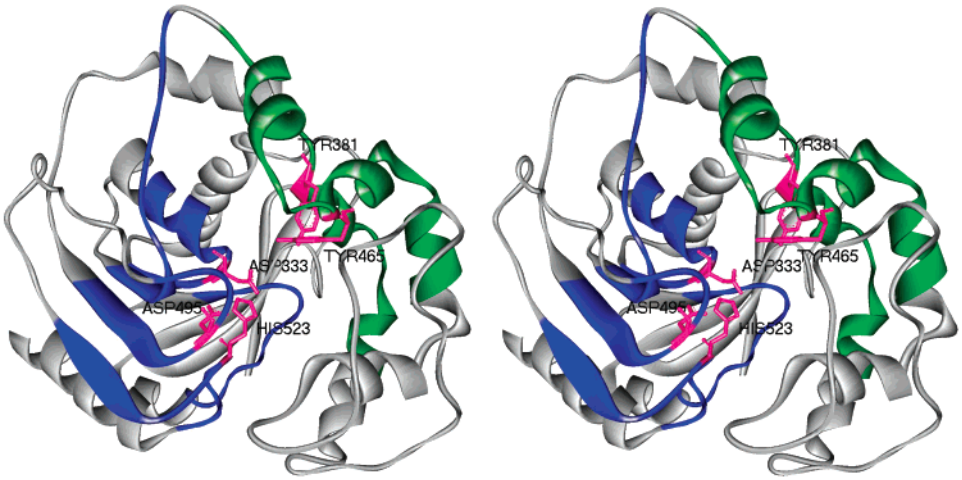


Figure 13. Mapping of the correlated and anti-correlated motions on the structure of the catalytic domain of sEH. Protein segments in green belongs to the A-family whereas the B-family is colored blue. All residues of the same color show correlated intra-domain motions, whereas the two families exhibit inter-domain anti-correlated overall motion. Residues colored magenta belong to the active site.

To make up a collectively moving domain these segments have to all move in an overall correlated manor. Careful inspection of Figure 12 reveals that Tyr465 belongs to a collectively moving domain, family A, consisting of amino acid residues 291−300, 372−396, and 453−467, because these three segments show significant positive super-correlation coefficients with respect to all other segments within the family (yellow-red islands are found in Figure 12). The other residue involved in the enzymatic control of NAC formation is Asp333. A similar analysis reveals, that Asp333 belongs to a segment consisting of residues 327−343. Again, it is revealed by inspecting Figure 12 that a collectively moving domain can be extracted from the data, forming family B, consisting of residues 260−272, 327−343, 354−365, 489−499, and 516−525. The most important finding connected to the analysis of Figure 12, is that any pair of amino acid segments having one segment from family A and the other from family B shows negative SCC-(*i*,*j*)'s, thus indicating that the two collectively moving families on a larger scale move oppositely. To visualize the two collectively moving families, the structure of the reaction zone is depicted in stereo in Figure 13 with the two domains colored differently.

Figure 13 reveals that the five segments of amino acids belonging to the B-family (colored blue) are found in close proximity to one another in the 3D-structure, and also that the three segments from the A-family (colored green) are found close in space to one another forming one collectively moving domain. One of the identified domains belongs to the cap-region, whereas the other is located around the site of the catalytic triad. From this finding, it is proposed that the described collective domain movements can be associated with the larger scale anti-correlated movement between the cap region and the region of the catalytic triad, thereby for instance imposing compression on the active site and forcing the reaction to proceed, because the analysis of the MD trajectory indicate that more NACs are produced in the resulting compressed *tight* structure. This shows a way that enzyme dynamics can be coupled to enzyme catalysis, similar to what was found based on NMR-experiments for cyclophilin A.[53] In MD simulations of protein dynamics, examination of correlated and anti-correlated motions will soon tend to be routine. Thus, to understand the importance and involvement of such phenomena it is a requirement that sufficient examples are available experimentally and theoretically. Recently, Hammes-Schiffer et al. have similarly presented results derived from mixed quantum/classical MD simulations that point toward the existence of a network of coupled promoting motions for dihydrofolate reductase and liver alcohol dehydrogenase.[54]

Conclusions and Perspectives

In this study, we have outlined new mechanistic details of the enzyme catalyzed hydrolysis of 1S,2S-*trans*-methylstyren oxide (tMSO) by soluble epoxide hydrolase (sEH). Compared to similar enzymes such as microsomal epoxide hydrolase and haloalkane dehologenase, sEH exhibits a lower pH-optimum, 6.5−7.8[55] compared to 8.4−9.0[15,19] and 9.0,[56] respectively, for the two other enzymes. This can be rationalized from the MD simulations of the enzyme−substrate complex to be due to the profound preference of having the active site His523 as a protonated histidinium ion which is more like to happen at a lower pH even though the residue is buried inside an enzyme. When this residue is protonated and forms hydrogen bonds to both active site aspartates, Asp333 and Asp495, the formation of NACs was observed much more often than without the His523−Asp333 hydrogen bond. On the basis of the analysis of the trajectories, it was proposed that the hydrogen bond is necessary to keep the active site together, and keep the nucleophile, Asp333, from forming hydrogen bonds to other hydrogen bond donors which would prevent it from attacking the epoxide. The most convincing experimental technique for collecting evidence for the charge state of His523 is a neutron diffraction study. A revised reaction mechanism accounting for the structural MD-derived features is outlined Scheme 2. This modified mechanism is similar to the mechanisms proposed for the hydrolysis of epichlorohydrin by *Agrobacterium radiobacter*

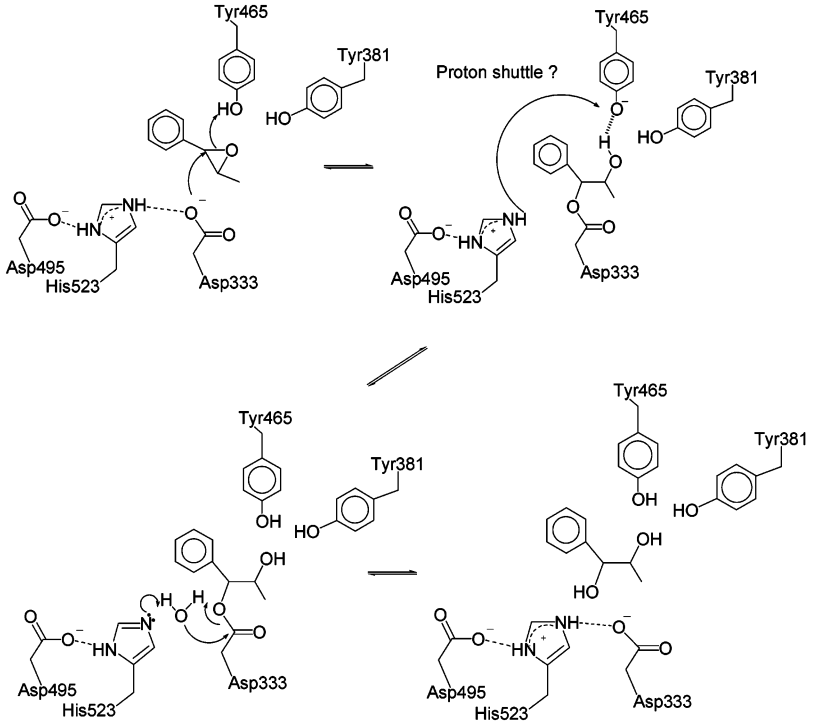
(53) Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D. *Science* **2002**, *295*, 1520−1523.

(54) (a) Billeter, S. R.; Webb, S. P.; Argawal, P. K., Iordanov, T.; Hammes-Schiffer, S. *J. Am. Chem. Soc*. **2001**, *123*, 11 262−11 272. (b) Agarwal, P. K.; Billeter, A. R.; Ravi Rajagopalan, P. T.; Benkovic, S. J.; Hammes-Schiffer, S. *Proc. Natl. Acad. Sci*. **2002**, *99*, 2794−2799.

(55) Meijer, J.; Depierre, J. W. *Eur. J. Biochem*. **1985**, *150*, 7−16.

(56) Guenther, T. M.; Oesch, F. *J. Biol. Chem*. **1983**, *258*, 15054−15061.

Scheme 2



AD1,[15b] and for the hydrolysis reactions carried out by mEHs[20,57] in the sense that a hydrogen bond is present between His523 and Asp495 throughout the catalytic cycle.

The major difference between the mechanism proposed by Argiriadi et al.[18] (Scheme 1) and this revised mechanism (Scheme 2) is, apart from the Asp495−His523 hydrogen bond, that the proton shuttle is found at an earlier stage before the water molecule for the second reaction step becomes activated and not as the last step in the catalytic cycle. This is in accordance with the kinetic data that reveal the hydrolysis step as rate-determining.[58] Further work is in progress to evaluate questions concerning the hydrolysis step of the reaction, e.g., identify proton shuttle pathway and describe the binding of the charged tyrosinate, if present.

Overall, the study has revealed that the job of the enzyme in this hydrolysis reaction compared to the reaction in the gas and solution phases is (*i*) to rotate the phenyl group to open up sterically the electrophilic site of the epoxide, (*ii*) position the nucleophile so only one of the aspartate oxygens is setup for attack. The latter was also found to be a feature of the related haloalkane dehalogenase.[43] The enzyme accomplishes this by making hydrogen bonds from Asp333(OD1) to the Trp334 backbone NH, and between Asp333(OD2) and His523(Nϵ2). These two hydrogen bonds are present throughout the molecular dynamics simulation. And finally, (*iii*) the enzyme most likely plays an important role in controlling the regioselectivity by formation of a hydrogen bond from Tyr465 to the substrate.

The regio- and enantioselctivities of the reaction were studied by Hammock et al. and they found that the 1S,2S-enantiomer of tMSO is more readily hydrolyzed than the enantiopode, 1R,2R.[59] Furthermore, they found that the benzylic carbon atom of the epoxide ring is most likely to be the site of nucleophilic attack of Asp333. Specifically, the regioselectivity of the reaction has been measured for hydrolysis of tMSO by various sources of sEH (mouse, rabbit, and *Aspergillus terreus*) to be between 95% and 100% for attack at the benzylic position of the epoxide ring (C1 of Figure 3).[59] The results of the molecular dynamics simulations presented in this study are in fine agreement with this finding, because among the identified NACs, 75% will lead to attack at the benzylic position. It is important here to remember that NAC formation is a purely ground-state phenomenon—thus, in the Michaelis complex the enzyme is able to distinguish the two potential sites of nucleophilic attack, and set up reactive conformations that will carry on to the reaction pathway with the expected lowest activation energy.[60] Once formed, the NACs associated with attack at epoxide C1 and C2 will most likely not proceed to associated transition states with identical activation energies. To further quantify the regioselectivity of the reaction, calculations are currently being performed to estimate these energy barriers. It is particularly interesting to note that in the *tight* active site the regioselectivity is measured to be 97%, as 1056 NACs are identified for attack at C1, whereas only 35 are found leading to C2 attack. In the *loose* active site, only very limited regioselectivity is measured, because attacks at the two positions are equally likely (470 and 461 identified NACs for C1 and C2 attacks, respectively). We have thus been able to get a hint of what might be the enzymatic secret for controlling the regioselectivity of the reaction. Namely, the formation of a hydrogen bond from Tyr465 to the substrate oxygen will ensure a high regioselectivity. It will be very interesting to gain further insights into the control of the regioselectivity by simulating the mutated enzymes, Y465F and/or Y381F, with bound substrate, to see

(57) Tzeng, H.-F., Laughlin, L. T.; Armstrong, R. N. *Biochemistry* **1998**, *37*, 2905−2911.
(58) Tzeng, H.-F.; Laughlin, L. T.; Lin, S.; Armstrong, R. N. *J. Am. Chem. Soc.* **1996**, *118*, 9436−9437.
(59) Willamson, K. C.; Morrisseau, C.; Maxwell, J. E. Hammock, B. D. *Tetrahedron Asymmetry* **2000**, *11*, 4451−4462.
(60) Lin, B.; Whalen, D. L. *J. Org. Chem.* **1995** , *59*, 1638−1641.

how NAC formation is influenced by this change of the active site structure. Also, it would be very interesting to know experimentally how the regioselectivity of the reaction is influenced by these mutations. Our calculations suggest that Y465F may show a noticeable influence on the regioselectivity because the essential hydrogen bond from Tyr465 to the epoxide oxygen cannot form in this mutant. Interestingly, it was found experimentally that such mutations in fact influences the enantioselectivity of the hydrolysis of styrene oxide and *para*-nitrostyrene oxide catalyzed with *Agrobacterium radiobacter* AD1.[17a] No experimental reports of the effect of mutations on regioselectivity have to the best of our knowledge emerged yet.

The finding of two overall anti-correlated collectively moving domains that structurally are located on opposite sites of the active site in sEH suggests the possibility that the dynamics of sEH can influence the catalytic reaction by e.g., applying compression on the active site. To our knowledge, this is the first time where the effects of enzyme dynamics on catalysis[49] has been outlined to possibly being caused by a anti-correlated domain movement **across** the active site. Further work is in progress to explore deeper the proposal of enzyme dynamics being coupled to catalysis e.g., through anti-correlated domain movements across the active site and to unravel the scope of its transferability to other enzymatic reactions.

Acknowledgment. The work was supported by grants from the National Institute of Health (SK37DK09171-36) and from the Novo Nordisk Foundation. B.S. thanks the Danish Natural Science Research Council for a 3-year Steno-grant and for a travel grant. Computations were made possible through allocations of computer time at UCSB's Origin 2000, which is partially funded by grants from the NSF (CDA96-01954) and Silicon Graphics Inc, at the National Partnership for Advanced Computational Infrastructure, NPACI, at the San Diego Supercomputer Center through NSF cooperative agreement ACI-9619020, and at the Origin 2000 located at Aarhus University, Denmark. Thanks are expressed to E. Y. Lau, K. Kahn, J. Luo, and B. B. Iversen for fruitful discussions and suggestions.

Supporting Information Available: A list with the assignments of the charge state of histidine residues is included together with a residue topology description and added force constants to model the substrate tMSO. Figures illustrating the RMSD of Cα atoms of the four active site complexes as well as a plot of the cross-correlation matrix for the full production time dynamics are provided. The calculation of SCC(*i*,*j*) is outlined. This material is available free of charge via the Internet at http://pubs.acs.org.

JA021021R