JOC Article

Theoretical Investigation of the Origins of Catalysis of a Retro-Diels−Alder Reaction by Antibody 10F11

Andrew G. Leach,*,† K. N. Houk,*,‡ and Jean-Louis Reymond§

AstraZeneca Pharmaceuticals, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG, U.K., Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095-1569, and Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

houk@chem.ucla.edu

Received November 12, 2003

The antibody 10F11 catalyzes a retro-Diels−Alder reaction that forms HNO. Deductions about the mechanism of catalysis were made by Reymond, Baumann et al. from X-ray crystal structures and from kinetic measurements for mutated antibodies. We report a study of these reactions with quantum mechanical methods and a study of the substrate and transition state binding to the active site of the antibody 10F11 using density functional theory and empirical docking methods. We have quantitated the likely contributions to catalysis of three residues identified as possible causes of catalysis: Trp H104, Phe H101, and Ser H100. Trp H104 can make a significant contribution to catalysis through dispersive interactions (π-stacking aromatic−aromatic stabilization). On its own, Phe H101 makes only a small contribution to catalysis. When both aromatic residues are present, they act cooperatively and can make greater contributions to catalysis than expected for each residue alone. Ser H100 and the backbone NH of Phe H101 are expected to act through hydrogen bonding to speed up the reaction, but our calculations suggest that they make only a small contribution to catalysis. Reymond's studies suggest that the hydrogen-bonding network may be mediated through a water molecule in the binding site.

Introduction

Reymond and co-workers recently reported that the catalytic antibody 10F11 accelerates a retro-Diels−Alder reaction.[1] This reaction liberates HNO and is the reverse of a Diels−Alder reaction in which a nitroso compound acts as dienophile. The transition states of the Diels−Alder reaction of HNO and various substituted nitroso compounds with dienes have previously been studied.[2] On the basis of crystallographic studies, Hugot et al. have recently proposed that shape complementarity and an active-site tryptophan residue may make important contributions to catalysis.[3] A series of antibody mutants revealed that an "aromatic sandwich" formed by the tryptophan and a phenylalanine is an important catalytic motif.[4] A hydrogen bond to a binding-site serine hydroxyl was also found to contribute to catalysis.

We have now investigated the antibody catalysis of this reaction, employing quantum mechanical calculations on models designed to represent the functionality present in the binding site of the antibody. A theoretical model for catalysis, a so-called theozyme,[5] can be used to evaluate the magnitude of particular interactions between the substrate or the transition state and an active-site residue. The results show that the active site tryptophan in the absence of the phenylalanine can contribute 1−2 kcal/mol to catalysis. The active-site phenylalanine can contribute <0.5 kcal/mol but may be detrimental to catalysis in the absence of the tryptophan. Together, the tryptophan and phenylalanine may contribute 2−4 kcal/mol to catalysis: these groups act cooperatively as catalysts. The active site serine is calculated to have little catalytic effect.

Background

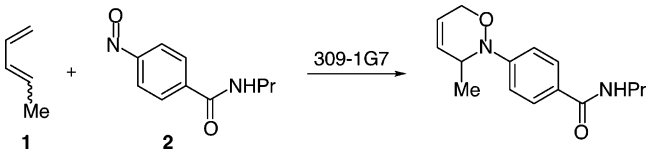
Although some evidence for an enzymatically accelerated reaction has been presented, the Diels−Alder reaction is not a reaction typically utilized by nature.[6] This makes the Diels−Alder reaction an attractive target for the preparation of nonnatural catalysts. Catalytic antibodies that accelerate a range of Diels−Alder reactions have been prepared. Among the catalytic antibodies accelerating Diels−Alder reactions, 1E9[7] is an especially efficient catalyst.[8]

† AstraZeneca Pharmaceuticals.
‡ University of California.
§ University of Bern.

(1) Bensel, N.; Bahr, N.; Reymond, M. T.; Schenkels, C.; Reymond, J.-L. *Helv. Chim. Acta* **1999**, *82*, 44−52.
(2) Leach, A. G.; Houk, K. N. *J. Org. Chem.* **2001**, *66*, 5192.
(3) Hugot, M.; Bensel, N.; Vogel, M.; Reymond, M. T.; Stadler, B.; Reymond, J.-L.; Baumann, U. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9674−9678.
(4) Reymond, J.-L. Unpublished results.
(5) Tantillo, D. J.; Chen, J.; Houk, K. N. *Curr. Opin. Chem. Biol.* **1998**, *2*, 743−750.
(6) (a) Tremblay, M. R.; Dickerson, T. J.; Janda, K. D. *Adv. Synth. Catal.* **2001**, *343*, 577−585. (b) Reymond, J.-L. *Top. Curr. Chem.* **1999**, *200*, 59−93.
(7) Hilvert, D.; Hill, K. W.; Nared, K. D.; Auditor, M.-T. M. *J. Am. Chem. Soc.* **1989**, *111*, 9261−9262.
(8) Kim, S. P.; Leach, A. G.; Houk, K. N. *J. Org. Chem.* **2002**, *67*, 4250−4260.

10.1021/jo035669d CCC: $27.50
Published on Web 05/05/2004

SCHEME 1



1 **2**

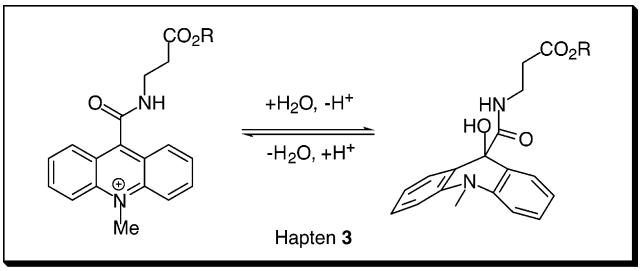
A number of the antibody catalysts for Diels−Alder reactions have been the subject of theoretical studies. These investigations have sought to elucidate the contribution of the various features of these antibodies to catalysis, with the ultimate goal of understanding antibody and enzyme catalysis. For instance, a theoretical investigation of 1E9 employing docking,[9] molecular dynamics, and linear interaction energy calculations found that the antibody has a high degree of shape complementarity to the transition state and provides a hydrophobic pocket with a well placed leucine residue whose motion is considerably damped on going from the reactant complex to the transition state indicating a more favorable enthalpic interaction. An aromatic−aromatic interaction between the transition state and a tryptophan side chain also contributes to catalysis.

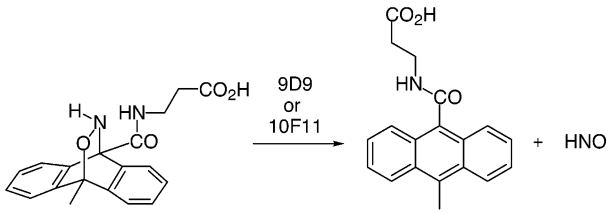
A study of the exo-Diels−Alderase antibody 13G5[10] showed that a number of residues could undergo specific interactions with both substrates and transition states.[11] Endo transition states are not able to bind in such a way as to benefit from the available hydrogen-bonding residues. In one stereochemical arrangement of the exo transition states, one of the residues (Asn L91) had a more stabilizing effect on the substrates than on the transition state and thus would have a decelerating effect on the reaction leading to a specific enantiomer of the product.

In a recent study of the antibody 39A11,[12] a theozyme model was developed,[13] employing model residues to mimic important binding-site residues. Further calculations employing solvation corrections (through use of the CPCM method) showed that the dielectric constant of the medium had a large effect on the endo−exo selectivity with endo being more strongly favored in the gas phase than in water. The antibody, which provides a less polar environment than water, should therefore favor the endo product.

Pandit and co-workers extended this class of antibodies by preparing the antibody 309-1G7 which catalyzes the hetero-Diels−Alder reaction between *cis*- and *trans*-piperylene (**1**) and nitroso compound **2**. These demonstrated respectable levels of acceleration with k_{cat}/k_{uncat} values of 2618 and 1204 for *cis*- and *trans*-piperylene, respectively.[14] A survey of typical binding constants revealed that a typical catalytic antibody will have a k_{cat}/k_{uncat} of ∼1000.[15]

SCHEME 2



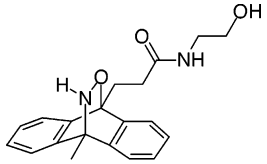


4 **5**

Reymond and co-workers have reported a number of catalytic antibodies which catalyze the retro-Diels−Alder reaction of the Diels−Alder adducts of nitroso compounds. The first antibody in the series, 9D9, was elicited against the hapten, **3**, and catalyzed the retro-Diels−Alder reaction of **4** with a k_{cat}/k_{uncat} of 406.[16] 9D9 was identified by ascertaining its tight binding of the hapten **3**. This indicates that it is likely to bind the transition state of the retro-Diels−Alder reaction if the hapten is a faithful replica of the transition state.

In an extension of this work, Reymond and co-workers exploited the fluorescence of the anthracene product of the reaction, **5**, to screen antibodies directly for catalysis, thereby allowing selection of the antibody with the optimal desired property, catalytic effect, rather than hapten binding.[1] This fluorescence screening permitted the identification of the catalytic antibody 10F11, which has k_{cat}/k_{uncat} equal to 2500. Further kinetic studies of 10F11 provided a range of key data for the antibody. K_M (approximately equal to the dissociation constant for the substrate) is 255 μM, corresponding to a free energy of binding of 4.9 kcal/mol. K_{tx} (dissociation constant for the transition state, equal to $k_{uncat}K_M/k_{cat}$)[17] is 102 nM, which corresponds to a binding free energy of 9.5 kcal/mol, a value that is 4.6 kcal/mol greater than for the substrate.

To investigate the specificity of the antibody, the catalysis of the reaction of the regioisomeric substrate, **6**, was studied. It was found that the catalysis was weaker; $k_{cat}/k_{uncat} = 190$, but the K_M of 107 μM (corresponding to free energy of binding of 5.4 kcal/mol) indicates that substrate **6** is bound more tightly than **4**, while the K_{tx} of 413 nM (a free energy of binding of 8.5 kcal/mol) shows that the transition state is bound less tightly for **6** than **4**, by about 1 kcal/mol.



6

(9) Chen, J.; Deng, Q.; Wang, R.; Houk, K. N.; Hilvert, D. *ChemBioChem* **2000**, *1*, 255−261.

(10) Yli-Kauhaluoma, J. T.; Ashley, J. A.; Lo, C.-H.; Tucker, L.; Wolfe, M. M.; Janda, K. D. *J. Am. Chem. Soc.* **1995**, *117*, 7041−7047.

(11) Cannizzaro, C. E.; Ashley, J. A.; Janda, K. D.; Houk, K. N. *J. Am. Chem. Soc.* **2003**, *125*, 2489−2506.

(12) Braisted, A. C.; Schultz, P. G. *J. Am. Chem. Soc.* **1990**, *112*, 7430−7431.

(13) Zhang, X.; Deng, Q.; Yoo, S.; Houk, K. N. *J. Org. Chem.* **2002**, *67*, 9043−9053.

(14) (a) Meekel, A. A. P.; Resmini, M.; Pandit, U. K. *J. Chem. Soc., Chem. Commun.* **1995**, 571−572. (b) Resmini, M.; Meekel, A. A. P.; Pandit, U. K. *Pure Appl. Chem.* **1996**, *68*, 2025−2028.

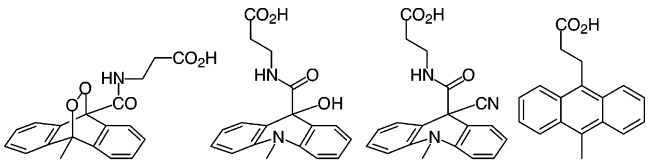


FIGURE 1. Ligands used in crystallization with Antibody 10F11.

Four crystal structures of 10F11 have been reported, with the four species **7**, **8**, **9**, and **10** bound in the active site.[3] Compound **7** was viewed as a substrate mimic, **8** and **9** were considered transition state mimics, and **10** was chosen to mimic **5**, the product of the reaction (Figure 1). Crystallography suggested that the antibody has a higher shape complementarity to the transition state than to either the substrate or the product. A tryptophan residue, H104, was proposed to contribute to catalysis. Antibody 9D9 has a tyrosine at H104, and the binding site has diminished shape complementarity to the transition state.

The most recent study of 10F11 reported the catalysis by a number of mutant antibodies.[4] The initial studies focused on a recombinant antibody that is found to have the Phe L22 exchanged for a Ser in 10F11. The recombinant antibody had similar kinetic parameters to 10F11 and an identical binding site array. This monomutant of 10F11 was considered the "wild type" reference for the mutation studies.

Replacing Trp H104 with other aromatic residues diminished the catalytic effect of the antibody, but replacing Trp H104 with nonaromatic residues almost destroyed the catalysis; the mutant with Ala at H104 showed no measurable catalysis. Substituting Phe H101 with Ala decreased k_{cat} by a factor of 10; K_M shows 2 times poorer substrate binding. These studies show the contribution of these two residues, observed to form an aromatic sandwich in the crystal structures, to catalysis. The crystal structures also revealed that a triangle of hydrogen bonding species may contribute to catalysis: the backbone NH of Phe H101, the side chain OH of Ser H100, and a binding-site water molecule. A mutant in which Ser H100 was changed to Ala showed a 5-fold diminution of catalytic effect. Further studies demonstrated that catalysis could be enhanced by mutating Leu L101 to Phe. This was postulated to improve the catalysis by causing Trp H104 to be pressed more closely against the bound species and by supplementing the "aromatic sandwich". Studies in which the solvent was varied for the uncatalyzed reaction showed that there is little effect of solvent polarity on the retro-Diels−Alder reaction of **4**. The reaction was found to have a similar rate in water and in benzene, suggesting that a nonpolar hydrophobic environment presented by the presence of aromatic rings was not sufficient to speed up the reaction.

A theoretical study of the transition states of the Diels−Alder reactions of nitroso compounds has been reported.[2] Building on this, additional transition-state calculations for structures related to these and a series of docking calculations have been performed.[5]

Computational Methods

B3LYP/6-31G* calculations employing the Gaussian 98[18] suite of programs were used to model the structures of substrates **4** and **6** and the transition states **TS4** and **TS6**. These stationary points were characterized by frequency calculations. ChelpG charges were calculated and used to parametrize the structures for docking studies using AUTODOCK 3.0.[19] Docking calculations employed the genetic algorithm implemented in AUTODOCK 3.0. The crystal structure of antibody 10F11 with the inhibitor **7** bound provided the protein coordinates used in these docking studies. The interaction energy between each ligand and the antibody in each binding orientation was calculated by AUTODOCK. Each docking study generated 120 structures that were clustered into groups of similar binding orientation (as judged by RMSD differences between atomic positions). Initial docking runs with **7** showed that the crystal structure docking was reproduced with AUTODOCK. The docking identified the binding modes available for the two stereoisomers of each substrate and transition state (nitrogen is a chiral center in these computational studies since it does not invert). The structures were docked using both a rigid and a flexible docking procedure. In the flexible procedure, five of the side chain dihedral angles were permitted to vary, and the side chain C−N amide bond was restrained to keep the OCNH dihedral at 180°. Only results for the flexible docking runs will be reported; all rigid docking produced binding modes resembling that adopted by the inhibitor **7**.

Subsequent theozyme calculations employed the model substrate **11** and transition state **TS11** in which the C−N and C−O bonds were constrained to the corresponding lengths in **4** (1.513 and 1.474 Å) and **TS4** (2.023 and 2.250 Å), respectively. Solvation single points employing the CPCM polarizable conductor-like model,[20] IEFPCM (integral equation formalism PCM),[21] and self-consistent isodensity PCM (SCIPCM)[22] methods were performed in Gaussian 03.[23] Geometry optimization including solvation was performed using the IEFPCM method for which analytical frequencies are available. The catalytic effect of Trp H104 was tested with quantum mechanical calculations using structures taken from the docking studies. Here, the Trp was replaced by an indole in the same position, and the substrate and transition state were modeled as **11** and **TS11**. The relative positions of the indole and docked molecule were constrained (by fixing three intermolecular distances),

(15) Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. *Angew. Chem., Int. Ed.* **2003**, *42*, 4872−4897.

(16) Bahr, N.; Güller, R.; Reymond, J.-L.; Lerner, R. A. *J. Am. Chem. Soc.* **1996**, *118*, 3550−3555.

(17) Wolfenden, R.; Snider, M. J. *Acc. Chem. Res.* **2001**, *34*, 938−945.

(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.

(19) AUTODOCK 3.0: Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(20) (a) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995. (b) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669−681.

(21) (a) Cancès, M. T.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032. (b) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 253. (c) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151.

(22) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem.* **1996**, *100*, 16098.

and the remaining structure was optimized with B3LYP/6-31G*. MP2/6-31G* single-point energy evaluations were used because B3LYP/6-31G* is known to neglect dispersion effects that are likely to play a key role in this aromatic−aromatic interaction.[24] MP2 is preferred for modeling weak intermolecular interactions.[25] The catalytic effect of Phe H101 and Ser H100 was estimated in a similar fashion, using the structures obtained from docking, and replacing the serine by *N*-formylethanolamine and constraining three intermolecular distances. The combined catalytic effect of Trp H104 and Phe H101 was estimated in a similar fashion with six constrained distances. For the interactions with aromatic species, PM3 geometry optimization[26] was also employed with the resulting structures being used for MP2 single-point energy evaluations.

All B3LYP, PM3, and MP2 calculations were performed in Gaussian 98.[18]

Results and Discussion

The substrate **4** and transition state **TS4** were optimized using B3LYP/6-31G*. The structures obtained are shown in Figure 2. The regioisomeric substrate **6** and transition state **TS6** were also optimized. These structures show that the transition states for the Diels−Alder reaction with this anthracene-based diene are more synchronous than the corresponding transition states with simple dienes such as butadiene, in which the partial C−N bond is typically 2.0 Å but the C−O bond is typically 2.6−2.7 Å.[2] The activation energy for the retro-Diels−Alder reaction of **4** was calculated to be 24.1 kcal/mol (with no vibrational corrections) and for **6**, 22.5 kcal/mol.

It is known that the retro-Diels−Alder reaction of nitroso compounds can be suppressed by water and can show sensitivity to the nature of the surrounding medium.[27] To test how the polarity of the binding site might influence the rate of the reaction, solvation single-point energy evaluations were performed on the model structures **11** and **TS11** using the CPCM, SCIPCM, and IEFPCM models as implemented in Gaussian 03.[23,28] These structures had previously been optimized with the C−N and C−O bonds constrained to be the same as those

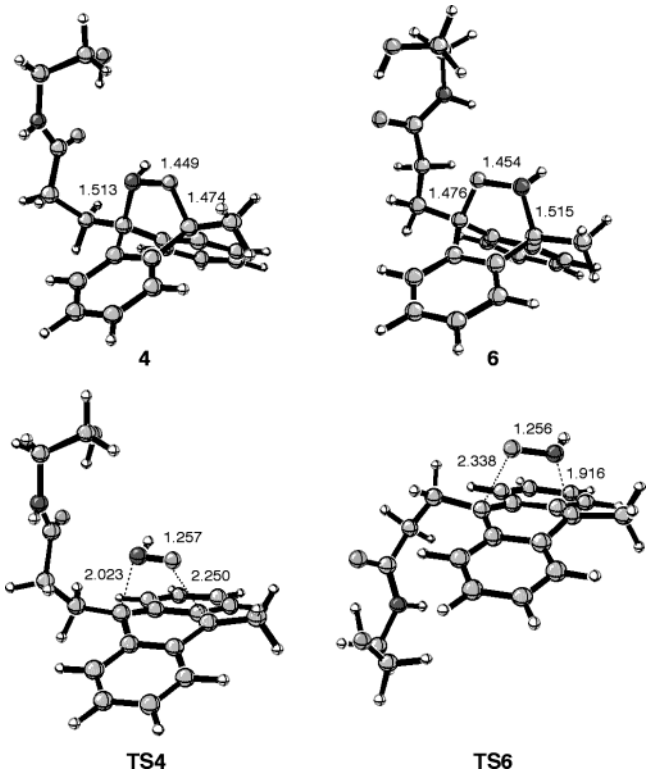


FIGURE 2. B3LYP/6-31G*-optimized structures of **4**, **TS4**, **6**, and **TS6**. Distances are given in Å.

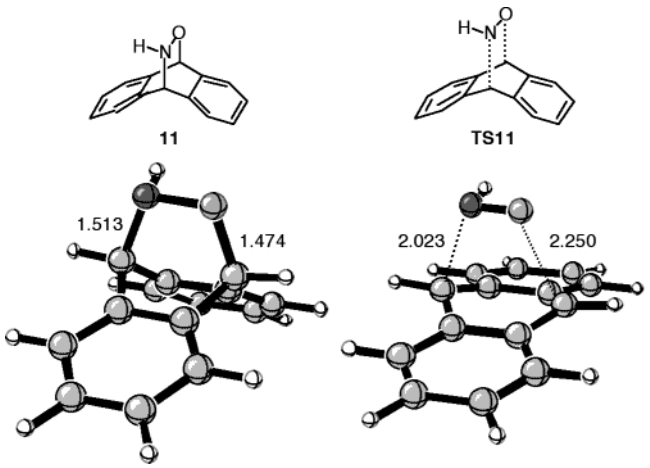


FIGURE 3. Model substrate **11** and transition state **TS11** used in many of the model calculations for the purpose of computational efficiency.

in the optimized structures **4** and **TS4** (Figure 3**)** while the remainder of the molecule was optimized using gas-phase B3LYP calculations. The gas phase ($\epsilon = 1$) barrier (with no zero-point correction included because frequency calculations are impractical once solvation is included) was calculated to be 20.8 kcal/mol. The barrier in water ($\epsilon = 78.1$) was predicted to be from 21.7 to 22.5 kcal/mol based on solvation single-point calculations (Table 1, column 2). Further studies employed the same structure as **11** but without the constraint. In this case, the C−N distance was 1.510 in the substrate and 1.823 Å in the transition state. The C−O distance was 1.466 and 2.438 Å in the same structures. This gave a gas-phase activa-

(23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.2; Gaussian, Inc.: Pittsburgh, PA, 2003.

(24) (a) Hobza, P.; Sponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315. (b) Ujaque, H.; Lee, P. S.; Houk, K. N.; Hentemann, M. F.; Danishefsky, S. J. *Chem. Eur. J.* **2002**, *8*, 3423.

(25) For recent examples, see: (a) Hobza, P.; Riehn, C.; Weichert, A.; Brutschy, B. *Chem. Phys.* **2002**, *283*, 331−339. (b) Wang, Y. L.; Hu, X. C. *J. Am. Chem. Soc.* **2002**, *124*, 8445−8451. (c) Kovács, A.; Szabó, A.; Nemcsok, D.; Hargittai, I. *J. Phys. Chem. A* **2002**, *106*, 5671−5678. (d) Raymo, F. M.; Bartberger, M. D.; Houk, K. N.; Stoddart, J. F. *J. Am. Chem. Soc.* **2001**, *123*, 9264−9267.

(26) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220.

(27) Wijnen, J. W.; Engberts, J. B. F. N. *Liebigs Ann.* **1997**, 1085−1088.

(28) It was found that different CPCM energies are obtained for the default cavity size implemented in Gaussian 98.

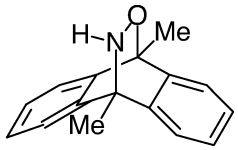
TABLE 1. Activation Barriers with No Vibrational Corrections (in kcal/mol) for the Retro-Diels−Alder Reaction of 11[a]

method	gas-phase model structures including C−N and C−O distances constrained to the same as in full structures **4** and **TS4**	gas-phase model structures optimized with no constraints	aqueous phase structures optimized with no constraints using the IEFPCM method
gas phase	20.8	20.5	
CPCM	22.5	21.6	21.3
SCIPCM	21.7	21.0	21.0
IEFPCM	22.4	21.6	21.4

[a] Three different geometries were employed: a gas-phase geometry in which the C−O and C−N distances in **11** and **TS11** were constrained to be the same as in **4** and **TS4**, a fully optimized gas-phase equivalent of **11** and **TS11**, and an aqueous phase optimized equivalent of **11** and **TS11**. Single-point energies using three different solvation models are listed.

tion barrier of 20.8 kcal/mol and a barrier in water of 21.0−21.6 kcal/mol (Table 1, column 3). Finally, the same structures were optimized to an aqueous phase structure using the IEFPCM model. This method predicted an activation barrier of 21.4 kcal/mol. CPCM and SCIPCM single-point calculations on the solvent optimized structures predicted almost identical activation barriers (Table 1, column 4). After solvation optimizations, the C−N distance in the transition state was 1.839 Å and the C−O distance 2.547 Å. This is more asynchronous than the gas-phase transition state. The more polar asynchronous transition state is stabilized by the high dielectric constant of the aqueous phase. The C−N and C−O distances in the substrate change much less upon aqueous optimization, they are 1.511 and 1.473 Å in solution and 1.510 and 1.466 Å in the gas phase.

The antibody hydrophobic binding pocket will have dielectric constant much less than that in the surrounding water. Values around $\epsilon = 4$ are often measured or proposed.[29] Calculations indicate a negligible 0.5−1.0 kcal/mol reduction of activation energy by attenuating the dielectric constant from 78.5 to 4. In agreement with the small calculated effect of dielectric constant, the rate of decomposition of **12** in $CDCl_3$ was found by Reymond et al. to be almost identical to that of both **4** and **6** in aqueous buffer (4×10^{-6} s^{-1} for **12**, 9.8×10^{-6} s^{-1} for **4**, and 7.1×10^{-6} s^{-1} for **6**). Reymond et al.[4] attribute the lack of solvent effect in this case to the irreversible nature of the reaction that contrasts with the reversibility of the reaction of the cyclopentadiene adduct studied by Engberts et al.[27]



12

The B3LYP structures **4** and **TS4** were also used for a docking study. The crystal structure obtained with the inhibitor **7** was used to provide coordinates for the antibody, because its geometry was closest to that of both the substrate, **4**, and transition state, **TS4**. The docking calculations revealed that the two stereoisomeric substrates *R*-**4** and *S*-**4** (nitrogen is a chiral center in these computational structures) and transition states *R*-**TS4** and *S*-**TS4** bind in well-defined ways that resemble the binding modes adopted by the inhibitor **7** in the crystal structure (Figure 4). The binding mode shown for each species in Figure 4 accounted for a large number of the docked structures obtained from a docking run with 120 distinct starting points. In each case, the structure shown was also that calculated to have the best binding energy apart from for the *S* stereoisomers, in which it was the second best. In these cases, the lowest energy structure was located infrequently and had the side chain protruding deeper into the protein rather than out toward solution. The docking statistics are summarized in the first part of Table 2 and clearly show that the two transition states adopt very well-defined binding modes. By contrast, the substrates bind rather more flexibly and weakly. Eight of the other clusters located during these docking runs correspond to variations in the conformation of the side chain, and three (including that already mentioned) involve the docking of the molecule with the side chain buried in the protein. All of these binding modes may play a part in the binding of these various structures with the antibody, but the structures shown in Figure 4 are those likely to be the productive complexes involved in catalysis.

The regioisomeric substrates *R*-**6** and *S*-**6** and the corresponding transition states, also optimized with B3LYP/6-31G*, were docked into the antibody structure in the same way described above. The binding modes, illustrated in Figure 5, were selected by the same means as before. The docking statistics for these structures show that the docking is less specific and that docking of both substrates and transition states for these regioisomers is weaker than for the regioisomer known to be experimentally favored (**4**). The alternative clusters in this case comprise different side chain conformations, inverted structures with buried side chains and also structures in which a rotation by 120° has taken place such that the NO bridge is interchanged with one of the aromatic rings. In these rotated structures, the NO bridge is oriented either toward Trp H104 or toward Ser H35 and Ser H50.

Docking shows, in agreement with experimental observation, that the transition states, *R*- and *S*-**TS4** bind most tightly and in well-defined binding modes. The corresponding substrates *R*- and *S*-**4** also bind well in similar binding modes. The regioisomeric transition states and substrates *R*- and *S*-**TS6** and **6** on the other hand are calculated to bind less strongly and show a range of alternative binding modes; here the two transition states bind more strongly than the substrates, and therefore, catalysis should still be expected. The two

(29) See, for example: (a) Muegge, I.; Schweins, T.; Warshel, A. *Proteins: Struct., Func., Genet.* **1998**, *30*, 407−423.
(30) Olson, M. A.; Reinke, L. T. *Proteins: Struct., Func., Genet.* **2000**, *38*, 115−119.

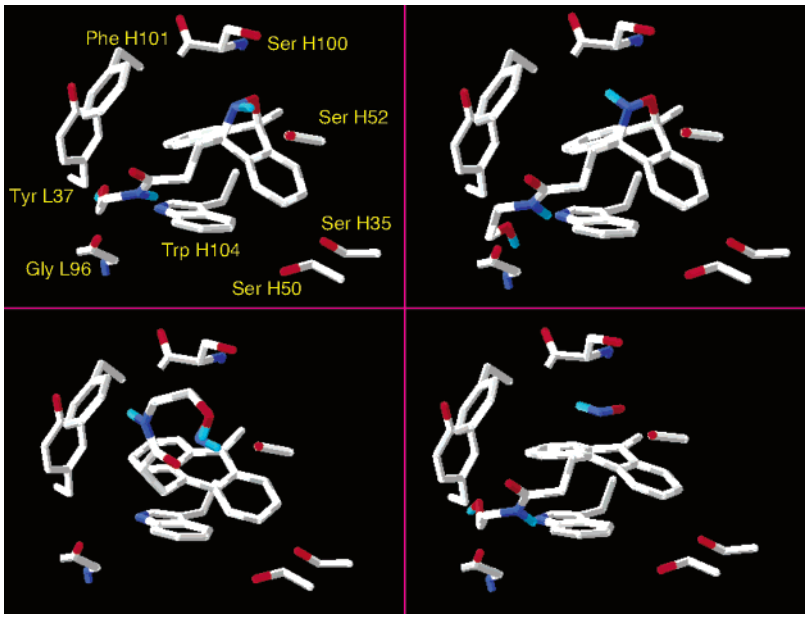


FIGURE 4. Binding modes for *R*-**4** (top left), *R*-**TS4** (bottom left), *S*-**4** (top right), and *S*-**TS4** (bottom right) obtained by AUTODOCK. The structure shown was selected on the basis of high frequency of location by AUTODOCK and by having the best, or second best (*S* stereoisomers), calculated binding free energy.

TABLE 2. Docking Statistics for AUTODOCK Docking of Substrates and Transition States Corresponding to Substrate 4 (Top Half of Table) and Substrate 6 (Bottom Half of Table)

docked structure	binding free energy (best for that cluster) for clusters corresponding to the structure shown in Figure 4 or 5[a]	binding free energy and frequency in parentheses for other clusters found during docking (with docked energies[b] within 1 kcal/mol of the structure illustrated in Figure 4 or 5, listed in order of docked energy)
R-**4**	−13.1 (46)	−12.9 (10), −12.9 (23), −12.5 (2), −12.5 (32), −11.9 (7)
R-**TS4**	−13.8 (117)	-
S-**4**	−13.1 (87)	−13.2 (1), −12.2 (13), −12.3 (11)
S-**TS4**	−13.9 (114)	−14.8 (2), −13.6 (1), −13.5 (2)
R-**6**	−12.2 (15)	−11.9 (3), −11.7 (1), −12.0 (32), −11.5 (3), −11.4 (17), −11.2 (38), −11.5 (1)
R-**TS6**	−12.0 (104)	−12.0 (1), −11.5 (11), −11.6 (2)
S-**6**	−12.1 (22)	−11.5 (1), −11.8 (40), −11.4 (36), −11.2 (1), −12.0 (1), −11.5 (6), −11.1 (6), −10.9 (1)
S-**TS6**	−13.3 (112)	−13.2 (1), −13.0 (1)

[a] Frequency (out of 120 clusters per run) is shown in parentheses. [b] The docked energy neglects the penalty of conformational changes in the ligand and refers only to protein−ligand interaction energies.

values of K_M for the regioisomeric substrates **4** and **6** suggest that **6** should bind more strongly than **4** in contrast to docking calculations. The error in the docking energies calculated by AUTODOCK is such that the values are only correct to within a few kcal/mol. Consequently, although some trends are suggested by these values, quantitative comparisons cannot be made.

The docked structures clearly reveal that two residues make intimate contact with the substrates and transition states: Trp H104 and Ser H100. Reymond et al.'s mutation studies highlight the importance of a second aromatic residue Phe H101.[4] Evaluations of the possible contributions to catalysis by these three residues were made. First, structures obtained from the docking studies were excised, and constrained optimizations were performed using B3LYP/6-31G*. MP2/6-31G* single-point calculations were subsequently performed on the optimized structures, because MP2 is preferred over B3LYP for modeling weak intermolecular interactions.[25] The Trp H104 side-chain was modeled with indole. The Phe H101 side chain was modeled with benzene. For Ser H100, the residue was modeled as *N*-formylethanolamine. This models not only the side chain hydroxyl of Ser H100 but also the backbone NH of the adjacent Phe H101, both of which are proposed to be contributors to catalysis.[4] In all cases, three intermolecular distances were fixed to maintain the amino acid model and the ligand in approximately the orientation present in the binding site. The substrate and transition state were also simplified to the model compounds **11** and **TS11**. The combined effect of Trp H104 and Phe H101 was probed in a study in which both model side chains were used and six intermolecular distances constrained. The known weakness of B3LYP for modeling intermolecular interactions, particularly aromatic−aromatic interactions,[24] prompted a second geometry optimization method to be employed for the aromatic residues, within the constraints of computational time. The semiempirical PM3 method was used with the same intermolecular constraints. MP2/6-31G* single-point energies were obtained for these PM3 geometries, and the energetics thus calculated were compared to those with B3LYP geometries. B3LYP/6-

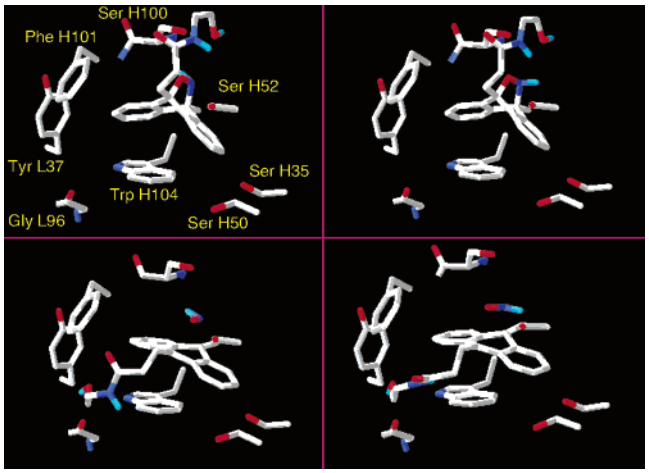


FIGURE 5. Binding modes for *R*-**6** (top left), *R*-**TS6** (bottom left), *S*-**6** (top right), and *S*-**TS6** (bottom right) obtained by AUTODOCK. The structure shown was selected on the basis of being located with high frequency by AUTODOCK and by having the best, or second best (*R*-**TS6**), calculated binding free energy.

31G* should provide reasonable values for the polar intermolecular interaction with *N*-formylethanolamine, and so in this case, only one geometry optimization method was employed; both MP2 single-point energies and B3LYP energies are reported and analyzed. The B3LYP energies and PM3 values of the interaction energies are reported in the Supporting Information. The optimized structures and distances that were used as constraints are shown in Figure 6 for the indole complexes, Figure 7 for the benzene complexes, Figure 8 for the complexes with both indole and benzene, and Figure 9 for the *N*-formylethanolamine complexes. By way of reference, the solvation energies of the substrate computed by CPCM, SCIPCM, and IEFPCM are 12.0, 5.9, and 11.9 kcal/mol, respectively, and those for the transition state are 11.0, 5.4, and 12.8 kcal/mol.

All species are computed to undergo attractive interactions with each amino acid model. The key to whether a particular residue is expected to be catalytic is that the transition state **TS11** should be stabilized more than the substrate **11** by interaction with the model side chain. In each of the following tables, the energy of interaction of **11** and **TS11** with each model side chain is listed, along with the amount by which the transition state is stabilized more than the substrate, $\Delta\Delta E$, such that positive values indicate a residue computed to be catalytic and negative values a residue that is anticatalytic.

Table 3 lists the computed energies of interaction with indole modeling Trp H104. This shows that both stereoisomers are computed to undergo catalytic interactions with this residue. B3LYP and PM3 geometries both give energies of interaction which suggest a contribution to catalysis of ∼1 to 2 kcal/mol for Trp H104 in the absence of any other effects. This is a significant proportion of the 4.6 kcal/mol catalysis found experimentally. As the substrate flattens on moving toward the transition state, the aromatic surface is better able to experience face-face interactions with the indole of Trp H104; enhanced dispersion and charge-transfer interactions result.

Table 4 lists the corresponding energies computed for the benzene model for Phe H101. This shows that the *R*

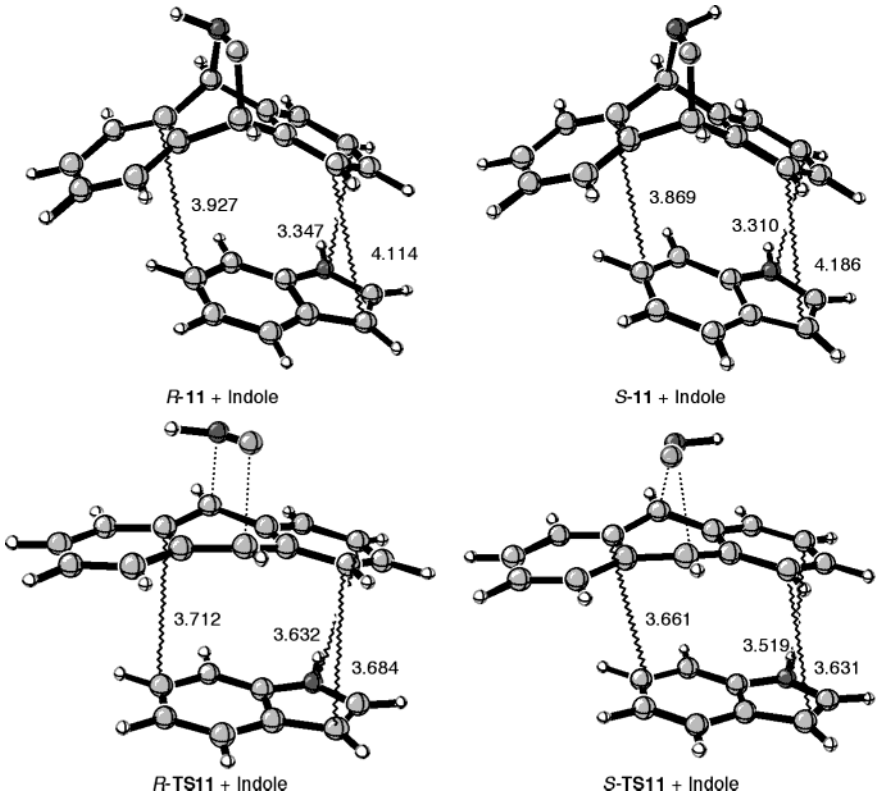


FIGURE 6. Complexes of *R*-**11**, *R*-**TS11**, *S*-**11**, and *S*-**TS11** with indole, optimized with constrained B3LYP/6-31G* calculations. The distances indicated are those that were used as constraints during the optimizations and were taken from the structures obtained by AUTODOCK (Figure 4).

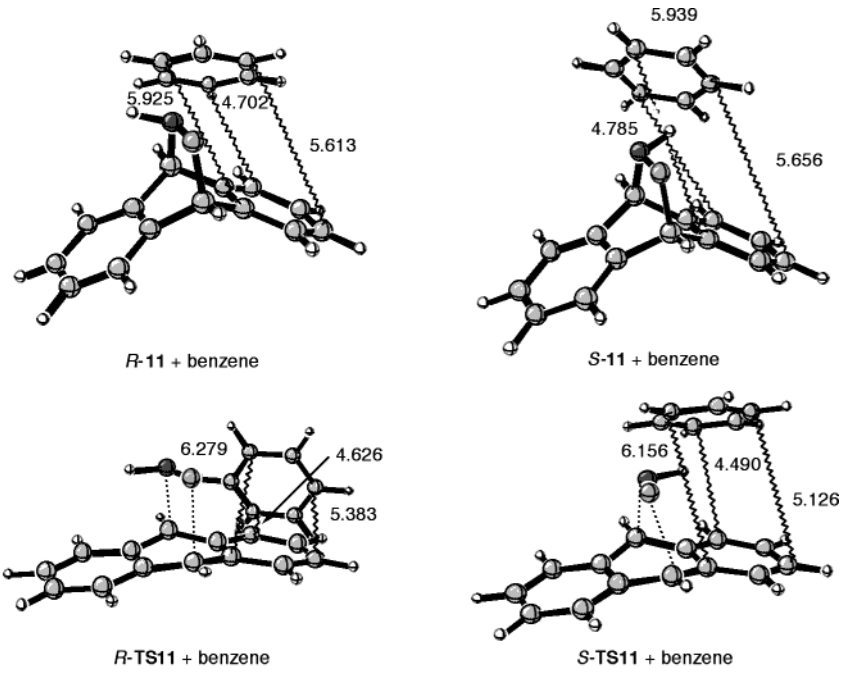


FIGURE 7. Complexes of *R*-**11**, *R*-**TS11**, *S*-**11**, and *S*-**TS11** with benzene, optimized with constrained B3LYP/6-31G* calculations. The distances indicated are those that were used as constraints during the optimizations and were taken from the structures obtained by AUTODOCK (Figure 4).

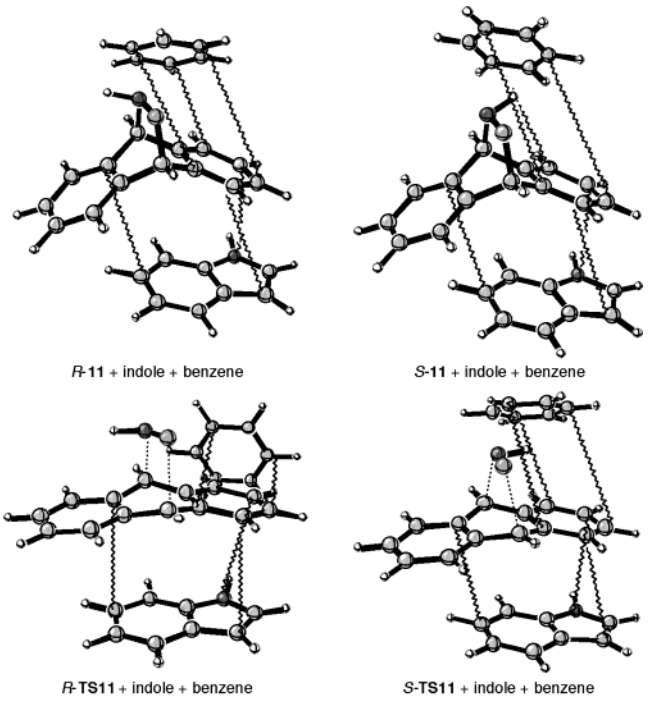


FIGURE 8. Complexes of *R*-**11**, *R*-**TS11**, *S*-**11**, and *S*-**TS11** with indole and benzene, optimized with constrained B3LYP/6-31G* calculations. The interatomic distances indicated are those that were used as constraints during the optimizations and were taken from the structures obtained by AUTODOCK (Figure 4), the value of each constrained distance is the same as shown in Figure 6 (distances to indole) and Figure 7 (distances to benzene).

stereoisomer is computed to undergo slightly anticatalytic interactions, while the *S* stereoisomer experiences slightly catalytic interactions. B3LYP and PM3 geometries yield differences in the scale of these contributions but suggest that at best Phe H101 contributes ~0.5 kcal/mol to catalysis in the absence of all other residues.

In Table 5, the effect of indole modeling Trp H104 interacting with **11** and **TS11** in the presence of a benzene model for Phe H101 is shown. In the case of the *R* stereoisomer, the sum of the individual interactions of the two models (from Tables 3 and 4) should be −0.3 and 0.0 kcal/mol according to the B3LYP and PM3 geometries, respectively. However, Table 5 shows that when both indole and benzene are present, the total interaction contributes +0.3 or +2.4 kcal/mol according to the two methods. These values are 0.6 and 2.4 kcal/mol higher than the sum of the individual interactions and suggest a cooperative effect of the two aromatic residues in the presence of one another. In the case of the *S* stereoisomer, the sum of the individual interactions is +1.4 and +2.4 kcal/mol for B3LYP and PM3 geometries, respectively. The values for the two model aromatic compounds in the presence of one another are 0.6 and 1.8 kcal/mol higher than these individual contributions, again showing a cooperative effect of these two residues. The cooperative effect of an "aromatic sandwich" is therefore computed to be between ~0.5 and 2 kcal/mol. Although Trp H104 may contribute 1−2 kcal/mol to catalysis and Phe H101 ~0.5 kcal/mol, the two together may contribute up to 4 kcal/mol and likely contribute around 3 kcal/mol. There is a cooperative effect of having both aromatic residues present and this may be extended further by adding more aromatic residues, as seen experimentally for the antibody in which Leu L101 is exchanged for Phe. The computations with benzene show that in general such nonpolar aromatic rings will have only a weak effect and are as likely to retard the reaction as accelerate it, and hence, the reaction is not observed to be faster in benzene than in

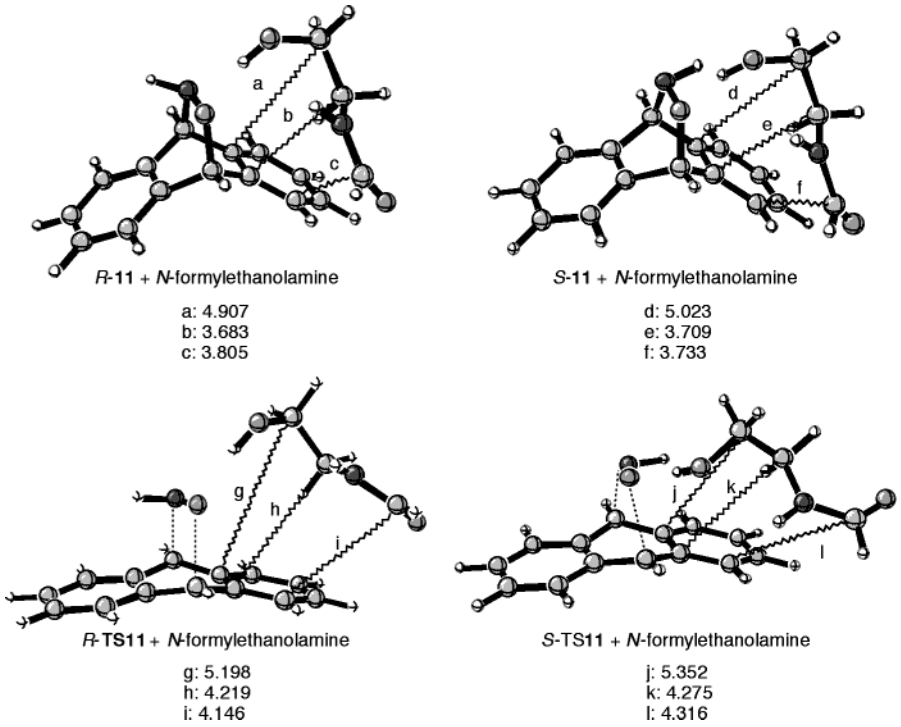


FIGURE 9. Complexes of *R*-**11**, *R*-**TS11**, *S*-**11**, and *S*-**TS11** with *N*-formylethanolamine optimized with constrained B3LYP/6-31G* calculations. The distances indicated are those that were used as constraints during the optimizations and were taken from the structures obtained by AUTODOCK (Figure 4).

TABLE 3. Interaction Energies between Indole and Both Model Substrate 11 and Model Transition State TS11 (Both Stereoisomers)[a]

structure	B3LYP	PM3
R-**11** + indole	7.7	6.7
R-**TS11** + indole	9.4	7.4
ΔΔE (*R*)	+1.7	+0.7
S-**11** + indole	8.0	6.5
S-**TS11** + indole	9.2	8.2
ΔΔ*E* (*S*)	+1.2	+1.7

[a] These are MP2/6-31G* single-point energies (in kcal/mol) on constrained B3LYP/6-31G* (second column) and PM3 (third column) geometries. ΔΔ*E* values indicate the extra stabilization of the transition state relative to the substrate, positive values indicate a catalytic interaction, and negative values an anticatalytic interaction.

TABLE 4. Interaction Energies between Benzene and Both Model Substrate 11 and Model Transition State TS11 (Both Stereoisomers)[a]

structure	B3LYP	PM3
R-**11** + benzene	5.1	3.5
R-**TS11** + benzene	3.1	2.8
ΔΔ*E* (*R*)	−2.0	−0.7
S-**11** + benzene	4.8	2.6
S-**TS11** + benzene	5.0	3.3
ΔΔ*E* (*S*)	+0.2	+0.7

[a] These are MP2/6-31G* single-point energies (in kcal/mol) on constrained B3LYP/6-31G* (second column) and PM3 (third column) geometries. ΔΔ*E* values indicate the extra stabilization of the transition state relative to the substrate, positive values indicate a catalytic interaction, and negative values an anticatalytic interaction.

nonaromatic solvents. Catalysis requires correctly positioned aromatic residues.

TABLE 5. Interaction Energies between Both Indole and Benzene and Both Model Substrate 11 and Model Transition State TS11 (Both Stereoisomers)[a]

structure	B3LYP	PM3
R-**11** + indole + benzene	12.9	11.4
R-**TS11** + indole + benzene	13.1	13.7
ΔΔ*E* (*R*)	+0.3	+2.4
S-**11** + indole + benzene	12.8	10.4
S-**TS11** + indole + benzene	14.8	14.6
ΔΔ*E* (*S*)	+2.0	+4.2

[a] These are MP2/6-31G* single-point energies (in kcal/mol) on constrained B3LYP/6-31G* (second column) and PM3 (third column) geometries. ΔΔ*E* values indicate the extra stabilization of the transition state relative to the substrate, positive values indicate a catalytic interaction, and negative values an anticatalytic interaction.

TABLE 6. Interaction Energies between *N*-Formylethanolamine and Both Model Substrate 11 and Model Transition State TS11 (Both Stereoisomers)[a]

structure	MP2	B3LYP
R-**11** + *N*-formylethanolamine	13.3	8.4
R-**TS11** + *N*-formylethanolamine	11.7	8.1
ΔΔ*E* (*R*)	−1.6	−0.3
S-**11** + *N*-formylethanolamine	12.0	7.2
S-**TS11** + *N*-formylethanolamine	11.4	7.7
ΔΔ*E* (*S*)	−0.7	+0.5

[a] These are MP2/6-31G* single-point energies (in kcal/mol) on constrained B3LYP/6-31G* geometries (second column) and B3LYP/6-31G* energies on the same geometries (third column). ΔΔ*E* values indicate the extra stabilization of the transition state relative to the substrate, positive values indicate a catalytic interaction, and negative values an anticatalytic interaction.

Table 6 shows the computed energies of interaction with *N*-formylethanolamine which models both the side chain hydroxyl of Ser H100 and the backbone amide NH

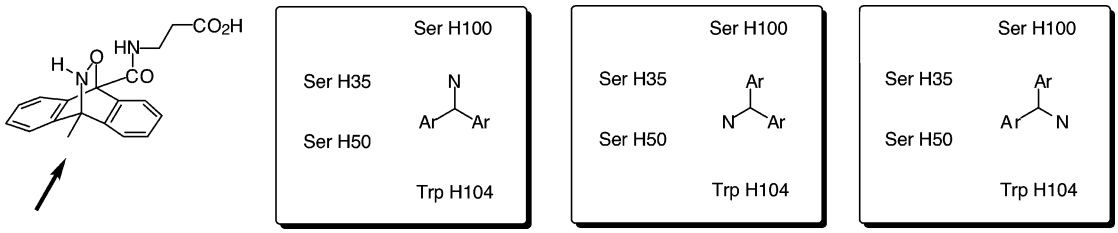


FIGURE 10. Schematic representations of alternative rotated binding orientations found in docking studies for substrate **6** in both stereoisomers. The rotated structures are represented as the view looking down the arrow shown in the left-hand picture of the substrate.

of Phe H101. The table shows that this interaction can be anticatalytic. In fact, only the B3LYP//B3LYP calculation for the *S* stereoisomer has a calculated catalytic effect, of only 0.5 kcal/mol. In every other case, a small anticatalytic interaction is computed, in line with ChelpG charges which show that the oxygen of the departing HNO bears a larger negative charge in the substrate (−0.22) than in the transition state (−0.15) and may therefore be expected to undergo anticatalytic hydrogen-bonding interactions. These results are in contradiction to the experimental observation that mutating Ser H100 to Ala leads to a 5-fold decrease in k_{cat} which corresponds to a decrease of nearly 1 kcal/mol in the $\Delta\Delta E$ value. It is likely that B3LYP energies are more appropriate for computing this interaction and that the *S* stereoisomer is the form through which the reaction proceeds (it is the preferred stereochemistry with the aromatic residues described above). The interaction may require a triangular hydrogen bonding network including a binding site water molecule, as suggested by Reymond et al.[4] The uncertainty in placement of hydrogen atoms in the crystal structure makes for some uncertainty in the calculations here, particularly for the side-chain hydroxyl, but would make the computation of the effect including the extra two hydrogens of the water molecule very difficult to carry out in a satisfactory fashion.

While the reactive regioisomer **4** docks preferentially in one orientation, the regioisomeric substrate **6** binds in a range of orientations (Table 2 and Figure 10). This makes substrate binding entropically more favorable and diminishes catalysis since **TS6** still prefers one docking mode. Several of these substrate binding modes are less catalytically active than the one orientation accessed by **4**, due to the loss of the favorable interactions of Trp H104, Phe H101, and Ser H100 and the entry of more water into the binding site. In the rotated complexes, the "aromatic sandwich" of Trp H104 and Phe H101 is not positioned well to undergo catalytic interactions with the π system of the substrate and transition state. Therefore, although substrate binding may increase due to the enhanced flexibility of binding, this is not catalytically important as many of these binding modes do not benefit from the catalytic effect afforded by the various residues described above.

Conclusions

The theoretical explorations here, combined with the extensive structural and kinetic studies by Reymond et al., indicate that antibody 10F11 catalyzes the retro-Diels−Alder reaction of **4** by several cooperative effects. The antibody provides a tryptophan side chain that undergoes a more stabilizing interaction with the transition state of the reaction than the substrate, due in part to the flattening of the molecule along the reaction coordinate. A phenylalanine side chain undergoes interactions that on their own make only a small contribution to catalysis. When these two aromatic residues act together, there is a cooperative effect which enhances the catalytic effect by up to 2 kcal/mol. A serine side chain hydroxyl and phenylalanine backbone NH interact with the substrate and each other in such a way as to be only weakly catalytic, thanks to the more polar nature of the substrate than the transition state. This interaction may be mediated by a binding site water molecule which was not included in these calculations and may be a key part of a triangular hydrogen bonding network which must all be present for catalysis to be effective. Generally, the binding modes of the *S* stereoisomer are computed to be the catalytically most effective and to have the higher binding energy and it may be that, even though this stereoselectivity is not observable in either the product or substrate due to rapid inversion at nitrogen, the antibody nevertheless acts in a stereoselective manner. The less polar nature of the binding pocket compared to water may make a modest contribution to catalysis. The regioisomeric substrate binds in a range of orientations, some of which are not primed for catalysis such that stronger substrate binding and weaker transition state binding ensues. This agrees with the qualitative conclusions from an analysis of crystal structures, and with the quantitative conclusions presented from kinetic studies of mutated antibodies.

Acknowledgment. We are grateful to the National Institute of General Medical Sciences, National Institutes of Health, for financial support (K.N.H.) and to the National Computational Science Alliance for computer time. We thank AstraZeneca and the UK Fulbright commission for a fellowship to A.G.L.

Supporting Information Available: Tables of interaction energies for substrates with models for binding-site side chains and Cartesian coordinates of all optimized geometries. This material is available free of charge via the Internet at http://pubs.acs.org.

JO035669D