Control of the *exo* and *endo* Pathways of the Diels-Alder Reaction by Antibody Catalysis

Veronique E. Gouverneur, K. N. Houk, Beatriz de Pascual-Teresa, Brett Beno, Kim D. Janda, Richard A. Lerner

Catalytic antibodies that control the reaction pathways of the Diels-Alder cycloaddition have been generated. One antibody catalyzes the favored *endo* and the other the disfavored *exo* pathway to yield the respective *cis* and *trans* adducts in enantiomerically pure form. A comparison of the x-ray structure of the hapten with the calculated geometry of the transition structure showed that [2.2.2] bicyclic compounds are excellent mimics of the transition state of the Diels-Alder reaction. To achieve catalysis and the high degree of stereoselectivity shown here, the antibody must simultaneously control the conformation of the individual reactants and their relation to each other. In the case of the disfavored process, binding energy must be used to reroute the reaction along a higher energy pathway. The rerouting of reaction pathways has become a major focus of antibody catalysis and other disfavored reactions can be expected to be catalyzed so long as the energy barrier is not extreme. The energy requirements needed for absolute control of all of the stereoisomers of many Diels-Alder reactions fall in the energy range (~20 kilocalories per mole) deliverable by antibody binding.

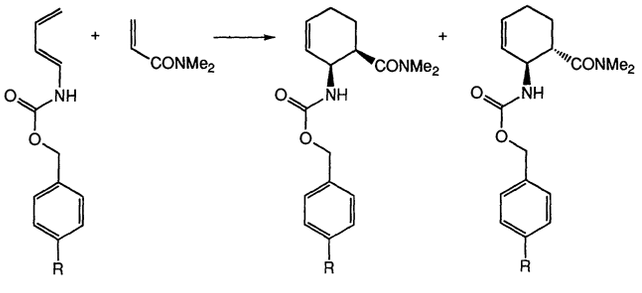
The Diels-Alder transformation is often said to be the most important carbon-carbon bond-forming process in chemistry (*1*). This transformation formally belongs to the larger class of pericyclic reactions and is a [4π + 2π] cycloaddition of a diene and dienophile. In view of its central role in the theory and practice of organic chemistry, there has been much interest in the development of catalytic methods for improving its rate or selectivity or both (*2*). For many applications, the value of the Diels-Alder reaction relies only upon the degree of stereocontrol that can be exercised. Currently, the requirements are stringent in that the absolute as well as the relative configuration must be considered for most synthetic pathways where the Diels-Alder reaction is used as a key step. In the cycloaddition between a monosubstituted diene and dienophile, up to eight isomers can be formed. In order to secure only one adduct out of the eight possible isomers, the reaction must be regio-, diastereo-, and enantioselective. Extensive efforts have focused on control of the *endo/exo* and absolute configuration of the reaction products (*2*). The most difficult problem yet to be solved concerns control of the Diels-Alder reaction pathways to yield otherwise disfavored products in enantiomerically pure form (*1, 2*).

Antibodies, with their specific and chiral binding pockets, should be ideal catalysts for completely controlling the stereochemical outcome of the Diels-Alder cycloaddition. However, it is not only necessary to catalyze the reaction and control its stereochemistry, but disfavored reaction pathways must be simultaneously facilitated. Already, researchers have generated catalytic antibodies for the Diels-Alder reaction (*3–5*), and recently we showed that antibodies can catalyze a disfavored chemical process (6). Here, we combine these elements to generate a pair of antibodies where one catalyzes the formation of the favored *endo* and the other the disfavored *exo* Diels-Alder adduct. In each case, the catalyzed reaction proceeds with precise control of the absolute stereochemistry.

The reactions we studied were the cycloadditions between *trans*-1-*N*-acylamino-1,3-butadiene **1a** and **1b** (*7*) and *N*,*N*-dimethylacrylamide (Fig. 1). The resulting adducts belong to a family of compounds useful as synthetic intermediates in the total synthesis of alkaloids (8). Without a catalyst, diene **1a** cleanly undergoes cycloaddition with dimethylacrylamide at 110°C in toluene to afford a mixture of the stereoisomeric *endo* (*cis*) and *exo* (*trans*) adducts **2a** and **3a** in a ratio of 66:34. The same reaction starting with the water-soluble diene **1b** (9), when performed in phosphate-buffered saline (PBS) at 37°C, gives an 85:15 mixture of the *endo* adduct **2b** and the *exo* adduct **3b** (*10*). Both reactions are completely regioselective, giving only the "*ortho*" regioisomers (*11*).

Features of the transition state. To design haptens that will induce catalytic antibodies, one needs an appreciation of the critical stereoelectronic features of the transition state of the reaction to be catalyzed. We used ab initio molecular orbital calculations to locate the transition structures for the Diels-Alder reaction between acrylamide and *N*-(1-butadienyl)-carbamic acid. Calculations were executed with the restricted Hartree-Fock method with the GAUSSIAN 92 program (*12*).

A conformational study of the reactants was carried out. For acrylamide, two possible conformers for rotation around "a" (Fig. 2) were calculated. At the RHF/6-31G* level there is a 1.9 kcal/mol preference for the *s-cis*-acrylamide. This preference for a *cis* conformation is in agreement with previous ab initio studies on acrolein and acrolein derivatives (*13*). Steric effects favor this geometry, and it should be even more highly favored in the transition state because it has a lower lowest unoccupied molecular orbital (LUMO) energy than the *s-trans* conformer. The alkene, *N*-vinylcarbamic acid (Fig. 2), was then studied as a model for the diene. The four possible conformations for the rotation around "a" and "b" were studied. There is a 2 to 5



1a R = COOMe
1b R = COO⁻Na⁺

2a R = COOMe
2b R = COO⁻Na⁺

3a R = COOMe
3b R = COO⁻Na⁺

Fig. 1. Diels-Alder reaction selected for the control of the diastereo- and enantioselectivity by antibody catalysis.

V. E. Gouverneur, K. D. Janda, and R. A. Lerner are in the Departments of Molecular Biology and Chemistry, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037. K. N. Houk, B. de Pascual-Teresa, and B. Beno are in the Department of Chemistry and Biochemistry, University of California–Los Angeles, 405 Hilgard, Los Angeles, CA 90024.

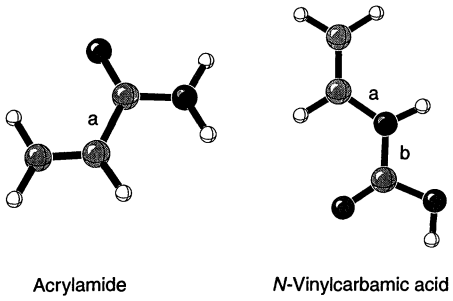


Fig. 2. Compounds used in model studies.

kcal/mol preference at the RHF/6-31G* level for the conformation shown in Fig. 2, relative to the other conformers.

Transition structures for the reaction of acrylamide and *N*-(1-butadienyl)-carbamic acid in the lowest energy conformations were located at the RHF/3-21G level. Four transition structures were considered, corresponding to the four possible stereo- and regioisomers (*ortho endo, ortho exo, meta endo*, and *meta exo*). The energies are summarized in Table 1. There is a large preference (2 to 4 kcal/mol) for the *ortho* as compared to the *meta* transition structures. This result is in agreement with our experimental work on the reaction of *trans*-*N*-acylamino-1,3-diene **1** and dimethylacrylamide because the *meta* regioisomer could not be detected in the crude mixture of products. In related studies, a high regioselectivity that was rationalized by using frontier molecular orbital concepts, was also encountered (>98% *ortho* adducts) (*14*).

The predicted stereoselectivity is also very high and overestimates experimental results. The diene and dienophile approach each other in nonparallel planes with all of the =C–H bonds extending out of the two planes. This geometry allows better overlap between the terminal carbons of each reagent. Substitutions of an electron-withdrawing group on the dienophile and an electron-donating group on the diene necessarily make the transition state structures asynchronous (*14*) (Fig. 3). The calculations predict that the energy barrier for the favored *endo* and disfavored *exo* processes differ by 1.9 kcal/mol, and thus the disfavored process falls easily within the binding energy deliverable by an antibody. These calculations agree with the stereoselectivity observed for our experimental system where a high *endo* preference is found (85% *endo* adduct starting with diene **1b**). This observation is consistent with other results (*15*). The *endo* preference can be attributed to secondary orbital interactions. This reaction is ideally suited for antibody catalysis because the product ratio is under kinetic control yielding mostly the thermodynamically higher energy *endo* adduct.

Table 1. Calculated activation energies of the transition structures with respect to reactants. Relative energies are given in parentheses.

Transition structure	Activation energy (kcal/mol)	
	RHF/3-21G	6-31G*//3-21G
Ortho, endo	27.30 (0.0)	40.80 (0.0)
Ortho, exo	28.84 (1.5)	42.70 (1.9)
Meta, endo	29.82 (2.5)	42.88 (2.1)
Meta, exo	30.95 (3.6)	43.94 (3.1)

Hapten design. Properly designed haptens should mimic critical features of the transition state of the reaction to be catalyzed. In the Diels-Alder reaction studied here, we have the special problem that control of the relative and absolute stereochemistry requires generation of antibodies appropriately tailored to stabilize only one out of the four possible boat-shaped transition states (Fig. 4).

Haptens **A** and **B** were designed to selectively mimic the transition states of the reaction for either the *endo* or *exo* approach of the dienophile (Fig. 5). Because the reaction proceeds through a highly ordered transition state for which ΔS is typically on the order of −30 to −40 entropy units, the use of antibodies as "entropic traps" is particularly attractive (*16*). The bicyclic structure for both haptens should generate antibodies that overcome the large translational and rotational entropic barriers by binding the two reactants together in the combining site. The comparison of the calculated transition state structure with the x-ray crystallographic structure of compound **5-endo**, which is a precursor to hapten **A**, shows that the bicyclo[2.2.2]octene moiety is a good mimic of the transition state of the cycloaddition (Fig. 6). The structure includes a boat-shaped cyclohexene ring that mimics satisfactorily the pericyclic transition state (*17*). The part of the structure that resembles the diene is locked into a good mimic of the required *s-cis* conformation which the substrate should, therefore,

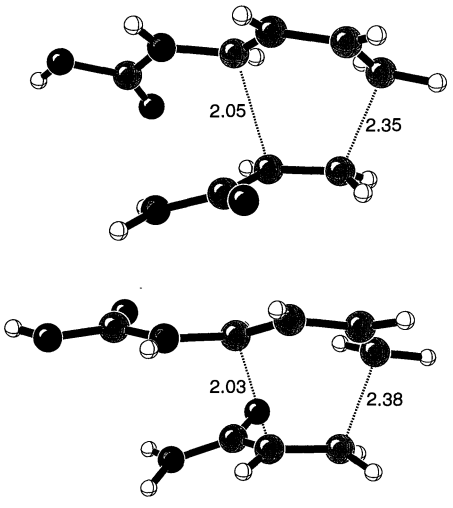


Fig. 3. Structures optimized at the RHF/3-21G level for the *ortho-endo* and *ortho-exo* transition states.

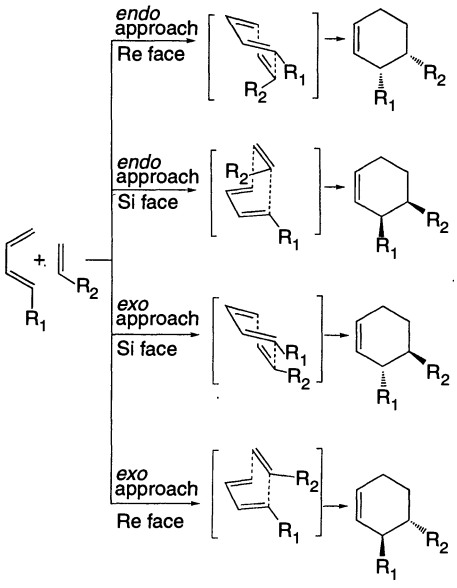


Fig. 4. Geometrical features of the four transition states that control the *endo*/*exo* enantioselectivity of the Diels-Alder transformation.

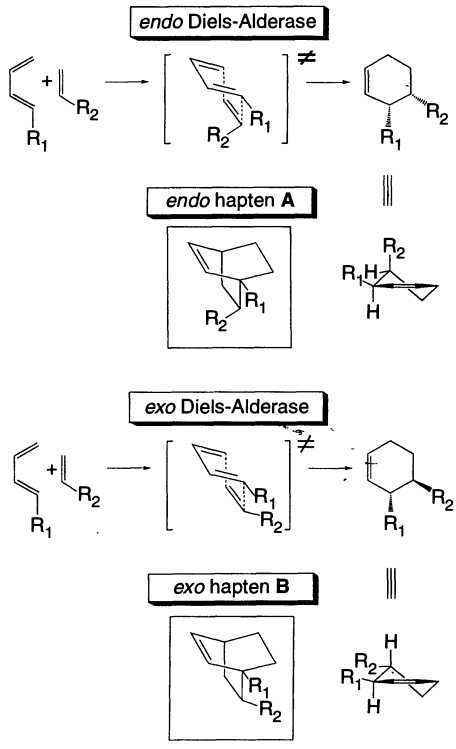


Fig. 5. Schematic representations of the *endo* and *exo* approaches and the corresponding haptens **A** and **B**.

be forced to adopt in the antibody combining site.

Calculations predict that the diene adopts a nearly planar *s-cis* conformation as the dihedral angle C3–C4–C5–C6 is equal to 1.1°. This quasi-planar conformation is perfectly mimicked by the hapten as this same dihedral angle is equal to 0.7°, according to the x-ray data. Considering angle C5–C6–C1 (100.3°), we can deduce from the calculation that the approach of the diene toward the dienophile is not parallel but has a deviation of 10.3°. This deviation was found to be 19.7° in the hapten (C5–C6–C1 = 109.7°). Interestingly, the optimal orientation of the amido substituent of the dienophile toward the diene predicted by the theoretical calculations is satisfactorily mimicked by the hapten. This is deduced by comparing the dihedral angle C5–C6–C1–C9 and the angle C1–C9–O1 that according to the calculations are equal to −73.1° and 123.0° and for the hapten −71.0° and 121.9°. This geometrical feature is important for the *endo* hapten, as this orientation allows attractive interactions responsible for the lower energy of activation for the *endo* transition state over the *exo* (*13*).

In contrast to the calculated asynchrony of the transition state, our haptens mimic synchronous transition states. Our calculations show that in the *endo* transition state the C–C bonds being formed have unequal lengths of 2.05 and 2.35 Å, a difference of 0.30 Å. The *exo* transition state is slightly more asynchronous; the C–C bonds being formed have lengths of 2.03 and 2.38 Å, a difference of 0.35 Å (Fig. 3). In our hapten, the corresponding C–C bond lengths are fixed at 1.55 Å. Although this geometrical difference may influence the rate enhancement achieved, it should have no effect on the control of the diastereo- and enantioselectivity. In general, the haptens mimic a more compact transition state structure, which should allow induction of antibodies that force approximation of the diene and dienophile, thereby increasing their reactivity.

The stereochemical constraints required to control the diastereoselectivity of the reaction have also been designed into the bicyclic system. Hapten **A** mimics the *endo* approach in the transition state (Fig. 5) in that the amido group of the dienophile is oriented toward the π orbitals of the diene and thus should generate antibodies that catalyze the formation of the *cis* adduct. Conversely, in the *exo* transition state (Fig. 5), the reference substituent of the dienophile is oriented away from the π system. This approach is mimicked by hapten **B**, which should generate antibodies whose binding sites stabilize the *exo* transition state and thus catalyze the formation of the corresponding *trans* adduct.

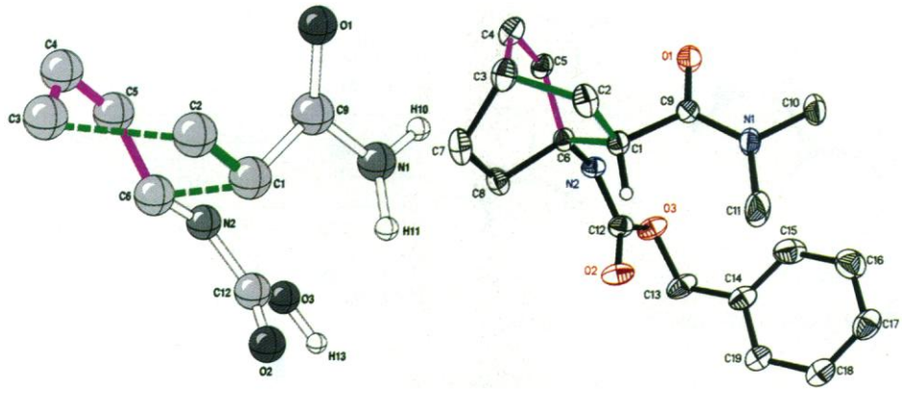


Fig. 6. The *endo* transition state structure (**left**), orientated with respect to the ORTEP plot of the x-ray structure of the *endo* hapten (**right**).

Control of the absolute stereochemistry of the reaction depends on the extent to which an antibody can regulate which face of the dienophile is approached by the diene. Because each antibody has a distinct chiral environment, we expected that the enantiofacial discrimination should be tightly regulated during the cycloaddition process to yield only one enantiomer.

Finally, the hapten has design features that should help minimize product inhibition in that its structure corresponds to the higher energy boat conformation of the product. Consequently, as the reaction products adopt the lower energy twisted-chair conformation, they should be released from the antibody that was not programmed to recognize this shape (Fig. 5) (*18*).

Synthesis of the transition state analogs and generation of antibodies. Haptens **A** and **B** have been synthesized according to the scheme shown in Fig. 7. The key step of the synthesis is the thermal cycloaddition of the cyclic diene **4** and dimethylacrylamide. This reaction affords a mixture of *endo* and *exo* adducts that are cleanly separated by column chromatography. The stereochemistry of each adduct was determined by nuclear magnetic resonance (NMR) spectroscopy (*19*), and an x-ray crystallography study was undertaken on the *endo* adduct to assign unambiguously its stereochemistry (Fig. 6).

Each hapten was coupled to the carrier proteins bovine serum albumin (BSA) and keyhole limpet hemocyanin (KLH). We immunized 129 G_{IX}+ mice with the KLH conjugates. Antibodies were generated by standard protocols (*20*). Monoclonal antibodies (mAbs) were purified from ascites fluid by ammonium sulfate precipitation (SAS), anion exchange chromatography on DEAE Sephacel, and affinity chromatography on a protein G–Sepharose column and were studied for binding to the hapten coupled to BSA (*21*). Twenty-two mAbs were obtained from immunization with hapten **A** and 25 from hapten **B**. The antibodies were assayed with substrates **1b** and *N*,*N*-dimeth-

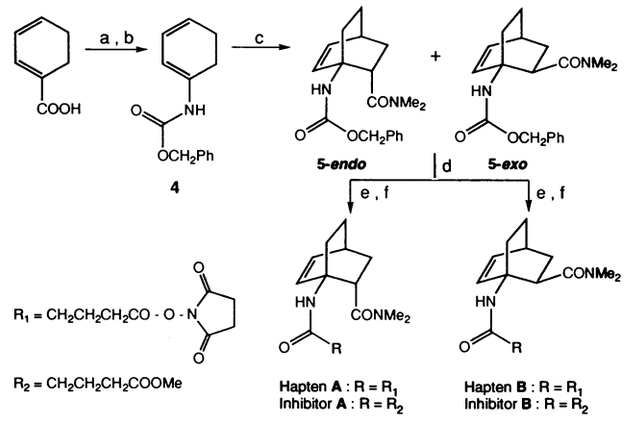


Fig. 7. Reagents and conditions for the synthesis of haptens **A** and **B** and inhibitors **A** and **B**. (a) 1.2 equivalents (eq) ethylchloroformate, 1.2 eq diisopropylethylamine, dry acetone, 0°C, 30 min, then 2.4 eq sodium azide, 0°C, 15 min; (b) 0.8 eq benzyl alcohol, dry toluene, reflux, 45 min, 62%; (c) 5 eq dimethylacrylamide, dry toluene, reflux, 50 hours, 71%; (d) separation of the *endo* and *exo* isomers by column chromatography by using a 7/3 mixture of ethylacetate/hexane; and (e) 1.1 eq iodotrimethylsilane, dry chloroform, room temperature, 2 hours. (f) Synthesis of the haptens: 1.2 eq diisopropylethylamine, 1.2 eq 5-((2,5)-dioxo-1-pyrrolidinyl)-5-oxopentanoyl chloride, room temperature, dry dichloromethane, 1 hour (60% for hapten **A**, 76% for hapten **B**); synthesis of the inhibitors: 1.2 eq diisopropylethylamine, 1.2 eq methyl-4(chloroformyl)-butyrate, room temperature, dry dichloromethane, 1 hour (72% for inhibitor **A**, 74% for inhibitor **B**).

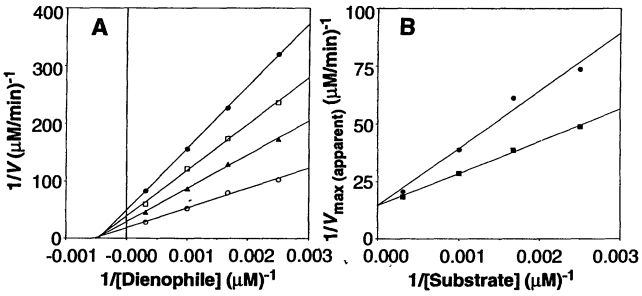


Fig. 8. (**A**) Double reciprocal plot of initial velocities of *endo* adduct **2b** formation in the presence of 7D4. Diene **1b** concentrations were fixed (●, 400 μM; □, 600 μM; ▲, 1000 μM; and ○, 3200 μM), whereas the concentrations of the dienophile were varied over the same range. The symmetrical plot with fixed dienophile and varied diene concentrations was also constructed. (**B**) Replot of the apparent maximum velocity (V_{max}) values from the Lineweaver-Burk plots versus fixed substrate concentrations (■, diene and ●, dienophile) to obtain the true V_{max} value as $1/(y$ intercept) and Michaelis constant (K_m) values as $-1/(x$ intercept). Catalysis of *exo* adduct **3b** formation by 22C8 was characterized in the same manner, and the resulting data were plotted and interpreted in a fashion identical to that shown.

ylacrylamide for production of either the *endo* **2b** or *exo* adduct **3b**. The cycloaddition reactions were performed at 37°C in PBS, pH 7.4. The formation of products was monitored by reversed-phase high-performance liquid chromatography (HPLC) under conditions that allowed separation of the *exo* and *endo* adducts (*22*).

Two mAbs (13D4 and 7D4) derived from immunization with hapten **A** and four mAbs (22C8, 27E4, 14F2, and 8B11) derived from immunization with hapten **B** were found to catalyze exclusively the formation of either the *endo* or *exo* adducts, respectively. Antibodies 7D4 and 22C8 appeared to provide the greatest rate enhancement and hence were chosen for further characterization.

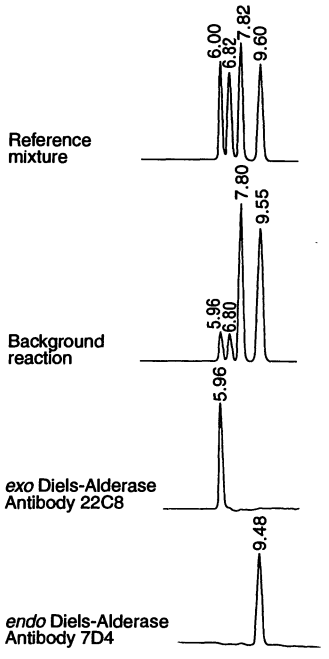


Fig. 9. A DAICEL Chiralpak AD column was used with an isocratic program 70/30; hexane (1.5% trifluoroacetic acid)/isopropanol, 1 ml/min, 240 nm. The retention times for the four isomers were as follows: 6.00 min and 6.82 min (*trans* isomers) and 7.82 min and 9.60 min (*cis* isomers). The HPLC traces 3 and 4 were obtained by using the following reaction conditions: trace 3, 500 μM diene, 500 μM dienophile, 135 μM antibody 22C8, 65 hours, 12 μM of **3b** formed; trace 4, 500 μM diene, 500 μM dienophile, 89 μM antibody 7D4, 20 hours, 36 μM of **2b** formed.

Catalysis and stereoselectivity. The kinetics of both antibody-catalyzed reactions were determined by measuring the differences in initial rates between the catalyzed and background reactions. All reactions were carried out in 10 mM PBS, 160 mM NaCl, pH 7.4, at 37°C. Catalyzed reactions were run in the presence of 20 μM antibody 7D4 or 22C8. Initial velocities were determined by following the formation of *endo* and *exo* adducts by HPLC within <10% reaction completion. Catalysis of both the *endo* and *exo* Diels-Alder reactions by antibodies 7D4 and 22C8 were examined as random bireactant systems. Lineweaver-Burk plots were constructed by holding one substrate at a fixed concentration while varying the concentration of the other. From Fig. 8, we see that for the *endo* Diels-Alderase 7D4, one ligand (either the diene or dienophile) has no effect on the binding of the other. A similar pattern was observed for the *exo* Diels-Alderase 22C8. These data provide an accurate estimate of the binding of the diene or dienophile to 7D4 and 22C8 either in the presence or absence of the other substrate. The Michaelis constant (K_m) values for diene **1b** and the acrylamide dienophile were determined to be 960 and 1700 μM for mAb 7D4 and 700 and 7500 μM for mAb 22C8, respectively. The catalytic rate constant (k_{cat}) values of 3.44×10^{-3} and 3.17×10^{-3} min^{-1} were obtained for antibodies 7D4 and 22C8, respectively. The second-order rate constants for the background reactions, k_{uncat}, were found to be 7.15×10^{-4} and 1.75×10^{-4} M^{-1} min^{-1} for the *endo* and *exo* Diels-Alder reactions, respectively. A measure of the entropic gain of these antibody-catalyzed reactions can be estimated by the ratio k_{cat}/k_{uncat}. Here, the "effective molarity" was determined to be 4.8 M for mAb 7D4 and 18 M for mAb 22C8, respectively. Thus, a better rate acceleration is achieved for the more difficult chemical process.

To ensure that catalysis was being carried out in the antibody binding pocket, inhibition studies were performed with inhibitors **A** (*endo*) and **B** (*exo*) (synthesis and structures in Fig. 7), which are congruent to the immunizing haptens. Both reactions were strongly inhibited by stoichiometric amounts of these inhibitors when working at approximately K_m for each substrate.

The enantioselectivity of the cycloaddition reaction was studied with a normal-phase HPLC column of chiral support (Fig. 9). The HPLC trace of the reference mixture of *cis* and *trans* adducts shows that the enantiomers of both *endo* and *exo* adducts were all simultaneously separable on the column (trace 1). The background reaction (trace 2) gives an *endo*:*exo* mixture of the adducts in a ratio of 85:15. Antibody 7D4 (trace 4) and antibody 22C8 (trace 3) catalyze the formation of the *cis* (*endo*) and *trans* (*exo*) adduct, respectively, with >98% enantioselectivity.

Antibody catalysis of disfavored reactions. In this study, we have generated antibodies that catalyze the Diels-Alder transformation and control its stereochemistry. The catalysis of the *exo* pathway is of special interest and represents another example of the use of an antibody's binding energy in a highly specific way to reroute a chemical reaction to facilitate an otherwise disfavored process. The key to the successful generation of these catalysts is the ability to program the antibody molecule to address simultaneously multiple parameters expected along the reaction pathway. Thus, there is concerted control of the approximation and relative orientation of the reactants. The strategy described here should be general and can be used for any cycloaddition involving an acyclic diene by using haptens that bear the appropriate substituents. The only requirement is that the reaction should not face an energy barrier greater than the ~20 kcal/mol of binding energy deliverable by an antibody. The next challenge is to control the regiochemistry of this Diels-Alder reaction, which should be feasible for reactions with substituents similar to those studied here because our calculations indicate that they are disfavored by only 2 to 4 kcal/mol (Table 1). If this can be done, we can imagine the generation of a series of antibodies where each one is a selective catalyst for only one of the 32 possible outcomes of a cycloaddition reaction using four sets of isomeric reactants.

REFERENCES AND NOTES

1. O. Diels and K. Alder, *Justus Liebigs Ann. Chem.* **460**, 98 (1928); O. Diels, J. H. Blom, W. Koll, *ibid.* **443**, 242 (1925); B. M. Trost, I. Fleming, L. A. Paquette, *Comprehensive Organic Synthesis* (Pergamon, Oxford, 1991), vol. 5, pp. 316; F. Friguelli and A. Tatichi, *Dienes in the Diels-Alder Reaction* (Wiley, New York, 1990), and references therein; W. Carruthers, *Cycloaddition Reactions in Organic Synthesis* (Pergamon, Oxford, 1990), and references therein; G. Desimoni, G. Tacconi, A. Barco, G. P. Pollini, *ACS Monograph No. 180* (American Chemical Society, Washington, DC, 1983), and references cited therein.
2. For recent reviews, see U. Pindur, G. Lutz, C. Otto, *Chem. Rev.* **93**, 741 (1993); H. B. Kagan and O. Riant, *ibid.* **92**, 1007 (1992), and references therein.
3. D. Hilvert, K. W. Hill, K. D. Nared, M-T. M. Auditor, *J. Am Chem. Soc.* **111**, 9261 (1989).
4. A. C. Braisted and P. G. Schultz, *ibid.* **112**, 7430 (1990).
5. C. J. Suckling, M. C. Tedford, L. M. Bence, J. I. Irvine, W. H. Stimson, *Biorg. Med. Chem. Lett.* **2**, 49 (1992).
6. K. D. Janda, C. G. Shevlin, R. A. Lerner, *Science* **259**, 490 (1993).
7. L. E. Overman, G. F. Taylor, C. B. Petty, P. J. Jessup, *J. Org. Chem.* **43**, 2164 (1978).
8. For example: *dl*-pumiliotoxin, C. L. E. Overman and P. J. Jessup, *Tetrahedron Lett.* 1253 (1977); *J. Am. Chem. Soc.* **100**, 5179 (1978); *dl*-perhydrogephyrotoxin, L. E. Overman and C. Fukaya, *ibid.* **102**, 1454 (1980); *dl*-isogabaculine, S. Danishefsky and F. M. Hershenson, *J. Org. Chem.* **44**, 1180 (1979); tylidine, L. E. Overman, C. B. Petty, R. J. Doedens, *ibid.*, p. 4183; amaryllidaceae alkaloids, L. E. Overman *et al.*, *J. Am. Chem. Soc.* **103**, 2816 (1981); gephirotoxine, L. E. Overman, D. Lesuisse, M. Hashimoto, *ibid.* **105**, 16, 5373 (1983).
9. Diene **1b** was prepared by neutralizing the corresponding carboxylic acid suspended in PBS (sodium phosphate buffer) with one equivalent of NaOH.
10. All new adducts were characterized by NMR and mass spectrometry.
11. In the crude mixture of the cycloaddition, we could not detect any trace of the *meta* regioisomer by ^{1}H NMR (>98% *ortho* regioselective).
12. M. J. Frisch *et al.*, *Gaussian 92* (Gaussian, Pittsburgh, PA, 1992).
13. R. J. Loncharich, F. K. Brown, K. N. Houk, *J. Org. Chem.* **54**, 1129 (1989). All structures were fully optimized with the 3-21G basis set. Reactants were also fully optimized with the 6-31G* basis set. Vibrational frequencies for all transition structures were calculated at the RHF/3-21G level and shown to have one imaginary frequency. In addition, the energies were evaluated by single-point calculations with the 6-31G* basis set on 3-21G geometries.
14. K. N. Houk, R. J. Loncharich, J. Blake, W. Jorgensen, *J. Am. Chem. Soc.* **111**, 9172 (1989).
15. L. E. Overman, G. F. Taylor, K. N. Houk, I. N. Domelsmith, *ibid.* **100**, 3182 (1978).
16. M. I. Page and W. P. Jencks, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 1678 (1971).
17. F. Bernardi *et al.*, *J. Am. Chem. Soc.* **110**, 3050 (1989); F. K. Brown and K. N. Houk, *Tetrahedron Lett.* **25**, 4609 (1984).
18. This strategy to avoid product inhibition was first described by Braisted and Schultz in (*4*).
19. The stereochemistry was established on the basis of ^{1}H NMR analysis. For example: **5-exo**, δH_6 = 2.65 ppm and $^4J_{6,7a}$ = 3.5 Hz; **5-endo**, δH_6 = 3.12 ppm, and $^4J_{6,7a}$ = 0 Hz.
20. G. Kohler and C. Milstein, *Nature* **256**, 495 (1975); the KLH conjugate was prepared by slowly adding 2 mg of the hapten dissolved in 100 μl of dimethylformamide to 2 mg of KLH in 900 μl of 0.01 M sodium phosphate buffer, pH 7.4. Four 8-week-old 129GIX+ mice each received 2 intraperitoneal (IP) injections of 100 μg of the hapten conjugated to KLH and RIBI adjuvant (MPL and TDM emulsion) 2 weeks apart. A 50 μg IP injection of KLH conjugate in alum was given 2 weeks later. One month after the second injection, the mouse with the highest titer was injected intravenously with 50 μg of KLH-conjugate; 3 days later, the spleen was taken for the preparation of hybridomas. Spleen cells (1.0×10^8) were fused with 2.0×10^7 SP2/0 myeloma cells. Cells were plated into 30 96-well plates; each well contained 150 μl of hypoxanthine, aminopterin, thymidine–Dulbecco's minimal essential medium (HAT-DMEM) containing 1% nutridoma, and 2% bovine serum albumin.
21. E. Engvall, *Methods Enzymol.* **70**, 419 (1980).
22. A C_{18} column, (VYDAC 201TP54) was used with an isocratic program of 78/22; H_2O (0.1% trifluoroacetic acid)/acetonitrile, at 1.5 ml/min, 240 nm. The retention time for the *trans* isomer is 6.00 min and for the *cis* isomer is 6.60 min.

23. We thank J. Ashley for technical assistance in the kinetic analysis; L. Ghosez, K. B. Sharpless, K. C. Nicolaou, E. Keinan, and D. Boger for helpful discussions and comments on the manuscript; and R. K. Chadha for the x-ray crystallographic study. Supported in part by the National Science Foundation (CHE-9116377) and the A. P. Sloan Foundation (K.D.J.). We also thank Fundación Ramón-Areces (Spain) for a fellowship to B.P.T.

16 July 1993; accepted 3 September 1993

Detecting Subtle Sequence Signals: A Gibbs Sampling Strategy for Multiple Alignment

Charles E. Lawrence, Stephen F. Altschul, Mark S. Boguski, Jun S. Liu, Andrew F. Neuwald, John C. Wootton

A wealth of protein and DNA sequence data is being generated by genome projects and other sequencing efforts. A crucial barrier to deciphering these sequences and understanding the relations among them is the difficulty of detecting subtle local residue patterns common to multiple sequences. Such patterns frequently reflect similar molecular structures and biological properties. A mathematical definition of this "local multiple alignment" problem suitable for full computer automation has been used to develop a new and sensitive algorithm, based on the statistical method of iterative sampling. This algorithm finds an optimized local alignment model for *N* sequences in *N*-linear time, requiring only seconds on current workstations, and allows the simultaneous detection and optimization of multiple patterns and pattern repeats. The method is illustrated as applied to helix-turn-helix proteins, lipocalins, and prenyltransferases.

Patterns shared by multiple protein or nucleic acid sequences shed light on molecular structure, function, and evolution. The recognition of such patterns generally relies upon aligning many sequences, a complex, multifaceted research process whose difficulty has long been appreciated. This problem may be divided into "global multiple alignment" (*1, 2*), whose goal is to align complete sequences, and "local multiple alignment" (*2–11*), whose aim is to locate relatively short patterns shared by otherwise dissimilar sequences. We report a new algorithm for local multiple alignment that assumes no prior information on the patterns or their locations within the sequences; it determines these locations from only the information intrinsic to the sequences themselves. We focus on subtle amino acid sequence patterns that may vary greatly among different proteins.

Much research on the alignment of such patterns uses additional information to supplement algorithmic analyses of the actual sequences, including data on three-dimensional structure, chemical interactions of residues, effects of mutations, and interpretation of sequence database search results. However, such research, which has led to many discoveries of sequence relations and structure and function predictions [see (*12*) for a recent example], is laborious and requires frequent input of expert knowledge. These approaches are becoming increasingly overwhelmed by the quantity of sequence data.

A number of automated local multiple alignment algorithms have been developed (*2–11*), and some have proved valuable as part of integrated software workbenches. Unfortunately, rigorous algorithms for finding optimal solutions have been so computationally expensive as to limit their applicability to a very small number of sequences, and heuristic approaches have gained speed by sacrificing sensitivity to

C. E. Lawrence, S. F. Altschul, M. S. Boguski, A. F. Neuwald, and J. C. Wootton are with the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894. J. S. Liu is with the Statistics Department, Harvard University, Cambridge, MA 02138. C. E. Lawrence is also affiliated with the Biometrics Laboratory, Wadsworth Center for Labs and Research, New York State Department of Health, Albany, NY 12201.