Design and Synthesis of a Transition State Analog for the Ene Reaction between Maleimide and 1-Alkenes

Ari Yliniemelä,*,† Henrik Konschin,‡ Constantin Neagu,‡ Aarne Pajunen,‡ Tapio Hase,‡ Gösta Brunow,‡ and Olle Teleman*,§

Contribution from the Technical Research Centre of Finland, Chemical Technology, POB 1401, FIN-02044 VTT Espoo, Finland, Department of Chemistry, University of Helsinki, POB 55, FIN-00014 Helsinki, Finland, and Technical Research Centre of Finland, Biotechnology and Food Research, POB 1503, FIN-02044 VTT Espoo, Finland

Received October 31, 1994[@]

Abstract: An analog with high similarity to the transition state of the ene reaction between 1-alkenes and maleimide has been designed and synthesized. The reaction mechanism was studied using propene as a representative 1-alkene. The MP2/6-31G*//HF/6-31G* level of accuracy was used for the *ab initio* reaction modeling. Although the semiempirical AM1 method does not provide accurate reaction energetics, the transition state geometries were found to be in good agreement with the corresponding HF/6-31G* structures. The *endo* transition state of the reaction between 1-butene and maleimide was used for the analog design. The structures for both transition state and its putative analogs were optimized using the AM1 method. The transition state and its analogs were compared by rigid body field fitting. The X-ray structure of the analog suggests a reasonable agreement between the computational and the experimental results. The analog is now used as a hapten to obtain catalytic antibodies.

Introduction

The ene reaction, i.e. the reaction between an alkene with an allylic hydrogen (the "ene") and a compound containing an electron deficient multiple bond ("enophile"), is one of the simplest organic, thermal reorganization processes.[1] Mechanistically the ene reaction bears a certain resemblance to the Diels—Alder cycloaddition and to the [1,*j*] sigmatropic migrations of hydrogen. In fact, the [1,5] sigmatropic rearrangement of hydrogen is an internal ene reaction. The Diels—Alder reaction has been the subject of numerous theoretical studies.[2] Far less is known theoretically about the ene reaction, although its synthetic value has been widely recognized.[3] Notably, Houk and co-workers have taken an interest in calculating transition state structures of the reactions of ethylene and formaldehyde with propene[4] and the intramolecular ene reactions of 1,6-dienes with unactivated and activated enophiles.[5] The HF/3-21G level used in their calculations allowed reliable estimates of transition state structures, but the computation of reaction and, and in particular, activation energies calls for electron correlation treatment. Houk and Loncharik also calculated the 3-21G transition structures for the retroene reaction to determine the difference in energy between the *endo* and *exo* structure.[6] Uchimaru et al. have investigated the reaction between methyl

acrylate and propene in order to elucidate a possible *exo/endo* selectivity of the transition state.[7] The transition state found by Uchimaru et al. with split valence 3-21G and 6-31G basis sets suggests less asynchronous character in the bond reorganization as compared to the parent reaction between ethylene and propene. In accordance with the calculations made by Houk et al., Uchimaru and co-workers also found that the activation energy of the reaction is extremely sensitive to electron correlation effects. The activation energy decreases by about a factor of 2 at the MP2/6-31G*/6-31G level as compared to the uncorrelated HF value.

A long-standing concern for the correct description of the reaction mechanism and pathways of multibond reactions has been expressed by the Dewar group.[8] In particular, the question of synchronicity in the reaction has been discussed vigorously. Early MINDO/3 and MNDO calculations of kinetic isotope effects of various retroene reactions seemed to yield inconclusive results with respect to both activation parameters and mechanistic predictions.[9] The more recent AM1, PM3, and SAM1 semiempirical methods have been applied to ene- and retrocene-type reactions with considerably improved accuracy.[10,11] Nevertheless, deciding the level necessary for accurate description of a reaction energy surface may still be a tedious and demanding task.

The present investigation ties together techniques familar to theoretical chemistry with modern synthetical methods including antibody generation. We have studied the ene reaction of maleimide and 1-alkenes with both semiempirical and *ab initio*

---

* Correspondence should be addressed to either of these authors.
† Technical Research Centre of Finland, Chemical Technology.
‡ University of Helsinki.
§ Technical Research Centre of Finland, Biotechnology and Food Research.
@ Abstract published in *Advance ACS Abstracts,* April 15, 1995.
(1) Mikami, K.; Shimizu, M. *Chem. Rev.* **1992,** *92,* 1021.
(2) Li, Y.; Houk, K. N. *J. Am. Chem. Soc.* **1993,** *115,* 7478 and references therein.
(3) (a) Hoffman, H. M. R. *Angew. Chem.* **1969,** *81,* 597. (b) Snider, B. B. In *Comprehensive Organic Synthesis;* Trost, B. M., Fleming, I., Eds.; Pergamon: London, 1991; Vol. 5. (c) Snider, B. B. *Acc. Chem. Res.* **1980,** *13,* 426. (d) Oppolzer, W.; Snieckus, V. *Angew. Chem.* **1978,** *90,* 506.
(4) (a) Loncharich, R. J.; Houk, K. N. *J. Am. Chem. Soc.* **1987,** *109,* 6947. (b) Borden, W. T.; Loncharich, R. J.; Houk, K. N. *Annu. Rev. Phys. Chem.* **1988,** *39,* 213. (c) Thomas, B. E.; Loncharich, R. J.; Houk, K. N. *J. Org. Chem.* **1992,** *57,* 1354.
(5) Thomas, B. E., IV; Loncharich, R. J.; Houk, K. N. *J. Org. Chem.* **1992,** *87,* 1354.

(6) Loncharich, R. J.; Houk, K. N. *J. Am. Chem. Soc.* **1988,** *110,* 2089.
(7) (a) Uchimaru, T.; Tsuzuki, S.; Tanabe, K.; Hayashi, Y. *J. Chem. Soc., Chem. Commun.* **1989,** 1861. (b) Uchimaru, T.; Tsuzuki, S.; Tanabe, K.; Hayashi, Y. *Bull. Chem. Soc. Jpn.* **1990,** *63,* 1861.
(8) (a) Dewar, M. J. S. *J. Am. Chem. Soc.* **1984,** *106,* 209. (b) Dewar, M. J. S.; Olivella, S.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1986,** *108,* 5771. (c) Dewar, M. J. S.; Jie, C. *Acc. Chem. Res.* **1992,** *25,* 537.
(9) (a) Dewar, M. J. S.; Ford, G. P. *J. Am. Chem. Soc.* **1977,** *99,* 8343. (b) Brown, S. B.; Dewar, M. J. S.; Ford, G. P.; Nelson, D. J.; Rzepa, H. S. *J. Am. Chem. Soc.* **1978,** *200,* 7832.
(10) Dewar, M. J. S.; Jie, C.; Yu, J. *Tetrahedron* **1993,** *49,* 5003.
(11) Davies, A. G.; Schiesser, C. H. *Tetrahedron* **1991,** *47,* 1707.

Scheme 1. The Ene Reaction between Maleimide and Propene[a]
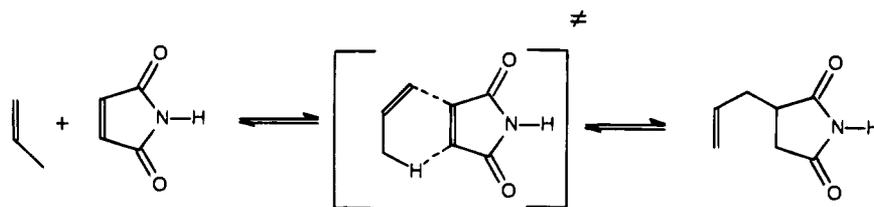
[a] The reaction was studied with both *ab initio* and semiempirical MO methods.

methods in order to characterize the transition state of the reaction. With an accurate knowledge of the transition state structure, the synthesis of suitable transition state analogs can be attempted, with the ultimate purpose to generate an antibody against the transition state. If such an antibody can be generated, a thorough catalytic tailoring of the reaction would be attainable. Recently, Gouverneur et al. showed how the *exo/endo* pathways of the Diels—Alder reaction may be controlled via catalytic antibodies, the haptens of which mimic the respective transition states of the reaction.[12] We will show that information obtained by computational means also aids the synthetic effort considerably in the case of the reaction between maleimide and 1-alkenes.

Results and Discussion

Reaction Mechanism and Transition State Structures. Despite the growing interest in ene reactions, that between maleimide and 1-alkenes has not, to our knowledge, been studied theoretically or experimentally. However, both experimental and computational data are available for the reaction between 1-alkenes and maleic anhydride.[10,13] The experiments suggest that the activation energy is independent of the length of the 1-alkene. Solvent effects for ene reactions have been reported,[14] but the solvent effect on the rate of the maleic anhydride reaction has been found to be negligible. These experimental facts presumably apply to the reaction of maleimide with 1-alkenes as well. It is thus reasonable to study the energetics and mechanism of that reaction computationally in the gas phase, using propene as a representative 1-alkene.

It has been argued that the *ab initio* studies of reactions should be carried out using split valence basis set with polarization functions and electron correlation correction throughout.[15] There is also evidence suggesting that pericyclic reactions can be studied reliably at a lower level.[16] In a previous work,[17] we studied the effect of the computational level, and it turned out that the HF/3-21G level geometry optimization combined with a MP2/6-31G* single point calculation gives an accurate picture of both mechanism and energetics of the ene reaction.

We examined the reaction between propene and maleimide (Scheme 1) using both *ab initio* and semiempirical methods. The *ab initio* calculations were done using the Gaussian92 software.[18] The AM1 model in the Mopac 6.0 and Ampac 5.0 programs was used for the semiempirical calculations.[19] To ensure that the calculated transition structures actually connect the reactants and the products, AM1 level IRC calculations were carried out with respect to the transition state structures

Table 1. The Energetics (kcal/mol) of the Reaction between Maleimide and Propene

Method	route	activation energy[a]	reaction energy
AM1	*endo*	32.51	−35.96
AM1	*exo*	32.44	−36.86
HF/3-21G	*endo*	42.79	−27.61
HF/3-21G	*exo*	44.94	−27.22
HF/6-31G*	*endo*	54.45	−23.55
HF/6-31G*	*exo*	56.07	−23.72
MP2/6-31G*//HF/3-21G	*endo*	20.07	−31.44
MP2/6-31G*//HF/3-21G	*exo*	22.13	−30.75
MP2/6-31G*//HF/6-31G*	*endo*	19.97	−31.22
MP2/6-31G*//HF/6-31G*	*exo*	22.10	−30.65
MP2/6-31G*//HF/6-31G*[b]	*endo*	19.65	−27.22
MP2/6-31G*//HF/6-31G*[b]	*exo*	21.66	−26.76
experiment[13]		21.5	

[a] The experimental activation energy is for the reaction between maleic anhydride and propene. [b] Corrected for zero-point vibrational energy. The zero-point energies were taken from the HF/6-31G* calculations.

optimized with the same method.[20] The activation and reaction energies, calculated at different levels of accuracy, are given in Table 1.

The MP2 corrected 6-31G* activation energies are about 35 kcal/mol smaller than the corresponding Hartree—Fock results. Both Houk and Uchimaru have made similar observations.[4,7] The MP2/6-31G* activation energies we recently calculated for the reaction between 1-alkenes and maleic anhydride,[17] using HF/3-21G optimized structures, are within the experimental accuracy. This excellent agreement is partially fortuitous, but in this particular case, the fact that the remaining approximations cancel out lends credibility to the geometry optimization. The second-order perturbation treatment is known to account for about 90% of the basis set correlation energy,[21] but it is also known that MP2 calculations overcorrect the HF result in the case of activation energies.[22] Accepting this fact together with the incompleteness of the 6-31G* basis set, the combination MP2/6-31G* still represents a reasonable compromise between economy and accuracy, at least for the ene and similar reactions.

The difference between the *endo* and *exo* activation energies is 2.1 kcal/mol, favoring the *endo* route. This value contains only the activation enthalpy. The value indicating the actual preference between the routes is the free energy of activation.

(12) Gouverneur, V. E.; Houk, K. N.; de Pascual-Teresa, B.; Beno, B.; Janda, K. D.; Lerner, R. J. *Science* **1993**, *262*, 204.

(13) Benn, F. R.; Dwyer, J. *J. Chem. Soc., Perkin Trans. 2* **1977**, 533.

(14) Desimoni, G.; Faita, G.; Righetti, P. P.; Sfulcini, A.; Tsyganov, D. *Tetrahedron* **1994**, *50*, 1821.

(15) Dewar, M. J. S. *Int. J. Quantum Chem.* **1992**, *44*, 427.

(16) (a) Houk, K. N. *Pure Appl. Chem.* **1989**, *4*, 643. (b) Spellmeyer, D. C.; Houk, K. N. *J. Am. Chem. Soc.* **1988**, *110*, 3412. (c) Jensen, F.; Houk, K. N. *J. Am. Chem. Soc.* **1987**, *109*, 3139.

(17) Yliniemelä, A.; Konschin, H.; Pietilä, L.-O.; Teleman, O. *J. Mol. Struct.* (*Theochem*), in press.

(18) For the *ab initio* calculations Gaussian92 (Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Robb, M. A.; Binkley, J. S.; Gonzales, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R. L.; Kahn, L. R.; Stewart, J. J. P.; Topiol, S.; Pople, J. A.; Gaussian Inc.: Pittsburgh, PA, 1992) was used on a CRAY X-MP 432 supercomputer and on a Convex C3840 server at CSC.

(19) The AM1 model in the Mopac 6.0, on a Silicon Graphics Indigo 2 workstation, and in the Ampac 5.0 (1994 Semichem, 7128 Summit, Shawnee, KS 66216), on a Convex C3840 server, was used for the semiempirical calculations.

(20) Stewart, J. J. P.; Davis, L. P.; Burggraf, W. *J. Comp. Aid. Mol. Design* **1987**, *8*, 1117.

(21) Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359.

(22) Bachrach, S. M. *J. Org. Chem.* **1994**, *59*, 5027.
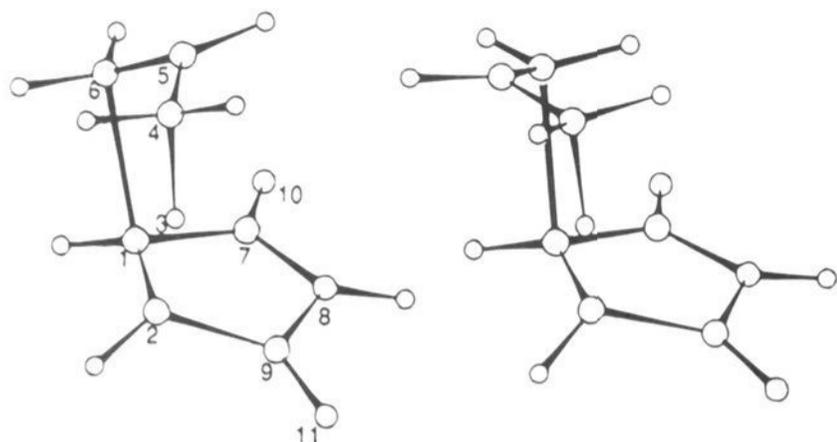
Figure 1. The *endo* (left) and *exo* transition states of the ene reaction between propene and maleimide.

Table 2. The Bond Lengths between the Heavy Atoms and the Migrating Proton of the *Endo* and *Exo* Transition States of the Reaction between Propene and Maleimide[a]

bond	*endo* ts	*exo* ts
C1−C2	1.407	1.405
C1−C6	2.018	2.079
C2−H3	1.532	1.488
C4−H3	1.304	1.324
C4−C5	1.403	1.401
C5−C6	1.386	1.389
C2−C9	1.484	1.490
C1−C7	1.506	1.503
C7−N8	1.375	1.378
N8−C9	1.392	1.389
C7−O10	1.190	1.189
C9−O11	1.189	1.189

[a] The structures were calculated at the HF/6-31G* level of accuracy.

Table 3. The Hybridization of the Heavy Atoms Forming the Bicyclic Ring Structure Studied with Respect to *Endo* and *Exo* Transition States[a]

atom	*endo* ts	*exo* ts
C1	2.30	2.27
C2	2.18	2.22
C4	2.23	2.22
C5	2.07	2.07
C6	2.17	2.16
C7	2.00	2.00
N8	2.03	2.02
C9	2.00	2.00

[a] Calculated at the HF/6-31G* level of accuracy.

To take entropy into account, we adopted the absolute entropies from the vibrational frequency calculations with respect to the HF/6-31G* optimized transition structures. At 600 K the *endo* and *exo* transition state entropies are 89.660 and 90.106 cal mol^{-1} K^{-1}, including translational and rotational contributions. Even at 600 K these entropies only amount to a difference in free energy of 0.27 kcal/mol (neglecting phase transitions). This difference is too small to change the preferred reaction route and the *endo* route is, in consequence, the favored one.

The synchronicity of the reaction was studied by analyzing the HF/6-31G* optimized transition structures (Figure 1). In both routes, the hybridization of C1 and C2 and the bond lengths indicate that the C−C bond formation slightly precedes the proton migration step (Tables 2 and 3).[23] The *endo* route is the more asynchronous one. Despite this small asynchronicity, the reaction mechanism still is clearly a concerted one.

The transition state structures, calculated at the AM1, 3-21G, and 6-31G* levels, were found to be very similar (Figure 2). Bond lengths and superimposition of the structures by fitting

(23) The equation sp^n = −1/cos α was used for the hybridization calculations. *n* is the hybridization number of an atom, α is the mean of the angles between the bonds connected to that atom.
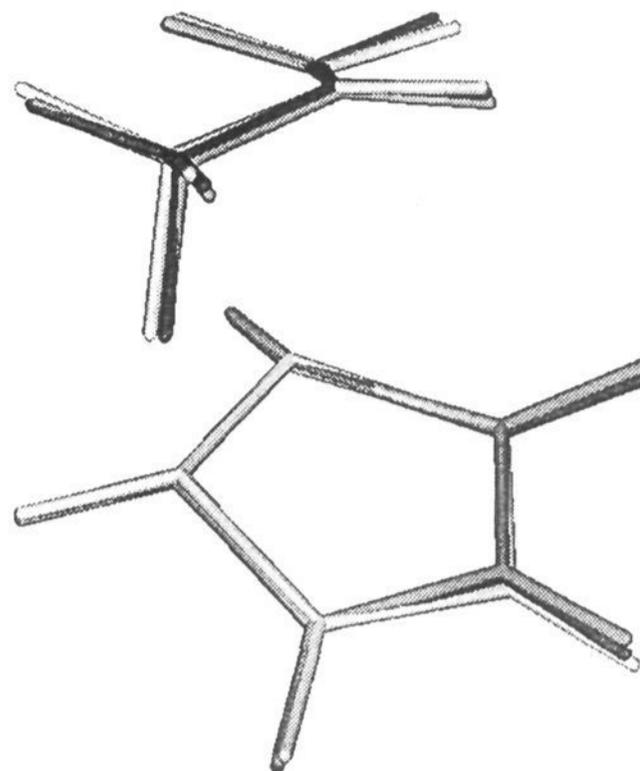
Figure 2. The *endo* transition state of the reaction between propene and maleimide, optimized at AM1 (white), 3-21G (grey), and 6-31G* (dark grey) levels of accuracy. The structures were superimposed using the rigid body, least-squares fitting. The geometry of the transition state was found to be independent on the computational method.

the carbons and the migrating proton forming the "reactive" 6-membered ring were used for comparison. The unsigned mean deviation (UMD) of the bond lengths between the 6-31G* and 3-21G transition states was less than 0.007 Å. The comparison between the 6-31G* and AM1 structures gives a UMD of less than 0.025 Å. The superimposition of the 3-21G transition states with the respective 6-31G* structures yields root-mean-square (RMS) deviations less than 0.02 Å. The AM1 structures, superimposed onto the 6-31G* transition states, still give RMS deviations below 0.06 Å. These results indicate that the transition state geometry is not, in practice, dependent on the method. The pyramidality of the nitrogen was found to be characteristic for the *ab initio* transition states. We attached a dummy atom (Du) to the nitrogen of the transition states and fixed the Du−N8−H angle to 90.0°. The AM1, 3-21G, and 6-31G* angles of Du−N8−C7 (equal to Du−N8−C9) were 90.1°, 92.3°, and 95.3°, respectively. It is known that semiempirical methods do not always treat nitrogen correctly.[10] However, this difference is not remarkable.

Design and Synthesis of the Hapten. Pauling suggested that molecules mimicking a reaction transition state could be successful inhibitors for the catalyzing enzyme.[24] Such analogs should have high affinity to the binding site of the enzyme and are thus good leads for derivatization. The concept of similarity and its applications to drug design have been studied extensively over the years.[25] The comparison of sterics, electron densities, and electrostatic potentials has been proposed. In addition to drug design, the use of similarity may aid the design of haptens to obtain catalytic antibodies.[26]

The activation energy for the reaction between maleimide and 1-alkenes is relatively high, particularly with a view to antibody catalysis. Therefore, the design of the hapten has to be done carefully. Here, we demonstrate the use of standard

(24) Pauling, L. *Chem. Eng. News* **1946**, *24*, 1375.

(25) See, for example: (a) Cooper, D. L.; Mort, K. A.; Allan, N. L.; Kinchington, D.; McGuigan, C. *J. Am. Chem. Soc.* **1993**, *115*, 12615. (b) Cioslowski, J.; Fleischmann, E. D. *J. Am. Chem. Soc.* **1991**, *113*, 64. (c) Meyer, A. Y.; Richards, W. G. *J. Comp. Aid. Mol. Design* **1991**, *5*, 427.

(26) (a) Richards, W. G.; Hodgin, E. E. *Chem. Br.* **1988**, *24*, 1141. (b) Tramontano, A.; Janda, K. D.; Lerner, R. A. *Science* **1986**, *234*, 1566 and references therein.

Scheme 2. The Structures of the Putative Transition State Analogs[a]

ts   a   b   c

d   e   f   g

h   i   j   k

l   m   n   o
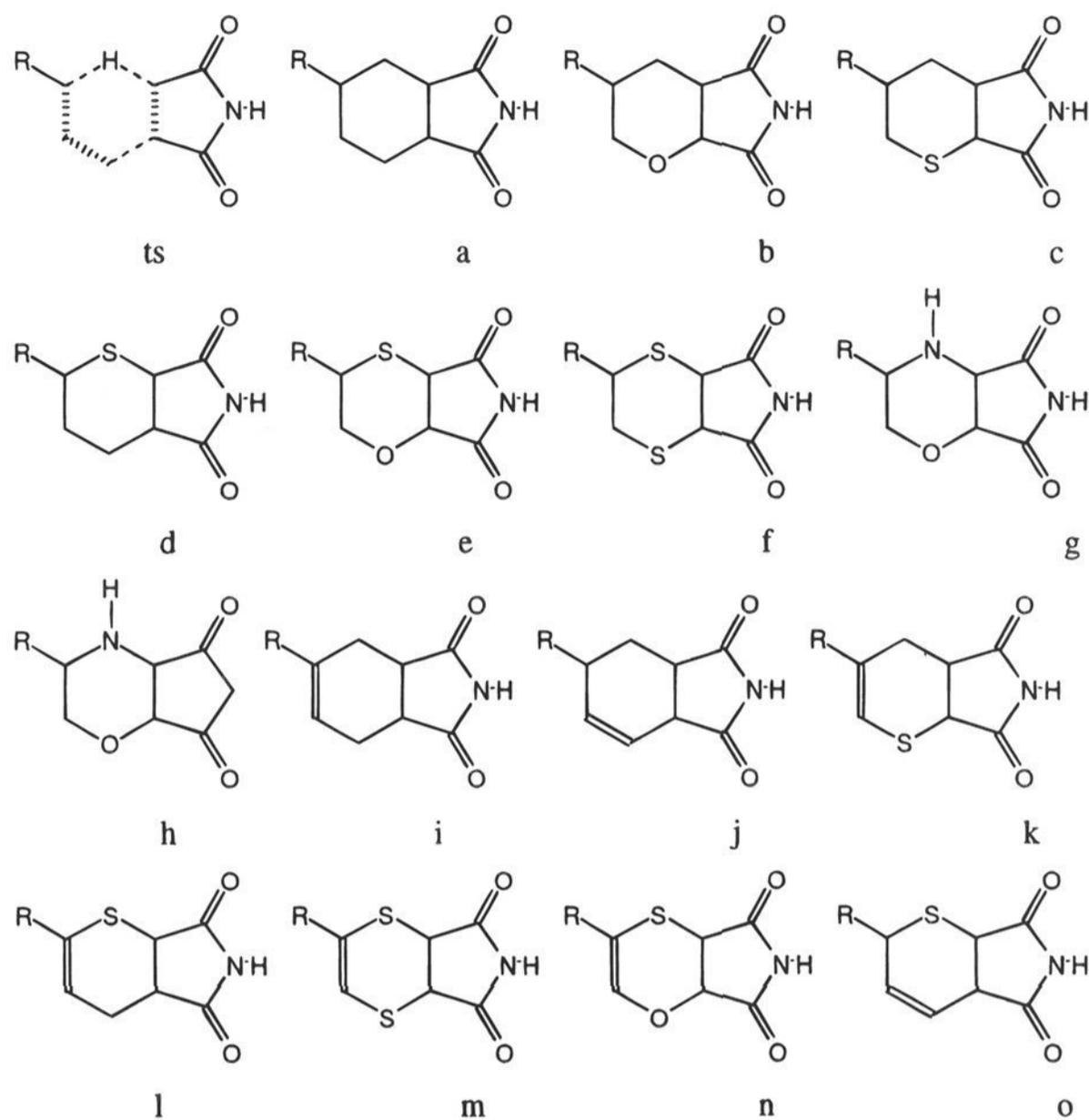
[a] The similarity of the structures was compared with the *endo* transition state of the reaction between butene and maleimide using the field fit option of the CoMFA module of SYBYL software.

drug design software for designing analogs with high similarity to the transition state. The synthesized analog is now used as a hapten for antibody generation.

To use propene as a model for the 1-alkenes is a reasonable approximation in view of the experimental results described above and reduces the computational effort considerably. For antibody generation, haptens have to be linked to a carrier protein via a side chain between the transition state analog and the protein. The 4-position of the transition state (Figure 1), i.e. the propene C3, is an obvious location for the linker because it is the place where the chain of 1-alkenes departs from the 6-membered ring of the transition state. We extended the 1-alkene reactant to 1-butene to account for the side chain orientation. As the AM1 transition structures were found to sterically agree closely with the corresponding HF/6-31G* transition states, we optimized the *endo* transition state of the reaction between 1-butene and maleimide at the AM1 level (Figure 3). As part of the transition state 1-butene has two possible orientations and we optimized the preferred transition state (*endo*) for both orientations of the butene C4 methyl group. With the methyl pointing in the "*exo*" direction, the AM1 activation energy (32.43 kcal/mol) was very close to that for the reaction with propene (32.51 kcal/mol). Placing the methyl group in the "*endo*" direction gave an activation energy of 31.22 kcal/mol and thus may indicate a moderate stabilization.

We intuitively designed 15 putative transition state analogs (Scheme 2). For modeling purposes, the side chain connecting the transition state analog to the carrier protein was truncated to a methyl. Both stereoisomers possessing a methyl group connected to a saturated carbon, as well as all analogs, were
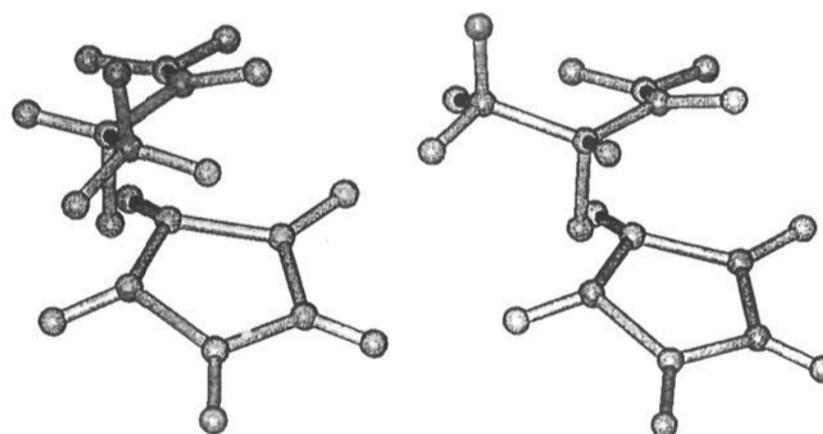
Figure 3. The AM1 *endo* transition state of the reaction between butene and maleimide was used as a basis of the design of the transition state analogs. The *endo* transition state has two forms, depending on the orientation of the terminal methyl group of the butene part.

optimized with AM1. For each candidate, the conformer with the best steric similarity between the bicyclic ring structure and the transition state was chosen visually for further studies.

There were three reasons for not basing the choice of conformer on the Boltzmann distribution and conformational energies. First, AM1 conformational energies are not very accurate and to calculate them at a higher level is prohibitive in terms of computer time. Second, solvent and physiological environment effects cannot be properly accounted for. Finally, even a non-optimal conformer will, if not too high above the optimal conformer, be somewhat populated. It will therefore produce antibodies on immunization but also extract antibodies when panning phage display libraries. In all cases the chosen conformer had an AM1 energy at most a couple of kcal/mol

Scheme 3. The Synthesis of the Haptens[a]
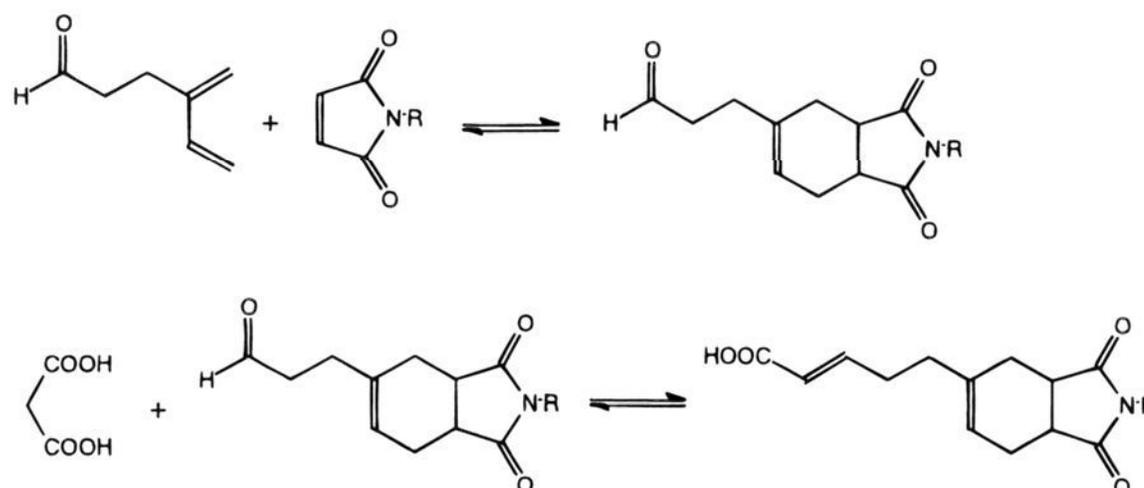
[a] The synthesis was carried out in two steps: a Diels—Alder reaction was followed by Knoevenagel condensation. R = H, CH$_3$

Table 4. The Similarity between the 15 Putative Analogs and the *Endo Transition*[a]

	separation of the lattice points (Å)				
	2.000	1.000	0.500	0.333	0.250
a	2.72 (3)	3.50 (4)	3.77 (3)	3.90 (3)	3.95 (3)
b	3.85 (14)	4.16 (13)	4.54 (13)	4.64 (12)	4.70 (12)
c	3.28 (7)	3.81 (10)	4.09 (7)	4.18 (7)	4.26 (7)
d	2.27 (1)	2.85 (1)	3.33 (1)	3.46 (1)	3.52 (1)
e	3.54 (10)	3.74 (8)	4.29 (11)	4.36 (8)	4.41 (8)
f	2.90 (4)	3.52 (5)	3.86 (4)	3.97 (4)	4.00 (4)
g	3.21 (6)	3.34 (2)	4.02 (5)	4.11 (5)	4.16 (5)
h	3.77 (13)	3.58 (6)	4.27 (8)	4.37 (9)	4.42 (9)
i	2.34 (2)	3.36 (3)	3.63 (2)	3.83 (2)	3.80 (2)
j	3.48 (9)	3.92 (11)	4.27 (8)	4.42 (11)	4.48 (11)
k	3.14 (5)	3.79 (9)	4.28 (10)	4.39 (10)	4.47 (10)
l	3.66 (12)	3.70 (7)	4.05 (6)	4.17 (6)	4.20 (6)
m	3.41 (8)	4.05 (12)	4.57 (14)	4.68 (14)	4.73 (14)
n	4.50 (15)	4.67 (15)	5.12 (15)	5.27 (15)	5.32 (15)
o	3.60 (11)	4.07 (14)	4.47 (12)	4.65 (13)	4.63 (13)

[a] The table lists field fit values obtained by rigid body field fit as implemented in the SYBYL software. The compounds are ranked according to similarity. The rankings are given in parentheses.

above the lowest conformer, and our choice of molecule was vindicated by the experimental data discussed below.

The quantitative comparison of transition states with putative analogs was carried out in two stages. First, we did a rigid-body, least-squares fitting between the transition states and the putative analogs. The nine heavy atoms forming the bicyclic rings and the carbon of the methyl group were superimposed on the respective atoms of the transition states. It turned out in all cases that the smallest sterical deviation was found for the transition state with the methyl group pointing to the *endo* direction. That particular transition state was chosen for further studies.

The final comparison was made using the field fit option of the CoMFA module in the Sybyl molecular modeling software.[27] We carried out several comparisons in a grid box of 10 × 10 × 10 Å^3, with 2.0-, 1.0-, 0,5-, 0.333-, and 0.25-Å separations between the lattice points. The field differences converge with decreasing grid spacing (Table 4), even if the field fit analysis appears somewhat sensitive to the spacing between lattice points.[28] A grid spacing of 0.5 Å seems an obvious compromise between accuracy and economy. The analysis gives the following trends: Oxygen, as compared to carbon, in the 6-position (for numbering, see Figure 1) is clearly disfavored. This observation, to some extent, applies to sulfur. In the

(27) (a) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959. (b) Allen, M. S.; Tan, Y-C; Trudell, M. L.; Naraynan, K.; Schindler, L. R.; Martin, M. J.; Schultz, C.; Hagen, T. J.; Koehler, K. F.; Codding, P. W.; Skolnick, P.; Cook, J. *J. Med. Chem.* **1990**, *33*, 2343.
(28) Calder, J. A.; Wyatt, J. A.; Frenkel, D. A.; Casida, J. E. *J. Comp. Aid. Mol. Design* **1993**, *7*, 45.
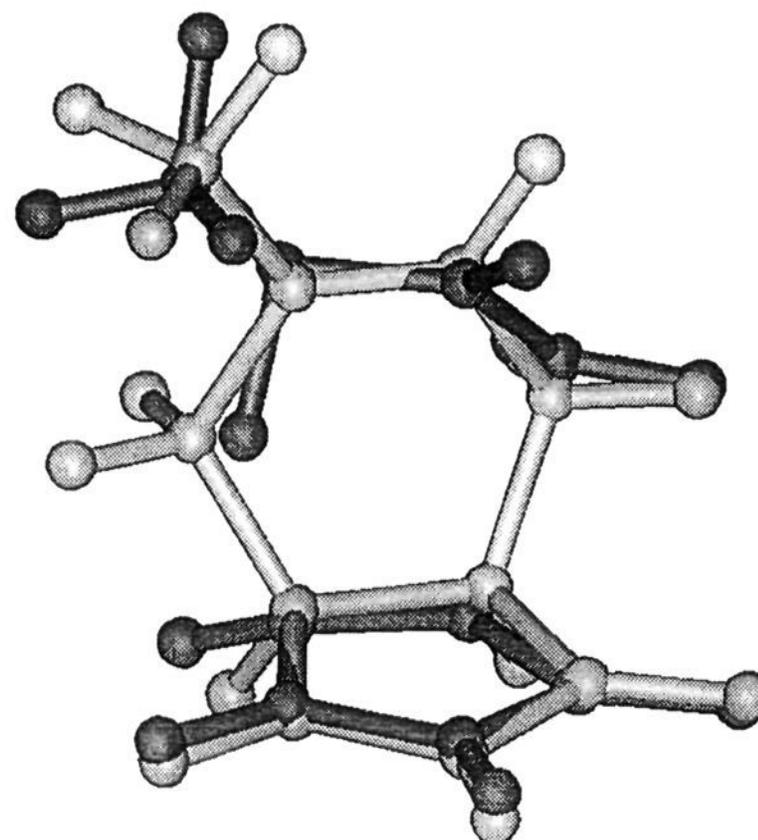
Figure 4. The analog i (light grey), chosen for synthesis, superimposed on the *endo* transition state (dark grey) of the reaction between 1-butene and maleimide.

3-position, sulfur, as compared to carbon, is favored, if the six-membered ring is saturated. Nitrogen at the 3-position also seems to be favored.

The analog i was chosen for synthesis. In addition to the results of the field fit, our decision also was based on the ease of synthesis and appropriate stability in the physiological environment. Moreover, the hybridization of C4 and C5 in the 6-31G* transition state of the reaction between propene and maleimide suggests double bond rather than single bond character. As a final criterion, the C4—C5 double bond reduces the number of synthesized stereoisomers, and the number of possible ring conformations is only two. The compound d has a small similarity advantage over compound i but was not chosen because its 6-membered ring possesses multiple iso- and conformers.

In the case of the analog i, pent-2-enoic acid was used as a suitable linker. The double bond between the two carbons rigidifies the side chain and thus keeps the bicyclic ring at a reasonable distance from the carrier.

The hapten was synthesized in two steps: The first step, a Diels—Alder reaction between maleimide and 4-methylene-5-hexenal,[29] was followed by Knoevenagel condensation between the formed aldehyde and malonic acid (Scheme 3).[30] The same procedure was repeated using *N*-methylated maleimide.

(29) Shea, K. J.; Burge, L. D. *J. Org. Chem.* **1988**, *53*, 318.
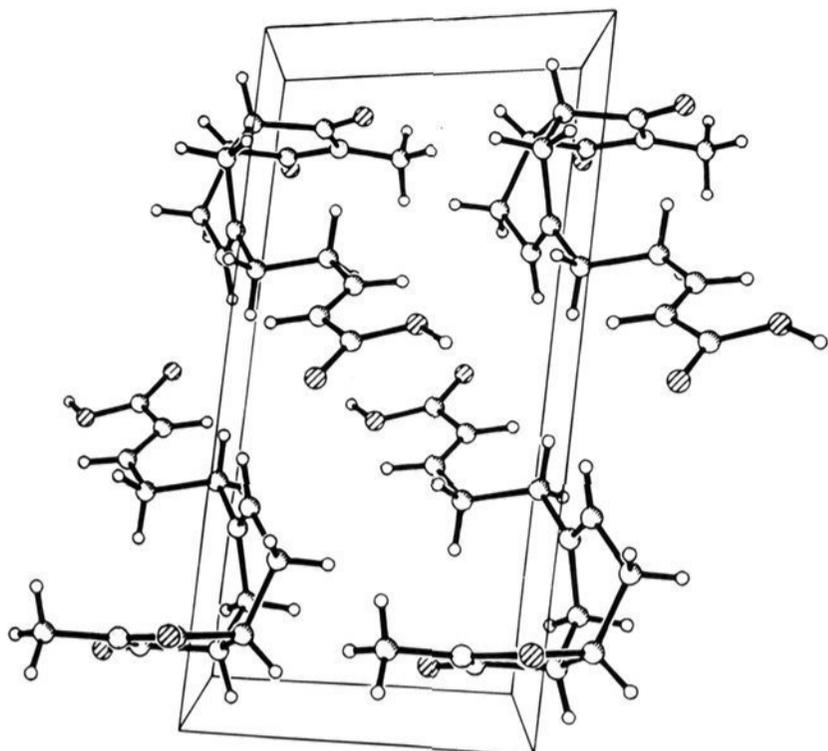
Figure 5. The X-ray structure of the *N*-methyl derivative of the hapten used for the immunization.
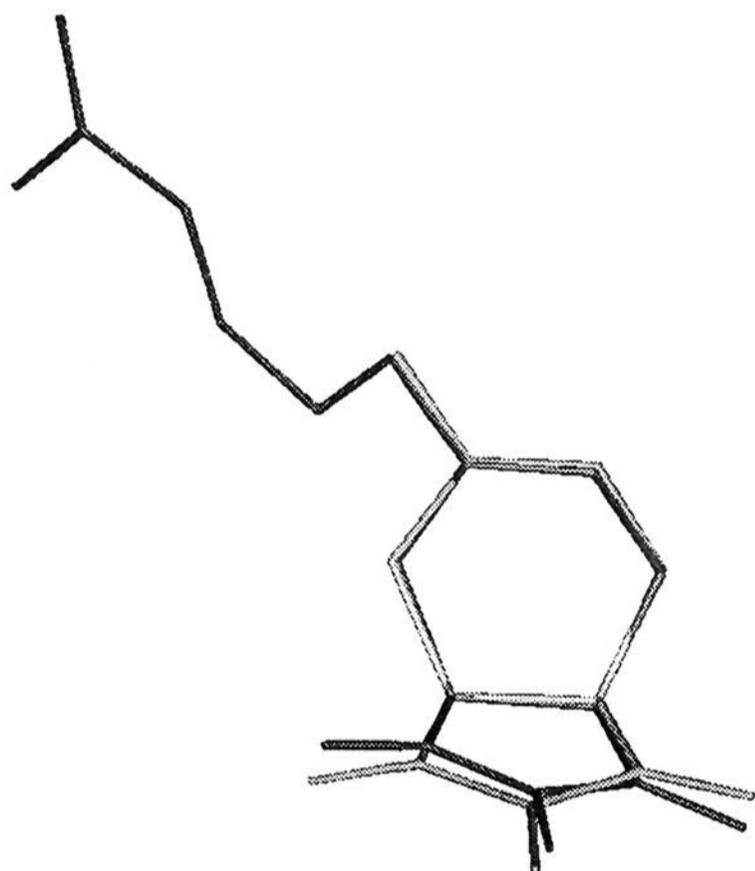
Figure 6. The X-ray structure (dark grey) superimposed on the AM1 structure (light grey) of *N*-methylated compound i. For clarity, only the heavy atoms are shown.

X-ray crystallography (Figure 5) of the *N*-methylated hapten shows the expected carboxylic acid dimer structure. The side chain exists in an extended conformation, suggesting a reasonable distance between the carrier protein and the transition state analog.

The AM1 structure of *N*-methylated compound i was compared with the crystal structure by superimposing the 9 heavy atoms forming the bicyclic rings of the molecules (Figure 6). The 6-membered rings overlap very well. The 5-membered ring of the crystal structure is slightly deformed, in contrast with the essentially planar 5-membered ring of the calculated AM1 structure. A possible explanation for the deformation is crystal packing. To check this, we first carried out a HF/6-31G* level optimization with respect to the AM1 optimized structure. The *ab initio* structure was found to be essentially

similar to the AM1 structure, with a planar 5-membered ring. A superimposition of the *ab initio* structure with the crystal structure gave an RMS deviation not better than in the case of the semiempirical structure. To study the effect of crystal packing on the geometry of the 5-membered ring, we carried out an AM1 optimization with respect to the X-ray structure, followed by an analysis of vibrational frequencies with the same method. The lowest frequency corresponded to the twisting of the 5-membered ring, thus explaining its deformation.

Conclusion

Using *ab initio* reaction modeling and standard drug design tools, we have designed and synthesized a transition state analog for use as a hapten for antibody generation. The approach presented is a rational way to design transition state analogs and can thus be applied to both hapten and enzyme inhibitor design. When studying transition states, we feel that the computational level should be calibrated against experimental data. This is necessary to establish a detailed reaction mechanism, and a precise 3-dimensional transition state structure. Once the transition state has been found, analog design is rather straightforward. In our case, superficially similar analogs were first visually inspected, followed by quantitative comparison to the transition state. Because of the rather high activation energy we used a quantitative approach in the hapten design, to ensure that the hapten mimics the transition state as well as possible. It is too early to say whether the stabilization of the transition state, caused by the antibody, is enough to conquer the high energy barrier. We feel, however, that the approach presented here is useful for those who want to apply transition state analogs in their research.

Experimental Section

Computational Methods. (a) Reaction Modeling. The reactions of maleimide with propene and butene were studied using molecular orbital methods. The Mopac 6.0 and Ampac 5.0 programs were used for the semiempirical calculations and Gaussian92 software was used for the *ab initio* calculations. The semiempirical optimizations were carried out using the AM1 model and the *precise* ($\times 100$) convergence criterion. The *ab initio* optimizations were done at the RHF/3-21G and RHF/6-31G* levels. The stability of the RHF solutions was verified by carrying out UHF level single point calculations with respect to the transition state structures. The geometries of the reactants, transition states, and products were fully optimized. The nature of the stationary points was assessed by vibrational frequency analysis. To verify that the calculated transition states connect the reactants and products, AM1 level intrinsic reaction coordinate following (IRC) calculations were done with respect to the transition states optimized with the same method. Correlation energy corrections were done by carrying out MP2/6-31G* single point calculations with respect to the RHF/3-21G and RHF/6-31G* optimized stationary points.

(b) Hapten Design. The structures of the 15 transition state analogs were optimized using the AM1 method. The optimized structures were compared with the AM1 *endo* transition state of the reaction between 1-butene and maleimide in two steps. First, a rigid-body least-squares fitting was done by superimposing the methyl group and the nine atoms forming the bicyclic rings. With the result of this superimposition as a starting point, a rigid-body field fit was done using the CoMFA module of the Sybyl 6.0.3 software (Tripos Associates Inc., 1699 S. Hanley Road, Suite 303, St. Louis, MO) on a Silicon Graphics Indigo 2 workstation. The field fit was done in a $10 \times 10 \times 10$ Å^3 grid box with separations of 2.0, 1.0, 0.5, 0.33, and 0.25 Å between lattice points. In addition to Sybyl, InsightII software (Biosym Technologies, 9685 Scranton Road, San Diego, CA 92121-2777), on a Silicon Graphics Indigo 2 workstation, was used as a graphical interface. Figures 2−4 and 6 were drawn using the Scarecrow program.[31]

(30) Vijn, R. J.; Hiemstra, H.; Kok, J. J.; Knotter, M.; Speckamp, W. N. *Tetrahedron* **1987**, *43*, 5019.

(31) Laaksonen, L. *J. Mol. Graphics* **1992**, *10*, 33.

Synthesis. Melting points were obtained on an Electrothermal melting point apparatus in open capillary tubes and are uncorrected. Mass spectra were run on a JEOL JMS-SX102 instrument and NMR spectra on a Varian Gemini 200-MHz spectrometer. Yields refer to chromatographically and spectroscopically (^{1}H NMR) homogeneous materials, unless otherwise stated.

3-(1,3-Dioxo-2,3,3a,4,7,7a-hexahydro-1*H*-isoindol-5-yl)propanal (1a). Maleimide (485 mg, 5 mmol) and 4-methylene-5-hexenal[29] (551 mg, 5 mmol) were dissolved in 2.5 mL of dry benzene. After 20 h at ambient temperature, the solvent was removed under vacuum and the crude product recrystallized from pentane−ethyl acetate (6:1) to afford **1a** (740 mg, 71%): mp 82 °C; ^{1}H NMR (200 MHz, CDCl$_3$) δ 2.33 (m, 8 H, aliphatic), 3.12 (m, 2 H, H-3 and H-3a), 5.55 (m, 1 H, H-6), 9.68 (s, 1 H, CHO); ^{13}C NMR (50.3 MHz, CDCl$_3$) δ 24.4 (t), 27.6 (t), 29.6 (t), 40.7 (t), 41.3 (d), 41.6 (d), 121.5 (d), 138.9 (s), 180.5 (s), 180.6 (s), 202.0 (d); HRMS for C$_{11}$H$_{13}$NO$_3$, calcd 207.0895, found 207.0904.

3-(1-Methyl-1,3-dioxo-2,3,3a,4,7,7a-hexahydro-1*H*-isoindol-5-yl)propanal (1b). Using the above procedure with *N*-methylmaleimide (555 mg, 5 mmol) and 4-methylene-5-hexenal (551 mg, 5 mmol) afforded **1b** (974 mg, 88%) as an oil: ^{1}H NMR (200 MHz, CDCl$_3$) δ 2.39 (m, 8 H, aliphatic), 2.93 (s, 3 H, NCH$_3$), 3.07 (m, 2 H, H-3 and H-3a), 5.55 (m, 1 H, H-6), 9.68 (s, 1 H, CHO); ^{13}C NMR (50.3 MHz, CDCl$_3$) δ 24.5 (q), 25.4 (t), 27.6 (t), 29.6 (t), 39.5 (t), 40.0 (d), 41.48 (d), 121.5 (d), 138.8 (s), 179.9 (s), 180.6 (s), 201.8 (d); HRMS for C$_{12}$H$_{15}$NO$_3$, calcd 221.2560, found 221.2574.

5-(1,3-Dioxo-2,3,3a,4,7,7a-hexahydro-1*H*-isoindol-5-yl)pent-2-enoic Acid (2a). The aldehyde **1a** (606 mg, 3 mmol) and malonic acid (420 mg, 4 mmol) were dissolved in pyridine (2 mL). After 3 h at 90 °C, the reaction mixture was poured into a 2 N HCl solution (30 mL) and extracted with ether (3 × 20 mL). The combined organic layers were washed with brine (2 × 25 mL) and filtered, and the filtrate was dried (Na$_2$SO$_4$) and evaporated in vacuo. Recrystallization of the residue from ethyl acetate−ether (1:1) gave **2a** (480 mg, 63%): mp 160 °C, ^{1}H NMR (200 MHz, CDCl$_3$) δ 2.22 (m, 6 H, aliphatic), 2.52 (m, 2 H, CH$_2$), 3.08 (m, 2 H, H-3 and H-3a), 5.58 (m, 1 H, H-6), 5.78 (d, J = 15.59 Hz, 1 H, =CH), 6.82 (dt, J = 15.59, 13.18 Hz, 1 H, =CH), 10.05 (br s, 1 H, COOH); ^{13}C NMR (50.3 MHz, CDCl$_3$) δ 23.9 (t), 27.2 (t), 29.8 (t), 35.4 (t), 40.3 (d), 40.8 (d), 121.0 (d), 122.2

(d), 138.8 (s), 148.2 (d), 168.4 (s), 180.5 (s), 180.7 (s); HRMS for C$_{13}$H$_{15}$NO$_4$, calcd 249.1001, found 249.0987.

5-(2-Methyl-1,3-dioxo-2,3,3a,4,7,7a-hexahydro-1*H*-isoindol-5-yl)pent-2-enoic Acid (2b). Using the above procedure with the aldehyde **1b** (664 mg, 3 mmol) and malonic acid (420 mg, 4 mmol) afforded **2b** (590 mg, 75%) as white crystals: mp 90 °C (from ethyl acetate−pentane, 2:1); ^{1}H NMR (200 MHz, CDCl$_3$) δ 2.40 (m, 8 H, aliphatic), 2.93 (s, 3 H, NCH$_3$), 3.07 (m, 2 H, H-3 and H-3a), 5.55 (m, 1 H, H-6), 5.75 (d, J = 15.59 Hz, 1 H, =CH), 6.92 (dt, J = 15.59, 13.18 Hz, 1 H, =CH), 10.5 (br s, 1 H, COOH); ^{13}C NMR (50.3 MHz, CDCl$_3$) δ 24.5 (q), 25.4 (t), 27.5 (t), 30.1 (t), 35.6 (t), 39.5 (d), 40.0 (d), 121.5 (d), 121.7 (d), 139.0 (s), 151.3 (d), 171.9 (s), 180.5 (s), 180.7 (s); HRMS for C$_{14}$H$_{17}$NO$_4$, calcd 269.2932, found 269.2909.

X-ray Analysis of 2b. X-ray crystallography was performed on a 0.2 × 0.2 × 0.4 mm^3 crystal with Mo Kα radiation (λ = 0.71069 Å) on a Rigaku AFC7S diffractometer. The triclinic (space group P_1) cell parameters are a = 7.975(3) Å, b = 13.169(4) Å, c = 6.854(3) Å, α = 94.03(3)°, β = 111.28(3)°, γ = 95.51(4)°, V = 663.3(4) Å^3, Z = 2. Data were collected at 213 K using the $\omega/2\Theta$ scan technique. A total of 2405 unique reflections were collected. Lorentz−polarization and decay corrections were applied to the data. The structure was solved by the direct-methods program SHELXS-86 with full-matrix least-squares refinement.

Supplementary Material Available: Tables of positional parameters and anisotropic thermal parameters for (**2b**) and hydrogen coordinates and isotropic displacement parameters for **1** (5 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA943541V