Computer Chemistry

DOI: 10.1002/anie.200603293

Computer-Aided Rational Design of Catalytic Antibodies: The 1F7 Case**

S. Martí, J. Andrés, E. Silla, V. Moliner, I. Tuñón,* and J. Bertrán*

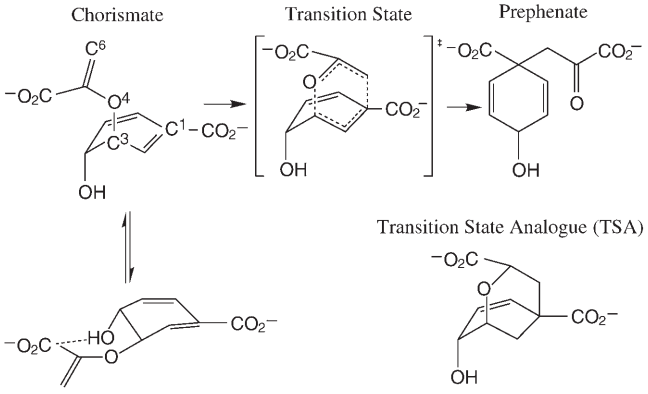
Two main strategies have been raised in recent years to improve the catalytic power of catalytic antibodies (CA):[1] the rational-design approach consists of direct mutation of residues on selected specific positions on the active site of the protein; and directed evolution in vitro consists of search of sequence space, iterative cycling of variation, and selection.[2,3] All of these strategies can be combined as proposed by some other authors.[4–6]

Herein we show a computer-aided rational-design protocol to improve the efficiency of CAs, in particular the 1F7, which catalyzes the chorismate to prephenate rearrangement. The information derived from theoretical and computational-chemistry techniques has allowed the proposal of mutations at the active site of the 1F7 that have previously never been proposed and that enhance the rate constant of this chemical step.

The usual starting point of rational design is the X-ray crystal structure of the protein together with a stable molecule that resembles the structure of the transition state (TS) of the chemical reaction (transition-state analogue, TSA); however, this structure does not correspond to the real protein–substrate TS. Two serious drawbacks appear: first, the specific interaction patterns established between the substrate in its TS and the residues of the active site do not exactly match that found in the TSA; and second, the static picture obtained from X-ray crystallography techniques does not reflect the flexibility of the protein. These limitations lead to a partial knowledge of the complete molecular mechanism of the catalyzed chemical reaction inside the protein,[2,7] which would need to be completed with additional strategies.

From a theoretical point of view, the combination of methods and techniques of quantum mechanics to treat bond-forming and bond-breaking processes, together with molecular-mechanics methods (QM/MM),[8] that include the whole solvent and protein environment effects in the simulations, can be used to study enzyme reaction mechanisms.[9] This methodology combined with molecular-dynamics simulation techniques allow the exploration of relevant configurations in the system in the different states and incorporates the flexibility of the protein during the course of the chemical reaction. Thus, all the specific substrate–protein interactions that are established in the active site of the protein and, in particular, at the TS that controls the chemical step can be analyzed.[10] The knowledge of this pattern of interactions will provide the clues to decide which residues of the active site, a CA, or a protein scaffold should be replaced to better stabilize the TS and to enhance the rate constant of the chemical step of a full catalytic process.[11] Furthermore, the free-energy profile can be traced from reactants to products, through the corresponding TS,[12,13] rendering theoretical predicted barriers comparable with experimental data.

Herein, we show that this methodology can be used as a computer-aided rational-design protocol to overcome some of the limitations of standard rational-design techniques and that it is being tested for the chorismate to prephenate simple metabolic reaction (see Scheme 1). This reaction is accelerated more than a millionfold by chorismate mutase (CM) enzymes, which are a key step in the shikimate pathway for



Scheme 1. The molecular mechanism of the chorismate to prephenate rearrangement and details of the oxabicyclic TSA originally used to elicit the immune response that generated the 1F7 CA.

[*] Dr. S. Martí, Prof. J. Andrés, Prof. E. Silla,[+] Dr. V. Moliner
Departament de Ciències Experimentals
Universitat Jaume I
Box 224, Castellón (Spain)
Fax: (+34) 964-728-066
E-mail: moliner@exp.uji.es
Dr. I. Tuñón
Departament de Química Física
Universidad de Valencia
46100 Burjasot (Spain)
Fax: (+34) 963-544-564
E-mail: Ignacio.Tunon@uv.es
Prof. J. Bertrán
Departament de Química
Universitat Autònoma de Barcelona
08193 Bellaterra (Spain)
[+] Permanent address:
Departament de Química Física
Universidad de Valencia (Spain)

[**] This work was supported by DGI project BQU2003-04168-C03, BANCAIXA project P1·1B2005-13, and Generalitat Valenciana project GV06/152 and GV06/21. We acknowledge the Servei d'Informàtica of the Universitat Jaume I for providing us with computer assistance.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

biosynthesis of the aromatic amino acids in bacteria, fungi, and higher plants.[14,15] CMs from different organisms, such as *Bacillus subtilis* (BsCM)[16] or *E. coli* (EcCM),[17] exhibit similar kinetic properties, although they may share little sequence similarity. Furthermore, a CA with modest chorismate mutase activity was prepared against a TSA of CM[18,19] (see Scheme 1), and its three-dimensional structure was determined at 3.0-Å resolution.[20]

The free-energy profiles obtained for the chorismate to prephenate rearrangement in aqueous solution, in the two chorismate mutase enzymes (BsCM and EcCM), and in the 1F7 CA are depicted in Figure 1 (details of the calculations

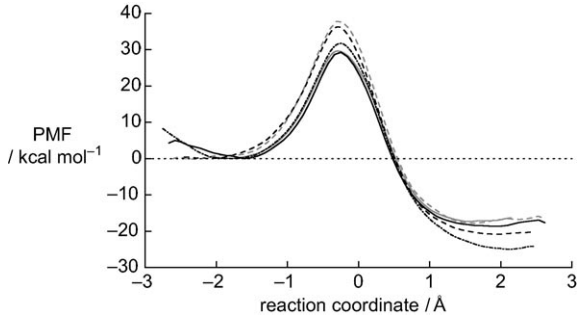


Figure 1. Free-energy profiles (in terms of potential of mean force, PMF) for the chorismate to prephenate rearrangement obtained in the different environments: BsCM (——), EcCM (——), 1F7 (----), 1F7 (N33S) mutant (–·–), and in aqueous solution (----). The reaction coordinate is the antisymmetric combination of the interatomic distances of the breaking and forming bonds, C3···O4 and C1···C6, respectively.

are given in the Experimental Section), whereas our best estimation of the free-energy barriers are listed in Table 1. The resulting profiles are in accordance with the expected results: the catalytic efficiency of the 1F7 appears between the catalytic power of both enzymes (which are very close to each other) and the reaction studied in solution.

The most remarkable result in Table 1 is that the computed catalytic power of the different tested proteins ($\Delta\Delta G^{\neq}_{\text{theor.}}$) is in very good agreement with that of experimental ($\Delta\Delta G^{\neq}_{\text{expt}}$). This fact validates the employed methodology, giving encouragement for its use in obtaining a deeper insight into the catalysis in the enzymes and 1F7. Further to free-energy barriers, Figure 1 can be used to identify the

Table 1: Theoretical free-energy barriers (in kcal mol^{-1}) for the uncatalyzed chorismate to prephenate rearrangement compared with the catalyzed reaction by BsCM, EcCM, 1F7, and 1F7 (N33S) mutant CA.

	Thermal rearrangement	BsCM	EcCM	1F7	1F7 (N33S)
$\Delta G^{\neq}_{\text{theor.}}$	29.3[a]	20.6[a]	20.9	27.5	23.0
$\Delta\Delta G^{\neq}_{\text{theor.}}$	0.0	−8.7	−8.4	−1.8	−6.3
$\Delta G^{\neq}_{\text{expt}}$	24.5[b]	15.4[b]	17.2[c]	21.6[d]	–
$\Delta\Delta G^{\neq}_{\text{expt}}$	0.0	−9.1	−7.3	−2.9	–

[a] Values are taken from reference [13]; [b] Values are taken from reference [16]; [c] Values are taken from reference [17]; [d] Values are taken from reference [18].

position of the TS and the Michaelis complex along the reaction coordinate, defined as the antisymmetric combination of the breaking and forming bonds, C3−O4 and C1−C6, respectively. Considering that the values of the reaction coordinate of the different TSs are very close (see Figure 1), if the difference of the reaction coordinates between the TS and the Michaelis complex is small, the pre-equilibrium of the substrate (see Scheme 1) will be displaced towards the chairlike structure of the chorismate. In this regard, although the free-energy profile is rather flat in the minimum region, it can deduce a direct relationship between the difference in the reaction coordinate at the Michaelis complex and the TS and the value of the free-energy barrier.[21] The smallest differences are obtained in the enzymatic processes (1.4 Å and 1.5 Å for the BsCM and EcCM, respectively), whereas the largest difference are obtained in the solvent environment (2.3 Å). An intermediate value of 1.8 Å was determined for 1F7, thus fitting the order in the free-energy barriers. Roughly speaking, although water molecules fit to the substrate structure in solution, the protein induces conformational changes in the substrate.

The analysis of the averaged structures obtained in the different biological systems allows determination of which interactions favor the stabilization of the TS. Figure 2 presents the averaged interaction energy, electrostatic and van der Waals contributions, of individual residues with the substrate at the corresponding TSs (snapshots of the BsCM, EcCM, and the 1F7, representative of the TS, are provided in the Supporting Information). The first conclusion that can be derived from these figures is that in both enzymes, BsCM and EcCM, the favorable interactions take place through the positively charged residues (mostly arginine residues) with the two negatively charged carboxylate groups of the substrate and the negative charge that develops on the ether oxygen.[10,22–26] Those interactions with negatively charged amino acids are, in general, nonfavourable. Furthermore, the pattern of interactions obtained in both enzymes for the TS complex is generally quite similar to the one deduced from the X-ray protein–TSA complexes (see, for instance, reference [20]). Concerning the 1F7, the magnitude of all the interactions is dramatically smaller than in the enzymes except for the interaction established with ArgH95, which presents similar values to the enzymatic ones. It seems that the 1F7 does not properly interact with the two carboxylate groups of the substrate, leaving them partially exposed to the solvent. This can be confirmed by the strong interactions established with the water molecules that are accessible to the cavity (see Figure 2c), which are much smaller than in either of the two CMs (Figure 2a,b). This result suggests that the substrate fits better in the enzyme active sites than in the CA pocket, which is in agreement with the previous observation concerning the reaction coordinate values in the Michaelis complex: much closer to the TS in the CMs than in the 1F7 case.

It is also important to point out that the pattern of interactions obtained in the TS of the 1F7 is not equal to the TSA–CA structure determined by the X-ray diffraction study (for a graphical analysis, see figures provided in the Supporting Information). The experimentally obtained structure of

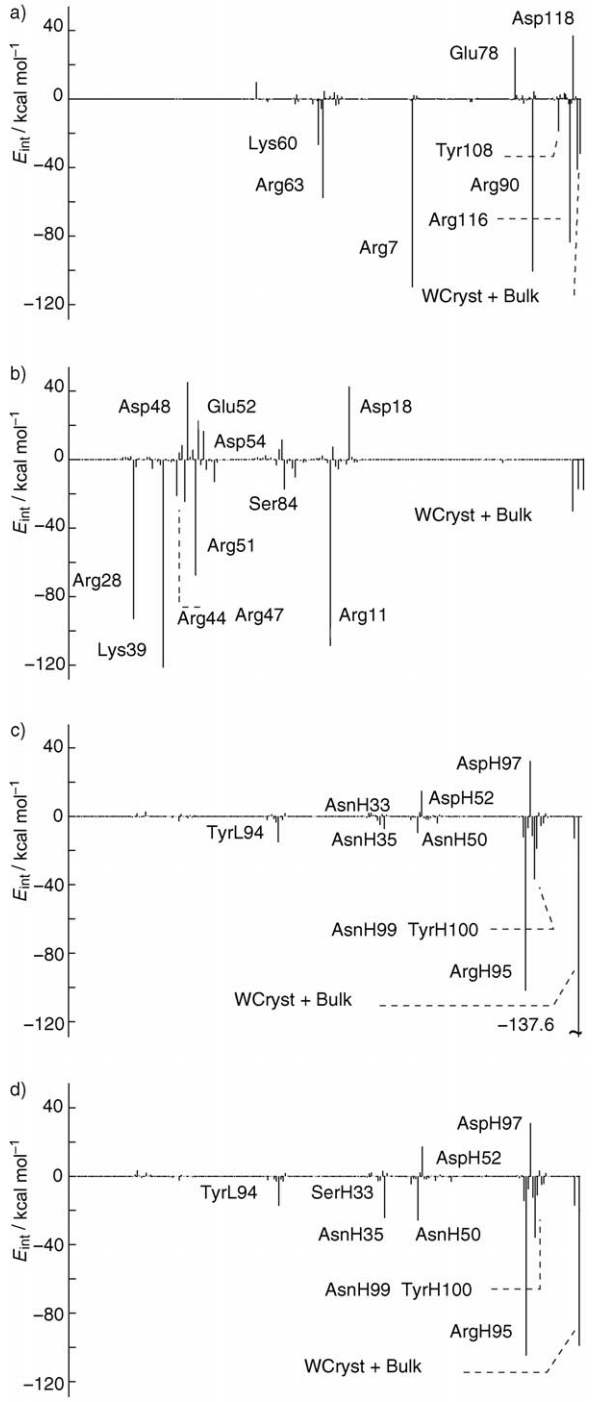
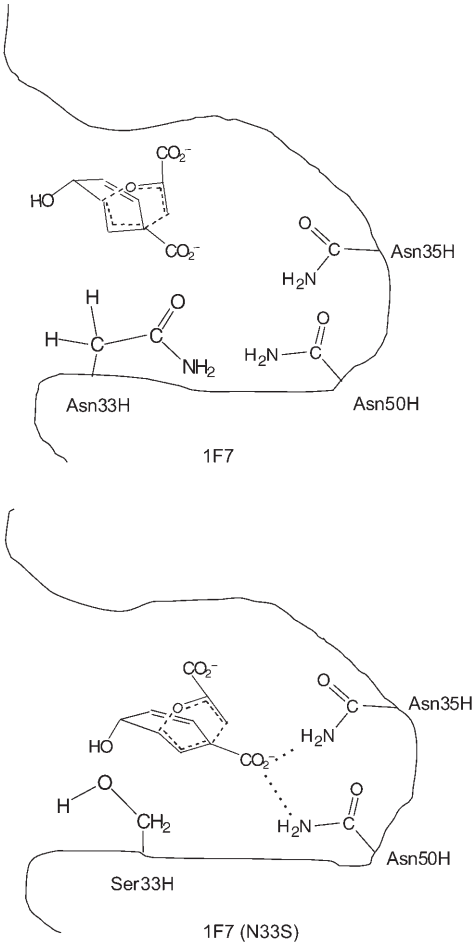


Figure 2. Contributions of individual amino acid residues (ordered along the *x* axis) to the TS interaction (in kcal mol^{-1}) of the a) BsCM, b) EcCM, c) 1F7, and d) 1F7 (N33S) mutant. E_{int} = substrate–protein interaction energy, WCryst + Bulk = crystallization and bulk water molecules.

the CA–TSA complex is appreciably different to the TS structures located in the CA active site. Thus, for instance, AsnH33 presents a noticeably different orientation in the TS–CA with respect to the TSA–CA complex: although the hydrogen atoms of the amino group interact with the hydroxy group of the inhibitor in the later, in the TS–CA complex, there are steric interactions between this residue and the aliphatic hydrogen atoms of the substrate that impede optimum positioning of the substrate into the cavity. As a result, strong interactions between the carboxylate group and amino acid residues of the inner part of the cavity are prevented, as is depicted in Scheme 2. The low capability of



Scheme 2. The substrate–protein interactions in the active site of the 1F7 and the 1F7 (N33S) CAs at their respective TSs.

the 1F7 to enhance the rate constant of the chorismate to prephenate rearrangement can be understood from this analysis. The strong stabilizing interactions observed in the enzyme between both carboxylate groups and the protein are not reproduced by the immune-system process when eliciting antibodies against a stable molecule that resembles, but is not equal to, the TS of the desired chemical transformation.

From the conclusions obtained by comparing BsCM and EcCM with 1F7 (see above), we can propose and check mutations that improve the efficiency of the 1F7 CA. Thus, we changed the AsnH33 with a serine residue that would facilitate a better accommodation of the substrate in the cavity of the CA owing to its smaller size, presumably enhancing the interactions of the substrate with the residues located in the inner part of the cavity. Once this mutation was carried out, the free-energy profile of the new 1F7-N33S, also

presented in Figure 1, was obtained by using the same procedure as in the previous calculations. This mutation yields a noticeable decrease in the free-energy barrier, in comparison with the PMF obtained for 1F7. The corresponding activation free energy reported in Table 1 is 4.5 kcal mol^{-1} lower than the original 1F7 CA and only 2.4 kcal mol^{-1} above the most efficient BsCM enzyme. This diminution would imply an increase in the rate constant by a factor of 10^3 at room temperature, compared with 1F7 CA. Stabilization of the TS, as a consequence, preferentially selects and optimizes those reactant conformers that resemble the TS, thereby displacing the pre-equilibrium to the reactive reactant conformers. The reaction-coordinate difference between reactants and TS is now 1.6 Å, a value closer to the enzymes than to that calculated for the 1F7. The analysis of the substrate–protein interactions in the TS, presented in Figure 2 d, reveals that our predictions have been confirmed, at least in the computed structures and energies; the new CA presents a more favorable pattern of interactions than the 1F7. The most important effect of the mutation is the extra space generated in the cavity, allowing the ring of the substrate to slightly rotate and its carboxylate groups to optimize the interactions with the available residues between the TS and the cavity (see Scheme 2). In particular, the interactions established between the carboxylate group and residues such as AsnH35 and AsnH50 are stronger in the mutated CA. Simultaneously, the interactions with the water molecules are reduced, which is similar to the situation observed in the CMs. The steric hindrance of the AsnH33 with the substrate in 1F7 prevents this movement, whereas in the mutated CA, a combination of the smaller size of the residue and the weaker interaction established with the substrate facilitates a more favorable relative orientation in the CA cavity, thus reducing the free-energy barrier of the chemical step.

It has been suggested that the limited structural diversity of the immune system imposes inherent limitations on catalytic efficiency.[27] The present work shows how our methodology, combined with other experimental strategies, may be used to determine whether the antibody scaffolds are evolutionary dead ends or can be further improved, as seems to be the case for 1F7. The study of TS–protein complexes, which have been demonstrated not to be equal to the TSA–protein structures obtained from experimental techniques, can be used to decide which residues should be changed in the active site of the CA to reduce the free-energy barrier of the catalyzed chemical transformation. Computer-aided rational design might be used, not only as a first step for directed laboratory evolution experiments, but also to shed some light on the divergent evolution of enzyme superfamilies.

Experimental Section

Our computational studies started from the X-ray crystal structures of BsCM,[16] EcCM,[17] and 1F7,[20] all of which contained the inhibitor or the product in the active site. The geometries of the TSA (in the active site of the EcCM and 1F7) or the product (in the BsCM) were then modified to resemble the gas-phase transition structure of the chorismate to prephenate rearrangement, and the full systems were then placed in a simulation water box of 79.5 Å on the side. In the case of 1F7 and 1F7-N33S, the part of the protein that lies outside the box was removed and the boundary residues were kept frozen in subsequent simulations. The substrate was described by using quantum mechanics at the AM1 level,[28] whereas for the rest of the system we employed the OPLS-AA[29] and TIP3P[30] force fields with a cut-off radius for the nonbonded interactions of 14.5 Å. By using the DYNAMO program,[31] the systems were relaxed and the corresponding transition structures were located and characterized, including the full environment. The systems were equilibrated at 300 K by using the NVT ensemble and the Langevin–Verlet integrator. Potentials of mean force were obtained as a function of the reaction coordinate defined as the antisymmetric combination of the C3···O4 and C1···C6 distances. A parabolic potential with a force constant of 2500 kJ mol^{-1} Å^{-1} was used to restrain this coordinate at particular values in the range between the reactants and products in a series of 70 simulation windows. Each of these windows consisted of 5 ps of equilibration and 10 ps of production, with a time step of 1 fs. WHAM[32] was then used to obtain the full probability distribution function. The free-energy barriers were taken as the PMF difference between the maximum (the TS) and the minimum (reactants). These values were corrected to take into account the deficiencies of the AM1 method, adding the gas-phase difference between the B3LYP/6-31G* energy barrier, as implemented in the Gaussian03 package of programs,[33] and the AM1. The same simulation protocol was also used for the aqueous solution reaction. The 1F7-N33S system was constructed manually to convert AsnH33 into Ser and equilibrating the system over 500 ps before tracing the corresponding PMF. Average properties were obtained by means of 500-ps simulations in the reactant and transition states. Individual contributions of each residue to the interaction energy were computed by using the polarized wave function of the substrate and averaged during the last 100 ps.

Received: August 11, 2006
Revised: September 9, 2006
Published online: November 24, 2006

Keywords: catalytic antibody · chorismate–prephenate rearrangement · computer chemistry · enzymes · quantum mechanics/molecular mechanics

[1] *Catalytic Antibodies* (Ed.: E. Keinan), Wiley-VCH, Weinheim, **2005**.
[2] C. Gustafsson, S. Govindarajan, R. Emig, *J. Mol. Recognit.* **2001**, *14*, 308 – 314.
[3] K. L. Morley, R. J. Kazlauskas, *Trends Biotechnol.* **2005**, *23*, 231 – 237.
[4] T. M. Penning, J. M. Jez, *Chem. Rev.* **2001**, *101*, 3027 – 3046.
[5] L. Yuan, I. Kurek, J. English, R. Keenan, *Mol. Biol. Rep.* **2005**, *69*, 373 – 392.
[6] J. Shanklin, *Curr. Opin. Plant Biol.* **2000**, *3*, 243 – 248.
[7] Z. Shao, F. H. Arnold, *Curr. Opin. Struct. Biol.* **1996**, *6*, 513 – 518.
[8] A. Warshel, H. Levitt, *J. Mol. Biol.* **1976**, *103*, 227 – 249.
[9] V. Moliner, A. J. Turner, I. H. Williams, *Chem. Commun.* **1997**, 1271 – 1272.
[10] M. Roca, S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, J. Bertrán, *Chem. Soc. Rev.* **2004**, *33*, 98 – 107.
[11] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, J. Bertrán, *Angew. Chem.* **2005**, *117*, 926 – 931; *Angew. Chem. Int. Ed.* **2005**, *44*, 904 – 909.
[12] M. J. Field, *A Practical Introduction to the Simulation of Molecular Systems*, Cambridge University Press, Cambridge, **1999**.
[13] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, J. Bertrán, M. J. Field, *J. Am. Chem. Soc.* **2001**, *123*, 1709 – 1712.

[14] E. Haslam, *Shikimic Acid: Metabolism, Metabolites*; Wiley, New York, **1993**.
[15] U. Weiss, J. M. Edwards, *The Biosynthesis of Aromatic Compounds*, Wiley, New York, **1980**.
[16] Y. M. Chook, H. Ke, W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8600 – 8603.
[17] Y. M. Lee, P. A. Karplus, B. Ganem, J. Clardy, *J. Am. Chem. Soc.* **1995**, *117*, 3627 – 3628.
[18] D. Hilvert, K. D. Nared, *J. Am. Chem. Soc.* **1988**, *110*, 5593 – 5594.
[19] D. Hilvert, S. H. Carpenter, K. D. Nared, M. T. M. Auditor, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 4953 – 4955.
[20] M. R. Haynes, E. A. Stura, D. Hilvert, I. A. Wilson, *Science* **1994**, *263*, 646 – 652.
[21] N. A. Khanjin, J. P. Snyder, F. M. Menger, *J. Am. Chem. Soc.* **1999**, *121*, 11831 – 11846.
[22] S. V. Taylor, P. Kast, D. Hilvert, *Angew. Chem.* **2001**, *113*, 3408 – 3436; *Angew. Chem. Int. Ed.* **2001**, *40*, 3310 – 3335.
[23] P. Kast, M. Asif-Ullah, N. Jiang, D. Hilvert, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 5043 – 5048.
[24] A. Kienhöfer, P. Kast, D. Hilvert, *J. Am. Chem. Soc.* **2003**, *125*, 3206 – 3207.
[25] Y. S. Lee, S. E. Worthington, M. Krauss, B. R. Brooks, *J. Phys. Chem. B* **2002**, *106*, 12059 – 12065.
[26] B. Szefczyk, A. J. Mulholland, K. E. Ranaghan, W. Sokalski, *J. Am. Chem. Soc.* **2004**, *126*, 16148 – 16159.
[27] A. C. Backes, K. Hotta, D. Hilvert, *Helv. Chim. Acta* **2003**, *86*, 1167 – 1174.
[28] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902 – 3909.
[29] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225 – 11236.
[30] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926 – 935.
[31] M. J. Field, M. Albe, C. Bret, F. Proust-de Martin, A. J. Thomas, *J. Comput. Chem.* **2000**, *21*, 1088 – 1100.
[32] G. M. Torrie, J. P. Valleau, *J. Comput. Phys.* **1977**, *23*, 187 – 199.
[33] M. J. Frisch et al., Gaussian 03, Revision A.1, Gaussian, Inc.; Pittsburgh PA, **2003**.

