Comparison of Experimental and Theoretical Structures of a Transition State Analogue Used for the Induction of Anti-Cocaine Catalytic Antibodies

Edward C. Sherer,[†] Ginger Yang,[‡] Gordon M. Turner,[†] George C. Shields,*[,†] and Donald W. Landry*[,‡]

The Department of Chemistry, Lake Forest College, Lake Forest, Illinois 60045, and The Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York 10032

Received: April 16, 1997; In Final Form: September 3, 1997[⊗]

Comparison of the crystal structure of a transition state analogue that was used to raise catalytic antibodies for the benzoyl ester hydrolysis of cocaine with structures calculated by ab initio, semiempirical, and solvation semiempirical methods reveals that modeling of solvation is crucial for replicating the crystal structure geometry. Both SM3 and SM2 calculations, starting from the crystal structure TSA I, converged on structures similar to the crystal structure. The 3-21G(*)/HF, 6-31G*/HF, PM3, and AM1 calculations converged on structures similar to each other, but these gas-phase structures were significantly extended relative to the condensed phase structures. Two transition states for the hydrolysis of the benzoyl ester of cocaine were located with the SM3 method. The gas phase calculations failed to locate reasonable transition state structures for this reaction. These results imply that accurate modeling of the potential energy surfaces for the hydrolysis of cocaine requires solvation methods.

Introduction

Cocaine has been used by over 40 000 000 Americans since 1980 and addiction afflicts at least 2 000 000[1] with disastrous medical and social consequences.[2,3] Despite extensive effort, no antagonist to the reinforcing effect of cocaine has been identified. An alternative to receptor-based approaches to cocaine addiction would be to interfere with the delivery of cocaine to the central nervous system so that a dose of cocaine no longer had a reinforcing effect. Immunological approaches to substance abuse originally were proposed[4a] in the 1970s and are under renewed consideration.[4b,4c] However, simple binding antibodies are limited since high doses of the abused substance can irreversibly bind and inactivate the available circulating antibodies. Thus we developed anti-cocaine catalytic antibodies, a novel class of artificial enzymes,[5−7] with the capacity to bind and degrade cocaine and thereby become available for further binding.[8,9]

The ability of a given molecule to elicit a catalytic antibody stems from the structural correspondence between an enzyme−transition state complex and an antibody binding to the corresponding transition state analogue.[10] A catalytic antibody stabilizes the transition state structure (TS) relative to the ground state structure and lowers the energy of activation for the intended reaction. The art of generating catalytic antibodies lies in the design of the hapten, which must be chemically stable, yet mimics the structural and electronic properties of the transition state of the modeled reaction.

Our most active catalytic antibody for the hydrolysis of the benzoyl ester of cocaine enhances the rate by 23 000-fold.[9] The transition state analogue (TSA) which elicited this catalytic antibody is based on the phosphonate monoester structure, a structure used extensively to obtain artificial esterases.[5,7,11,12] In this paper we report the crystal structure of this TSA and compare the experimental results with calculated structures at semiempirical and ab initio Hartree−Fock (HF) levels of theory. In addition, we describe transition state structures for the catalysis of cocaine's benzoyl ester, the TS which the phosphonate TSA is designed to mimic. The successes and failures of various computational methods are discussed. Comparison of the results of the TSA and TS calculations along with the crystal structure results allows assessment of the different levels of theory that can be used in a computational approach to TSA design.

Methods

X-ray Crystallography. The transition state analogue (TSA) was prepared according to the method published.[9] Ethyl acetate (5 mL) was added dropwise to a solution of the hydrobromide salt of the TSA (50 mg) in 0.5 mL of methanol until the mixture turned slightly turbid. A few drops of methanol were added to clarify the solution. The vial was loosely capped and kept at room temperature. Large crystals appeared after 1 day. Because the crystals were deliquescent, all further operations were performed under a dry, inert atmosphere. After the removal of supernatant, the crystals were rinsed with ethyl acetate (1 mL) and dried under vacuum. The crystal was mounted under a stream of N_2.

Crystal data were collected at room temperature on a Rigaku AFC5R diffractometer with graphite monochromated Cu Kα radiation and a 12 kW rotating anode generator. The structure was solved by Patterson techniques, and absolute configuration was determined by refining the Flack parameter. The hydrogen atoms on the nitrogens and in the hydroxyls attached to the phosphorus were determined from electron density difference maps. All other hydrogen atoms were placed at calculated positions and were not refined. Full details of structure determination, as well as atomic coordinates, are available as Supporting Information.

Quantum Mechanics. We have used AM1,[13] PM3,[14] SM2,[15] and SM3[16] semiempirical methods along with 3-21G(*)/HF[17] and 6-31G*/HF[18] ab initio methods, implemented within the SPARTAN 4.1 software package,[19] to calculate the structures of the TSAs. Initially, semiempirical methods and the 3-21G-(*)/HF method were used, starting with both crystal structures

* Corresponding authors.
[†] Lake Forest College.
[‡] Columbia University College of Physicians and Surgeons.
[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1997.

S1089-5639(97)01320-0 CCC: $14.00

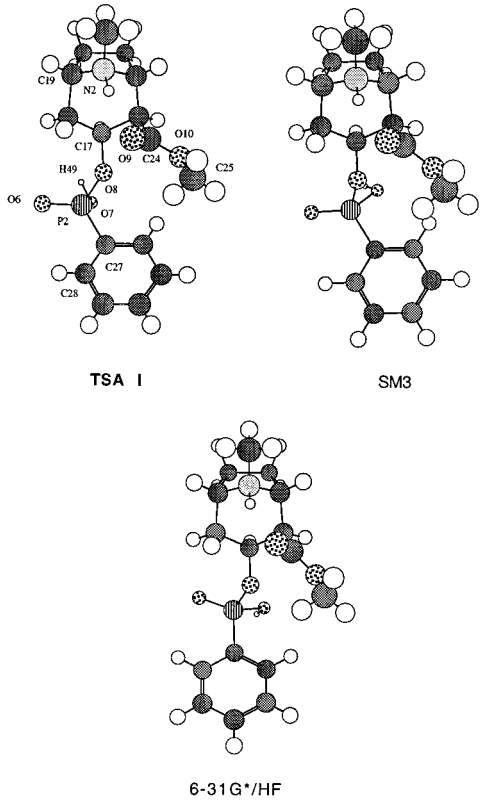

Figure 1. Crystal Structure of a transition state analogue that was used to raise catalytic antibodies for the benzoyl ester hydrolysis of cocaine, TSA I, and computed structures obtained with the SM3 and 6-31G*/HF methods.

as input. The 6-31G*/HF calculation was started from the 3-21G(*)/HF global minimum structure. In addition semiempirical conformational searches were performed within SPARTAN, starting from both crystal structures, to ensure that the lowest energy conformer was located. Extensive searches of the potential energy surface to locate transition states for the hydrolysis of the benzoyl ester of cocaine were performed using the AM1, PM3, 3-21G/HF, and SM3 methods.

Comparison of electrostatic potentials between calculated TSAs and the crystal structure were performed in the following manner. First, the different molecules were superimposed based on their molecular volumes. Then the best fit pair of structures as determined by molecular volume were used to compare the two molecular electrostatic potentials, computed with the MNDO method, grid point by grid point. The similarity between the MNDO electrostatic potentials of two different molecules was determined as the root mean square (rms) difference between the two grids. The grid resolution was 0.13 Å between any two consecutive grid points. This two step procedure was necessary as direct superposition of the electrostatic potentials for two different molecules was unreliable within SPARTAN.

Results

Figure 1 displays the best TSA structure determined by single-crystal X-ray diffraction, TSA I, along with its computed molecular structures obtained by geometry optimizations using the SM3 and 6-31G*/HF computational methods. The SM2 and SM3 methods are the continuum model solution methods developed by Cramer and Truhlar for the AM1 and PM3 semiempirical Hamiltonians. In their methodology, the solute−solvent system is simplified by modeling the solvent as an infinite medium characterized by the same dielectric constant as the bulk solvent. This dielectric continuum represents a time-averaged solvent environment, and these continuum methods are useful for examining the effect of the solvent on the solute.[15,16] Figure 2 displays transition state structures for the first step in the alkaline hydrolysis of the benzoyl ester of the protonated tropane nitrogen form of cocaine.

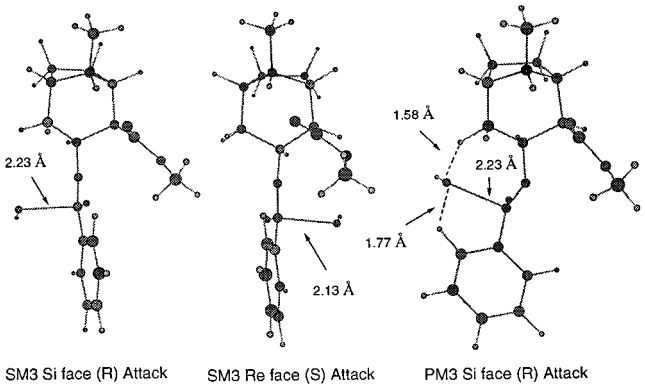

Figure 2. Transition state structures for the first step in the alkaline hydrolysis of the benzoyl ester of the protonated tropane nitrogen form of cocaine. Reasonable transition state structures were obtained with the SM3 method for attack of OH^- on the Si face of the benzoyl carbonyl of cocaine, leading to the *R* stereoisomer intermediate, and for attack on the Re face leading to the *S* stereoisomer intermediate.

Discussion

Comparison of X-ray and Calculated Structures. The X-ray structures reveal that the experimental TSA is positively charged, with an oxygen and a hydroxide attached to the central phosphorus and with the tropane nitrogen protonated. Thus the TSA used for the hapten conforms to the protonated tropane nitrogen form of cocaine, which is in equilibrium with the free cocaine base in aqueous solution, in a ratio of 15:1 at physiological pH. Both stereoisomers were crystallized, TSA I, shown in Figure 1, and TSA II, where O6 and O7-H49 switch positions. TSA II has a bromine anion and an unidentified, positionally disordered, solvent molecule near the phenyl ring. The solvent molecule was initially assigned as ethyl acetate; however, it broke apart upon refinement and several of the atoms had unrealistic temperature factors. After analysis of several difference maps, only four atoms could be located and refined. Two of the atoms were assigned to be methanol, and the other two were assigned to be water oxygens, as the distance between the atoms was too long for a reasonable C−O bond distance. Because most of the largest remaining peaks in the final electron density difference maps are in this area, close to the phenyl group of TSA II, and because the carbon atoms of the TSA II phenyl ring had higher temperature factors than all other atoms in both TSAs, TSA I was identified as the best molecular structure in this crystal. In addition, both gas phase and solvated semiempirical calculations resulted in substantial changes in the position of the phenyl ring for TSA II.

A wide body of evidence suggests that computed structures containing C, N, O, and P atoms are generally quite accurate when the 6-31G* basis set is used in a Hartree−Fock (HF) calculation.[20,21] Comparison of the X-ray structure TSA I with calculated structures using AM1, PM3, SM2, SM3, 3-21G(*)/HF, and 6-31G*/HF levels of theory reveal that although the ab initio methods accurately model bond distances (as do AM1 and PM3, though to a lesser degree of accuracy) only geometry optimization using solvation methods accurately captures the overall shape of the condensed phase crystal structure. All of the gas-phase calculations (AM1, PM3, 3-21G(*)/HF, 6-31G*/HF) result in structures such as the 6-31G*/HF structure

displayed in Figure 1. The difference between the crystal structure TSA I and the gas-phase structures is revealed by analysis of two torsion angles. The $C_{17}-O_8-P_2-C_{27}$ torsion angle is 172° in TSA I, but ranges from −170° to −176° for the gas-phase structures, a change of 12−18°. In addition, the $O_8-P_2-C_{27}-C_{28}$ torsion angle is 135° in TSA I, but ranges from 109 to 125° for the gas-phase structures, a change of 10−26°. The result is a more extended structure for all of the gas phase models, as the phenyl ring points directly away from the tropane nitrogen. Solvation and crystallization favor a conformation that is less extended, allowing the phenyl group to be in closer proximity to the methyl ester of the TSA.

A consideration of bond distances reveals that the 6-31G*/HF structure has only three bonds that exceed three esd (esd: estimated standard deviations) from the crystal structure values: O_9-C_{24}, $O_{10}-C_{25}$, and $C_{22}-C_{23}$. With the SM3 method the same three bond distances are outside three esd, along with three others: O_7-P_2, O_8-P_2, and N_2-C_{19}.

Comparison of the heavy atoms in the TSA reveal that the SM3 structure has an rms difference from TSA I of 0.30 Å, the SM2 structure calculated from the crystal structure has an rms difference from TSA I of 0.39 Å, while the 6-31G*/HF structure has an rms difference from TSA I of 0.53 Å. The corresponding rms values for the PM3 and AM1 calculations are 0.61 and 0.75 Å. Similarly, comparison of the electrostatic potentials, as outlined in Methods, for each structure yields rms difference values of 3.2, 3.3, and 4.4 kcal/mol for the SM3, SM2, and 6-31G*/HF TSA structures relative to TSA I. The corresponding values for the PM3 and AM1 calculations are 5.0 and 5.5 kcal/mol. It is clear that the use of solvation methods are esential for success in obtaining accurate geometries for the TSA.

The question arises, what is it about this molecule that requires a dielectric representation of solvent to adequately model the crystalline environment? Starting the SM3 calculation from the PM3 or 6-31G*/HF minimum structures yields the same result, returning the benzoyl ring to the less extended structure shown in Figure 1. It is reasonable to presume that the electrostatic interactions in the positively charged hapten are responsible for the conformational differences found by gas phase and aqueous phase calculations, with the large electrostatic charge in the gas-phase resulting in the extended structure, while the screening of the dielectric solvent allows for a different conformation. To further examine the influence of electrostatic interactions for this system, we changed the charge in our model by (i) removing the proton H49 from O7, making a zwitterionic structure (see Figure 1) and by (ii) removing the hydrogen from the tropane N, making a nonzwitterionic structure. Both of these two neutral models were geometry-optimized with the PM3 and SM3 methods.

The results of the zwitterionic calculation (removal of H49) revealed that there is virtually no change in the coordinates in the SM3 calculation; the rms deviation between the heavy atoms of the SM3 neutral structure and the TSA I crystal structure is 0.29 Å. The PM3 method improved only slightly, with the rms deviation between the neutral PM3 structure and the TSA I crystal structure now 0.51 Å. The gas-phase results are still linear and do not converge to the solvated structures. Removal of the hydrogen from the tropane nitrogen (see Figure 1) makes a neutral structure that is not a zwitterion. We have calculated the PM3 and SM3 geometry-optimized structures to evaluate the performance of the models for this nonzwitterionic neutral structure. The results are similar to the zwitterionic structure, with the SM3 structures converging on the previous SM3 results and the PM3 structures, starting from the SM3 minima, converging back to a similar structure as the original gas-phase result for TSA I. These results have important implications for modeling the cocaine system, as cocaine itself has a p*K*a of 8.65, and consideration of the protonation state of the tropane nitrogen is always required when attempting to model cocaine and its related haptens in physiologic (pH 7.4) solution. Since the predominant form of cocaine is protonated, formation of the transition state structure will make a zwitterionic species, and the results presented here imply that the transition state structure must be modeled with continuum methods. However, it is possible for an antibody to form around a TSA which does not have the tropane protonated, and the results reveal that this case must be modeled by solvation methods as well.

Transition State Calculations. The nucleophilic hydroxide ion can approach from either the Si or Re face of the carbonyl, as assigned according to the Cahn, Ingold, and Prelog nomenclature, forming two stereoisomer intermediates (*R* or *S*). We have previously calculated the TS for the attack of the hydroxide ion on the Si face of the neutral form of cocaine, using the PM3 method.[22] Extensive and careful searches of the TS potential surface for OH^- attack from both sides of the benzoyl carbon located only one TS structure, a structure which was stabilized by the well-known tendency of the PM3 Hamiltonian to underestimate repulsions between nonbonded atoms.[23−25] For this TS the nearest hydrogen from the benzoyl ring is separated by 1.709 Å from the incoming hydroxyl oxygen and the closest hydrogen from the tropane ring is 1.736 Å away from the same oxygen.

Our latest calculations searching for transition state structures have focused on the attack of OH− on the protonated form of cocaine, the species that predominates at physiological pH, since this is the form of cocaine that an antibody is most likely to encounter. Figure 2 displays the two transition state structures for attack at the benzoyl ester Si and Re faces, calculated with the SM3 method. In addition, this figure contains a PM3 TS (Si face) found from a careful search starting from the SM3 TS. The PM3 TS shows the same behavior as reported for the TS when OH^- attacks the free base of cocaine, with hydrogens stabilizing the TS structure. The SM3 structures on the other hand, are free from such distortions and were much easier to locate on the SM3 potential energy surface. We failed to locate any TS structures with AM1 and 3-21G(*)/HF calculations. Quite likely these results reveal that hydrolysis of cocaine in the gas-phase proceeds without an activation energy barrier, similar to results for methyl acetate and methyl benzoate ester systems.[26,27] For the model ester systems discussed previously, PM3 and 3-21+G/HF TS structures were located which yielded negative activation energies in the gas phase. These transition state structures become real activation energy barriers in solution, a consequence of the drastic lowering of energy for the OH^- anion when placed in a dielectric field. It has been shown previously that it is the desolvation of the hydroxyl anion which is responsible for the activation energy barrier for ester hydrolysis in solution.[28] The activation energies estimated from the SM3 calculations of OH^-, protonated cocaine, and the TS structures are 15.3 kcal/mol for attack of the Si face and 18.0 kcal/mol for attack at the Re face of cocaine (Figure 2), which is in good agreement with the experimental range of 12−15 kcal/mol for alkaline hydrolysis of esters in solution.[29,30] The success of the SM3 method in locating reasonable transition state structures is another indication that the cocaine system, with its high electrostatic charge, must be modeled with solvation methods in order to have an accurate representation of the entire potential energy hypersurface. Studies of additional

pathways for cocaine hydrolysis and comparison of calculated transition state structures with potential TSAs are underway.

Conclusions

Comparison of the crystal structure of a transition state analogue that was used to raise catalytic antibodies for the benzoyl ester hydrolysis of cocaine with structures calculated by ab initio, semiempirical, and solvation semiempirical methods shows that modeling of solvation is crucial. Both SM3 and SM2 calculations, starting from the crystal structure TSA I, converged on structures with rms deviations of 0.30 and 0.39 Å, respectively. In contrast, the 6-31G*/HF, PM3, and AM1 calculations converged on structures with rms deviations of 0.53, 0.61, and 0.75 Å respectively. The same trend was observed upon comparing electrostatic potentials. Clearly solvation must be included in geometry optimizations to obtain accurate structures of the phosphonate haptens used to induce catalytic antibodies for cocaine hydrolysis. In addition, transition state structures that TSA I is meant to model were located using solvation methods, allowing us to estimate the activation energy of cocaine hydrolysis of the benzoyl ester as 15−18 kcal/mol. These results imply that accurate modeling of the potential energy surfaces for hydrolysis of cocaine, transition state analogues, and the transition states themselves, should be modeled using solvation methods. We attribute the sensitivity of the cocaine system to a surrounding dielectric as a direct result of the high electrostatics of these molecules. This sensitivity results in one set of intramolecular interactions for an isolated molecule, which are then modulated by the effects of solvation.

Acknowledgment. This work was supported by Lake Forest College, an NSF ILI grant (G.C.S.), and by the Executive Office of the President of the United States of National Drug Control Policy (DWS).

Supporting Information Available: Crystal data, intensity measurements, details of structure solution and refinement, fractional atomic coordinates, anisotropic displacement parameters, bond lengths, bond angles, torsion angles, nonbonded contacts, structure factors, pdb files for TSA I and TSA II, plot of 2θ vs X-ray intensity, and Ortep drawings of TSA I and TSA II (71 pages). Ordering information is given on any current masthead page.

References and Notes

(1) *The National Drug Control Strategy*; Office of National Drug Control Policy, Executive Office of the President of the United States: Washington DC, 1996; 41−51.

(2) Cregler, L. L.; Mark, H. *N. Engl. J. Med.* **1986**, *315*, 1495−1500.

(3) Lesko, L. M.; Fischman, M. W.; Javaid, J. I.; Davis, J. M. *N. Engl. J. Med.* **1982**, *307*, 1153.

(4) (a) Bonese, K. F.; Wainer, B. H.; Fitch, F. W.; Rothberg, R. M; Schuster, C. R. *Nature*. **1974**, *252*, 708−710. (b) Rocio, M.; Cerrera, A.; Ashley, J. A.; Parsons, L. H.; Wirsching, P.; Koob, G. F.; Janda, K. D. *Nature* **1995**, *378*, 727−730. (c) Fox, B. S.; Kantak, K. M.; Edwards, M. A.; Black, K. M.; Bollinger, B. K.; Botka, A. J.; French, T. L.; Thompson, T. L.; Schad, V. C.; Greenstein, L.; Gefter, M. I., Exley, M. A.; Swain, P. A.; Briner, T. J. *Nature Medicine* **1996**, *2*, 1129−1132.

(5) Tramontano, A.; Janda, K. D.; Lerner, R. A. *Science* **1986**, *234*, 1566−1570.

(6) Benkovic, S. J. *Annu. Rev. Biochem* **1992**, *61*, 29−54.

(7) Pollack, S. J.; Jacobs, J. W.; Schultz, P. G. *Science* **1986**, *234*, 1570−1573.

(8) Landry, D. W.; Zhao, K.; Yang, G. X.-Q.; Glickman, M.; Georgiadis, T. M. *Science* **1993**, *259*, 1899−1901.

(9) Yang, G.; Chun, J.; Arakawa-Uramoto, H.; Gawinowicz, M. A.; Zhao, K.; Landry, D. W. *J. Am. Chem. Soc.* **1996**, *118*, 5881−5890.

(10) Lerner, R. A.; Benkovic, S. J.; Schultz, P. G. *Science* **1991**, *252*, 659−667.

(11) Tramontano, A.; Ammann, A. A.; Lerner, R. A. *J. Am. Chem. Soc.* **1988**, *110*, 2282−2286.

(12) Janda, K. D.; Benkovic, S. J.; Lerner, R. A. *Science* **1989**, *244*, 437−440.

(13) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3989.

(14) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−221; 221−264.

(15) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aid. Mol. Des.* **1992**, *6*, 629−666.

(16) Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **1992**, *13*, 1089−1097.

(17) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797−2803.

(18) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724−728.

(19) SPARTAN version 4.1; Wavefunction, Inc., 18401 Von Karman Ave., Suite 370, Irvine, CA, 92715, 1995.

(20) Hehre, W. J.; Radom, L.; Schleyer, P.v. R.; Pople, J. A. *Ab initio Molecular Orbital Theory*; John Wiley and Sons: New York, 1986.

(21) Hehre, W. J. *Practical Strategies for Electronic Structure Calculations.;* Wave function: Irvine, CA, 1995.

(22) Sherer, E. C.; Turner, G. M.; Lively, T. N.; Landry, D. W.; Shields, G. C. *J. Mol. Model.* **1996**, *2*, 62−69.

(23) Buss, V; Messinger, J.; Heuser, N. *QCPE Bull.* **1991**, *11* (1), 5.

(24) Jurema, M. W.; Shields, G. C. *J. Comput. Chem.* **1993**, 14, *89*−*104*.

(25) Csonka, G. I.; Angyan, J. G. *J. Mol. Struct. (THEOCHEM)* **1997**, 393, 31−38.

(26) Sherer, E. C.; Turner, G. M; Shields, G. C. *Int. J. Quantum Chem., Quantum Biol. Symp*. **1995**, *22*, 83−93.

(27) Turner, G. M.; Sherer, E. C.; Shields, G. C. *Int. J. Quantum Chem., Quantum Biol. Symp*. **1995**, *22*, 103−112.

(28) Dewar, M. J. S.; Storch, D. M. *J. Chem. Soc., Chem. Commun*. **1985**, 94−96.

(29) Amis, E. S.; Siegel, S. *J. Am. Chem. Soc.* **1950**, *72*, 674−679.

(30) Rylander, P. N.; Tarbell, D. S. *J. Am. Chem. Soc.* **1950**, *72*, 3021−3025.