JOC Article
pubs.acs.org/joc

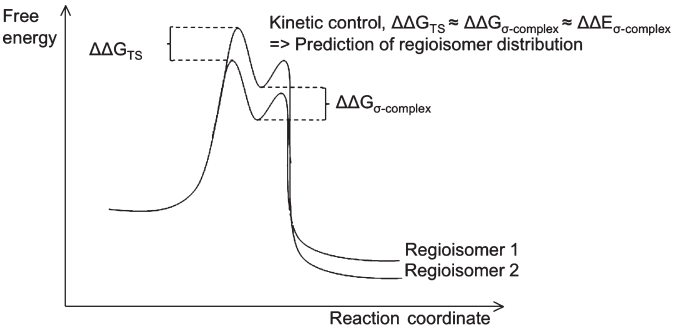
Validation of a Computational Model for Predicting the Site for Electrophilic Substitution in Aromatic Systems

Magnus Liljenberg,[†] Tore Brinck,*[,‡,§] Björn Herschend,[¶] Tobias Rein,[¶] Glen Rockwell,[¶] and Mats Svensson*[,⊥,∥]

[†]*Sweden Operations, AstraZeneca, and* [¶]*Pharmaceutical Development,* [⊥]*Medicinal Chemistry, AstraZeneca R&D, S-151 85 Södertälje, Sweden, and* [‡]*Department of Physical Chemistry, Royal Institute of Technology, S-100 44 Stockholm, Sweden*

tore@physchem.kth.se; mats.a.svensson@astrazeneca.com

Received February 18, 2010



We have investigated the scope and limitations of a method for predicting the regioisomer distribution in electrophilic aromatic substitution reactions that are under kinetic control. This method is based on calculation of the relative stabilities of the σ-complex intermediates using density functional theory. Predictions from this method can be used quantitatively for halogenations; it agreed to an accuracy of about 1 kcal/mol with experimental observations in 10 of the 11 investigated halogenation reactions. For nitrations, the method gave useful predictions for heterocyclic substrates. The method failed for nitration of monosubstituted benzenes, and we expect that more elaborate model systems, including explicit solvent molecules, will be necessary to obtain quantitatively useful predictions for such cases. For Lewis acid promoted Friedel−Crafts acylations, the method can be expected to give qualitatively correct predictions, that is, to point out the dominating isomer. For substrates where the regioisomeric outcome is highly dependent on the reaction conditions, the method can only be of qualitative use if the concentration of the free Lewis acid is high during the reaction. We have also compared the predictive capacity of the method to that of a modern reactivity index, the average local ionization energy, $I(\mathbf{r})$. The latter method is found to predict the regisolectivity in halogenations and nitrations qualitatively correctly if the positions for the $I(\mathbf{r})$ minima ($I_{S,min}$) are not too sterically hindered but fails for qualitative predictions of F−C reactions. The downscaled $I_{S,min}$ values also perform well for the quantitative prediction of regioisomer distributions of halogenations. The accuracy is slightly lower than that for the new method.

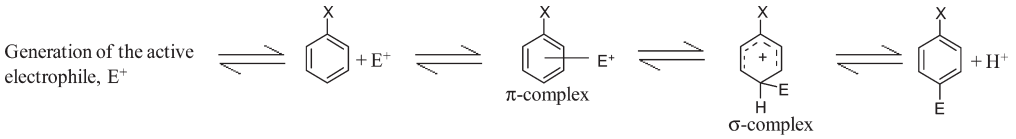
Introduction

The number of potentially applicable synthetic routes to a given target molecule (e.g., a drug candidate) increases rapidly with the level of complexity of the target structure. Thus, even for a relatively small target, the number of possible alternative "paper routes" can be large and, therefore, tools that can assist in prioritizing which route alternatives to focus on in the initial experimental evaluation are highly desirable. One aspect of this prioritization process is to estimate which levels of product selectivity can be expected in key steps of the different synthetic approaches, and, in this respect, computational chemistry is a powerful tool. A prerequisite, however, is that a sufficient level of accuracy can be obtained with a reasonable amount of computational

[§] Tel: +46 8 790 8210. Fax: +46 8 790 8207.
[∥] Tel: +46 8 553 28265. Fax: +46 8 553 88 92.

Published on Web 06/16/2010 DOI: 10.1021/jo100310v

SCHEME 1. General S_EAr Mechanism



resource; there must be a balance between precision and throughput to make a method practically useful. In this article, we describe an approach in which computational chemistry is used to estimate product selectivities for a very important type of synthetic transformation, electrophilic aromatic substitution (S_EAr).

We also compare the predictive capacity of the method to that of a modern reactivity index, the average local ionization energy, $I(\mathbf{r})$. The two methods have been applied to halogenations, nitrations, and Friedel−Crafts (F−C) acylations.

S_EAr Mechanistic Theory. S_EAr is a synthetically and industrially very important reaction type. The mechanism for this reaction has attracted great interest for more than 60 years, and it is now one of the most thoroughly studied classes of organic reactions from a mechanistic point of view.[1−3] There have been mechanistic proposals of both a stepwise mechanism and an alternative, concerted mechanism,[4,5] but there is now a large body of data supporting the stepwise reaction mechanism.[6−8] This, generally accepted, mechanism (Scheme 1) starts with generation of the active electrophile, E^+. Then there is an usually rapid, reversible complexation of the electrophile with the π-electron system of the aromatic ring, resulting in a so-called π complex. No position selectivity is associated with the π complex. In order for substitution to occur, a cationic intermediate, called a σ complex or a Wheland intermediate, is formed. In this intermediate, the carbon at the site of substitution is bonded to both the electrophile and its original ring hydrogen. The σ-complex intermediate is, in general, highly reactive because the stabilizing conjugation is broken. Formation of the σ complex can be reversible, but most often it is easier to eliminate a proton, in which case formation of the σ complex is essentially irreversible. In some cases, the intermediate σ complex has been observed experimentally,[9] and in some cases, it has been trapped.[10]

Halogenations are usually considered to be kinetically controlled, and the rate-limiting step is usually either formation of the active electrophile or formation of the intermediate σ complex.[3,11] The deprotonation step is usually very fast. Halogenations are often run in the presence of Lewis or Brönsted acids, and there exists a wide variety of halogenating agents, for example, X_2 (X = Cl, Br, I) and many "cationic halogen" compounds such as *N*-X-succinimide and hypochloric acid. Unlike in nitrations, there is no commonly identified active electrophile in halogenations. Br^+ is believed to be the active electrophile in uncatalyzed bromination.[2] When chlorine gas is used as the halogenating agent, the acid present is believed to assist in Cl−Cl bond breaking in a reactant−Cl_2 complex.[2] When hypochlorus acid (HOCl) is used in an acidic solution, kinetic studies indicate that two species, both Cl_2O and $[H_2OCl]^+$, work as active electrophiles.[2] The kinetics of halogenations is frequently complex.[12]

Like halogenations, nitrations are also usually considered to be kinetically controlled with the same rate-limiting step and a very fast deprotonation step. A wide variety of nitrating agents are available,[13] but the nitronium ion (NO_2^+) is commonly considered to be the active electrophile,[3,14,15] although there are cases where other species have been suggested.[16,17]

F−C reactions are also mainly considered to be kinetically controlled[3,11] but not to the same extent as nitrations and halogenations. There are numerous examples of thermodynamically controlled F−C reactions, but conditions such as the use of polyphosphoric acid and elevated temperatures are usually required.[18−22] The F−C acylation of benzene and benzene-d_6 has been shown to have a substantial primary kinetic hydrogen isotope effect,[23] indicating that the deprotonation step may be at least partially rate-limiting.

F−C reactions usually involve the reaction of an acyl halide or an acid anhydride, a Lewis acid promotor, and the aromatic substrate, but other acylating species and catalysts have been used, for example, zeolites.[24−27] Several species may function as the active electrophile. Direct kinetic

(1) Taylor, R. *Electrophilic Aromatic Substitution*; Wiley: New York, 1995.
(2) Carey, F.; Sundberg, R. *Advanced Organic Chemistry*, 4th ed.; Kluwer Academic/Plenum Publishers: New York, 2000.
(3) March, J. *Advanced Organic Chemistry*; McGraw-Hill: New York, 1984.
(4) Wheland, G. W. *J. Am. Chem. Soc.* **1942**, *64*, 900–908.
(5) Lowry, T. H.; Richardson, K. S. *Mechanism and Theory in Organic Chemistry*, 3rd ed.; Harper-Collins: New York, 1987; pp 623−640.
(6) Melander, L. *Ark. Kemi 2* **1950**, 213–292.
(7) Olah, G. A.; et al. *J. Am. Chem. Soc.* **1978**, *100*, 6299–6308.
(8) Olah, G. A.; Kuhn, S. J. *J. Am. Chem. Soc.* **1958**, *80*, 6535–6539.
(9) Olah, G. A.; Schlossberg, R. H.; Porter, R. D.; Mo, Y. K.; Kelly, D. P.; Mateescu, G. *J. Am. Chem. Soc.* **1972**, *94*, 2034–2043.
(10) Corey, E. J.; Barcza, S.; Klotmann, G. *J. Am. Chem. Soc.* **1969**, *91*, 4782–4786.
(11) Klumpp, G. W. *Reactivity in Organic Chemistry*; Wiley: New York, 1982; pp 227−378.
(12) Stock, L. M.; Baker, F. W. *J. Am. Chem. Soc.* **1962**, *84*, 1661–1667.
(13) Hoggett, J. G.; Moodie; R. B.; Penton, J. R.; Schofield, K. *Nitration and Aromatic Reactivity*; Cambridge University Press: London, 1971.
(14) Hughes, E. D.; Ingold, C. K.; Reed, R. I. *Nature* **1946**, *158*, 448–449.
(15) Hughes, E. D.; Ingold, C. K. *J. Chem. Soc.* **1950**, 2400–2684.
(16) Crivello, J. V. *J. Org. Chem.* **1981**, *46*, 3056–3060.
(17) Waller, F. J.; Barrett, A.; Braddock, D. C.; Ramprasad, D. *J. Chem. Soc., Chem. Commun.* **1997**, 613–614.
(18) Olah, G. A.; Kuhn, S. J.; Flood, S. H.; Hardie, B. A. *J. Am. Chem. Soc.* **1964**, *86*, 2203–2209.
(19) Agranat, I.; Shih, Y.-S.; Bentor, Y. *J. Am. Chem. Soc.* **1974**, *96*, 1259–60.
(20) Girdler, R. B.; Gore, P. H.; Thadani, C. K. *J. Chem. Soc. C* **1967**, *24*, 2619–2624.
(21) Andreou, A. D.; Bulbulian, R. V.; Gore, P. H. *J. Chem. Res. S* **1980**, *7*, 225.
(22) Levy, L.; Pogodin, S.; Cohen, S.; Agranat, I. *Lett. Org. Chem.* **2007**, *4*, 314–318.
(23) Olah, G. A.; Lukas, J.; Lukas, E. *J. Am. Chem. Soc.* **1969**, *91*, 5319–5323.
(24) Kasturi, T. R.; Damodaran, K. M. *Can. J. Chem.* **1969**, *47*, 1529–1535.
(25) Richard, F.; Carreyre, G.; Perot, G. *J. Mol. Catal. A: Chem.* **1995**, *101*, L167–L169.
(26) Kantam, M. L.; Ranganath, K. V.; Sateesh, M.; Kumar, K. B.; Choudary, B. M. *J. Mol. Catal. A: Chem.* **2005**, *225*, 15–20.
(27) Fromentin, E.; Coustard, J. M.; Guisnet, M. *J. Mol. Catal. A: Chem.* **2000**, *159*, 377–388.

measurements are difficult, however,[2] partly because the promotors form complexes with reactants, solvents, and products, and these complexes interconvert in an often unknown manner during the course of the reaction.[28] Thus, kinetic data seldom permit us to unambigously conclude the nature of the active electrophile in F−C reactions. Suggestions for active electrophiles are an acylium ion ($RC{\equiv}O^+$), the complex formed between the acyl halide and the Lewis acid catalyst [$RC(=O)X^+M^-X_n$], or the protonated acylium ion ($RC^+{\equiv}O^+H$).[2] Both acyl halide−Lewis acid complexes and acylium ions have been observed experimentally.[29,30]

S_EAr Prediction Models. Many efforts have been made to develop ways of predicting selectivity patterns in S_EAr reactions. Among the noncomputational models is one in which the prediction is based on the hyperfine coupling constants from electron spin resonance spectra.[31] Another approach has been to organize existing literature data on reactivity into generalized rules, either giving the rules explicitly, as in expert systems,[32] or implicitly from examples, forming neural networks.[33]

A number of theoretical studies have been carried out to give reactivity indices for the different positions in the aromatic systems in S_EAr reactions. One of the first was presented by Wheland in 1942, where he used localization energies of the reaction intermediate, with the calculations based on a molecular orbital treatment.[34] Other early efforts have been based both on localization energies[35,36] and on calculation of the π-electron densities and free valency.[37,38] Furthermore, the frontier molecular orbital theory, developed by Fukui,[39,40] as well as the theory of hard and soft acids and bases[41−44] has been applied as an index of reactivity in S_EAr reactions. Many of the more recent examples involve more sophisticated methods: semiempirical methods, Hartree−Fock, or density functional theory (DFT) methods.[45,46] Among them are approaches that are based on calculation of the protonation energies relative to benzene as a measure of the energy for activation,[47] charge densities,[48] free-electron superdelocalizability indices,[49] ionization energies,[50−52] and orbital interactions defined as reactive hybrid orbitals.[53]

Many of these methods are quite successful in making qualitatively correct predictions of the selectivity pattern in S_EAr reactions. Because of the very nature of these approaches, they cannot, however, take the structure of the transition state and its solvation into quantitative consideration. Furthermore, most of them lack the possibility of modeling the steric effects originating from the interaction between the aromatic substrate and the active electrophile.

One way to make quantitative predictions of the selectivity pattern in S_EAr reactions is obviously to calculate the potential energy profile in each case, including the transition states. Theoretical investigations of the detailed potential energy profile in the nitration of benzene have been made recently within the DFT framework.[54,55] The potential energy surface was studied in vacuo by Esteves et al.[54] and revealed a very complex mechanism, with numerous energy minima and transition-state structures. In another paper, nitration of benzene, phenol, and benzonitrile was analyzed theoretically.[56] In this work, solvent effects were considered using the PCM model. Rather surprisingly, the potential energy profile did not correctly predict the experimentally favored meta isomer for nitration of benzonitrile. Neither the energies for the transition state forming the intermediate σ complex nor the energies for the intermediate σ complex itself favored this isomer.[56]

For halogenations and F−C reactions, the exact nature of the active electrophile is often uncertain, and this would add to the difficulties in finding the potential energy profile for these cases. An additional disadvantage with this method, should one want to screen a large number of examples, is that the procedure is very time-consuming and difficult to automate. Easier methods for the quantitative prediction of the selectivity in S_EAr reactions would thus be highly valuable.

Proposed Method: Theoretical Support and Justifications. In this work, we use an alternative approach: we calculate the relative thermodynamic stability of each isomeric σ complex to predict the regioisomeric distribution. This approach has attractive features: (i) it avoids difficult transition-state optimizations and replaces them with optimizations to local minima, and (ii) structural similarities between isomers lead to error cancellations, which limits the need for highly accurate quantum mechanical methods. Our approach involves the following assumptions: First, the reaction is kinetically controlled, and the formation of the intermediate σ complex is the rate-determining step. Second, the energy differences between the isomeric transition states of the rate-determining step can be approximated with the energy differences between the corresponding intermediate σ

(28) DeHaan, F. P.; et al. *J. Org. Chem.* **1984**, *49*, 3959–3963.
(29) Glavincevski, B.; Brownstein, S. *J. Org. Chem.* **1982**, *47*, 1005–1007.
(30) Olah, G. A.; Comisarow, M. B. *J. Am. Chem. Soc.* **1964**, *86*, 5682–5683.
(31) Pedersen, E. B.; Pedersen, T. E.; Torssell, K.; Lawesson, S. O. *Tetrahedron* **1973**, *29*, 579–584.
(32) Bures, M. G.; Roos-Kozel, B. L.; Jorgensen, W. L. *J. Org. Chem.* **1985**, *50*, 4490–4498.
(33) Elrod, D. W.; Maggiora, G. M.; Trenary, R. G. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 477–484.
(34) Wheland, G. W. *J. Am. Chem. Soc.* **1942**, *64*, 900–908.
(35) Chung, A. L. H.; Dewar, M. J. S. *J. Chem. Phys.* **1965**, *42*, 756–766.
(36) Dewar, M. J. S.; Gleicher, G. J. *J. Am. Chem. Soc.* **1965**, *87*, 685–692.
(37) Coulson, C. A.; Longuet-Higgins, H. C. *Proc. R. Soc. London* **1947**, *A191*, 39; A192, 16.
(38) Coulson, C. A. *Discuss. Faraday Soc.* **1947**, *2*, 9–18.
(39) Fukui, K.; Yonezawa, T.; Shingu, H. *J. Chem. Phys.* **1952**, *20*, 722–725.
(40) Fukui, K.; Fujimoto, H. *Frontier Orbitals and Reaction Paths*; World Scientific: Singapore, 1997.
(41) Pearson, R. G. *Hard and Soft Acids and Bases*; Dowden, Hutchinson and Ross: Stroudsburg, PA, 1973.
(42) Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512–7516.
(43) Zhou, Z.; Parr, R. G. *J. Am. Chem. Soc.* **1990**, *112*, 5720–5724.
(44) Zhang, X.; Zhao, L.; Zhang, G. *Huaxue Yanjiu Yu Yingyong* **1996**, *8* (1), 47–51.
(45) Ess, D. H.; Jones, G. O.; Houk, K. N. *Adv. Synth. Catal.* **2006**, *348*, 2337−2361 and references cited therein.
(46) Minerva, T. *THEOCHEM* **2006**, *762*, 79−86 and references cited therein.
(47) Bader, R. F. W.; Chang, C. *J. Phys. Chem.* **1989**, *93*, 5095–5107.
(48) Bartel, H. G. *Z. Chem.* **1975**, *15*, 62–63.
(49) von Szentpaly, L. *Chem. Phys. Lett.* **1981**, *77*, 352–358.
(50) Sjöberg, P.; Murray, J. S.; Brinck, T.; Politzer, P. *Can. J. Chem.* **1990**, *68*, 1440–1443.
(51) Brinck, T. *J. Phys. Chem. A* **1997**, *101*, 3408–3415.
(52) Brinck, T. *Theoretical Organic Chemistry*; Parkanyi, C., Ed.; Elsevier Science BV: Amsterdam, The Netherlands, 1998.
(53) Hirao, H.; Ohwada, T. *J. Phys. Chem. A* **2003**, *107*, 2875–2881.
(54) Esteves, P. M.; de Carneiro, J. W.; Cardoso, S. P.; Barbosa, A. G. H.; Laali, K. K.; Rasul, G.; Prakash, G. K. S.; Olah, G. A. *J. Am. Chem. Soc.* **2003**, *125*, 4836–4849.
(55) Chen, L.; Xiao, H.; Xiao, X.; Gong, J. *J. Phys. Chem. A* **2003**, *107*, 11440–11444.
(56) Arrieta, A.; Cossio, F. P. *THEOCHEM* **2007**, *811*, 19–26.

complexes, in accordance with the Hammond postulate. Third, the relaxation of the structure upon solvation is similar for the different regioisomers, and this effect, therefore, will cancel out. Fourth, the entropy terms ($T\Delta S$) for reactions forming the different regioisomers will be very similar, and these terms therefore will also cancel out ($\Delta\Delta E \cong \Delta\Delta G \cong \Delta\Delta H$). The purpose of the work presented in this paper is to estimate the scope and limitations of this approach and to investigate how far in accuracy these assumptions can take us, in the areas where they are found to be valid.

The first assumption is supported by a large body of data, as discussed in a previous section of this Introduction. The method can, of course, not be applied to S_EAr reactions that are under thermodynamic control, like sulfonations or halogenations with iodine.[57]

The second assumption is our key assumption, and it is based on the applicability of the Hammond postulate for these types of S_EAr reactions. This postulate can be stated as follows: for any single reaction step, the geometry of the transition state for that step resembles the side to which it is closer in free energy.[58] This means that for similar reactions the more exothermic reaction will have the earlier transition state. One way to obtain the Hammond postulate in a more quantitative form is by using the Marcus equation.[59] Postulating equal parabolic curvature of the energy surfaces on both sides of any single reaction step, the Marcus treatment gives as a result that the reaction barrier can be decomposed into a parabolic component characteristic for the reaction type and a linear thermodynamic component due to the reaction energy.[60] Our second assumption follows from this.

There are many cases in which an approach like ours is bound to fail because the Hammond postulate is invalid there. For example, a comparison of the reactivity by applying the Hammond postulate to the reaction intermediate in acyl transfer reactions is likely to fail because the transition-state structure would vary over a series of reactions and lead to a change in the intrinsic reaction barriers.[61] In a multistep reaction with two possible regioisomers, the Hammond postulate can be invalid in such way that the major product arises from the minor intermediate. This has been observed in several reaction types (e.g., some asymmetric hydrogenations).[62,63] We noted a similar case in our investigation: under certain reaction conditions, the selectivity in the acetylation of naphthalene was reverted and the computationally most stable σ complex gave the minor product isomer.[64]

The S_EAr reaction type can be analyzed based on the Curtin−Hammett[65] principle, which states that the ratio of products formed from conformational isomers is not determined by the conformer population ratio but by the relative energies of the transition states. This principle is applicable for the reaction step between the π and σ complexes (Scheme 1) because the reaction barrier for formation of the σ complex is generally much higher than the barrier for the interconversion of the π-complex conformations.[2] Consequently, the ratio between the σ complexes does not depend on the energies of the π-complex conformations. Curtin−Hammett is generally not applicable for the last reaction step in S_EAr reactions. The reason is that the reaction barrier for the deprotonation step is not much higher than the reaction barrier for the conformational equilibration of the σ complexes. Thus, the conclusion is that the free energies of the transition states for formation of the σ complexes determine the regioselectivity and that these may be estimated from the free energies of the σ complexes.

For which of our three investigated S_EAr reaction types can we expect the best resemblance between the σ complex and the corresponding transition state for its formation? It is clear that the energy of the σ complex, in general, is higher than the energy of the reactants for all three reaction types.[2,3] A comparison of the reactivity between the three reaction types can be made, based both on the Hammett correlations[66] and on the partial rate factors.[67] These comparisons show that the electrophile involved in halogenations is usually less reactive than the electrophiles involved in nitrations and F−C reactions. When the Hammond postulate is applied, this implies that halogenations will usually have a later transition state than nitrations and F−C reactions. Thus, if the Hammond postulate is applicable for these S_EAr reactions, we would expect halogenations to show the best resemblance between the σ complex and the corresponding transition state for its formation. The proposed method also turns out to be the most successful for halogenations.

The third and fourth assumptions are based on the fact that we are using relative, not absolute, energies and comparing very similar chemical species. Quantum chemical methods, like DFT, are usually rather accurate for relative energies because errors tend to be systematic and at least partially cancel out during comparisons of similar systems.[60]

Methods and Procedure

Proposed Method. Calculations were carried out on all possible σ complexes including the actual electrophile. All systems have at least two distinct substitution sites, and some have as many as seven. The method can be described as follows. First, the geometries of each σ complex, without coordinated catalysts or promotors, are optimized in vacuo using the DFT functional B3LYP with a DZP quality basis set.[68,69] Second, if the resulting structures look reasonable, they are used as input files for solvent calculations in the relevant solvent, using the continuum model, called PBF, within the same software;[68] this calculation can be seen as an a posteriori energy correction. Third, the distribution of isomers follows the energy differences between the isomeric σ complexes via a Boltzmann distribution: the temperature used in this calculation is the one used experimentally in each specific reaction.

(57) Sykes, P. *A guidebook to mechanism in organic chemistry*, 5th ed.; Longman Inc.: New York, 1981.
(58) Hammond, J. *J. Am. Chem. Soc.* **1955**, *77*, 334.
(59) Marcus, R. A. *J. Phys. Chem.* **1968**, *72*, 891.
(60) Jensen, F. *Introduction to computational chemistry*, 2nd ed.; Wiley: New York, 2007.
(61) Um, I.-K.; Jeon, J.-S.; Kwon, D.-S. *Bull. Korean Chem. Soc.* **1991**, *12*, 406–410.
(62) Chan, A. S. C.; Pluth, J. J.; Halpern, J. *J. Am. Chem. Soc.* **1980**, *102*, 5952–5954.
(63) Schmidt, T.; Dai, Z.; Drexler, H.-J.; Hapke, M.; Preetz, A.; Heller, D. *Chem.—Asian J.* **2008**, *3*, 1170–1180.
(64) Dowdy, D.; Gore, P. H.; Waters, D. N. *J. Chem. Soc., Perkin Trans. 2* **1991**, 1149–1159.
(65) Curtin, D. Y. *Rec. Chem. Prog.* **1955**, *15*, 111.
(66) Rys, P.; Skrabal, P.; Zollinger, H. *Angew. Chem., Int. Ed. Engl.* **1972**, *11*, 874–883.
(67) Stock, L. M.; Brown, H. C. *Adv. Phys. Org. Chem.* **1963**, *1*, 35–154.
(68) *Jaguar*, version 7.0; Schrödinger, LLC: New York, 2007.
(69) The basis set used is called LACVP** in the *Jaguar* software and is a 6-31G** basis set with an effective core potential on heavy atoms.

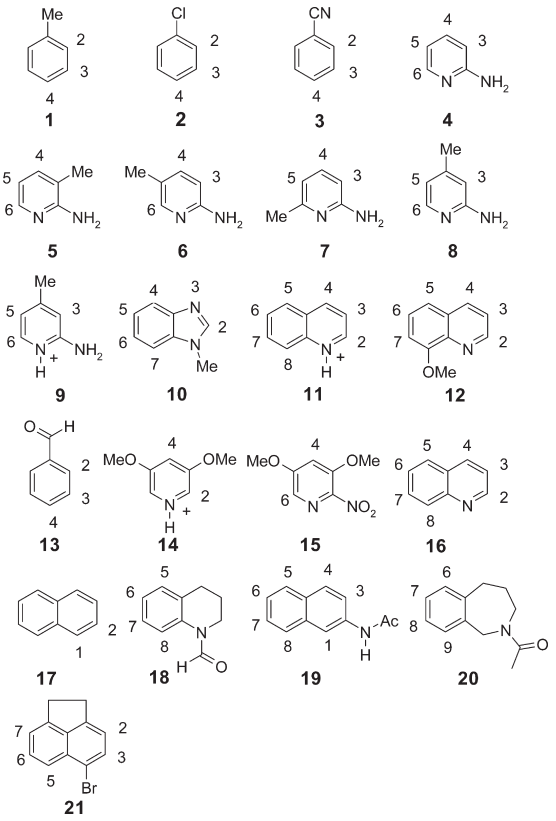


FIGURE 1. Schematic diagram of the structures studied in this work, also showing the labeling of positions.

There are cases in which it is necessary to investigate the energy differences between different conformations of the σ-complex isomers. In the F−C acetylations, for example, conformations that differ in the way the acyl group is rotated often differ by more than 1 kcal/mol. We did not perform any systematic conformational searches for the respective σ-complex isomers, and the set of conformational structures is by no means complete.[70] For this reason, we did not apply a Boltzmann summation but instead simply used the conformation with the lowest energy.

Average Local Ionization Energy. We have compared the proposed method to that of a modern reactivity index, the average local ionization energy $I(\mathbf{r})$. With this method, it is only necessary to perform calculations on the unreacted aromatic substrates shown in Figure 1. As a consequence, the method is a factor 5−10 faster than the proposed method in both computer and labor time. $I(\mathbf{r})$ is rigorously defined within the framework of Hartree−Fock or Kohn−Sham molecular orbital theory by

$$I(\mathbf{r}) = \sum_{i}^{\text{occ}} \frac{-\varepsilon_i \rho_i(\mathbf{r})}{\rho(\mathbf{r})}$$

where $\rho_i(\mathbf{r})$ is the electron density of the ith molecular orbital at the point $\mathbf{r}$ and ε_i is the orbital energy.[50] According to Koopmans theorem, $I(\mathbf{r})$ can be seen as the average energy needed to ionize an electron at a point $\mathbf{r}$ around a molecule. Minima in $I(\mathbf{r})$ on molecular surfaces ($I_{S,min}$) are indicative of sites susceptible to electrophiles. For example, $I_{S,min}$ of substituted benzenes have been shown to predict the regioselectivity and reactivity toward electrophilic aromatic substitution. Although $I(\mathbf{r})$ is usually obtained from gas-phase calculations, it has been shown to reflect the reactivity under solvated conditions. For example, excellent correlations between aqueous pK_a values and conjugate base $I_{S,min}$ have been obtained for both neutral and cationic acids.[71,72] A plausible rationale is that in solution electrostatic interactions are diminished and charge-transfer interactions, which are well reflected by $I(\mathbf{r})$, become dominating. In this work, we have computed $I(\mathbf{r})$ at molecular surfaces defined by the 0.001 atomic unit contour of the electron density using the *HS95* program.[51,52] *Gaussian 03* was used to obtain the input wave functions at the B3LYP/6-31G* level.[73]

During analysis of the regioselectivity data, it was observed that better predictions were obtained by scaling the relative $I_{S,min}$ values. The unscaled values gave average absolute deviations from experiment (using the same procedure and criteria as those in Table 1 for calculation of the average absolute deviation values) of 1.6 kcal/mol for halogenations, 2.0 kcal/mol for nitrations, and 2.2 kcal/mol for F−C acylations. The corresponding average absolute deviations with scaling factors of 0.5 and $^1/_3$ were 1.3, 1.6, and 2.1 kcal/mol and 1.1, 1.3, and 2.0 kcal/mol, respectively. Using a smaller scaling factor than $^1/_3$ gave a too low energy difference between the lowest and next-lowest regioisomer, in some cases where experiments have shown that only one isomer is formed (entries 4 and 9 in Table 2). The physical rational for the scaling is that rearrangement of the electron density to form a partial bond in a transition state requires much less energy than removal of an entire electron from the system.

Results and Discussion

The performances of the two methods are summarized in Table 1. We have chosen to treat separately the cases of reactions with only one experimentally determined isomer and the cases with two or more experimentally determined isomers. In order to measure the accuracy, we calculated the energy differences between the two experimentally determined major isomers and the corresponding differences for both the proposed method (after the solvent calculation) and for the $\Delta I_{S,min}/3$ method. For each reaction, we then calculated the deviation between the experimental energy difference and the calculated energy difference. These data sets were used to calculate the average absolute deviation as well as to perform a t test for halogenations.

The molecules used in our investigation are all taken from the open literature, and the labeling of the positions is shown schematically in Figure 1. The relative energies and regioisomer ratios of all possible σ complexes for the investigated systems are shown in Tables 2 (halogenations), 3 (nitrations), and 4 (F−C acylations), with the lowest-energy structure in each case taken as zero. The calculated values, as well as the experimentally determined isomer distributions, are given both as an energy difference, in kcal/mol and, in parentheses, as the corresponding regioisomeric ratio (%). The results from the $I(\mathbf{r})$ method are also included in the tables. The calculated isomer distributions have been adjusted for degenerate ortho and meta positions. The experimental values have in some cases been adjusted to reflect the ratio of the monosubstituted products (di- and polysubstituted products, in a

(70) Comparable conformational structures for the different isomers were investigated. It is likely that there are cases in which a larger set of conformations has to be included in order to obtain accurate results.

(71) Brinck, T.; Murray, J. S.; Politzer, P. *J. Org. Chem.* **1991**, *56*, 5012–5015.

(72) Brinck, T.; Murray, J. S.; Politzer, P.; Carter, R. E. *J. Org. Chem.* **1991**, *56*, 2934–2936.

(73) Frisch, M. J et al. *Gaussian 03*; Gaussian Inc.: Wallingford, CT, 2004.

TABLE 1. Accuracy of the Methods

	reactions with one experimentally determined isomer		reactions with two or more experimentally determined isomers	
	no. of reactions	no. of correct predictions[a]	no. of reactions	average absolute deviation (kcal/mol)
		Halogenations		
proposed method	5	4	6	0.4[b,c]
$\Delta I_{S,min}/3$	5	4	6	1.1[b,d]
		Nitrations		
proposed method	2	2	5	2.5[f]
$\Delta I_{S,min}/3$	2	0	5	1.3[d,f]
		F−C Acylations		
proposed method	1	1	6[e]	1.2[f]
$\Delta I_{S,min}/3$	1	0	6[e]	2.0[f]

[a]The prediction was counted as correct if the calculated energy difference between the experimentally observed isomer and the one closest in energy was ≥1.0 kcal/mol. [b]For entry 8 in Table 1, we assumed an equal reaction rate between substrates 8 and 9 (corresponding to the respective entries 8a and 8b in Table 1). [c]A *t* test calculation performed on the data set on the 95% confidence level gave a confidence interval of ±0.3 kcal/mol. According to this measure, the average absolute deviation for the $\Delta I_{S,min}/3$ method falls outside of the confidence limits of the new method, and the two methods can be said to perform significantly differently for halogenations. [d]For the cases in which we experimentally have more than one isomer and in which the $\Delta I_{S,min}/3$ method just pointed out one isomer, we assumed an energy difference of 3.5 kcal/mol to the next-largest isomer for the $\Delta I_{S,min}/3$ method. [e]The three entries 5, 8, and 10 in Table 3 have not been included (the three reactions with a low concentration of free Lewis acid during the reaction). [f]A *t* test calculation did not seem relevant here because the number of correct predictions were different for the two methods in the case of only one experimentally determined isomer.

TABLE 2. Modelling of S_EAr Halogenations[a]

entry	isomer	$\Delta I_{S,min}/3$	gas phase	solvent	experimental	entry	isomer	$\Delta I_{S,min}/3$	gas phase	solvent	experimental
		Chlorination of Toluene (**1**)[b]						Chlorination of 2-Amino-4-methylpyridine (**8**)[g,i]			
1	2	0.0 (74)	1.1 (23)	0.2 (59)	0.2 (56)	8a	3	1.1 (9)	2.9 (0.2)	1.8 (2.5)	0.0 (58)
	3		5.6 (0.0)	5.2 (0.0)			5	0.0 (91)	0.0 (100)	0.0 (98)	0.2 (42)
	4	0.2 (26)	0.0 (77)	0.0 (41)	0.0 (44)		6		*h*	*h*	
		Chlorination of Chlorobenzene (**2**)[c]						Chlorination of 2-Amino-4-methylpyridinium Ion (**9**)[g,j]			
2	2	0.5 (46)	2.9 (1.5)	*d*	2.2 (4)	8b	3	0.1 (45)	0.0 (100)	0.0 (87)	0.0 (58)
	3		8.5 (0.0)	*d*			5	0.0 (55)	3.2 (0.1)	0.9 (13)	0.2 (42)
	4	0.0 (54)	0.0 (99)	*d*	0.0 (96)		6		*h*	*h*	
		Chlorination of Benzonitrile (**3**)[e]						Bromination of 1-Methylbenzimidazole (**10**)[k]			
3	2		0.3 (54)	0.9 (17)	0.3 (34.1)	9	2		5.0 (0.0)	11.8 (0.0)	
	3	0.0 (100)[f]	2.7 (0.9)	0.0 (80)	0.0 (55.3)		4		1.2 (7)	3.7 (0.2)	
	4		0.0 (45)	1.5 (3)	0.6 (10.6)		5	0.0 (87)	0.0 (54)	4.7 (0.0)	(100)
		Chlorination of 2-Aminopyridine (**4**)[g]					6		0.2 (38)	0.0 (100)	
4	3	1.0 (11)	2.9 (0.2)	4.7 (0.0)			7	1.1 (13)	2.3 (1.0)	7.5 (0.0)	
	4		39.2 (0.0)	41.5 (0.0)							
	5	0.0 (89)	0.0 (100)	0.0 (100)	(100)			Bromination of Quinolinium Ion (**11**)[l]			
	6		*h*	*h*		10	2		25.9 (0.0)	23.1 (0.0)	
		Chlorination of 2-Amino-3-methylpyridine (**5**)[g]					3	2.9 (0.6)	8.6 (0.0)	3.3 (0.3)	
5	4		34.9 (0.0)	37.4 (0.0)			4,6,7		*h*	*h*	
	5	0.0 (100)[f]	0.0 (100)	0.0 (100)	(100)		5	0.5 (30)	0.3 (37)	0.6 (26)	0.0 (49)
	6		*h*	*h*			8	0.0 (70)	0.0 (63)	0.0 (73)	0.0 (51)
		Chlorination of 2-Amino-5-methylpyridine (**6**)[g]						Bromination of 8-Methoxyquinoline (**12**)[m]			
6	3	0.0 (100)[f]	0.0 (100)	0.0 (100)	(100)	11	2		19.8 (0.0)	21.1 (0.0)	
	4		32.1 (0.0)	36.8 (0.0)			3	4.8 (0.0)	20.8 (0.0)	20.9 (0.0)	
	6		*h*	*h*			4		16.6 (0.0)	21.2 (0.0)	
		Chlorination of 2-Amino-6-methylpyridine (**7**)[g]					5	0.0 (58)	0.0 (100)	0.0 (100)	(100)
7	3	1.0 (11)	1.8 (2.5)	0.8 (16)	0.9 (12)		6		*h*	*h*	
	4		*h*	*h*			7	0.2 (42)	3.9 (0.1)	6.8 (0.0)	
	5	0.0 (89)	0.0 (98)	0.0 (84)	0.0 (88)						

[a]Relative energies are given in kcal/mol; isomer distributions are given in %, in parentheses. The compound numbers in bold refer to Figure 1. The calculated isomer ratios have been adjusted for degeneracy. [b]Chlorine in acetic acid (0.1 M) at 25°C.[78] [c]Cl_2 (g) and chlorobenzene (g) over $FeCl_3$ and $AlCl_3$ catalyst. Temperature not stated.[79] [d]Gas-phase reaction. [e]Aqueous acetic acid and molecular chlorine, several months in the dark at rt.[80] [f]Only one reasonable position indicated by the software. [g]Cl_2 in 72% sulphuric acid at –33 °C.[75] [h]No stable intermediate found. [i]Reaction proceeding via the free base. [j]Reaction proceeding via the conjugated acid. [k]Br_2 (1 equiv, generated by electrolysis at a platinum electrode) with excess potassium bromide and acetate buffer at pH 5.5 in aqueous solution at 25°C.[81] [l]Bromine and silver sulphate in 98% sulphuric acid.[76] [m]NBS in chloroform at ambient temperature.[82]

maximum of 2% of the total amount of product, have then been neglected). We have used a temperature of 293 K for the Boltzmann distribution calculation in the cases in which the reaction temperature was not stated in the original paper.

TABLE 3. Modeling of S_EAr Nitrations[a]

entry	isomer	$\Delta I_{S,min}/3$	gas phase	solvent	experimental
		Nitration of Toluene (**1**)[b]			
1	2	0.0 (74)	2.5 (1.5)	2.8 (0.8)	0.0 (62)
	3		5.4 (0.0)	5.4 (0.0)	1.5 (5)
	4	0.2 (26)	0.0 (99)	0.0 (99)	0.0 (33)
		Nitration of Chlorobenzene (**2**)[c]			
2	2	0.0 (82)	3.7 (0.2)	4.6 (0.0)	0.0 (74.1)
	3		8.4 (0.0)	6.7 (0.0)	3.0 (0.4)
	4	0.5 (17)	0.0 (100)	0.0 (100)	0.2 (25.5)
		Nitration of Benzonitrile (**3**)[d]			
3	2		1.2 (11)	2.9 (0.5)	1.2 (11)
	3	0.0 (100)[e]	2.9 (0.6)	0.0 (79)	0.0 (89)
	4		0.0 (88)	0.8 (20)	
		Nitration of Benzaldehyde (**13**)[f]			
4	2		2.2 (1.5)	2.8 (0.6)	0.4 (31)
	3	0.0 (100)[e]	0.0 (90)	0.0 (99)	0.0 (67)
	4		0.9 (9)	2.4 (0.6)	1.5 (2)
		Nitration of 3,5-Dimethoxypyridinium Ion (**14**)[g]			
5	2	1.9 (3)	0.0 (100)	0.0 (100)	(100)
	4	0.0 (97)	11.9 (0.0)	17.8 (0.0)	
		Nitration of 2-Nitro-3,5-dimethoxypyridine (**15**)[h]			
6	4	0.0 (89)	21.2 (0.0)	23.2 (0.0)	
	6	1.3 (11)	0.0 (100)	0.0 (100)	(100)
		Nitration of Quinoline (**16**)[i]			
7	2			25.8 (0.0)	(0.0)
	3	2.9 (0.3)		7.9 (0.00004)	5.8 (0.0014)
	4			24.6 (0.0)	(0.0)
	5	0.5 (28)		0.7 (21)	0.0 (53)
	6			2.2 (1.5)	1.9 (1.5)
	7			6.4 (0.0006)	4.5 (0.015)
	8	0.0 (71)		0.0 (77)	0.1 (45)

[a]Relative energies are given in kcal/mol; isomer distributions are given in %, in parentheses. The compound numbers in bold refer to Figure 1. The calculated isomer ratios have been adjusted for degeneracy. [b]H_2SO_4/HNO_3 at 25 °C.[85] [c]H_2SO_4/HNO_3/HOAc at 25 °C for 3 h.[86] [d]HNO_3 and concentrated H_2SO_4 at 20 °C for 7 h.[87] [e]Only one reasonable position indicated by the software. [f]HNO_3 in TFA at 5 °C.[88] [g]HNO_3 and concentrated H_2SO_4 at 0 °C for 1 min (the reaction proceeds via the conjugate acid).[89] [h]HNO_3 and concentrated H_2SO_4 at 40 °C for 18 h.[89] [i]HNO_3 and concentrated H_2SO_4 at 0 °C. Geometry optimization made directly in the solvent (because no stable structures were found in vacuo).[90]

The selection of examples has been guided both by the theoretical interest of a given case and by the intention of spanning as wide as possible across the synthetically interesting space. In addition, examples have been rejected if the molecules were too large or if the isomer distribution was not determined properly. For the F−C reactions, we have limited ourselves to Lewis acid promoted reactions. We have also calculated the relative energies of some of the final products, in order to investigate kinetic versus thermodynamic control.[74]

Halogenations. The proposed method is most successful for halogenations, for both activated and deactivated systems, and gave quantitatively useful predictions in 10 of the 11 investigated examples. The relative stabilities of the final halogenated products do not correlate at all with the experimentally found isomer distributions, which support our assumption of kinetic control.[74] The solvation correction used in the method usually improves the prediction. Our results indicate that in the majority of halogenations the method gives quantitatively correct predictions within an accuracy of 1 kcal/mol.

(74) See the Supporting Information.

In the case of chlorination of 2-amino-4-methylpyridine, we have also made calculations on alternative intermediates, the σ complexes derived from the conjugated acid (Table 2, entry 8b), because the authors of the paper suggested that the reaction could, partly, take that route.[75] Our results show much better agreement with experimental data if the assumption is made that some chlorination actually takes place via the neutral form and some via the conjugate acid form. The conjugate acid form was also used in bromination of quinoline (Table 2, entry 10), because the authors claimed that "it seems almost certain" that the reaction proceeds via the quinolinium cation.[76] In the case of bromination of 1-methylbenzimidazole (Table 2, entry 9), the method did not give a correct prediction.[77]

$I(\mathbf{r})$ analysis is also quite successful for predicting the regioselectivity of halogenations. For 10 out of the 11 molecules, the lowest $I_{S,min}$ is found on the carbon atom that is most reactive according to experiment. Interestingly, this approach is capable of predicting the dominating isomer for the reaction of 1-methyl-benzimidazole (Table 2, entry 9), where the new method fails. The relative $I_{S,min}$ values scaled by $^1/_3$ ($\Delta I_{S,min}/3$) provide a good estimate of the isomer distribution for the majority of molecules. However, the overall agreement with experiment is slightly worse than that for the new method. For example, in the case of benzonitrile, $I_{S,min}$ values are only found in the meta position, whereas significant amounts of para and ortho products were observed experimentally. For chlorobenzene, the $\Delta I_{S,min}/3$ values are not able to reproduce the 4:96 ortho/para ratio observed in the gas phase but predict a ratio of 46:54, which is closer to what is expected in solution. This is in agreement with earlier observations that $I(\mathbf{r})$ reflects the reactivity in solution.[71,72] Rather surprisingly, we find that for quinoline, which can react to form seven different isomers, the predicted isomer distribution from $\Delta I_{S,min}/3$ is almost identical with that of the new method. The agreement with experiment is also good.

Nitrations. The situation for nitrations is more problematic. The method gives correct predictions for the large heterocyclic substrates (Table 3, entries 5−7) but is less

(75) Kress, T. J.; Moore, L. L.; Costantino, S. M. *J. Org. Chem.* **1976**, *41*, 93–96.
(76) De la Mare, P. B. D.; Kiamud-Din, M.; Ridd, J. H. *J. Chem. Soc.* **1960**, 561–565.
(77) We performed higher level calculations (CCSD) with increased basis sets and also tried more sophisticated solvent models as well as geometry optimization directly in the solvent, but the same general picture remained.
(78) O'Connell, J. L.; Simpson, J. S.; Dumanski, P. G.; Simpson, G. W.; Easton, C. *J. Org. Biomol. Chem.* **2006**, *4*, 2716–2723.
(79) Hillman, M.; Brown, J. D. WO-9008118 A1, 1990.
(80) Baciocchi, E.; et al. *J. Am. Chem. Soc.* **1972**, *94*, 7030–7034.
(81) Evans, D. J.; Thimm, H. F.; Coller, A. W. *J. Chem. Soc., Perkin Trans. 2* **1978**, 865–871.
(82) Gershon, H.; McNeil, M. W.; Schulman, S. G. *J. Org. Chem.* **1972**, *37*, 4078–4082.
(83) Sokolov, A. V. *J. Quantum Chem.* **2004**, *100*, 1–12.
(84) The failure of the $I(\mathbf{r})$ method for entry 5 in Table 3 could also have other reasons, except for steric ones. The $I(\mathbf{r})$ method cannot, for example, include effects like intramolecular stabilization between the substrate and the attacking electrophile. We thank a reviewer for calling our attention to this.
(85) Olah, G. A.; Overchuk, N. A. *Can. J. Chem.* **1965**, *43*, 3279–3293.
(86) Sparks, A. WO-3253045, 1966.
(87) Mekhtiev, S.; et al. *Azerb. Khim. Zh.* **1978**, *4*, 59–63.
(88) Östman, B.; Lindgren, I. *Acta Chem. Scand.* **1970**, *24*, 1105–1106.
(89) Johnson, C. D.; Katritzky, A. R.; Viney, M. *J. Chem. Soc. B* **1967**, 1211–1213.
(90) Crout, D. H.; Penton, J. R.; Schofield, K. *J. Chem. Soc. B* **1971**, 1254–1256.

TABLE 4. Modeling of F−C Reactions[a]

entry	isomer	$\Delta I_{S,min}/3$	gas phase	solvent	experimental
		Formylation of Toluene with Formyl Fluoride (**1**)[b]			
1	2	0.0 (74)	0.2 (57)	0.5 (42)	0.5 (43.3)
	3		4.5 (0.0)	4.5 (0.0)	1.7 (3.5)
	4	0.2 (26)	0.0 (43)	0.0 (58)	0.0 (53.2)
		Acylation of Toluene with $ClCH_2COCl$ (**1**)[c]			
2	2	0.0 (74)	2.3 (2)	2.0 (3)	1.6 (11.1)
	3		4.5 (0.0)	4.2 (0.1)	2.5 (2.3)
	4	0.2 (26)	0.0 (98)	0.0 (97)	0.0 (86.6)
		Benzoylation of Toluene with 2,4-Dichlorobenzoyl Chloride (**1**)[d]			
3	2	0.0 (74)	3.6 (0.4)	2.8 (1.5)	1.8 (8.8)
	3		5.6 (0.0)	4.1 (0.2)	3.5 (0.4)
	4	0.2 (26)	0.0 (100)	0.0 (98)	0.0 (90.8)
		Acetylation of Naphthalene with AcCl (**17**)[e,f]			
4	1	0.0 (100)[g]	0.0 (98)	0.0 (99)	0.0 (99)
	2		2.1 (2)	2.6 (0.8)	2.5 (1)
		Acetylation of Naphthalene with AcCl (**17**)[e,h]			
5	1	0.0 (100)[g]	0.0 (97)	0.0 (99)	1.5 (7)
	2		2.1 (2.5)	2.6 (1)	0.0 (93)
		Benzoylation of Naphthalene with C_6H_5COCl (**17**)[i]			
6	1	0.0 (100)[g]	0.0 (50)	0.0 (99)	0.0 (60−80)
	2		0.0 (50)	2.9 (0.7)	0.2−0.8 (20−40)
		Acetylation of *N*-Acetyltetrahydroquinoline with AcCl (**18**)[j]			
7	5		13.8 (0.0)	11.7 (0.0)	
	6	0.0 (50)	0.0 (100)	0.0 (100)	0.1 (44)
	7		11.9 (0.0)	10.5 (0.0)	0.0 (56)
	8	0.0 (50)	4.9 (0.0)	6.2 (0.0)	
		Acetylation of Acet-2-naphthalide with AcCl (**19**)[k]			
8	1	0.0 (30)	0.0 (100)	0.0 (100)	
	3	0.2 (21)	11.1 (0.0)	10.6 (0.0)	
	4		16.2 (0.0)	14.8 (0.0)	
	5		15.7 (0.0)	13.8 (0.0)	
	6	0.1 (25)	10.4 (0.0)	9.0 (0.0)	1.3 (8)
	7		17.9 (0.0)	16.3 (0.0)	
	8	0.1 (25)	9.0 (0.0)	7.6 (0.0)	0.0 (92)
		Acetylation of *N*-Acetyl-2,3,4,5-tetrahydro-1*H*-2-benzazepine with AcCl (**20**)[l]			
9	6	0.0 (69)	1.2 (5.5)	1.8 (5)	
	7		0.2 (27)	1.2 (12)	
	8	0.5 (31)	0.0 (36)	0.0 (81)	(>95)
	9		0.1 (31)	2.4 (2)	
		Acetylation of 4-Bromoacenaphthene with AcCl (**21**)[m]			
10	2	1.5 (5)	3.6 (0.1)	3.1 (0.2)	0.0 (50)[n]
	3	3.6 (0.1)	9.2 (0.0)	11.9 (0.0)	
	5	0.0 (85)	0.0 (100)	0.0 (100)	
	6		10.7 (0.0)	11.8 (0.0)	
	7	1.1 (10)	3.6 (0.1)	3.2 (0.2)	0.0 (50)[n]

[a]Relative energies are given in kcal/mol; isomer distributions are given in %, in parentheses. The compound numbers in bold refer to Figure 1. The calculated isomer ratios have been adjusted for degeneracy. [b]BF_3-promoted reaction with formyl fluoride in toluene (no other solvent) at −30 °C.[93] [c]$AlCl_3$-promoted reaction in CS_2 at 25 °C.[93] [d]$AlCl_3$-promoted reaction in nitromethane at 20 °C.[28] [e]$AlCl_3$-promoted reaction with AcCl and naphthalene in dichloroethane.[91] [f]AcCl (1.1 equiv) and naphthalene (1.0 equiv) in dichloroethane were added to a slurry of $AlCl_3$ (1.2 equiv) in dichloroethane at 0−5 °C.[91] [g]Only one reasonable position indicated by the software. [h]AcCl (1.0 equiv) and $AlCl_3$ (1.0 equiv) in dichloroethane were added to a solution of naphthalene in dichloroethane (1.0 equiv) at 25 °C.[91] [i]$AlCl_3$-promoted reaction with benzoyl chloride and naphthalene in dichloroethane at 25 °C (different stoichiometric relationships).[94] [j]Refluxing the substrate with AcCl (1.1 equiv) and $AlCl_3$ (2.3 equiv) in 1,2-dichloroethane.[95] [k]$AlCl_3$-promoted reaction in CS_2 at 0 °C.[96] [l]1.1 mol equiv of AcCl in the presence of $AlCl_3$ (2.3 mol equiv) in 1,2-dichloroethane at 50 °C for 4 h.[97] [m]$AlCl_3$-promoted reaction in nitrobenzene at −15 °C.[98] [n]The exact ratio between these two isomers was not stated in the paper, but from the description, it is clear that they were the only isomers found and that they were obtained in approximately equal amounts.

reliable for monosubstituted benzenes, for both activated and deactivated systems (Table 3, entries 1−4); it predicts a too high energy for the ortho isomer relative to the para/meta isomers in all cases with the result that the amount of the ortho isomer is underestimated. Still, for benzonitrile and benzaldehyde, substitution at the meta position is predicted to be dominating, in agreement with experiment. In the cases of toluene and chlorobenzene, however, the para isomers are incorrectly predicted to be dominating.

In order to understand the problems, we investigated the potential energy surface for nitration of benzonitrile by geometry optimizations in solution using the PCM method and one explicit solvent molecule (water). In this model, σ complexes were no longer found to be stable intermediates.

Instead, the addition of the nitro group was followed by proton transfer to the water molecule and a loss of the formed hydronium ion in a concerted process. Thus, this result indicates that it is not possible to correctly model the potential energy surface for nitration of benzonitrile without consideration of explicit solvent molecules. This is in agreement with the results of Arrieta and Cossio, who also found that an implicit solvent model gave the wrong regioselectivity.[56] Chen et al. demonstrated a significant solvent effect on the nitration of benzene,[55] and Sokolov did not, in a work using semiempirical methods, find any stable σ complexes for deactivated substrates in solvent.[83] We expect that, in general, more elaborate model systems, including explicit solvent molecules, will be necessary to correctly predict the structures and energies for the important stationary points in nitration of substituted benzenes.

$I(\mathbf{r})$ performs slightly better than the new method for the monosubstituted benzenes. In particular, it predicts a better balance between the ortho and para isomers for toluene and chlorobenzene. Generally, for benzenes with meta directors, such as benzonitrile and benzaldehyde, $I_{S,min}$ values are only found at the meta positions. Thus, $I(\mathbf{r})$ is not able to reproduce the contribution of the ortho isomers, which is highly significant (31%) for benzaldehyde. For quinoline, $I(\mathbf{r})$ gives a result similar to that of the new method. On the other hand, $I(\mathbf{r})$ fails for entries 5 and 6 in Table 3. In both cases, the lowest $I_{S,min}$ is found at the sterically hindered position between the two methoxy groups. These examples show that the $I(\mathbf{r})$ approach has problems in predicting the regioselectivity in systems where some of the most activated sites are sterically hindered.[84]

F−C Acylations. The predictions for Lewis acid promoted F−C reactions vary: of the 10 investigated cases, two were quantitatively correct (Table 4, entries 3 and 4), in five cases they could be used qualitatively (Table 4, entries 1, 2, 6, 7, and 9), and in three cases they were qualitatively incorrect (Table 4, entries 5, 8, and 10).

The regioselectivity in F−C reactions can be strongly dependent upon variables like the mode of addition of reagents, total concentration of reactants/promotor, and types of Lewis acid promotor and solvent.[91,92] This is probably one of the reasons that investigations of isomer distributions, at times, have given confusing and contradictory results.[92] In the case of acetylation of xylene, for example, the regioisomer distribution was independent of the reaction conditions like the mode of addition of the reagents.[91] In the case of acetylation of naphthalene,[91] the ratio of α- and β-substituted products was 99/1 when the acetylating agent and naphthalene in the solvent were added slowly to a slurry of $AlCl_3$ (Table 4, entry 4); in contrast, it was 7/93 when the acetylating agent and $AlCl_3$ in the solvent were added slowly to naphthalene (Table 4, entry 5). The 99/1 distribution was predicted well by our method (assuming kinetic control). The 7/93 distribution obtained from naphthalene, as well as the other cases where our method failed to make a qualitatively correct prediction, could be considerably better predicted if thermodynamic control was assumed.[74] It does not seem likely, however, that these particular F−C reactions really are under thermodynamic control because they were not run under conditions normally associated with thermodynamic control (e.g., polyphosphoric acid and elevated temperatures).

Several authors have investigated the dependence of the regioisomer distribution on the reaction conditions in detail. Both Friedman and Honour[91] and Dowdy, Gore, and Waters[64] concluded that the regioisomeric outcome can be highly dependent upon the concentration of free Lewis acid/free acylating reagent relative to the concentration of the substrate. Friedman and Honour treated several substrate examples, but they could not give a common mechanistic rationale for all of their investigated substrates.[91] Dowdy, Gore, and Waters focused solely on acetylation of naphthalene, and they concluded that, under certain reaction conditions (excess AcCl, corresponding to a low concentration of free Lewis acid), the deprotonation step can become rate-determining because of steric hindrance inherent in the naphthalene molecule.[64] We analyzed the results in Table 4 further to see if the concentration of free Lewis acid appeared to be an important factor. The model failed to make a qualitatively correct prediction for one of the acetylation modes of naphthalene, acetylation of acet-2-naphthalide, and for acetylation of 4-bromoacenaphthene (Table 4, entries 5, 8, and 10, respectively). In these three examples, $AlCl_3$ was added slowly to a solution containing the substrate, making the relative amount of free $AlCl_3$ low during the reaction. The model made a quantitatively excellent prediction in two cases: benzoylation of toluene and in one of the acetylation modes of naphthalene (Table 4, entries 3 and 4). In these two cases, the substrate or acylating agent was added slowly to a solution containing $AlCl_3$, making the relative amount of free $AlCl_3$ high during the reaction. No specific pattern regarding the free Lewis acid concentration was noted for the five intermediate cases in which the predictions could be used qualitatively.

The strong dependence of the regioisomeric outcome on the Lewis acid promotor (amount, mode of addition, etc.) and the different conformations of the attacking electrophile make F−C reactions more difficult to predict than halogenations. Not all F−C reactions gave reasonable structures for the σ complexes in our calculations. For example, acylation of toluene with EtCOCl as well as benzoylation of 1,3-dichlorobenzene with C_6H_5COCl gave structures that more resembled reaction complexes (π complexes). Because the Lewis acid promotor is not modeled within our method, we can hardly expect our approach to be generally applicable for all F−C reactions. It would seem that one of our assumptions, that formation of the σ complex is the rate-determining step, is invalid in at least one case (Table 4, entry 5)[64] and possibly also for the two other, naphthalene-like, cases in which the method failed (Table 4, entries 8 and 10). On the other hand, our assumption that the energy difference between the σ complexes is a good measure of the energy difference between the transition states for the rate-determining step seems to be reasonably valid when the relative

(91) Friedman, L.; Honour, R. J. *J. Am. Chem. Soc.* **1969**, *91*, 6344–6349.
(92) Jensen, F. R. *J. Am. Chem. Soc.* **1957**, *79*, 1226–1231.
(93) Olah, G. A.; Kobayashi, S. *J. Am. Chem. Soc.* **1971**, *93*, 6964–6967.
(94) Jensen, F. R. *J. Am. Chem. Soc.* **1957**, *79*, 1226–1231.
(95) Ishihara, Y.; Tanaka, T.; Goto, G. *J. Chem. Soc., Perkin Trans. 1* **1992**, 3401–3406.
(96) Leonard, N. J.; Hyson, A. M. *J. Am. Chem. Soc.* **1949**, *71*, 1961–1964.
(97) Ishihara, Y.; Tanaka, T.; Miwatashi, S.; Fujishima, A.; Goto, G. *J. Chem. Soc., Perkin Trans. 1* **1994**, 2993–2999.
(98) Nightingale, D. V.; Brooker, R. M. *J. Am. Chem. Soc.* **1950**, *72*, 5539–5543.

concentration of the free Lewis acid promotor is high during the reaction.

The $I(\mathbf{r})$ approach has problems similar to those of our new method for predicting the regioselectivity of F−C reactions. The overall performance is relatively poor, and only in a few cases is the method able to predict the dominating isomer. The $I(\mathbf{r})$ approach cannot be used qualitatively even if the relative concentration of the free Lewis acid promotor is high during the reaction.

Conclusions

The investigated method can, with good probability, be expected to predict regioisomer distributions in electrophilic aromatic halogenations quantitatively correctly. The method cannot be used in a general way for electrophilic aromatic nitrations. Lewis acid promoted F−C reactions can, with reasonable probability, be expected to be predicted qualitatively correctly. For substrates in which the regioisomeric outcome is highly dependent on the reaction conditions, the method can only be of qualitative use if the concentration of free Lewis acid is high during the reaction. The fact that the method works best for halogenations is in accordance with the Hammond postulate being applicable to the types of electrophilic aromatic substitution reactions investigated in this study. In terms of throughput, it should be noted that we, in general, can obtain an answer for a problem of an average complexity in one working day. Furthermore, the simplicity of our protocol makes it useful also for a nonexpert.

The $I(\mathbf{r})$ method does not need any additional solvent energy calculation, and the throughput is an order of magnitude higher than that for the proposed method. It can, with high probability, be expected to predict the regioselectivity in halogenations and nitrations qualitatively correctly, if the positions for the $I(\mathbf{r})$ minima ($I_{S,min}$) are not too sterically hindered. After the $I_{S,min}$ values are downscaled, the method can also quantitively predict regioisomer distributions in halogenations with reasonable probability. It cannot be used for qualitative predictions of F−C reactions.

Supporting Information Available: Results from the calculations of some selected final products, all of the optimized structures, and their gas-phase energies. This material is available free of charge via the Internet at http://pubs.acs.org.