signal from each pulse, superimposed on the first part of the FID. Fortunately, these base line distortions could be almost entirely eliminated by selectively removing the few pulse breakthrough data points from the beginning of the FID.[29] Oxygen-17 chemical shifts referenced to internal (1.0–4.0 M) or external acetone gave the same value in ppm (±1 ppm). Thus, changes in the ^{17}O resonance due to a solvent effect is negligible. Chemical shifts are presented in Table II relative to external water (acetone resonates 572 ppm downfield of external water). The proton and ^{13}C NMR spectra are referenced to tetramethylsilane.

(28) Harris, R. K., Ed., "NMR and the Periodic Table"; Academic Press: New York, 1978; pp 383-419.
(29) (a) Canet, D.; Goulon-Ginet, C.; Marchal, J. P. *J. Magn. Reson.* **1976**, *22*, 537. (b) Roth, K. *J. Magn. Reson.* **1980**, *38*, 65.

Acknowledgment. This research was assisted financially by Grant GM-21249 (to R.G.S.) from the Division of General Medical Sciences of the National Institutes of Health and the donors of The Petroleum Research Fund, administered by the American Chemical Society (to E. L.C.). We also thank Professor S. F. Nelsen for a sample of dioxetane **13**.

Registry No. **2b**, 6671-70-1; **2c**, 6786-21-6; **2d**, 52148-56-8; **3a**, 279-35-6; **3b**, 280-53-5; **3c**, 283-35-2; **3d**, 52965-57-8; **4a**, 4362-13-4; **4b**, 5703-46-8; **5a**, 628-37-5; **5b**, 29914-92-9; **5c**, 16642-57-2; **5d**, 110-05-4; **6b**, 1127-11-3; **7b**, 1127-10-2; **8c**, 7124-86-9; **9c**, 283-19-2; **9d**, 284-26-4; **10**, 67105-55-9; **11**, 19077-73-7; **12**, 512-85-6; **13**, 88510-82-1; **14**, 5718-73-0; bicyclo[4.2.2]deca-3,7,9-triene, 35733-48-3.

Computer-Assisted Mechanistic Evaluation of Organic Reactions. 11. Electrophilic Aromatic Substitution

Mark G. Bures, Barbara L. Roos-Kozel, and William L. Jorgensen*

Department of Chemistry, Purdue University, West Lafayette, Indiana 47907

Received April 23, 1985

CAMEO, an interactive computer program which predicts the products of organic reactions primarily through mechanistic reasoning, has been expanded to treat electrophilic aromatic substitution (EAS). The major types of EAS reactions, i.e., nitration, sulfonation, alkylation, acylation, and halogenation, have been implemented for substituted benzenes, polycyclic aromatic hydrocarbons, and heteroaromatic systems. Key considerations include perception and classification of electrophiles, determining the relative reactivities of aromatic compounds and the most reactive sites, and matching electrophile and substitution site reactivity. Integral to the treatment is the development of a structure/reactivity correlation equation for EAS reactions. The paper begins with a brief review of the important facets of EAS chemistry. A discussion of the implementation of EAS in CAMEO follows including sample EAS reaction sequences produced by the program.

I. Introduction

An interactive computer program, CAMEO, that predicts the products of organic reactions given starting materials and reaction conditions is under continued development. Algorithms have been implemented to cover base-catalyzed and nucleophilic processes[1] including some organometallic reactions,[2] organosilicon chemistry,[3] pericyclic reactions,[4] nucleophilic aromatic substitution,[5] and acid-catalyzed reactions.[6] The program has also been expanded to treat electrophilic aromatic substitution (EAS) reactions as described here. Naturally, this has required analysis of the key features governing reactivity for this class of processes. Therefore, attention has been focused on perception and classification of electrophiles and identification of the most reactive rings and of the most reactive sites for a given electrophile. Substituent effects are considered in detail and relative reactivity is gauged via structure/activity relationships. To begin, a brief overview of electrophilic aromatic substitution is provided followed by a discussion of the implementation of EAS chemistry in CAMEO. The paper concludes with a presentation and discussion of sample EAS sequences predicted by the program.

Table I. Examples from the Three Classes of Electrophiles

I (weak)	II (moderate)	III (strong)
$HC{\equiv}NH^+$	R_3C^+	NO_2^+
NO^+	RCH_2X	Br_2
ArN_2^+	RCO^+	Cl_2
$H_2C{=}NR_2^+$	RCOX	$BrOH_2^+$
	$R_2C{=}OH^+$	$ClOH_2^+$
	$RN{=}C{=}O$	SO_3
		RSO_2^+

II. Key Aspects of EAS

Electrophilic aromatic substitution is a class of reactions that has been particularly well studied, so only a brief review of the major points of EAS chemistry is presented here. The discussion focuses on electrophiles for EAS, relative reactivity of different aromatic systems, and identification of the most reactive sites of aromatic compounds. For a more detailed coverage of EAS chemistry, several excellent sources may be consulted.[7,8]

(1) Salatin, T. D.; Jorgensen, W. L. *J. Org. Chem.* **1980**, *45*, 2043.
(2) Salatin, T. D.; McLaughlin, D.; Jorgensen, W. L. *J. Org. Chem.* **1981**, *46*, 5284.
(3) Peishoff, C. E.; Jorgensen, W. L. *J. Org. Chem.* **1983**, *48*, 1970.
(4) (a) Burnier, J. S.; Jorgensen, W. L. *J. Org. Chem.* **1983**, *48*, 3923. (b) Burnier, J. S.; Jorgensen, W. L. *J. Org. Chem.* **1984**, *49*, 3001.
(5) Peishoff, C. E.; Jorgensen, W. L. *J. Org. Chem.* **1985**, *50*, 1056.
(6) McLaughlin, D.; Jorgensen, W. L. *J. Org. Chem.*, to be submitted.
(7) Reviews: (a) Norman, R. C.; Taylor, R. "Electrophilic Substitution in Benzenoid Compounds"; Elsevier: Amsterdam, 1965. (b) de la Mare, P. B. D.; Ridd, J. H. "Aromatic Substitution; Nitration and Halogenation"; Academic Press: New York, 1959.
(8) March, J. "Advanced Organic Chemistry"; McGraw Hill: New York, 1977; Chapter 11.

0022-3263/85/1950-4490$01.50/0

Table II. Relative Reactivities of Aromatic Compounds under Neutral to Weakly Acidic Conditions[a]

compound	log k_{rel} (halogenation)[b]
pyrrole	10
pyrene, $PhNR_2$, furan	8
PhOMe, selenophene	7
benzothiophene, thiophene	6
phenanthrene, fluorene, naphthalene	5
triphenylene	4
PhSR	3
biphenyl	2
PhR	1
PhH	0
PhX	−1
$PhCO_2R$	−2
$PhNO_2$, $PhNR_3^+$, PhCN	−5
pyridine	−10
diazenes	<−10

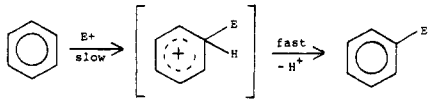
[a] References 11b and 13–25. [b] k_{rel} represents the rate of reaction relative to benzene.

Table III. Relative Reactivities of Aromatic Compounds under Strongly Acidic Conditions[a]

compound	log k_{rel} (nitration)[b]
pyrene	4
naphthalene, phenanthrene, triphenylene,	3
biphenylene, PhR, thiophene	1
PhH	0
PhX	−1
PhX_2, $PhMeNO_2$	−2
$PhCO_2H$	−4
oxazole	−5
isoxazole, PhCN, $PhCONH_2$	−6
quinoline, isoquinoline, cinnoline 2-oxide	−7
$PhNO_2$, $PhNR_3^+$, cinnoline, imidazole	−8
pyridine *N*-oxide, thiazole, pyrazole	−9
isothiazole	−12
pyridinium ion, diazines	<−12

[a] References 7b, 11b, 13, and 25–28. [b] k_{rel} represents the rate of reaction relative to benzene.

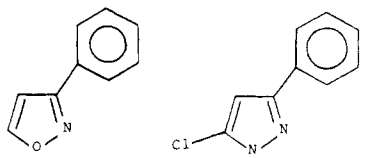
A. Electrophiles for EAS. Electrophilic aromatic substitution reactions occur for a wide range of electrophiles. Nevertheless, most EAS reactions follow the same mechanism. The mechanism features the creation of an arenium ion intermediate in the first step by attack of the electrophile (E^+) on an aromatic ring. Subsequent departure of the leaving group (usually a proton) forms the substituted system (eq 1).[9]



(1)

It is convenient to classify electrophiles according to the types of aromatic systems that they can substitute. For the present purposes, electrophiles may be grouped into three categories, examples of which are shown in Table I.[10] The relatively weak electrophiles of group I react only with strongly activated aromatic systems. The moderate electrophiles of group II, including ones commonly used in Friedel–Crafts alkylations and acylations, react readily with activated and halogenated rings but do not substitute deactivated rings. Group III consists of strong electrophiles, such as those derived from mineral acids, that can substitute successfully on nearly any aromatic ring, whether activated or deactivated. It should be noted that a great variety of precursors may lead to the electrophiles in Table I, particularly the acylium and carbenium ions. More thorough summaries of electrophiles for EAS and their generation can be found in several sources.[11]

B. Reactivity of Aromatic Systems. The next important consideration for EAS reactions is gauging the relative reactivity of different aromatic systems. This is necessary so that in case more than one aromatic ring is present, the most reactive ring can be identified. The factors that determine an aromatic ring's reactivity are (1) the type of parent system, aromatic hydrocarbon or heteroaromatic, (2) the number, electronic nature, and arrangement of substituents, and (3) the specific reaction. As can be seen from the data in Tables II and III, reactivity of aromatic compounds varies enormously. Table II illustrates the relative reactivity of representative compounds under neutral to weakly acidic conditions, while Table III lists corresponding data in strongly acidic media. For these two reactions, the orders of relative reactivity are similar, though the rate accelerations are compressed and the retardations expanded for nitration compared to halogenation. When there is more than one aromatic ring present, the most reactive individual ring is usually substituted. For example, both compound 1 and 2-phenylpyridine nitrate only on the phenyl ring.[12,29] However,



reaction conditions also affect the relative reactivity of a given aromatic system, particularly when protonation of a heteroatom is possible. Thus, **2** is brominated on the pyrazole ring but nitrated on the benzene ring.[29]

(9) Berliner, E. *Prog. Phys. Org. Chem.* **1964**, *2*, 253.
(10) Carey, F. A.; Sundberg, R. J. "Advanced Organic Chemistry Part A: Structure and Mechanisms"; Plenum Press: New York, 1977; pp 385–387.
(11) (a) Schofield, K. "Aromatic Nitration"; Cambridge University Press: Cambridge, 1980; pp 1–127. Olah, G. A. "Friedel-Crafts Chemistry"; Wiley: New York, 1973. (b) Bamford, C. H.; Tipper, C. F. H. "Comprehensive Chemical Kinetics"; Elsevier: Amsterdam, 1972; Vol. 13, pp 1–406.
(12) Hands, A. R.; Katritzky, A. R. *J. Chem. Soc.* **1958**, 1754.
(13) Stock, L. M.; Brown, H. C. *Adv. Phys. Org. Chem.* **1963**, *1*, 35.
(14) de la Mare, P. B. D.; *Acc. Chem. Res.* **1974**, *7*, 361.
(15) Stock, L. M.; Baker, F. W. *J. Am. Chem. Soc.* **1962**, *84*, 1661.
(16) Stock, L. M.; Brown, H. C. *J. Am. Chem. Soc.* **1960**, *82*, 1942.
(17) de la Mare, P. B. D.; Hassan, M. *J. Chem. Soc.* **1957**, 3004.
(18) de la Mare, P. B. D.; Hall, D. M.; Harris, M. M.; Hassan, M.; Johnson, E. A.; Klassen, N. V. *J. Chem. Soc.* **1954**, 4450.
(19) de la Mare, P. B. D.; Hall, D. M.; Harris, M. M.; Hassan, M.; Johnson, E. A.; Klassen, N. V. *J. Chem. Soc.* **1962**, 3784.
(20) Marino, G. *Tetrahedron* **1965**, *21*, 843.
(21) Marino, G. *Adv. Heterocycl. Chem.* **1971**, *13*, 235.
(22) Olah, G. A.; Kuhn, S. J.; Hardie, B. A. *J. Am. Chem. Soc.* **1964**, *86*, 1055.
(23) Clementi, S.; Linda, P.; Marino, G. *J. Chem. Soc. B* **1971**, 79.
(24) Linda, P.; Marino, G. *J. Chem. Soc. B* **1970**, 43.
(25) Joule, J. A.; Smith, G. F. "Heterocyclic Chemistry", 2nd ed.; Van Nostrand Reinhold: London, 1978.
(26) Katritzky, A. R.; Terem, B.; Scriven, E. V.; Clementi, S.; Tarhan, H. O. *J. Chem. Soc., Perkin Trans 2* **1975**, 1600. Katritzky, A. R.; Tarhan, H. O.; Terem, B. *J. Chem. Soc., Perkin Trans. 2* **1975**, 1620. Katritzky, A. R.; Clementi, S.; Tarhan, H. O. *J. Chem. Soc., Perkin Trans. 2* **1975**, 1624.
(27) Ingold, C. K. "Structure and Mechanism in Organic Chemistry", 2nd ed.; Cornell University Press: Ithaca, NY, 1969; p 290.
(28) Hoggett, J. G.; Moodie, R. B.; Penton, J. R.; Schofield, K.; "Nitration and Aromatic Reactivity"; Cambridge University Press: Cambridge, 1971.
(29) Paquette, L. A. "Modern Heterocyclic Chemistry"; W. A. Benjamin: New York, 1968.

Table IV. Relative Reactivities of Substituted Benzenes[a]

substituent	log k_{rel} (nitration)	substituent	log k_{rel} (nitration)
OCH_3	1.31	NHCOMe	-2.23
Me	1.03	$CH_2NMe_3^+$	-4.43
H	0.00	COMe	-4.45
CH_2CH_2OMe	0.11	CO_2H	-4.74
I	-0.76	NH_3^+	-6.79
F	-0.93	$CONH_2$	-6.87
CH_2OMe	-1.11	CN	-7.29
Br	-1.13	NO_2	-8.39
Cl	-1.18		
	Multiple Substituents		
1-NH_2-2-Cl-4-NO_2	2.88	1-Me-4-NO_2	-3.11
1,3-$(OMe)_2$	1.31	1,2-Cl_2	-3.15
1,3,5-Me_3	1.24	1-NO_2-4-Cl	-6.53
1,4-Me_2	1.02	1,3,5-Me_3-2,4-$(NO_2)_2$	-8.42
1-NH_2-4-NO_2	0.19	1-Me-2,4-$(NO_2)_2$	-10.41
1-NO_2-2,4-$(OMe)_2$	-1.50		

[a] Reference 26.

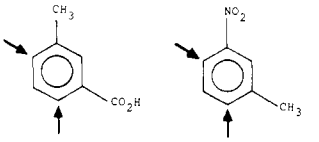
Polycyclic aromatic hydrocarbons (PAH) such as naphthalene, anthracene, phenanthrene, etc. are more reactive to electrophilic aromatic substitution than benzene due in part to the greater possibilities for delocalization of charge in the transition state. Moreover, for substituted systems of this type, the effect of a substituent is generally greatest on the ring that bears the substituent. For example, the substituted ring of 2-methylnaphthalene is activated relative to the nonsubstituted ring.[8]

Concerning heteroaromatic comounds, it is important to note first that some heteroaromatics do not undergo EAS reactions of any kind. Examples of systems inert to typical EAS conditions are oxadiazoles, thiadiazoles, and deactivated diazenes among others.[25,30] Furthermore, some heteroaromatics react at the heteroatoms rather than at carbon sites, and some heteroaromatics are prone to decomposition in strongly acidic media. For example, pyridine and pyrrole react with halogens and acyl chlorides at the nitrogen atom,[29] and when pyrrole is subjected to aqueous mineral acid, polymeric products result.[30] In general, the π-excessive heteroaromatics (furan, pyrrole, thiophene, etc.) are more reactive than benzene, while the π-deficient heteroaromatics (pyridine, diazines, etc.) are less reactive than their benzenoid counterparts.[25,30] More information on the reactivity of heteroaromatics is available in several excellent sources.[25,29,30]

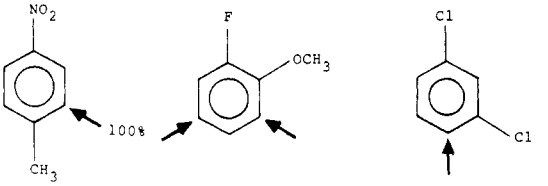
The data in Tables II and III also show that substituents play an important role in determining relative reactivity. This point is further illustrated by the results in Table IV for nitration of a variety of substituted benzenes.[26] The entries in the bottom part of the table show the effects of multiple substituents. The relative reactivities reveal the familiar trends characteristic of the electron-donating and -withdrawing abilities of the functional groups. Thorough coverage of substituent effects can be found in Topsom's review.[31]

C. Reactive Sites of Aromatic Systems. Some general observations can be made concerning the influence of substituents on the regiochemistry of electrophilic aromatic substitutions, though specific details of the reaction conditions can significantly affect product ratios. Examination of the isomer distributions for a wide range of typical reactions yields the following approximate distributions for the different classes of substituents: (1) strongly withdrawing groups such as NO_2, NR_3^+, and CN give an ortho:meta:para ratio of 5:90:5, (2) moderately withdrawing groups such as CO_2R and CO_2NR_2 give a ratio of 15:80:5, (3) halogens exhibit a ratio of 35:10:55, (4) alkyl groups yield a ratio of 50:0:50, and (5) strongly donating groups such as OR, SR, and NR_2 show a ratio of 20:0:80. The separation of a strongly withdrawing group from an aromatic ring by methylene groups reduces the meta preference. When one methylene group intervenes the ratio changes from approximately 5:90:5 to 20:60:20. Two or more intervening methylene groups cause predominantly para substitution.[28,32-34]

The substitution patterns for polysubstituted benzenes depend on whether the substituents reinforce a particular substitution pattern or work in opposition and on the directing strength of the substituents. Ortho/para-directing substituents usually dominate meta-directing substituents, as shown in the bromination of compounds **3** and **4** (the arrows indicate sites of bromination).[35] When

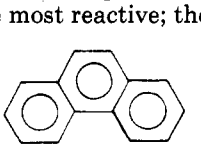


the directing effects of substituents reinforce each other, the expected isomers predominate, as in the halogenation of compound **5**.[35a] In a substituted benzene with two meta-directing groups, all possible meta products are usually formed.[7b] If both substituents are ortho/para directing, the substitution pattern depends on the relative strength of the directing ability of the substituents. Thus, for **6** the methoxy group dominates.



Steric factors also play a role in determining substitution patterns in EAS reactions. They are most important when the aromatic ring contains two ortho/para-directing substituents ortho or meta to each other. In both cases, substitution at the 4-position is the major product, as illustrated by compound **7**.[8] For a more thorough coverage of substituent effects, several reviews are available.[31,33]

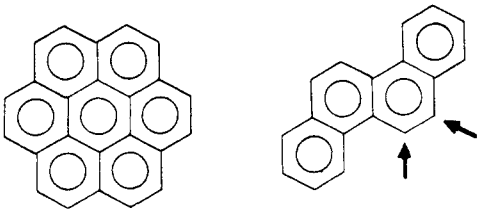
The most reactive sites for polycyclic aromatic hydrocarbons (PAH) may be predicted in an analogous fashion. The α-position of naphthalene is the kinetically preferred site for EAS because α-attack leads to a more charge-delocalized transition state and intermediate than attack at the β-position. However, sulfonation and Friedel–Crafts alkylation of naphthalene can produce mixtures of 1- and 2-isomers.[36] The central rings of anthracene and phenanthrene (**8**) are the most reactive; the 9-chloro, -bromo,



and -nitro derivatives of **8** are the major isomers produced.

(30) Newkome, G. R.; Paudler, W. W. "Contemporary Heterocyclic Chemistry"; Wiley: New York, 1982.
(31) Topsom, R. D. *Prog. Phys. Org. Chem.* **1967**, *12*, 1–20.
(32) Norman, R. O. C.; Radda, G. K. *J. Chem. Soc.* **1961**, 3610.
(33) Ferguson, L. N. *Chem. Rev.* **1952**, *50*, 47.
(34) Kovacic, P.; Hiller, J. J. *J. Org. Chem.* **1965**, *30*, 1581, 2871.
(35) (a) Scheufelen, A. *Liebigs Ann. Chem.* **1885**, *231*, 152. (b) Jacobsen, O. *Ber. Dtsch. Chem. Ges.* **1881**, *14*, 2347.
(36) Jackson, A. H. *J. Chem. Soc., Perkin Trans. 2* **1981**, 286.

Some polycyclic aromatic hydrocarbons (PAH) are symmetrical and have several equivalent sites of substitution, while others can be thought of as being composed of smaller PAH units and their substitution patterns predicted from these. For example, coronene (**9**) has 12 equivalent positions, and chrysene (**10**) substitutes at the indicated positions, which are analogous to the 9- and 10-positions of phenanthrene.[37]



The reactive sites of heteroaromatic compounds arise from the most stable arenium ion intermediate, as in substituted benzenes. Unsubstituted furan, thiophene, and pyrrole react predominantly α to the heteroatom.[25,38] Furthermore, the patterns for substituted five-membered monoheteroaromatics are as follows. An electron-donating substituent at C2 directs substitution to C5, while an electron-donating group at C3 directs substitution to C2. An electron-withdrawing substituent at C2 or C3 generally directs substitution to C5.[25,30] Benzo-fused analogues of the above compounds substitute α to the heteroatom for benzofuran and β for indole and benzothiophene.[21,23] An activating substituent on the heteroaromatic ring directs substitution to this ring for benzo-fused systems, while a deactivating group on the heteroaromatic ring results in substitution occurring on the benzene ring.[25,30] The presence of additional heteroatoms in the five-membered heteroaromatic ring substantially deactivates these compounds, making them much less reactive than the monoheteroatom species. Some of the more common heteroaromatics that fit this category are the imidazoles, oxazoles, thiazoles, pyrazoles, isoxazoles, and isothiazoles. For the 1,3-azoles C4 and C5 substitution are found, while for the 1,2-azoles C4 substitution usually dominates.[25] Not surprisingly, the benzo-fused analogues of the above compounds tend to substitute on the benzene ring, e.g., for benzothiazole.[29]

Heteroaromatic derivatives of benzene, such as pyridine, substitute β to the heteroatom. Activating groups at positions 2 or 4 in pyridine reinforce β-substitution.[29,30] Quinoline, isoquinoline, and the other benzo-fused six-membered heteroaromatics tend to substitute at the α positions in the benzene ring. The presence of additional heteroatoms, as in the diazenes, again causes substantial deactivation so there are few synthetically useful EAS reactions for such compounds.[25] More detailed coverage of EAS reactions of heteroaromatics can be found in the comprehensive book by Newkome and Paudler.[30]

III. Implementation of EAS Reactions in CAMEO

One of the major aims of the CAMEO project is the organization of the literature data into generalized rules governing reactivity.[1] The rules are then incorporated into the program, where their utility can be fully tested. In the present case, algorithms are required to deal with the topics in the previous section, i.e., electrophile classification, relative reactivity, and identification of the most reactive sites in aromatic compounds. This section summarizes the current algorithms for EAS and their implementation in the program. Only substitution of hydrogen, the predominant case, is being considered at present. A separate mechanistic module has been added to CAMEO for EAS; however, all of the mechanistic modules use primarily the same routines for the graphical input and display of structures and for the perception of many structural features of molecules.[1-6] A simplified flow chart showing the order of processing for EAS reactions in CAMEO is given in Figure 1.

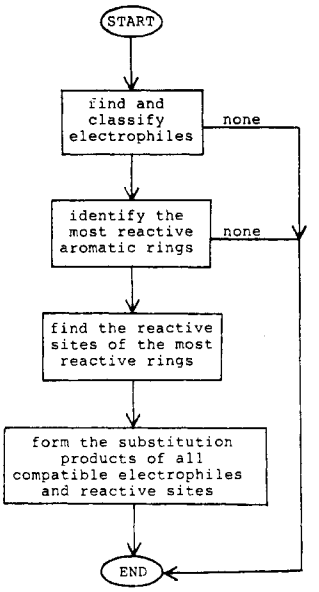


Figure 1. Overview of the mechanistic evaluation of EAS reactions.

A. Electrophile Perception. Electrophiles for EAS reactions are recognized in a stepwise manner in CAMEO. As each electrophile is found, it is rated and cataloged according to its reactivity. The index of reactivity for an electrophile, designated ETYP, varies from 1 for very weak electrophiles to 3 for very reactive electrophiles corresponding to the classification described above and illustrated in Table I. This rough measure of reactivity is used in conjunction with an estimate of the aromatic substrate's reactivity to gauge the likelihood of reaction as discussed below in section D. Also, only the most reactive electrophiles are considered for processing. That is, electrophiles with a higher ETYP will cause those with a lower ETYP to be ignored. In addition, among the moderate electrophiles, acylium ions are taken as more reactive than carbenium ions.

The types of electrophiles and the order in which they are perceived as follows. First the reactants are scanned for the presence of reactive electrophiles derived from mineral acids, such as NO_2^+ from HNO_3 and SO_3 from H_2SO_4. Electrophiles that are unsaturated cations such as acylium, diazonium, and iminium ions are recognized next, followed by electrophiles containing halogen–heteroatom bonds such as chlorosulfonic acid, SCl_2, sulfonyl halides, and molecular halogens. Most members of the latter group require catalysts, so the detection of Lewis acids is also performed at this point. Electrophiles featuring heteroatom–heteroatom single bonds such as thiocyanogen and peracids are identified next. The final class of electrophiles perceived is derived from compounds containing an aliphatic carbon singly bonded to a heteroatom including alkyl halides, tosylates, alcohols, and ethers. These percursors yield carbenium ions under typically acidic reaction conditions. Wagner–Meerwein rearrangements to more stable isomers are also considered

(37) Gore, P. H.; Kamounah, F. S.; Miri, A. Y. *Tetrahedron* **1979**, *35*, 2927.
(38) Cordell, G. A. *J. Org. Chem.* **1975**, *40*, 3161.

Table V. Parent Systems and Their Corresponding RN_0's under Acidic and Neutral Conditions

parent system	RN_0 acidic	RN_0 neutral
benzene	0.0	0.0
pyrrole	5.0	10.0
furan	4.0	8.0
thiophene	1.0	6.0
oxazole	−5.0	3.0
isoxazole	−5.5	3.0
imidazole	−8.0	4.0
thiazole	−9.0	1.0
pyrazole	−9.7	4.0
isothiazole	−12.0	1.0
pyridine	−12.0	−10.0
diazines	−13.0	*a*
triazoles	*a*	*a*
tri-, tetrazines	*a*	*a*

[a] Unreactive.

and performed here. Finally, it should be noted that if an electrophile can react intramolecularly to form a five-, six-, or seven-membered ring, then this electrophile is assumed to be the most reactive in its class.

B. Calculation of Relative Reactivity. One of the most important aspects of predicting the products of EAS reactions is determination of the relative reactivity of each aromatic ring present. The routine in CAMEO used for this purpose processes each ring individually. As stated earlier, the key factors that influence reactivity of aromatic rings are (1) the type of parent system, (2) the electronic nature and arrangement of substituents, and (3) the type of reaction. The relative reactivity of an aromatic ring is represented in CAMEO by an index termed RN (reactivity number). The RN is the log of the rate of reaction for the aromatic ring relative to benzene (RN = 0) under the same conditions. The influencing factors listed above have been dealt with in the computational procedure, as discussed below.

1. Each aromatic ring present is given a base reactivity number depending on the type of parent system for the ring. Substituted benzenes and benzene rings that are part of a polycyclic aromatic hydrocarbon are assigned a base reactivity number, RN_0, of 0. The RN_0 assigned to heteroaromatic rings is dependent on the number, type, and arrangement of heteroatoms present in the ring and on the type of reaction. π-excessive heteroaromatics, which are more reactive under acidic and neutral conditions than benzene, have values of RN_0 greater than 0. E.g, under acidic conditions the RN_0's for pyrrole, furan, and thiophene are 5.0, 4.0, and 1.0, respectively. π-deficient heteroaromatics and other heteroaromatics that are less reactive in acidic media than benzene are assigned negative values of RN_0. Some examples consistent with the data in Table III are as follows: pyridine (−12.0), diazines (−13.0), pyrazole (−9.7), and isoxazole (−5.5). The 1,2- and 1,3-azoles are more reactive under neutral conditions than benzene and are given values of RN_0 greater than 0, e.g., imidazole (4.0), oxazole (3.0), and thiazole (1.0).[25,30] Heteroaromatics that are not reactive to EAS are noted and assigned an appropriately negative value of RN_0 which causes them to be eliminated from consideration as reactive rings. A listing of parent systems and their corresponding RN_0 under acidic and neutral conditions is shown in Table V.

2. The next step in calculating the relative reactivity is to consider the effects of substituents including aromatic and nonaromatic ring fusions and to adjust the RN value accordingly. This is accomplished by first perceiving and cataloging the substituents on the ring according to their electronic nature, i.e., the substituents are grouped as electron donating, strongly electron withdrawing, etc. The presence of any aromatic and nonaromatic ring fusions is also noted at this time. Substituents that would be protonated under the reaction conditions, such as amines under acidic conditions, are recognized and placed among the set of withdrawing groups rather than in the group they would otherwise represent. The final reactivity number, RN, for each ring is then calculated using a structure/reactivity correlation equation, eq 2. The equation was

$$RN = RN_0 + 6.6NNIT + 3.3NARO + 1.7NFUS + 1.3NOS + 1.0NCARB - 0.2NSTER - 1.0NHALO - 2.2NDWD - (4.5NMWD) - (MULT \times NSWD) \quad (2)$$

developed by performing least-squares fits to reactivity data for nitration by HNO_3 at 25 °C in 94% H_2SO_4,[26] chlorination by Cl_2 at 25 °C in $PhNO_2/FeCl_3$,[11b] and isopropylation at 25 °C in $MeNO_2/AlCl_3$.[11b] The significance of the variables in the equation is discussed below.

The variables are as follows: NNIT, number of substituents with a nitrogen atom attached to the ring (excluding withdrawing groups such as nitro and nitroso), e.g., NH_2, NHR, NR_2, NHCOR, and NHAr; NARO, number of substituents which are aromatic rings (does not include aromatic ring fusions), e.g., Ph and Naph; NFUS, number of atoms in aromatic ring fusions, as in polycyclic aromatic hydrocarbons; NOS, number of substituents with an oxygen, sulfur, or phosphorus atom attached to the ring, e.g., OH, OR, OCOR, OAr, SR, and PR_2; NCARB, number of attached alkyl groups, e.g., Me, Et, etc. NSTER, number of sterically bulky alkyl groups, e.g., *i*-propyl, *tert*-butyl; NHALO, number of halogen substituents; NDWD, number of withdrawing groups separated from the ring by a heteroatom, e.g., NHCOR, OCOR, etc.; NMWD, number of moderately withdrawing substituents, e.g., COR, COOH, and COOR, and strongly withdrawing groups separated from the ring by a methylene group such as $CH_2NR_3^+$; NSWD, number of strongly withdrawing substituents, e.g., NO_2, CN, SO_3H, and NR_3^+. MULT is an adjustment factor for multiple groups. It is 7.4 for NSWD = 1, 5.7 for NSWD = 2, and 4.6 when NSWD = 1 and NCARB or NOS are not zero.

Table VI shows the values of the above variables along with the calculated RN and experimental rate data (log k_{rel}) for several representative aromatic compounds. As implied by eq 2, the RN's for multiply substituted compounds are calculated in an additive manner. This additivity of substituent effects holds for many aromatic compounds, while for other multiply substituted compounds adjustments must be made. For example, toluene, *p*-xylene, and mesitylene have nearly the same experimental log k_{rel} for nitration;[26] therefore the value of NCARB is adjusted from 2 or 3 to a value of 1 for these types of systems. Adjustments are made for other types of multiply substituted aromatic compounds that have sufficient experimental data available. At this time, NOS is set to 1 for dialkoxy- and dihydroxybenzenes, NOS = 0 for alkoxy- and hydroxypyridine, and NHALO is increased to 3 for dihalobenzenes. A plot of the experimental log k_{rel} vs. RN for 98 aromatic compounds, consisting of substituted benzenes, substituted heteroaromatics, and polycyclic aromatics, is shown in Figure 2.

The RN for each aromatic ring is calculated as discussed above and stored so that the set of the most reactive rings can be formed. This set consists of the ring with the greatest RN and those aromatic rings that have an RN value within one unit of the maximum RN. The processing of this section of the program is summarized in Figure 3. The set of reactive aromatic rings is then channeled to the

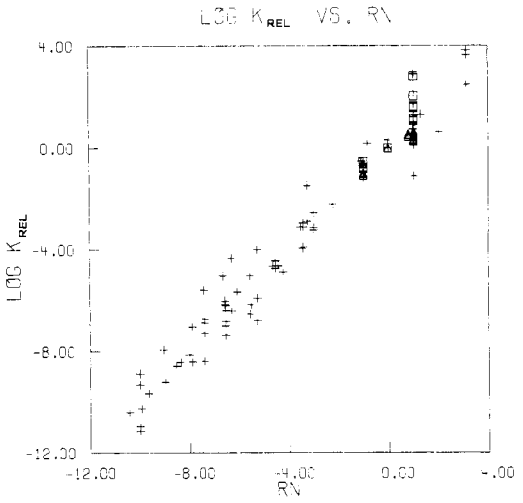


Figure 2. Plot of experimental log k_{rel} (ref 26) vs. calculated reactivity number for 98 aromatic compounds: (+) nitration; (▲) isopropylation; (□) chlorination.

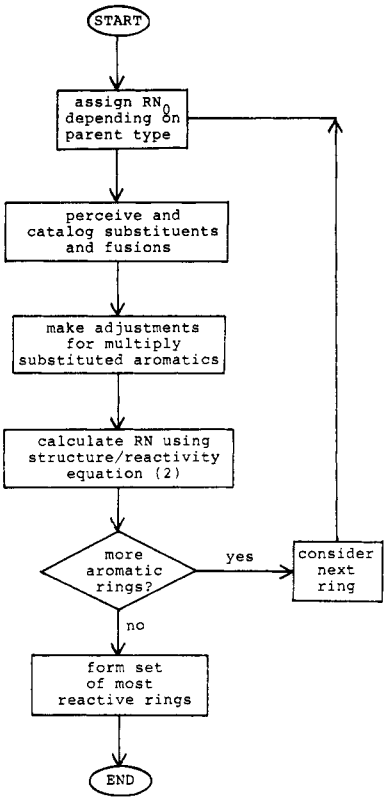


Figure 3. Outline of the calculation of the relative reactivity of aromatic rings.

next step in the evaluation of EAS reactions, perception of reactive sites, discussed below.

C. Perception of Reactive Sites. As in the calculation of relative reactivity, reactive sites are perceived by processing each ring individually. Only sites with an attached hydrogen need be considered. Also, the values of NFUS (number of aromatic ring fusions) and the other variables representing the ring's substituents are used to gauge the complexity of the ring. The routine that perceives the reactive sites is segmented to deal with the following broad classes of aromatic compounds: (1) substituted, isolated (nonfused) benzenoid rings, (2) substituted and nonsubstituted heteroaromatic rings, and (3) substituted and nonsubstituted polycyclic aromatic hydrocarbons. An overview of the perception of reactive sites is presented in Figure 4.

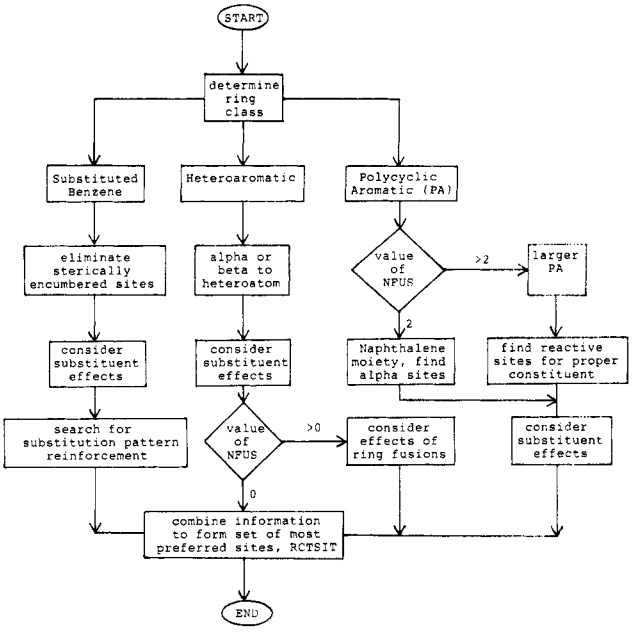
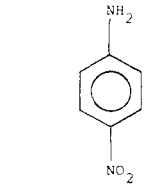
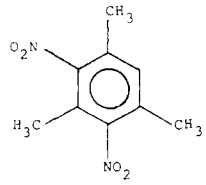
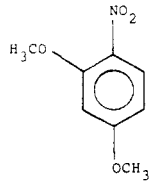
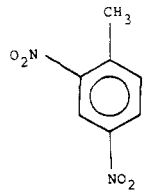
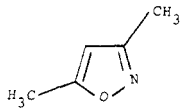
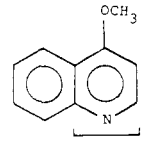
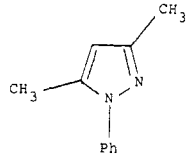


Figure 4. Overview of the perception of reactive sites.

1. Substituted Benzenes. If the value of NFUS for the ring under consideration is 0, the ring is isolated. If the ring is a heteroaromatic, it is evaluated as discussed in section 2 below. The first step in processing aromatic hydrocarbons is to remove from consideration sites that cannot be substituted such as those which are sterically inaccessible. These encumbered sites include sites ortho to a tertiary substituent, such as a *tert*-butyl group, and sites between two substituents with a 1,3-relationship. Formation of a 1,2,3-trisubstituted ring by addition to a 1,2-disubstituted system is also disfavored in comparison to attack at a more open site. However, if there are no alternatives, formation of sterically disfavored products is considered. Once these sites have been eliminated, the directing effects of the substituents on the ring are considered. The following order of relative directing strength is used to determine which substituents control the substitution pattern: NR_2 > OR > SR > Ar > R > X > moderately withdrawing (COOR, $CH_2NR_3^+$) > strongly withdrawing (NO_2, CN, NR_3^+). The viable sites ortho and para to donating groups and meta to withdrawing groups are found next. The directing order given above is then used to determine which substituents control the substitution pattern. Sites which are reinforced electronically by multiple substituents are also noted. The preceding information is then combined to form RCTSIT, the set of the most reactive sites. Some examples are provided in Table VII. For instance, in the second example in Table VII the ethyl groups control substitution, therefore sites 3, 4, and 6 are preferred electronically. Site 6 is eliminated due to steric inaccessibility. Site 3 is selected as the best possible reactive site because it is reinforced by both an ethyl group and the sulfonic acid group.

2. Heteroaromatics. The open sites α to the heteroatom in π-excessive heteroaromatics and the open sites β to the heteroatom in π-deficient heteroaromatics are found first. Next, the directing effects of the substituents (as outlined in section IIC) are taken into account. Once again, RCTSIT is formed by combining the preceding information to find the most preferred sites of substitution. If the heteroaromatic ring is part of a polycyclic aromatic compound, the directing effects of the ring fusions must be considered. For example, benzofuran substitutes as in furan; indole and benzothiophene, however, substitute β

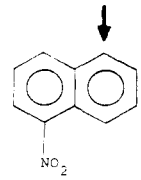
Table VI. Values of the Variables Used in Eq 2 for Several Aromatic Compounds

compound	RN_0	NIT	NOS	NARO	NFUS	NCARB	NHALO	NSWD	MULT	RN	log k_{rel}[a]
	0.0	1	0	0	0	0	0	1	7.4	−0.80	0.19
	0.0	0	0	0	0	3	0	2	5.7	−8.4	−8.42
	0.0	0	2	0	0	0	0	1	4.6	−2.0	−1.50
	0.0	0	0	0	0	1	0	2	5.7	−10.4	−10.41
	−5.5	0	0	0	0	2	0	0	0	−3.5	−3.95
	−12.0	0	1	0	2	0	0	0	0	−7.3	−6.19
	−9.7	0	0	1	0	2	0	0	0	−4.4	−4.89

[a] Relative rate for nitration taken from ref 26.

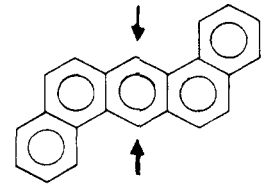
instead of α to the heteroatom.[21,23] The determination of the reactive sites for two heteroaromatics is also illustrated in Table VII.

3. Polycyclic Aromatics. When the value of NFUS for the ring under consideration is 2 or greater and there are no heteroatoms in the ring, a fused benzenoid ring is present. If NFUS equals 2 a naphthalene moiety exists. The α sites are stored in RCTSIT for unsubstituted naphthalenes, while the presence of a withdrawing group causes substitution at the α site farthest from the withdrawing group, as in the nitration of compound 11. If the



naphthalene is being sulfonated, RCTSIT is adjusted to contain β sites, as it has been shown that this is the preferred site of substitution.[36]

An NFUS of 3 or more indicates the presence of a larger polycyclic aromatic compound, such as chrysene or pyrene. The reactive sites for these larger PAH are found by considering them to be composed of smaller units such as naphthalene and anthracene, depending on the number and arrangement of the rings comprising the PAH. For example, for dibenz[*a*,*h*]anthracene (**12**) RCTSIT contains the atoms indicated by the arrows, which are analogous to the 9 and 10 positions of anthracene. A useful rule for



determining RCTSIT for larger PAH is that substitution

usually does not occur in 1,2,3,4-tetrafused rings, as also illustrated by compound **12**.[39] This notion is incorporated into the steric factor as in the last example in Table VII.

D. Formation of Products. Each potentially reactive ring is processed as described above to determine the set of reactive sites (RCTSIT) associated with it. The rings are then analyzed one at a time by a routine that decides if product formation is appropriate with the available electrophiles. Valid products are subsequently formed and displayed. All combinations of electrophiles and reactive sites are considered. A potential product is displayed only if the value of ETYP, the reactive strength of the electrophile, and the value of RN, the relative reactivity of the ring, are compatible. For example, if the value of ETYP is 2 (a moderately reactive electrophile), it may substitute only activated rings, i.e., those rings with an RN greater than 0. The values of ETYP and the corresponding RN's necessary to effect substitution are presented in Table VIII.

IV. Sample Sequences

Several illustrative examples of predictions made by the program are presented in Scheme I and discussed below. All products of EAS reactions predicted by CAMEO are shown in each case. Experimental yields are included when available. The nitration of the substituted benzene in example 1 illustrates that 1,2,3-trisubstituted products are usually not formed if another product, equally favorable electronically, is possible.[40]

A selective substitution of the more reactive ring (RN = 9.3 for the thiophene ring vs. RN = 3.3 for the benzene ring) is shown in example 2. Here the program predicts the formation of only one α-substituted product due to the reinforcement of that site by the conjugating group located at position 3. Substitution at position 5, reported as a minor product, is not predicted by the program.[41]

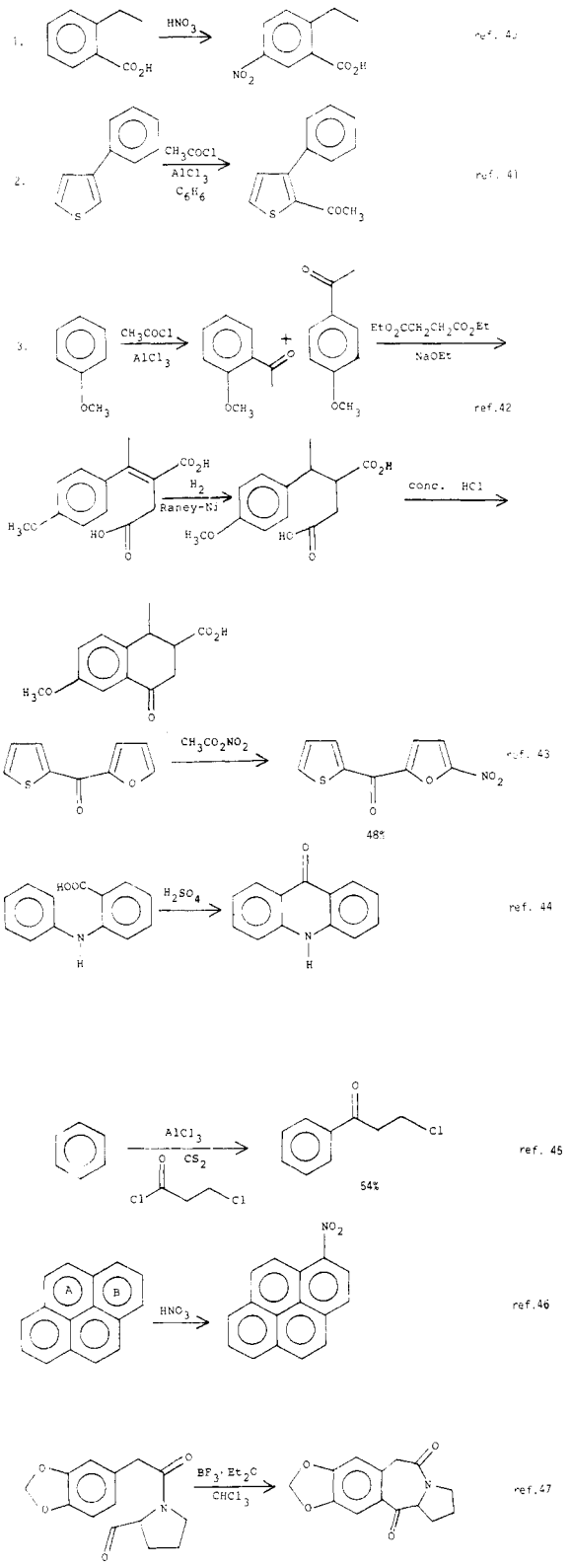
The third example illustrates a cyclization reaction, forming a substituted naphthalene, using two acylations along with an intermediate condensation step. The first acylation shows the expected ortho and para products; in practice the para isomer predominates and is used in the succeeding Stobbe condensation. The ring-closure step following the condensation is an example of an intramolecular Friedel–Crafts acylation. The reaction is especially facile for the formation of six-membered rings and provides the only predicted and reported product.[42]

The reactivity of the π-excessive heteroaromatics is sufficiently varied to allow selective reaction of one heteroaromatic ring in a multiheteroaromatic ring system, as shown in example 4. Here, although both heteroaromatic rings are deactivated by the bridging carbonyl, the higher reactivity of furan (RN_0 = 4.0) compared to thiophene (RN_0 = 1.0) accounts for the selectivity. The α product is expected for this substituted furan. This example also illustrates the program's ability to recognize acetyl nitrate as a source of NO_2^+. This mild nitrating agent is used in practice to limit decomposition of starting material which is more problematic with the traditional HNO_3/H_2SO_4.[43]

Polycyclic heteroaromatic compounds can also be made with intramolecular Friedel-Crafts acylations, as in the synthesis of acridone shown in example 5. This reaction also illustrates the program's ability to favor the intramolecular reaction in preference to a potential intermolecular sulfonation.[44]

The reaction in example 6 demonstrates electrophile selectivity. In the dihalo reagent two potential electrophilic sites exist; however, reaction is observed and predicted only from the more reactive acid chloride.[45]

Scheme I



(39) Clar, E. "Polycyclic Hydrocarbons"; Academic Press: New York, 1964.
(40) Giek, G. *Ber.* **1896**, *29*, 2537.
(41) Gjos, N.; Gronowitz, S. *Acta Chem. Scand.* **1970**, *24*, 99.
(42) Jacques, J.; Horeau, A. *Bull. Soc. Chim. Fr.* **1950**, 512.
(43) Gilman, H.; Young, R. V. *J. Am. Chem. Soc.* **1934**, *56*, 464.
(44) Patai, S. "The Chemistry of the Quinonoid Compounds"; Wiley: New York, 1974; Part 1.
(45) Hart, R. T.; Teble, J. *J. Am. Chem. Soc.* **1950**, *72*, 3286.

Table VII. Predicting the Reactive Sites of Aromatic Compounds by Combining the Influencing Factors

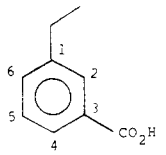
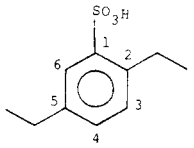
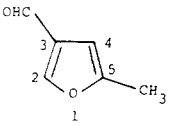
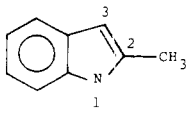
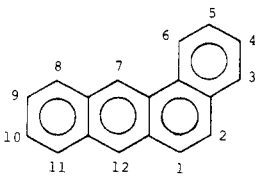
compd	preferred sites				reactive sites
	heteroatom influence	electronic effect	sterically preferred	electronic reinforcement	
		4, 6	4, 5, 6		4, 6
		3, 4, 6	3, 4	3	3
	2	2	2	2	2
	3	3	3	3	3
		1, 2, 7, 12	7, 12		7, 12

Table VIII. Values of RN and the Corresponding ETYP Necessary for Substitution To Occur

RN	ETYP
>2.0 (strongly activated ring)	1 (very weak electrophile)
>0.0 (nondeactivated ring)	2 (moderately reactive electrophile)
≥−12.0 (nearly all aromatic rings)	3 (very reactive electrophile)

The nitration of pyrene in example 7 illustrates the evaluation of reactive sites for a polycyclic aromatic hydrocarbon. The observed site of substitution is on ring B rather than ring A, due to the rule that 1,2,3,4-tetrafused rings are usually less reactive in EAS.

Electrophilic aromatic substitution reactions have also found utility in natural product synthesis, as illustrated by an intermediate step in the synthesis of the alkaloid cephalotaxine in example 8. Here, both carbonyls are potential electrophiles; however, the formation of the seven-membered ring is much more facile than four-membered ring formation, and again only one product is reported[47] and predicted.

V. Conclusion

The scope of the computer synthesis program, CAMEO, has been broadened by the implementation of modules that treat an important class of reactions, electrophilic aromatic substitution. The evaluation of these reactions requires consideration of electrophile reactivity, relative reactivity of aromatic rings, and the perception of reactive sites for a wide variety of aromatic systems. Necessary algorithms based on literature data were devised to address these issues and to permit the program to make sophisticated predictions on the outcomes of electrophilic aromatic substitutions.

Acknowledgment. Gratitude is expressed to the National Science Foundation for support of this work.

(46) Dewar, M. J. S.; Mole, T.; Wardord, E. W. T. *J. Chem. Soc.* **1956**, 3581.

(47) Auerbach, J.; Weinreb, S. M. *J. Am. Chem. Soc.* **1972**, *94*, 7172.