EXPERIENTIA

Vol. XII - Fasc. 5 Pag. 165–204 15. V. 1956

The Mechanism of Electrophilic Aromatic Substitution*

Recent Contributions by Kinetic Methods

By HCH. ZOLLINGER[1], Basle

I.–*Introduction*

Among the reactions of organic chemistry substitution reactions upon aromatic hydrocarbons are undoubtedly those of which the fundamentals have so far attracted the greatest amount of attention. When, at the beginning of the second half of the 19th century, the systematic study of the chemistry of carbon compounds began, the interest of organic chemists was to a great extent directed to these reactions, in which the hydrogen atoms of benzene, naphthalene and other aromatic hydrocarbons are relatively easily replaced by other groups. The observations made at that time served to develop fundamental principles which are still of significance at the present time, i.e. the *theory of the aromatic state*–initiated by the deduction of the formula of benzene by KEKULÉ (1865) – and the *rules of orientation* in aromatic compounds, the latter being associated more especially with the name of HOLLEMAN.

We do not intend to discuss more fully in this paper these two pillars of aromatic chemistry. We merely wish to mention that KEKULÉ's work has passed through several stages of development and has now become one of the roots of the application of the quantum theory to chemical problems (PAULING, COULSON, MULLIKEN, H. KUHN and many others). On the other hand, the rules of orientation have become (mainly due to INGOLD and HUGHES) the most important means of understanding the effect of substituents and side chains in all organic reactions, not only aromatic reactions.

In this review we wish to discuss a few conceptions connected with the *mechanism* of electrophilic aromatic substitution which have been collected very recently, and to summarize them so that the essential points can be understood even by those who are not familiar with the details of the subject. The present is a propitious time for this, because measurements of a kinetic character can now be carried out which enable definite statements to be made with regard to the course of such reactions.

In an electrophilic aromatic substitution an atom (or atomic grouping) bound to an aromatic ring is replaced by another atom (or atomic grouping), the newly introduced substituent being bound by an electron pair originating from the aromatic compound. In a narrower sense is meant the substitution of an aromatically linked hydrogen atom by an electrophilic reagent, as expressed by the general equation (1) in which Ar is

$$X^{\oplus} + \text{Ar-H} \rightarrow \text{Ar-}X + \text{H}^{\oplus} \qquad (1)$$

an aryl radical, and $X^{\oplus}$ is an electrophilic reagent.

At the present time well over a dozen such reactions are known. The most important are, nitration, sulphonation, the introduction of the nitroso group and the various halogenations. Organic residues can also be introduced as, for example, in Friedel-Crafts reactions and in azo coupling reactions. In all these processes the reagent is electrophilic, that is to say, it endeavours to combine additively with an atom having a relatively high electron density. Consequently, the more recent theoretical treatment of the rules of orientation is more especially concerned with calculating the electron densities at the individual carbon atoms of ring systems.

In principle the formulation of aromatic substitution (1) is very simple, but generations of chemists have been occupied in arriving at an exact explanation of the mechanism of the reactions.

The work carried out hitherto can be divided into two groups: On the one hand, the mechanism of the exchange of H for X has been investigated, namely the order of sequence in time of the additive combination of the electrophilic reagent and the release of the pro-

---

* From a lecture held at University College, London, May 2, 1956.

[1] Research Department of the Dyestuff Division of CIBA Ltd., Basle, and Institute of Dye Chemistry, University of Basle.

ton. In the following discussion we shall refer to this process as the "substitution proper".

On the other hand, it has been found that the question of the forms of equilibrium in which X and the aromatic compound are reactive at all is of decisive importance. The earlier theories of aromatic substitution have in the main neglected these preceding reactions and equilibria. Generally speaking the electrolyte theory based on the study of equilibria and the ion concept have been applied in organic chemistry rather late. Before dealing with the mechanism of the substitution proper, we must deal with the equilibria preceding the substitution reactions with reference to a few typical examples.

II.–*Equilibria Preceding Substitution Reactions*

Apart from a few exceptions, electrophilic aromatic substitutions are generally performed in a medium having a high dielectric constant. This promotes ionisation. In nearly all cases it must be assumed that at least the electrophilic reagent is present in the reaction medium in two or more forms which are in equilibrium with one another.

Thus, before it can be decided which form participates in the substitution proper, these systems of equilibrium prior to the substitution must be investigated. This applies both to the electrophilic component and to the aromatic compound to be substituted. The circumstances are especially favourable for this investigation when the reactions can be performed in a highly dilute aqueous solution, which is the case, for example, with *azo coupling*. As is known, so-called coupling components, i.e. phenols and aromatic amines (as well as a few other classes of aromatic and pseudo-aromatic compounds), readily react with diazo-compounds to form azo compounds. A correct result in the investigation of the reactive form was first obtained with the amines: **WISTAR** and **BARTLETT**[2] have shown that it is the free amine and not the ammonium ion that participates in the reaction. In analogous manner **PÜTTER**[3] was able to show that in couplings with phenols the active particle is the phenolate ion. Several groups of workers have contributed to the experimental determination of the reactive form of the diazo-compound[4]: So far only the diazonium ion has proved to be reactive in aqueous solution. In this connection, **WITTWER** and **ZOLLINGER** made a surprising and generally unexpected discovery in the course of their investigation of the diazo equilibrium system (2): The diazo-hydroxide, which is a compound that has been extensively studied ever since the classic investigations of A. **HANTSCH** (although is was never isolated), cannot exist at all in an aqueous medium:

$$\underset{\mathrm{I}}{\mathrm{Ar{-}\overset{\oplus}{N}{\equiv}N}} \underset{K_1}{\overset{+\mathrm{OH}^\ominus}{\rightleftarrows}} \underset{\mathrm{II}}{\mathrm{Ar{-}N{=}N{-}OH}} \underset{K_2}{\overset{+\mathrm{OH}^\ominus}{\rightleftarrows}} \underset{\mathrm{III}}{\mathrm{Ar{-}N{=}N{-}O^\ominus}} + \mathrm{H_2O} \qquad (2)$$

When an acid solution of a diazonium salt (I) is rendered alkaline, direct dissociation of the metastable diazo-hydroxide (II) leads to the diazotate (III). Such abnormal acid-base equilibria ($K_2 > K_1$) had previously been observed with a few substances for the first time by **SCHWARZENBACH**[5].

Substitutions that cannot be carried out in dilute aqueous solution needed much more extensive work to determine the reactive particles. The classical example of such an investigation is the work of **INGOLD**, **HUGHES** and co-workers[6] on the equilibria preceding the *nitration*. It was shown that in general the reactive form is the nitronium ion $\overset{\oplus}{\mathrm{NO_2}}$, as had been suggested 50 years earlier by **VON EULER**[7]. It has now been established that the formation of the nitronium ion takes place in the following stages:

$$2\,\mathrm{HNO_3} \rightleftarrows \mathrm{H_2\overset{\oplus}{N}O_3} + \mathrm{NO_3^\ominus} \qquad (3)$$

$$\mathrm{H_2\overset{\oplus}{N}O_3} \xrightarrow{\text{slow}} \overset{\oplus}{\mathrm{NO_2}} + \mathrm{H_2O} \qquad (4)$$

In connection with azo coupling it has been found that the most acidic form of the electrophilic component is the most reactive. The conditions are analogous with the nitrating reagents. The more acidic the nitrating particle, the more reactive it is. In fact, **HUGHES**, **INGOLD** *et al.* have shown that not only the nitronium ion itself, but also particles of the formula $\mathrm{NO_2}X$ are capable of taking part in the substitution proper. The more basic the residue X ($\mathrm{NO_2}X$ being correspondingly less acidic as a whole), the lower is its reactivity. This can be demonstrated by the series

$$\overset{\oplus}{\mathrm{NO_2}} > \mathrm{NO_2}\cdot\overset{\oplus}{\mathrm{O}}\mathrm{H_2} > \mathrm{NO_2}\cdot\mathrm{NO_3} > \mathrm{NO_2}\cdot\mathrm{OOCC_6H_5} > \mathrm{NO_2}\cdot\mathrm{OH}\ .$$

---

[2] R. WISTAR and P. D. BARTLETT, J. Amer. chem. Soc. *63*, 413 (1941).

[3] R. PÜTTER, Angew. Chem. *63*, 188 (1951).

[4] R. WISTAR and P. D. BARTLETT, J. Amer. chem. Soc. *63*, 413 (1941). – C. R. HAUSER and D. S. BRESLOW, J. Amer. chem. Soc. *63*, 418 (1941). – R. HUISGEN, Liebigs Ann. Chem. *574*, 184 (1951). – HCH. ZOLLINGER, Helv. chim. Acta *36*, 1723 (1953). – C. WITTWER and HCH. ZOLLINGER, Helv. chim. Acta *37*, 1954 (1954).

[5] G. SCHWARZENBACH, Helv. chim. Acta *26*, 420 (1943).

[6] C. K. INGOLD, E. D. HUGHES with G. A. BENFORD, E. L. BLACKALL, R. J. GILLESPIE, V. GOLD, J. GRAHAM, E. S. HALBERSTADT, E. R. A. PEELING, and G. H. WILLIAMS, J. chem. Soc. *1938*, 929; *1950*, 2400–2552; *1952*, 28, and further papers.

[7] H. VON EULER, Ann. *330*, 286 (1903), footnote 7. At the same time and independently of the work of INGOLD's school the existence of the nitronium ion was proved by J. CHÉDIN, Chim. et Ind. *56*, 7 (1946), and by F. H. WESTHEIMER and M. S. KHARASCH, J. Amer. chem. Soc. *68*, 1871 (1946).

The reverse is the case with the nucleophilic substrate (the aromatic compound to be substituted), where, as has been shown in investigations with phenol and amine coupling components, it is the basic form that takes part in the substitution.

It would be obvious to suppose that in other aromatic substitutions the various forms of electrophilic reagents could be arranged in analogous series of decreasing acidity. Taking *chlorination* as an example of a halogenation reaction, the chlorine atom can in principle participate in the substitution proper through the agency of a series of compounds, of which the following order may be given:

$$Cl^{\oplus} > Cl\cdot\overset{\oplus}{O}H_2 > Cl\cdot Cl > Cl\cdot OH\,.$$

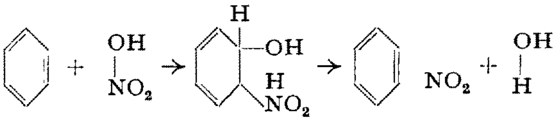
Although on purely theoretical grounds the correctness of this series can hardly be doubted, for the purposes of modern organic chemistry the matter cannot be left there. The existence and reactivity of each of these particles must be proved *experimentally*. In the case of chlorination this is of particular interest as far as the chlorinium ion, $Cl^{\oplus}$, is concerned. It might be supposed that this particle—a chlorine atom having only 6 electrons in the M-shell—would be extremely unstable, because this incomplete outermost shell (in contradistinction, for example, to that of the nitrogen atom of the nitronium-ion) cannot undergo mesomeric stabilisation by the electrons of vicinal atoms in the same particle[8]. Starting from the fundamental work of DE LA MARE, VERNON, and HUGHES[9], SWAIN and KETLEY[10] succeeded quite recently in finding direct proof of the existence of the chlorinium ion by comparing the rates of chlorination in light and heavy water (H_2O and D_2O, respectively).

We hope that by these three examples we have shown the importance that attaches to the investigation of the equilibria and reactions prior to substitution. Interesting results have also been obtained with other electrophilic substitutions, but these cannot be discussed within the scope of the present survey.

III.–*The Substitution Proper*

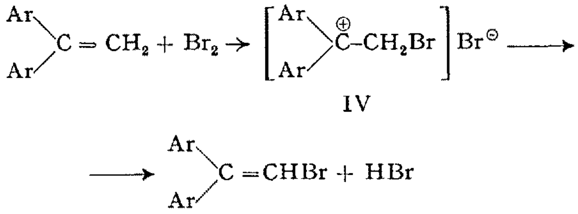
As has been mentioned above, in the last 9 decades the mechanism of substitution has been studied much more extensively than the question of the forms of the electrophilic and nucleophilic components participating in the substitution proper. This may be difficult for us to understand at the present time because irrefutable statements concerning the mechanisms of the reactions can, after all, be made only with knowledge of the nature of the particles participating in the reactions.

We shall not deal here with the historical development of the earlier hypotheses of aromatic substitution. Most of them are based on certain analogies between addition reactions at the double bonds of olefines, on one hand and aromatic substitutions, on the other. The nitration of benzene for example, was represented as follows:



(5)

However, it has never been possible to produce any proof whatever for the existence of the intermediate stage. Therefore a substantial advance was made when PFEIFFER and WIZINGER[11] succeded in 1928 in isolating an intermediate product IV in the bromination of 1:1-di-(*p*-dimethylaminophenyl)-ethylene. Thus, the reaction of this olefine did not lead to the addition of two bromine atoms with saturation of the double bond, but led to the replacement of a hydrogen atom by a bromine atom. Subsequently, a large number of analogous reactions with ethylene derivatives were observed, more especially by WIZINGER and his collaborators[12]. PFEIFFER and WIZINGER proposed the hypothesis that substitution in benzene and other aromatics also takes place according to the same mechanism:



$$\begin{matrix}Ar\\ Ar\end{matrix}\!\!>C=CH_2 + Br_2 \rightarrow \left[\begin{matrix}Ar\\ Ar\end{matrix}\!\!>\overset{\oplus}{C}-CH_2Br\right]Br^{\ominus} \longrightarrow \begin{matrix}Ar\\ Ar\end{matrix}\!\!>C=CHBr + HBr \quad (6)$$

What appeared most important in those days was that a salt-like intermediate IV, i.e. a particle of ionic character, was proved to exist. It was at least unusual for organic chemists at that time to contemplate an ionic reaction, notwithstanding that some investigators had pointed out such possibilities very much earlier. Apart from H. VON EULER who, as has been mentioned, had discussed the nitronium ion a quarter of a century before, it was more especially LAPWORTH[13] who had published as early as 1901 his ideas on benzene substitution, which not only agreed with the mechanism of olefine substitution according to PFEIFFER and WI-

---

[8] See calculations of the equilibrium $Cl_2 \rightleftarrows Cl^{\ominus} + Cl^{\oplus}$ by R. P. BELL and E. GELLES, J. chem. Soc. *1951*, 2734.

[9] P. B. D. DE LA MARE, C. A. VERNON and E. D. HUGHES, Research *3*, 192 (1950). – P. B. D. DE LA MARE, A. D. KETLEY, and C. A. VERNON, J. chem. Soc. *1954*, 1290.

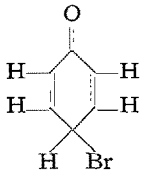
[10] C. G. SWAIN and A. D. KETLEY, J. Amer. chem. Soc. *77*, 3410 (1955).

[11] P. PFEIFFER and R. WIZINGER, Liebigs Ann. Chem. *461*, 132 (1928).

[12] See the summary by R. WIZINGER in Chimia *7*, 273 (1953).

[13] A. LAPWORTH, J. chem. Soc. *1901*, 1265.

ZINGER, but even anticipated the results of the most recent experimental research. LAPWORTH suggested that, when phenol is brominated, molecular bromine (Br_2) dissociates to a small extent into $Br^{\oplus}$ and $Br^{\ominus}$, and that the $Br^{\oplus}$ ion reacts with a phenolate ion arising from the dissociation of phenol accompanied by the formation of the intermediate V (and analogous compounds in the case of substitution in *ortho*-position and at the oxygen).



V

The hypothesis proposed by LAPWORTH completely corresponds, not only with regard to the particles taking part in the substitution proper, but also with regard to the mechanism of the stages, to the process of substitution which has now been proved experimentally.

After this discussion of the historical basis of the theory of aromatic electrophilic substitution, we will now deal with the recent work which has supplied the experimental basis for the two-stage mechanism of this type of reaction. Primarily it has to be shown that in the first stage (*7a*) the electrophilic reagent $X^{\oplus}$[14] additively combines with the aromatic compound, and that only thereafter is the hydrogen atom to be substituted split off as a proton, while the binding electron pair is retained by the aromatic compound (*7b*):

$$X^{\oplus} + \text{Ar–H} \rightarrow X\text{–}\overset{\oplus}{\text{Ar}}\text{–H} \qquad (7a)$$

$$X\text{–}\overset{\oplus}{\text{Ar}}\text{–H} + B \rightarrow \text{Ar–}X + \text{H}B^{\oplus} \qquad (7b)$$

Since, for chemical purposes, free protons have an immeasurably short life, it is supposed that the release of the proton is immediately followed by its additive combination with a base *B*. This process is further discussed below.

It is first necessary to consider whether these two stages cannot take place simultaneously. As the electrophilic reagent approaches the aromatic compound, the proton detaches itself simultaneously and attaches itself to the base. Similar mechanisms are known, for example in nucleophilic aliphatic substitution. This single-stage termolecular, so-called S_E3, mechanism (8)[15], can be excluded for a variety of reasons.

$$X^{\oplus} + \text{Ar–H} + B \rightarrow \underset{\text{Transition state}}{(X^{\oplus} \ldots \text{Ar} \ldots \text{H} \ldots B)} \rightarrow$$

$$\rightarrow \text{Ar–}X + \text{H}B^{\oplus} \qquad (8)$$

Thus, in nitrations the solvent effects preclude such a mechanism[16] and in the case of azo coupling reactions the rate of reaction of the substitution proper is not affected by variation in the concentration of the base B[17], normally. This is not consistent with the S_E3 mechanism, but is in accord with the two-stage bimolecular S_E2 mechanism (*7a*)–(*7b*), if the second stage–the protolysis–proceeds considerably more rapidly than the first stage. In that case, according to the laws of chemical kinetics, the rate of reaction is influenced only by variation in the concentration in the first stage, and not by such variation in the subsequent stage.

The most important and methodically elegant support for the two-stage mechanism has been supplied by MELANDER[18] by his investigation of electrophilic substitution in aromatic compounds which contain in the position where the substitution is performed an isotope of hydrogen, namely radioactive tritium (T) having the atomic weight 3. With a whole series of benzene derivatives he showed that the hydrocarbon containing tritium is nitrated and brominated at the same speed as the corresponding ordinary protium compound. MELANDER'S findings have been confirmed for the nitration of deuterated benzenes[19]. If the elimination of hydrogen is associated with the rate-determining part of the reaction (S_E3 mechanism), or if in the two-stage mechanism the second stage proceeds more slowly than the first, a so-called *kinetic isotope effect* would be expected.

It will be appropriate first to refer briefly to the causes of kinetic isotope effects in general. As soon as an A–B linkage is split up in a rate-determining reaction stage and A or B is replaced by a heavier isotope, the reaction proceeds more slowly. This phenomenon has been observed experimentally with a large number of isotopes, for example, of carbon, of nitrogen and above all of hydrogen. The term "kinetic isotope effect" denotes the ratio of the rate constants of the reactions with the corresponding isotopes in question. The magnitude of the effect can be calculated with a surprisingly high degree of approximation with the aid of the theory of absolute reaction rates (transition state theory)[20]. In fact, it was on this basis that CREMER and

---

[14] X is generally a cation, although it may be a neutral particle or an anion. Examples of neutral particles have been mentioned in the preceding chapter and in the case of an anion, see HCH. ZOLLINGER, Helv. chim. Acta *36*, 1723 (1953) "2:5-disulpho-diazobenzene".

[15] Cf. C. K. INGOLD, *Structure and Mechanism in Organic Chemistry* (G. Bell and Sons Ltd., London 1953), p. 279.

[16] E. D. HUGHES, C. K. INGOLD, and R. I. REED, J. chem. Soc. *1950*, 2400.

[17] HCH. ZOLLINGER and C. WITTWER, Helv. chim. Acta *35*, 1209 (1952). – HCH. ZOLLINGER, Helv. chim. Acta *38*, 1597 (1955).

[18] L. MELANDER, Arkiv Kemi *2*, 213 (1950).

[19] W. M. LAUER and W. E. NOLAND, J. Amer. chem. Soc. *75*, 3689 (1953). – T. G. BONNER, F. BOWYER, and G. WILLIAMS, J. chem. Soc. *1953*, 2650; *1955*, 2358.

[20] S. GLASSTONE, K. J. LAIDLER, and H. EYRING, *The Theory of Rate Processes* (McGraw-Hill Book Comp., New York 1941).

POLANYI[21] predicted the existence of the isotope effects before any experimental results were available.

The magnitude of the isotope effect depends on the element concerned, on the bond, the temperature and other factors. Thus, a reaction in which the bond of a hydrogen atom is broken in the rate-determining stage, proceeds about 5–8 times more slowly when protium (H) is replaced by deuterium (D) and about 20–30 times more slowly when it is replaced by tritium (T). With elements of higher atomic weight the calculated and observed isotope effects are of a considerably smaller order. Thus, for example, in the case of carbon, a ^{13}C-compound reacts only a few per cent more slowly than the compound containing ordinary ^{12}C. The result is similar when ^{14}N is replaced by heavy nitrogen ^{15}N.

From among the various physical causes of the isotope effect we will mention here only the most important, namely the zero-point energy of the bond involved. In the quantum theory this term is used to designate the lowest energy level of the vibrations of a bond, because even at the absolute zero-point it cannot be lowered. Even at room temperature the bulk of all bonds is still in this lowest energy state. The zero-point energy is proportional to the vibration frequency which latter is in turn inversely proportional to the square root of the reduced mass. In the case of a hydrogen atom bound to a relatively heavy organic molecule, a good degree of approximation for its reduced mass can be arrived at by taking the atomic weight 1 and correspondingly 2 can be taken in the case of deuterium. Thus, the zero-point energy of the deuterium bond is smaller by the factor $\sqrt{2}$. This value can be confirmed experimentally from the ratio of the frequencies of the two bonds in the infra-red region of the spectrum.

The theory of absolute reaction rates considers the progress of any reaction from the standpoint of energetics. The system of the starting materials is regarded as a pool of potential energy, the level of which corresponds substantially to the zero-point energies of the molecular bonds taking part. During the reaction the system passes an energy peak before falling to the level of the final products. The state of the reacting particles at this maximum is called the *transition state* or *activated complex*. It is an energetically labile state and can therefore never be chemically isolated or fixed. Nevertheless, the atoms and molecules taking part are in a well-defined arrangement, which in simple cases can be calculated, though not measured. The term "Reaktionsknäuel" used in the German terminology, is thus misleading and should in our opinion be dropped. The difference in energy between the transition state and the energy of the starting system (E_a^H and E_a^D respectively) is the main factor determining the reaction rate. These relationships are represented in Figure 1 in which the energy of the system of a protolysis and of a deuterolysis is plotted against the so-called reaction coordinate[22] serving as a measure of the progress of the reaction. This graph clearly shows that the activating energy in a deuterolysis (E_a^D), owing to the difference between the zero-point energies ($E_0^D < E_0^H$), is greater than that in a protolysis (E_a^H). The isotope effect can be calculated as a first approximation from the relation (9).

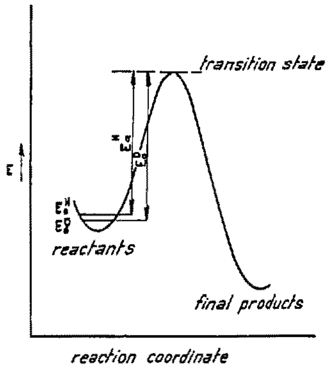


Fig. 1.—The variation in potential energy during the dissociation of a C–H bond and a C–D bond.

$$\frac{k_H}{k_D} \cong e^{\frac{E_0^H - E_0^D}{RT}} \qquad (9)$$

For further factors governing the isotope effect we must refer the reader to more detailed literature[23].

IV.–*Proof of the Kinetic Isotope Effects in Electrophilic Aromatic Substitutions*

The absence of an isotope effect in the nitrations and brominations reported by MELANDER proves by exclusion the correctness of the S_E2 mechanism: The single-stage termolecular reaction (8) can be excluded, so that the two-stage process is the most likely.

By the experimental discovery of kinetic isotope effects in the sulphonation of bromobenzene by BERGLUND-LARSSON and MELANDER[24] and the kinetic mathematical analysis of isotope effects in azo coupling reactions according to ZOLLINGER[25] it has recently been possible to confirm clearly the correctness of the two-stage mechanism. At the present time we are therefore able to exclude in all probability the possi-

---

[21] M. POLANYI and E. CREMER, Z. physik. Chem. [B] *19*, 443 (1932).

[22] See, for example, P. D. BARTLETT in H. GILMAN, *Organic Chemistry*, Vol. III (John Wiley & Sons, New York 1953), p. 8.

[23] J. BIGELEISEN, J. chem. Phys. *17*, 675 (1949); Science *110*, 14 (1949). – K. B. WIBERG, Chem. Reviews *55*, 713 (1955); and more especially L. MELANDER, Arkiv Kemi 2, 213 (1950).

[24] U. BERGLUND-LARSSON and L. MELANDER, Arkiv Kemi *6*, 219 (1953); Abstracts XIVth International Congress of pure and applied Chemistry *1955*, p. 82.

[25] HCH. ZOLLINGER, Helv. chim. Acta *38*, 1597, 1617, 1623 (1955).

bility of the substitution taking place due to a third mechanism not as yet discussed. The question of principle, namely to what extent it is permissible to generalize the knowledge gained from these model reactions will be briefly dealt with at the end.

Berglund and Melander have shown that the sulphonation of bromobenzene proceeds 2·5 times more slowly when the protium atom is replaced by tritium as the reactive atom. The interpretation of this experimentally very interesting discovery was uncertain for two reasons: On one hand, the mechanism of sulphonation including the equilibria prior to the reaction, has been only partially investigated and, on the other, when H is replaced by T, a much greater isotope effect would be expected, as mentioned above.

In view of this abnormally small effect Berglund and Melander put forward the hypothesis that this two-stage mechanism may be a system having a quasi-stationary intermediate (steady state), and this could be conclusively confirmed by the above mentioned discovery and the evaluation of the isotope effect in azo coupling reactions.

Normally, *no* isotope effects can be detected in azo coupling reactions. Thus, for example, 2-methoxy-diazobenzene reacts with 1-naphthol-4-sulphonic acid at the same rate as with 2-D-1-naphthol-4-sulphonic acid. The position is different in the coupling of 4-chloro-diazobenzene with 2-naphthol-6:8-disulphonic acid. In this case, when the reactive protium atom of the naphthalene nucleus is replaced by deuterium, isotope effects can be detected which agree well with theoretical expectations ($k_H/k_D = 6{\cdot}55$). The same reaction is very strongly catalyzed by a variety of bases, for example pyridine. But the reaction rate is not a linear function of the concentration of the base, as would be expected in the S_E3 mechanism. Especially noteworthy, however, is the observation that the magnitude of the isotope effect decreases as the concentration of the base increases.

All these observations agree with an *S_E2 mechanism having a quasi-stationary intermediate state.* This means that in the course of the reaction there is never a *temporary enrichment* of the intermediate $X\text{–}\overset{\oplus}{\text{Ar}}\text{–H}$ of the substitution process (7*a*)–(7*b*), and that (apart from short periods during the initial and final phases of the process) it is always present in a *weak* and *constant* concentration. According to Bodenstein[26] this is possible if the reserve reaction of the first stage is taken into account (10*a*–10*b*).

$$X^{\oplus} + \text{Ar–H} \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} X\text{–}\overset{\oplus}{\text{Ar}}\text{–H} \qquad (10a)$$

$$X\text{–}\overset{\oplus}{\text{Ar}}\text{–H} + B \xrightarrow{k_2} \text{Ar–}X + \text{H}B^{\oplus} \qquad (10b)$$

Pütter[3] was the first to point out, in the special case of the azo coupling reaction, the importance of the reverse reaction in giving a qualitative interpretation of the action of bases in catalyzing coupling reactions.

The reaction system (10*a*)–(10*b*) is expressed by the kinetic equation (11),

$$\frac{d\,[\text{Ar–}X]}{dt} = \frac{k_1 \dfrac{k_2}{k_{-1}}\,[B]}{1 + \dfrac{k_2}{k_{-1}}\,[B]}\,[\text{Ar–H}]\,[\text{X}^{\oplus}] \qquad (11)$$

in which k_2 is the only constant which when the reactive protium atom is replaced by deuterium, is diminished by the above mentioned theoretical factor of 5–8. If the reverse reaction of the first stage could be disregarded (k_{-1} is very small), it would follow from (11) that the reaction rate would be practically unaffected by a change in the values of k_2 and $[B]$, since, when k_{-1} is very small

$$1 + \frac{k_2}{k_{-1}}\,[B] \cong \frac{k_2}{k_{-1}}\,[B]\ .$$

The reaction would be one having no isotope effect and not catalyzed by bases. The opposite is the effect when the reverse reaction of the first stage is very rapid (i.e. k_{-1} is large). In the extreme case the reaction rate is linearly proportional to the concentration of the base and to the constant k_2. The isotope effect is then large.

By varying the three values k_2, k_{-1} and $[B]$ all the possible transitions between the two extreme cases can be obtained consequently, with the *same* mechanism reactions are included in which there is a large, a small or no isotope effect in the total reaction. This has been demonstrated by varying k_2 and $[B]$[27]. It can be proved mathematically that notwithstanding the high degree of dependence of the experimentally determined isotope effect of the total process on the concentration of the base, the isotope effect in the second stage (which is the only relevant one) has in all cases a practically constant value. It corresponds to the theory:

$$k_{2H}/k_{2D} = 6{\cdot}4 \pm 0{\cdot}3\ .$$

The variation in magnitude of the isotope effect connected with variation in the acidity of the diazo-component (which causes a variation of k_2) can also be interpreted in an analogous manner, although measurement or observation alone conveys an impression of processes of a different nature.

Whereas in the above mentioned papers the changes in k_2 and $[B]$ were experimentally investigated and mathematically analyzed, a few unpublished experiments concerning the magnitude of k_{-1} will now be

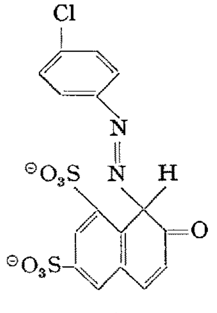
---

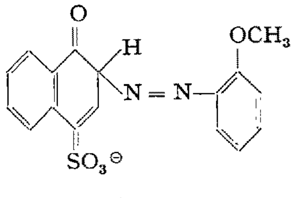
[26] M. Bodenstein, Z. physik. Chem. *85*, 329 (1913).

[27] Hch. Zollinger, Helv. chim. Acta *38*, 1617 (1955).

briefly referred to. What determines the rate of the reverse reaction of the first stage?

It is above all the steric conditions in the intermediate $\overset{\oplus}{\text{H–Ar}}\text{–}X$. Let us consider the typical cases of azo coupling referred to above having a large isotope effect and those having no isotope effect (VI and VII respectively).

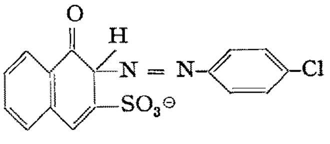

VI $\frac{k_H}{k_D} = 6{\cdot}55$


VII $\frac{k_H}{k_D} = 1{\cdot}0$

It will be seen that only in the compound VI is the diazoresidue introduced sterically hindered by the voluminous sulpho group. It follows therefore that compound VI has the greater "tendency" to eliminate the obstructing diazo substituent; kinetically this has the effect of accelerating the reverse reaction of the first stage (k_{-1}). Some examples are known, although little attention has been given to them, of cases in which steric hindrance not only retards but also accelerates reactions[28]. In the case we are considering this gives rise according to (11) to the isotope effects and the base effects mentioned above.

The correctness of this conclusion can be proved by the isotope effect in the coupling of 4-chlorodiazobenzene with 1:3-naphthol sulphonic acid; under otherwise identical reaction conditions the isotope effect is somewhere between the two above mentioned extremes ($k_H/k_D = 3{\cdot}0$)[29].

With the aid of models of molecules it is easy to show that in the intermediate (VIII) of this reaction

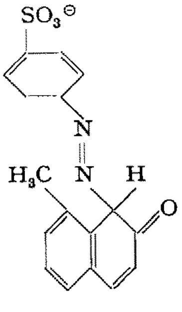

VIII

the sulpho group and the diazo group obstruct each other less than in compound VI. The constant k_{-1} in the coupling with 1:3-naphthol sulphonic acid therefore has a lower value than the corresponding constant in the coupling with the 2-naphthol-6:8-disulphonic acid.

Thus, the relationship between all three factors (k_2, k_{-1}, $[B]$) and the isotope effect in azo coupling reactions is explained experimentally. It will be understood now why BERGLUND-LARSSON and MELANDER obtained such a remarkably small isotope effect when H replaced by T in the sulphonation of bromobenzene. As MELANDER has observed[30], the steric hindrance in the intermediate product of the sulphonation is much lower than that in VI. The k_{-1} value of the sulphonation is therefore smaller and moreover, the intermediate product of the sulphonation is more acidic, so that k_2 is probably greater than in an azo coupling reaction. Consequently the isotope effect of the total reaction is small; at first sight only it does not correspond with expectations.

For mathematical reasons it is unfortunately possible to calculate only the constant k_1 and the quotient k_2/k_{-1}, but not k_2 and k_{-1} separately. However, it has been shown[31] that the rate of protolysis (k_2) is very high. Whether it is lower in reaction which exhibits isotope effects due to the proximity of sulpho groups, than in a "normal reaction", because the negative charge on the $SO_3^{\ominus}$ residue acts electrostatically to retard considerably the release of the proton H cannot at present be stated experimentally.

However, it is certain that this factor is *not the sole cause* of the isotope effect. This follows from the fact that coupling reactions with the use of naphthalene derivatives containing non-ionic substituents that cause steric hindrance in the intermediate product are equally catalyzed by bases, for example, in azo coupling with 8-methyl-2-naphthol in which compound IX is the intermediate product[32]. In this case the methyl group cannot be supposed to have an electrostatic action, but there is no doubt that steric hindrance exists.


IX

The great importance of the steric phenomena in the intermediate product is also apparent from a close study of the effect of bases in catalyzing the protolysis stage (10*b*) and (12). It has been observed[33] that the release of the proton from the intermediate product

---

[28] See C. K. INGOLD, Experientia Supplementum II (XIVth Intern. Congress for Pure and Applied Chemistry 1955), p. 73.

[29] Unpublished measurements by R. ERNST.

[30] Private communication.

[31] HCH. ZOLLINGER, Helv. chim. Acta *38*, 1608 (1955).

[32] Unpublished experiments made by O. STAMM.

[33] HCH. ZOLLINGER, Helv. chim. Acta *38*, 1624 (1955).

proceeds *via* a transition state X, in which the base particle is involved.

$$X\text{-}\overset{\oplus}{\text{Ar}}\text{-H} + B \rightarrow \underset{\text{X}}{(X\text{-Ar}\ldots\text{H}\ldots B)^{\oplus}} \rightarrow X\text{-Ar} + \text{H}B^{\oplus} \qquad (12)$$

The relative rate of the second stage (expressed by the quotient k_2/k_{-1}) can be determined from the nature of the base B. It is found that not only in an acid medium but also in a weakly alkaline medium (pH < 10) practically only H_2O molecules and no hydroxyl ions act as bases. Investigation of the activity of substituted pyridine bases has shown that the protolysis is strongly inhibited by α-substituted pyridines. As far as we are aware this is the first published case of steric hindrance in a proton transfer reaction.

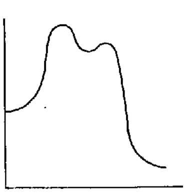
This thorough quantitive analysis of isotope effects in azo coupling reactions has given us the means for finding further aromatic substitutions having measureable isotope effects. Thus, for example, the bromination of 2-naphthol-6:8-disulphonic acid exhibits such an effect[34] the magnitude of which ($k_H/k_D = 2{\cdot}0 \pm 0{\cdot}6$) according to preliminary experiments agrees with theoretical expectations[35]. We have no doubt that, when the requisite steric conditions of the intermediate product are taken into consideration, further substitutions having isotope effects will be found by non empirical methods.

V.–*The Intermediate of the Substitution*

The kinetic investigations described above in connection with derivatives of benzene and naphthalene have enabled conclusive experimental proof to be given of the two-stage substitution mechanism as envisaged by LAPWORTH and expressly postulated by PFEIFFER and WIZINGER. MELANDER's nitrations having no isotope effect, the azo couplings either having or not having an isotope effect of the tritium atom, and the diaryl-ethylene substitutions of PFEIFFER and WIZINGER *via* intermediates capable of being isolated are all based on one and the same mechanism.

The differences in reaction behaviour are not inconsistent with the latter statement, but are due to *differences in energetics*. This becomes clear when we plot the energies against the reaction coordinates in the same manner as is described above for the differences in the behaviour of C–H and C–D bonds. In the Figures 2 to 6 are shown the progress of a substitution without an isotope effect and with a small and with a large isotope effect, of a diaryl-ethylene substitution, and of the single-stage termolecular S_E3 reaction, respectively. It will be noted principally that all these types of reaction, except the hypothetical S_E3 mechanism, exhibit *two* energy maxima (transition states). Whether a substitution is accompanied by an isotope effect depends on the relative energies of these two transition states: An isotope effect occurs only when the second energy maximum is not substantially lower than the first.

Graphs showing the variation in potential energy during the course of a few typical substitutions


2

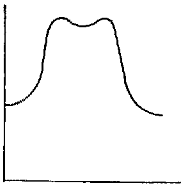
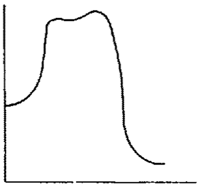

3

Fig. 2.—An aromatic substitution having no isotope effect (normal case): Nitration of benzene derivatives (MELANDER).

Fig. 3.—An aromatic substitution having a small isotope effect: Sulphonation of bromobenzene (BERGLUND-LARSSON and MELANDER).


4

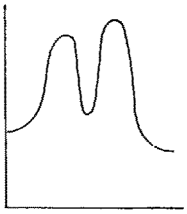

5

Fig. 4.—An aromatic substitution having a large isotope effect: Azo coupling of 2-naphtol-6:8-disulphonic acid (ZOLLINGER).

Fig. 5.—A diaryl-ethylene substitution yielding an intermediate capable of being isolated: Bromination of 1:1-di-(para-dimethylaminophenyl)-ethylene (PFEIFFER and WIZINGER).

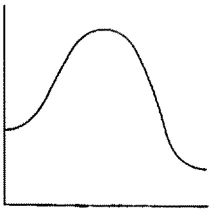


Fig. 6.—A hypothetical, termolecular single-stage substitution mechanism S_E3.

The benzene and the naphthalene reactions differ from those of diaryl-ethylenes in the stability of the intermediates. The addition product of an electrophilic reagent with a diaryl-ethylene is incomparably more stable than the corresponding addition product formed with a compound of the benzene series, because it is a *mesomerically stabilized carbonium ion*. This reduces its tendency to repel a proton; that is to say, its acidity is lower or, putting it conversely the basic

---

[34] Unpublished experiments.

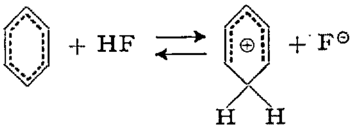
[35] The isotope effect must be smaller than that in the corresponding azo coupling, because the intermediate product of the bromination reaction is more acidic, so that k_2 is greater.

character of diaryl-ethylenes is much more pronounced than that of the aromatics[36].

In the intermediate of a benzene substitution the aromatic state of the starting material, which is particularly favourable from the standpoint of energetics, is no longer present. In the intermediate of the benzene substitution 2 of the original 6 π-electrons, which are responsible for the stability of the aromatic state, have been used up in binding the electrophilic reagent, whereby the system loses stability.

Nothing definite is as yet known concerning the arrangement of the atoms in this intermediate. A conclusive statement can hardly be made on the basis of kinetic considerations alone. The suggestion of HUGHES and INGOLD[37] that the H-atom and the introduced substituent X lie, respectively behind and in front of the plane of the aromatic ring is probably true. In this case a C-atom having 3 planar, aromatic bonds becomes a C-atom having 4 tetrahedrally arranged aliphatic bonds. Expressed in terms of the quantum theory this constitutes a transition from sp^2 into sp^3 hybridized electron orbitals. Recent calculations of MULLIKEN *et al.*[38] suggest that the stability of the intermediate is increased by the π-orbitals being overlapped by the σ-electrons of the C–H and C–X bonds.

Although in the classical chemical sense the intermediate must be looked upon as a very unstable particle (the concept of chemical stability depending on whether or not chemical compound can be isolated or detected by chemical means), it is stable according to exact thermodynamic definition because it corresponds to an energy ***minimum***. It is thus an intermediate proper and not a transition state. It appears important to us to stress this, because in some literature (even in very recent papers) this species is called a transition state.



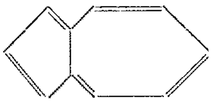
(13)

XI

As has been briefly mentioned above, this can be proved by the *non-linear* dependence of the reaction rate on the concentration of the base[39]. The intermediate of the substitution corresponds to the compound which is formed by the additive combination of an acid with an aromatic compound. This process has been studied very thoroughly by several authors. As an example there may be mentioned the formation of the benzenium ion XI from benzene and hydrofluoric acid (13)[40].

This is a typical equilibrium reaction with components that are stable in the thermodynamic sense of the word.

BROWN and his co-workers who studied the additive combination of hydrohalic acids and aluminium halides and boron halides with aromatic hydrocarbons[41] were able to prove, on the one hand, that such equilibria lead to ionic compounds of the benzenium ion type which, owing to the linking of the proton introduced to *one* carbon atom through a σ-bond, they called a *σ-complex*. On the other hand, according to BROWN, there is a linear relationship between the formation of the σ-complex of different aromatics with hydrofluoboric acid ($BF_3 \cdot HF$) and the halogenation, i.e., an aromatic substitution in these hydrocarbons. This proves that in the substitution an analogous type of intermediate is formed. Consequently, the constants of the equilibrium (13) corresponding to an acid-base relation serve as a measure of the stability of the intermediate of the substitution. The more readily a hydrocarbon additively combines with a proton (i.e., the more basic it is), the more stable is the corresponding benzenium ion and the greater is its reactivity in aromatic substitutions. This explains, for example, the readiness with which azulene (XII) reacts. HEILBRONNER, PLATTNER *et al.*[42] have investigated and discussed very thoroughly the basic character of azulene.



XII

However, under certain conditions BROWN and BRADY have been able to obtain with HCl addition reactions of a different kind which they associate with the binding of the hydrogen chloride by the π-electrons of the aromatic compound. This compound is called a π-complex; it corresponds to the compound formerly postulated as the intermediate stage in an aromatic

---

[36] The basicity of diaryl-ethylene has been measured and theoretically evaluated by V. GOLD, B. W. V. HAWES, and F. L. TYE, J. chem. Soc. *1952*, 2167, 2172, 2181, 2184.

[37] W. A. COWDREY, E. D. HUGHES, C. K. INGOLD, S. MASTERMAN, and A. D. SCOTT, J. chem. Soc. *1937*, 1257. – R. H. GILLESPIE, E. D. HUGHES, C. K. INGOLD, D. J. MILLEN, and R. I. REED, Nature *163*, 599 (1949).

[38] N. MULLER, L. W. PICHETT, and R. S. MULLIKEN, J. Amer. chem. Soc. *76*, 4770 (1954).

[39] It must be pointed out that the occurrence of isotope effects *alone* or catalysis by means of bases *as such* does not preclude the S_E3 mechanism and thus the conception of the intermediate compound as being a transition state.

[40] D. A. McCAULEY and A. P. LIEN, J. Amer. chem. Soc. *73*, 2013 (1951). – M. KILPATRICK and F. E. BUBORSKY, J. Amer. chem. Soc. *75*, 577 (1953). To show that the π-electrons are not localized they were not represented in the formulae of equation (13) as double bonds but, in accordance with more recent practice, by dotted lines.

[41] H. C. BROWN with J. D. BRADY, H. W. PEARSALL, and W. J. WALLACE, J. Amer. chem. Soc. *71*, 3573 (1949); *74*, 192, 3570 (1952); *75*, 6265, 6268 (1953).

[42] E. HEILBRONNER, PL. A. PLATTNER, S. WEBER, M. SIMONETTA, and L. H. CHOPARD-DIT-JEAN, Helv. chim. Acta *35*, 1036, 1049, 2170 (1952).

substitution by DEWAR[43]. Accordingly, Brown's investigations disprove Dewar's substitution mechanism. MULLIKEN[44] advanced theoretical reasons for the existence of σ- and π-complexes which he calls inner and outer complexes, respectively. The question whether π-complexes play a part in other reactions of aromatic chemistry (e.g. in rearrangements) is beyond the scope of the present survey, and we therefore will not enter into this here.

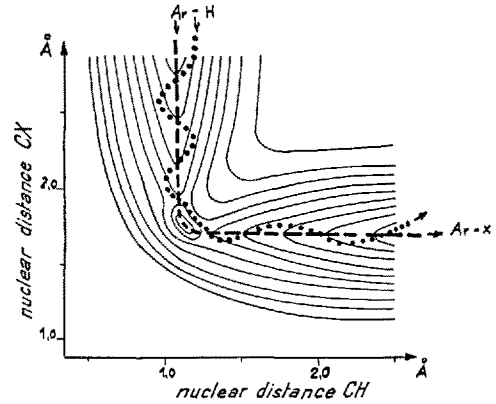


Fig. 7.—The course of an aromatic substitution

$X^{\oplus} + \text{Ar-H} \rightarrow \text{Ar-}X + \text{H}^{\oplus}$

(Graphical representation of the variation of potential energy as a function of the distances between the atoms)*.

* The models chosen for the lengths of the bonds were benzene for Ar-H and chlorobenzene for Ar-X (C-H = 1·08 Å, C-Cl = 1·69 Å; in the intermediate product: C-H = 1·09 Å, C-Cl = 1·76 Å, corresponding to the distances between the atoms in aliphatic bonds).

However a series of very recent papers of GOLD and SATCHEL[45] should be of immediate interest in the future investigation of the mechanism of substitution proper. They suggest that in the exchange of hydrogen isotopes in aromatics (a reaction which, according to INGOLD[46], is also an electrophilic aromatic substitution) the nucleophilic substrate and the proton are in a rapidly established equilibrium with an addition compound supposed to have the character of a π-complex. This non-classical compound is then supposed to undergo rearrangement into a σ-complex in a stage that determines the reaction rate, and from which complex the isotope exchange product is formed. With respect to this interesting investigation of the isotope exchange reaction, we consider that it would be premature to deal with the question how far Gold and Satchell's mechanism for the simplest electrophilic aromatic substitution is applicable in other cases.

However, as the most important outcome of all this work concerning the aromatic intermediate state, it has been established that this intermediate, although short-lived, is still in principle a stable product[47]. We would like to call attention in this connection to a problem which so far as we know, has not yet been investigated, namely that for an individual molecule it is not possible in the extreme case to distinguish between an intermediate product and a transition state. For this purpose we regard the substitution process in a three-dimensional representation of the so-called potential energy surfaces (Fig. 7). On the abscissa and the ordinates are plotted, respectively, the distances between the atoms in the original C-H bond and in the C-X bond to be formed. In this system of coordinates the energies of the reaction system corresponding to the aforesaid distances between the atoms are plotted in a manner similar to contour lines. In the normal case the energy levels, corresponding to the path leading through the "valley" over the two "summits" (= transition states), are traversed in the course of the substitution (– – –). However, a path such as represented by the dotted curve is not precluded in view of the thermal impulses which the system receives from the outside. Under these conditions the energy minimum of the intermediate may be by-passed. On the basis of thermodynamic considerations it follows that this process occurs to an appreciable extent when the depth of the energy trough of the intermediate (in a system containing one molecular proportion of reactive substance) does not exceed RT, which is 0·6 kilocalorie per mol at room temperature. The depth of the trough can be calculated only when the larger of the two constants k_{-1} and k_2 is known. Since, as has been mentioned above, the investigations carried out hitherto enable only the *ratio* $k_{-1}:k_2$ to be determined, it is not yet possible to examine the process of "by-passing the intermediate" experimentally.

VI.–*Conclusions*

Reviewing these recent contributions of reaction kinetics to the mechanism of electrophilic aromatic substitution, it is first pointed out that certain individual reactions of compounds, some of which are rather out of the ordinary, have been investigated very thoroughly. This is followed by discussions of the possibility of generalizing them. However, any wider application of the results obtained in a model case involves the risk of generalizing unjustifiably such approximate solutions as were permissible in the case of the model reaction. From the standpoint of exact and strictly logical scientific thought this method of working is unsatisfactory. This situation, as we have endeav-

---

[43] M. J. S. DEWAR, J. chem. Soc. *1946*, 406, 777; *The Electronic Theory of Organic Chemistry* (Oxford University Press, 1949), p. 163.

[44] R. S. MULLIKEN, J. phys. Chem. *52*, 801, 821 (1952).

[45] V. GOLD and D. P. N. SATCHELL, Nature *176*, 602 (1955); J. chem. Soc. *1955*, 3609, 3619, 3622.

[46] C. K. INGOLD, *Structure and Mechanism in Organic Chemistry* (G. Bell & Sons Ltd., London 1953), p. 303.

[47] The term "metastable" is often used.

oured to show with reference to this reaction mechanism as an example, can be improved by a very thorough working out of the model reactions and the application of a *series* of special investigations to the general case, but it cannot be changed fundamentally.

Whereas only a few decades ago the chemistry of carbon compounds was a self-sufficient almost sovereign science, it is now realized that we cannot make progress without very close cooperation with other fields. The study of the substitution reactions is a good example for this development in organic chemistry. Although, in the present paper, reference has been made primarily to concepts borrowed from physical chemistry and theoretical physics, there are a large number of other branches of natural science which must be resorted to in connection with other problems of organic chemistry. In most cases it is not merely a question of borrowing parts of general fundamentals, but rather of introducing the whole scientific potential of some special subject. Only if that is done, can such a cooperation become a process of solving problems on scientific principles of full value to all the branches of science involved.

I want to thank Dr. HCH. LABHART, Privatdozent at the University of Basle, for his valuable suggestions and discussions. Dr. E. J. KOLLER has been kind enough to read the manuscript of the present paper.

Zusammenfassung

Die Untersuchung der elektrophilen aromatischen Substitution hat in den letzten Jahren in zwei Richtungen, die in dieser Arbeit zusammengefasst und diskutiert werden, Fortschritte gemacht.

Zuerst wird an einigen Beispielen die Bedeutung der Reaktionen und Gleichgewichte, welche der eigentlichen Substitution vorgelagert sind, dargestellt. Die reaktionsfähigste Partikel eines elektrophilen Reagens ist seine sauerste Form; auf der andern Seite reagiert das nukleophile Substrat – der Aromat – mit der grössten Geschwindigkeit, wenn das Reaktionsmedium ein vorgelagertes Gleichgewicht dieser Komponente auf die basische Seite verschiebt.

Im weiteren werden eine Reihe kinetischer Arbeiten, die sich mit dem Mechanismus der eigentlichen Substitution, das heisst dem Austausch des Wasserstoffatoms beschäftigen, diskutiert. Durch die grundlegenden Untersuchungen von MELANDER konnte ein direkter Nachweis dafür erbracht werden, dass die Protonabspaltung normalerweise nicht zur geschwindigkeitsbestimmenden Stufe der Reaktion gehört. MELANDER zeigte, dass ein Benzolderivat, welches an der Reaktionsstelle ein Tritiumatom trägt, gleich rasch nitriert und bromiert wird wie die normale Protiumverbindung. Bei der Sulfierung dagegen wurde ein kleiner kinetischer Isotopeneffekt beobachtet. In jüngster Zeit sind nun grössere Isotopeneffekte bei Azokupplungen gefunden worden. Bei diesen Reaktionen liess sich auch eine allgemeine Basenkatalyse nachweisen. Alle diese Beobachtungen stehen im Einklang mit einem zweistufigen Mechanismus mit stationärem Zwischenprodukt. Ergänzt durch einige unveröffentlichte Messungen lässt sich dies durch eine kinetisch-rechnerische Analyse aller bisher bekannten Untersuchungen zeigen.

Eine Reihe von Arbeiten sprechen dafür, dass dieses Zwischenprodukt die gleiche Struktur hat wie die sogenannten σ-Komplexe, welche H. C. BROWN in Gleichgewichten von Aromaten und gewissen Lewis-Säuren untersucht hat. Das Zwischenprodukt ist eine metastabile Verbindung und kein Übergangszustand. In diesem Zusammenhang wird auf die Grenze der Unterscheidbarkeit von metastabilen Verbindungen und Übergangszuständen hingewiesen.