

# Ein System zur Dokumentation chemischer Reaktionen

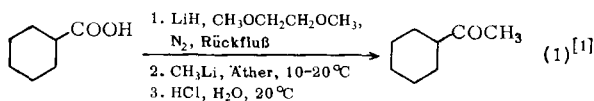
Von Oswald Schier, Wolfgang Nübling, Walter Steidle und Jacques Valls<sup>[\*]</sup>

*Die Reaktionenkartei des Dokumentationsringes der chemisch-pharmazeutischen Industrie arbeitet mit Lochkarten. Ausgangs- und Endprodukt werden mit einem Fragment-Code verschlüsselt, desgleichen werden die bei der Reaktion veränderten Bindungen notiert. Hilfsstoffe und physikalische Daten lassen sich ebenfalls einspeichern. Bei mehrstufigen Reaktionen werden sowohl die einzelnen Schritte als auch mehrere Stufen gleichzeitig registriert.*

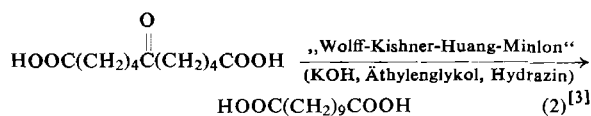
## 1. Einleitung

Die Registrierung chemischer Reaktionen bereitet wesentlich größere Schwierigkeiten als die Registrierung chemischer Stoffe. Eine Substanz ist ein selbständiger Begriff, der meist durch Formel oder Name oder beides definiert ist. Bei einer Reaktion dagegen setzen sich in der Regel mehrere Stoffe unter bestimmten Bedingungen miteinander unter Bildung neuer Substanzen um.

Eine chemische Reaktion umfaßt eine Vielzahl von Sachverhalten: die Ausgangsstoffe, die Reaktionsprodukte, die reagierenden Bindungen, die angewendeten Reagentien und die Reaktionsbedingungen. Eine einfache Reaktion wie (1) weist bereits eine beträchtliche Zahl von Begriffen auf, die bei der Registrierung berücksichtigt werden müssen.



Eine derartige Umsetzung ist nicht nur schwierig zu registrieren, sondern oft auch schon schwierig zu benennen. In vielen Fällen behilft man sich damit, daß man eine Reaktion mit den Namen ihrer Erfinder bezeichnet<sup>[2]</sup> und nicht nach der strukturellen Veränderung, die sie bewirkt [Reaktion (2)].



<sup>[\*]</sup> Dr. O. Schier  
CIBA AG  
CH-4000 Basel 7 (Schweiz)  
Dr. W. Nübling  
E. Merck  
61 Darmstadt 2  
Dr. W. Steidle  
Knoll AG  
67 Ludwigshafen  
Dr. J. Valls  
Roussel-Uclaf S.A.  
Romainville (Frankreich)

[1] T. M. Bare u. H. O. House, Org. Syntheses 49, 81 (1969).

[2] K. G. Wagner: Autoren-Namen als chemische Begriffe. Verlag Chemie, Weinheim 1951; A. R. Surrey: Name Reactions in Organic Chemistry. Academic Press, New York 1954; J. E. Gowan u. T. S. Wheeler: Name Index of Organic Reactions, Longmans, London 1960; H. Krauch u. W. Kunz: Reaktionen der organischen Chemie. Hüthig-Verlag, Heidelberg 1966.

Diese Schwierigkeiten bei der Dokumentation von Reaktionen sind einer der Gründe, warum die Konzeption eines guten Syntheseweges nicht allein an Erfahrung und Intuition des Chemikers hohe Anforderungen stellt. Entscheidend ist oft auch seine Fähigkeit, die geeigneten Teilschritte in der Literatur zu finden.

Bei einem so komplexen Netzwerk von Sachverhalten, wie es eine chemische Reaktion darstellt, versagen die konventionellen Registrierungsmethoden. Nur die modernen Methoden der Literaturdokumentation können zum Ziel führen. Schon vor Jahren hat u.a. Ziegler<sup>[4]</sup> ein Verfahren beschrieben, um chemische Reaktionen auf Flächenlochkarten zu erfassen.

Nach der Lösung von Problemen auf dem Gebiet der chemisch-medizinischen Dokumentation<sup>[5]</sup> stellte sich der Dokumentationsring der chemisch-pharmazeutischen Industrie die Aufgabe, eine Reaktionenkartei zu entwickeln. Wir berichten im folgenden über das Ergebnis dieser Arbeiten. Nach umfangreichen Erprobungen soll das System mit einem Grundstock von 20000 Referaten allgemein angeboten werden.

## 2. Die Dokumentationsmethode bei der Reaktionenkartei

Die Wahl der Dokumentationsmethode wird von der Menge des zu registrierenden Materials und der Zahl der neben- und miteinander vorkommenden Begriffe bestimmt. Bei einer Reaktionenkartei liegt das Minimum bei rund 20000 aufzunehmenden Einzelreaktionen: So umfassen allein die bisher erschienenen 24 Bände von W. Theilheimers „Synthetic Methods of Organic Chemistry“<sup>[6]</sup> mehr als 22000 Reaktionsbeispiele neben zahlreichen ergänzenden Literaturzitaten. Eine Reaktionenkartei muß mit einem jährlichen Zuwachs von mindestens 2000 Zitaten rechnen.

Eine Kartei dieses Umfangs läßt sich nur schwer als Standkartei führen und ist auch mit Sicht-, Rand- und

[3] L. J. Durham, D. J. McLeod u. J. Cason, Org. Syntheses, Coll. Vol. IV, 510 (1963).

[4] H. J. Ziegler, J. chem. Documentation 6, 81 (1966).

[5] a) Vgl. z.B. M. Scheublein, W. Steidle, W. Germann, W. Nübling u. R. Wilhelmi, Naturwissenschaften 55, 362 (1968); b) W. Nübling u. W. Steidle, Angew. Chem. 82, 618 (1970); Angew. Chem. internat. Edit. 9, 8 (1970).

[6] W. Theilheimer: Synthetic Methods of Organic Chemistry – Synthetische Methoden der Organischen Chemie. S. Karger Basel, New York, Band 1–24 (1946–1970).

Flächenlochkarten schlecht zu bewältigen. Eine Standardkartei käme mit der Vielzahl der pro Reaktion zu beachtenden Sachverhalte nicht mehr zurecht. Sichtlochkarten können pro Karte nur einen Begriff und eine begrenzte Zahl von Arbeiten verzeichnen. Bei Rand- und Flächenlochkarten sind die Möglichkeiten auch von der maschinellen Seite her begrenzt, da der Sortierapparat nur 250–500 Karten faßt und daher bei einem umfangreichen Material untragbar oft gefüllt werden muß.

So stellten wir uns die Frage, ob der Code für eine Reaktionskartei auch für Maschinenlochkarten oder nur für Computer geeignet sein sollte. Die gebräuchliche Lochkarte enthält 960 Lochpositionen, eine Zahl, die für eine Reaktionskartei ausreicht. Mit dem zu erwartenden Kartenanfall kommen andere Lochkartensortiermaschinen gut zurecht: So sortiert eine Maschine des Typs IBM-108 stündlich rund 60000 Lochkarten nach bis zu 12 Fragestellungen.

Für die Reaktionskartei des Dokumentationsringes wurde ein Code entwickelt, der sowohl für Maschinenlochkarten als auch für Computer geeignet ist. Mit der Verwendung von Lochkarten haben wir bewußt auf einige Möglichkeiten verzichtet, die ein ausschließlich auf Computer ausgelegter Code manchmal bietet. Ausschlaggebend war jedoch der Gedanke, daß eine Reaktionskartei einem möglichst breiten Kreis von Interessenten zugute kommen soll. Kartenbenutzer, die über Computer verfügen, können selbstverständlich den Inhalt der Lochkarten auf Band, Platte oder Trommel übertragen.

### 3. Die registrierten Sachverhalte bei der Reaktionskartei

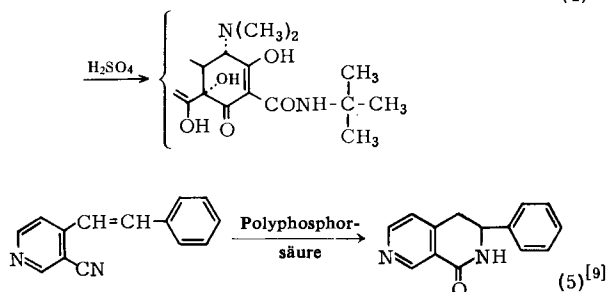
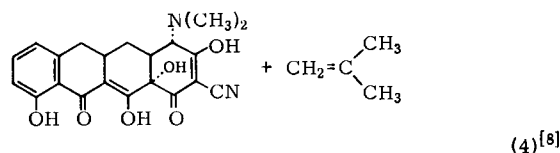
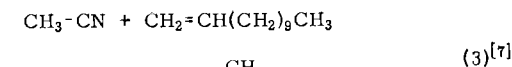
Am Beginn jeder Dokumentation steht die Entscheidung über die zu registrierenden Sachverhalte. Selbstverständlich sollte alles Wichtige aufgenommen und alles Unwichtige ferngehalten werden. Auf diese allgemeinen Überlegungen folgt die Entscheidung über den anzuwendenden Code.

#### 3.1. Ausgangsstoff und Reaktionsprodukt

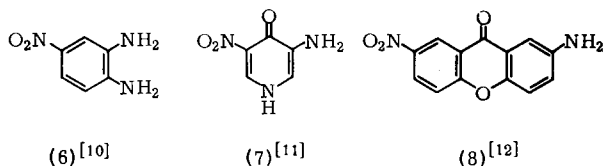
##### 3.1.1. Allgemeine Überlegungen

Wie wichtig ist die ausführliche Registrierung von Ausgangsstoff und Reaktionsprodukt in einer Reaktionskartei? Dazu ist festzustellen, daß der Chemiker bei der Anwendung einer Reaktion oft nach Substanzen sucht, welche mit dem von ihm bearbeiteten Stoff Strukturmerkmale gemeinsam haben. Die in der Literatur beschriebene Synthese z.B. eines substituierten Amids aus einem Nitril kann auch für die Darstellung eines Amids oder Lactams einer ganz anderen Substanzklasse wertvoll sein [Reaktion (3)–(5)].

In den Reaktionen (3) bis (5) wird nur nach den reagierenden und hergestellten Funktionen gefragt. Oft werden aber auch Teile des Moleküls verlangt, die



nicht reagieren. Bei der Suche z. B. nach der selektiven Reduktion von Dinitro-Verbindungen zu Nitroaminen wird eine zweite, nicht reagierende Nitrogruppe vorausgesetzt. Die aus Dinitro-Verbindungen mit derselben Methode – Einwirkung von Ammoniumsulfid – erhaltenen Nitroamine [Beispiele (6)–(8)] können dabei ganz verschiedenen Substanzklassen angehören.



Die früher erwähnte allgemeine Frage „Amide aus Nitrilen“ könnte ebenfalls spezifischer ein aliphatisches oder heterocyclisches Nitril, wie in den Beispielen (3) und (5), oder ein Nitril der Tetracyclin-Reihe, wie in (4), verlangen.

Diese Fragen betreffen reagierende, aber auch von der Reaktion nicht betroffene Teile des Moleküls. Es ist also notwendig, in einer Reaktionskartei die Strukturen der Substanzen möglichst vollständig zu erfassen. Dabei sind aber nicht die Gesamtstrukturen, sondern jene allgemeinen Strukturen und Teilstrukturen wesentlich, welche Anfragen der geschilderten Art ermöglichen.

##### 3.1.2. Die Verschlüsselung

Für die Verschlüsselung von Ausgangsstoff und Reaktionsprodukt wurde der 1958–1960 vom Dokumentationsring entwickelte Code herangezogen (vgl. <sup>[5]</sup>).

[7] T. Clarke, J. Devine u. D. W. Dicker, J. Amer. Oil Chemists' Soc. 41, 78 (1964).

[8] C. R. Stephens, J. J. Beereboom, H. H. Rennhard, P. N. Gordon, K. Murai, R. K. Blackwood u. M. Schach von Wittenau, J. Amer. chem. Soc. 85, 2643 (1963).

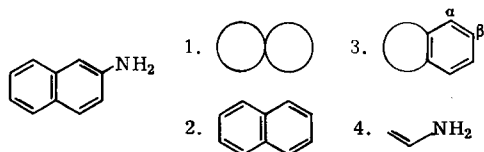
[9] J. M. Bobbitt u. R. E. Doolittle, J. org. Chemistry 29, 2298 (1964).

[10] K. P. Griffin u. W. D. Peterson, Org. Syntheses, Coll. Vol. III, 242 (1955).

[11] W. H. Crowe, J. chem. Soc. (London) 127, 2028 (1925).

[12] S. Dhar, J. chem. Soc. (London) 109, 744 (1916).

In diesem Code wird die zu verschlüsselnde Substanz in Teilstrukturen zerlegt. Man erhält so gerade jene Bausteine, die für eine Reaktionskartei wichtig sind. So enthält die Verschlüsselung von  $\beta$ -Naphthylamin die Angaben 1–4:



1. ein System von zwei miteinander kondensierten Ringen,
2. zwei Benzol-Ringe, die miteinander kondensiert sind,
3. eine  $\beta$ -Substitution an einem kondensierten Benzol-Ring,
4. eine primäre Amino-Gruppe an einer Doppelbindung.

Diesen Angaben entsprechen sechs Lochungen auf einer Lochkarte. Andere, beliebig substituierte Naphthylamine erhalten dieselben Lochungen (mit Ausnahme der  $\beta$ -Substitutionsangabe) und weitere Lochungen für die hinzugekommenen zusätzlichen Begriffe.

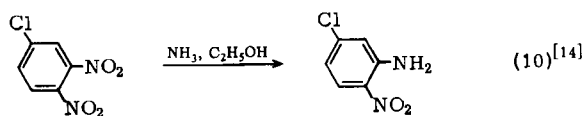
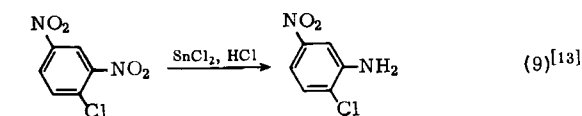
Verschlüsselung und Fragestellung sind einfach. Selbst bei dieser Zerlegung in Teilstrukturen arbeitet der Code nämlich mit jedem Chemiker geläufigen Bausteinen und verwendet weder Code-Nummern noch Buchstabenkombinationen.

Der Code benötigt für Ausgangsstoff und Reaktionsprodukt je 27 Spalten einer Maschinenlochkarte (je 324 Lochpositionen).

### 3.2. Die veränderten Bindungen

#### 3.2.1. Allgemeine Überlegungen

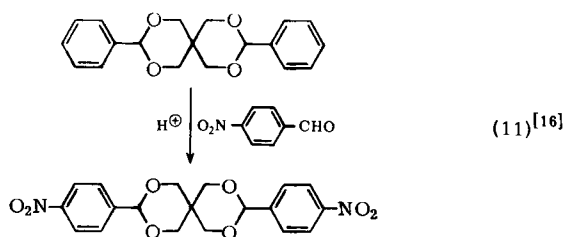
Die Registrierung von Ausgangsstoff und Reaktionsprodukt reicht für die Erfassung einer Reaktion im allgemeinen nicht aus. Selbst bei einfachen Umsetzungen kann man – nur aus den Strukturen und ohne Angabe der Reagentien – die Art der Reaktion oft kaum erkennen. So ist bei Reaktion (9) die Nitrogruppe zur Aminogruppe reduziert, bei (10) aber durch die Aminogruppe ersetzt worden, zwei völlig verschiedene Reaktionen also.



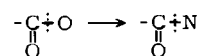
[13] A. Claus u. A. Stiebel, Ber. dtsch. chem. Ges. 20, 1379 (1887).

[14] A. Laubenheimer, Ber. dtsch. chem. Ges. 9, 1826 (1876).

Bei der Umacetalisierung (11) könnte ein Strukturvergleich zur Annahme führen, es handle sich um eine Nitrierung [\*].



Die Angabe der veränderten Bindungen ist aber auch aus anderen Gründen notwendig: sie erlaubt z.B. eine Klassifizierung von Reaktionen. Umsetzungen wie Ester  $\rightarrow$  Amid, Ester  $\rightarrow$  Lactam, Ester  $\rightarrow$  Hydrazid, Ester  $\rightarrow$  Acylharnstoff etc. werden durch die Formulierung



zusammengefaßt. Die Angabe der Bindungsänderung ist dabei auch unabhängig davon, ob sie in einer acyclischen Struktur oder in einem Ring erfolgte.

#### 3.2.2. Die Verschlüsselung

Die im Ausgangsstoff gelösten und im Reaktionsprodukt gebildeten Bindungen werden in einfacher Weise erfaßt. So tritt die Reaktion  $\text{H}-\text{N} \rightarrow \text{C}-\text{N}$  (z.B. die Alkylierung eines Amins) im Code mit den beiden Lochpositionen „H–N gelöst“ und „C–N gebildet“ auf.

In vielen Fällen ist nicht nur die veränderte Bindung, sondern auch ihre Umgebung interessant. Deshalb wird im Code z.B. bei der Spaltung einer C–H-Bindung unterschieden, ob es sich um C–H in den Strukturen  $\text{>C}=\text{C}-\text{H}$  (z.B. an Benzol),  $\text{>C}=\text{C}-\text{C}-\text{H}$  (z.B. Allyl),  $-\text{C}(=\text{Y})-\text{C}-\text{H}$  (z.B.  $\alpha$ -Position bei Ketonen, Säuren etc.) oder um eine andere C–H-Bindung handelt.

Zur Ergänzung der Bindungsangaben registrieren weitere Positionen des Codes die Bildung, Spaltung, Verengung oder Erweiterung von Ringen, ferner Reaktionen unter Umlagerung, die Überführung eines Racemats in optisch aktive Verbindungen, stereospezifische und asymmetrische Synthesen und vieles andere mehr.

Für die genannten Angaben sind in der Lochkarte für die Reaktionskartei des Dokumentationsringes acht Spalten (96 Lochpositionen) vorgesehen.

[\*] Es sind Versuche beschrieben worden [15], chemische Reaktionen durch Computer-Analyse der strukturellen Veränderungen zu erfassen. Eigene Untersuchungen zeigen, daß der Computer für derartige Analysen noch eine Reihe anderer Daten benötigt, wenn er nicht zu falschen Analysen gelangen soll; vgl. die Reaktionen (10) und (11).

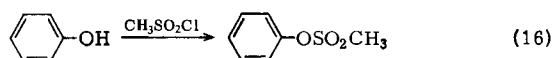
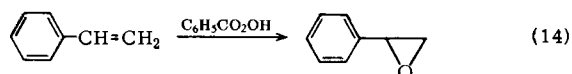
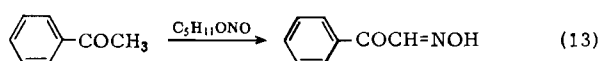
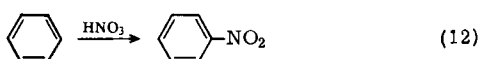
[15] J. E. Armitage, J. E. Crowe, P. N. Evans, M. F. Lynch u. J. A. McGuirk, J. chem. Documentation 7, 209 (1967).

[16] E. Bograchov, J. Amer. chem. Soc. 72, 2268 (1950).

### 3.3. Die Hilfsstoffe

#### 3.3.1. Allgemeine Überlegungen

An einer organisch-chemischen Reaktion sind in der Regel mehrere Substanzen beteiligt, die man mit „Ausgangsstoff“, „Reagens“, „Lösungsmittel“ und anderen Namen bezeichnet. Der Unterschied zwischen „Ausgangsstoff“ und „Reagens“ spiegelt in vielen Fällen nur die Zugänglichkeit und Geläufigkeit einer Substanz im Vergleich mit einer anderen wider. So ist in den Reaktionen (12) bis (16) das Reagens jener Stoff, der über dem Reaktionspfeil steht. In den beiden letzten Beispielen ist aber derselbe Stoff einmal „Ausgangsstoff“ und einmal „Reagens“.



Um Fehler zu vermeiden, sollte eine Dokumentation ihre Begriffe in der dem Chemiker geläufigen Art definieren. Man kommt am ehesten zum Ziel, wenn man als „Hilfsstoffe“ (Reagentien, Lösungsmittel etc.) die folgenden Substanzen bezeichnet:

anorganische Stoffe [vgl. (12), (15)],

organische Stoffe, welche keine C-Atome an das betrachtete Reaktionsprodukt liefern [vgl. (13), (14)], etwa 30 weitere organische Stoffe wie Methyljodid, Methanol, Acetanhydrid, Mesylchlorid, die dann als Hilfsstoff gelten, wenn sie mit einem organischen Stoff reagieren [vgl. (16)].

Die Trennung in Ausgangsstoffe und Reagentien hat von der Reaktion her ihre Berechtigung. Es wäre sinnlos, in einer organisch-chemischen Reaktionenkartei Reagentien wie Salzsäure, Salpetersäure etc. immer wieder als Ausgangsstoff mit gelösten Bindungen etc. zu registrieren.

#### 3.3.2. Die Verschlüsselung

Die Hilfsstoffe werden in der Reaktionenkartei des Dokumentationsringes in zehn Spalten der Lochkarte (120 Lochpositionen) erfaßt. Um allgemeine Anfragen zu gestatten, sind darin Oberbegriffe wie Oxidations-, Reduktions- und Kondensationsmittel, Katalysator, Inhibitor, biologisches Hilfsmittel und andere enthalten. Der Code erlaubt ferner Angaben darüber, ob die Reaktion im sauren, basischen und/oder gepufferten Milieu abläuft.

Die Verschlüsselung der Hilfsstoffe geschieht auf einfache Art und Weise. Häufig vorkommende Hilfs-

stoffe wie Alkohol, Aceton, Pyridin,  $\text{LiAlH}_4$ ,  $\text{HCl}$ , Eisen werden direkt angegeben, andere durch Kombination von Teilstrukturen, so Bleiacetat als Metallacetat + Blei.

### 3.4. Physikalische Daten, sonstige Angaben

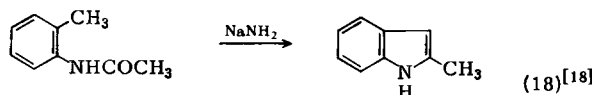
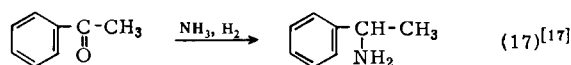
Zusätzlich zu den reagierenden und im Reaktionsmedium anwesenden Stoffen kommt den physikalischen Daten der Reaktion Interesse zu, da sie den Verlauf einer Umsetzung entscheidend beeinflussen können. Dazu gehören Druck, Temperatur, Einwirkung von Licht u.a. Die genannten Daten werden in einer Spalte der Lochkarte in expliziter Form angegeben.

Der Code gestattet ferner die Kennzeichnung von Übersichtsarbeiten sowie von Publikationen, die sich ausführlich mit dem Mechanismus einer Reaktion auseinandersetzen.

### 3.5. Die einstufige Reaktion

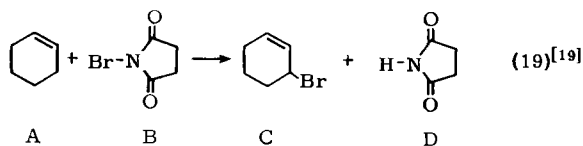
#### 3.5.1. Allgemeine Überlegungen

Bei der Konzeption einer Synthese wird im allgemeinen ein Ausgangsstoff betrachtet, der zu einem Reaktionsprodukt führen soll.



Dies entspricht der früher besprochenen Einteilung Ausgangsstoff / Hilfsstoff / Reaktionsprodukt.

Auch bei einer Reaktion von zwei organischen Stoffen zu zwei Reaktionsprodukten ist von den vier möglichen Kombinationen oft nur eine von Interesse. So ist in Reaktion (19) nur die Umsetzung des Cyclohexens zum präparativ isolierten 3-Bromcyclohexen ( $\text{A} \rightarrow \text{C}$ )



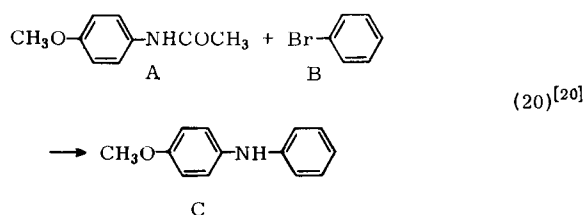
für die Registrierung interessant. Das Reagens Bromsuccinimid muß zwar erfaßt werden, seine Reaktionen zu Succinimid ( $\text{B} \rightarrow \text{D}$ ) oder gar zu 3-Bromcyclohexen ( $\text{B} \rightarrow \text{C}$ ) sind aber kaum interessant. Die Umsetzung Cyclohexen  $\rightarrow$  Succinimid ( $\text{A} \rightarrow \text{D}$ ), die vierte mögliche Kombination, ist gänzlich uninteressant.

[17] J. C. Robinson jr. u. H. R. Snyder, Org. Syntheses, Coll. Vol. III, 717 (1955).

[18] C. F. H. Allen u. J. Van Allan, Org. Syntheses, Coll. Vol. III, 597 (1955).

[19] K. Ziegler, A. Späth, E. Schaaf, W. Schumann u. E. Winkelmann, Liebigs Ann. Chem. 551, 80 (1942)

Anders zu beurteilen sind dagegen Reaktionen vom Typ (20). Hier sind die Umsetzungen  $A \rightarrow C$  und  $B \rightarrow C$  von Interesse und müssen registriert werden.



Bei (19) ist A, bei (20) sind A und B „Ausgangsstoff“; das Bromsuccinimid in (19) ist dagegen „Hilfsstoff“ im Sinne der früher gegebenen Definition. In beiden Reaktionen sind nur die präparativ isolierten Stoffe C als Reaktionsprodukte interessant. Damit ist der Rahmen für die zu registrierenden Reaktionen abgesteckt: Jede Umsetzung eines „Ausgangsstoffes“ zu einem präparativ isolierten Reaktionsprodukt sollte erfaßt werden.

### 3.5.2. Die Verschlüsselung

Auf jeder Lochkarte wird nur die Reaktion *eines* Ausgangsstoffes (verschlüsselt in Spalte 1–27 der Lochkarte) zu *einem* Reaktionsprodukt (Spalte 28–54) angegeben. Sind mehrere derartige Kombinationen möglich, wie z.B. in Reaktion (20), so werden zwei oder mehrere Lochkarten angelegt, die untereinander durch dieselbe Referat-Nummer verbunden und dadurch auch gemeinsam abfragbar sind.

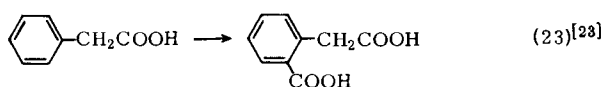
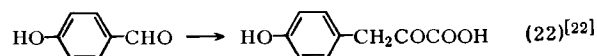
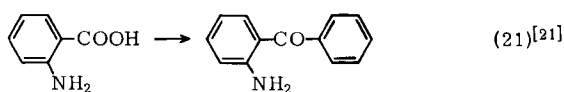
Die veränderten Bindungen und einige ergänzende Sachverhalte werden in den Spalten 55–62 gelocht. Nach der für Zwecke des einzelnen Benutzers freigehaltenen Spalte 63 werden in den Spalten 64–73 die Reagentien und Lösungsmittel (Hilfsstoffe) angegeben. Die physikalischen Bedingungen sind in Spalte 74 eingetragen.

Die Spalten 75–80 enthalten die Referat-Nummer eingelocht, welche Lochkarte und Referat verbindet.

## 3.6. Die mehrstufige Reaktion

### 3.6.1. Allgemeine Überlegungen

Die anschließend genannten Reaktionen (21) bis (23) sind Synthesen, die mit guten Ausbeuten in zwei oder mehreren Stufen durchgeführt werden. Sie sind als



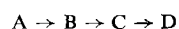
[20] F. E. King, T. J. King u. J. H. M. Muir, J. chem. Soc. (London) 1946, 5.

[21] H. J. Scheifele jr. u. D. F. DeTar, Org. Syntheses, Coll. Vol. IV, 34 (1963).

solche hochinteressant. Trotzdem sind sie, nach üblichen Methoden registriert, meist nicht in der angegebenen Form zu finden, da häufig nicht die Gesamtreaktion, sondern jede Stufe für sich verzeichnet wird. Eine Reaktionenkartei darf nicht in diesen Fehler verfallen.

### 3.6.2. Die Verschlüsselung

In der Reaktionenkartei des Dokumentationsringes werden bei mehrstufigen Reaktionen nicht nur die einzelnen Stufen registriert, sondern es werden zusätzlich zwei und mehrere Stufen zusammengefaßt. Die Reaktionssequenz



erscheint in Form mehrerer Lochkarten, welche die Teilreaktionen  $A \rightarrow B$ ,  $B \rightarrow C$ ,  $C \rightarrow D$ , aber auch  $A \rightarrow C$ ,  $A \rightarrow D$  und  $B \rightarrow D$  enthalten. Auf diese Art werden Reaktionen registriert und recherchierbar, die einstufig völlig ausgeschlossen wären. Wieviele Überlegungen durch eine derartige Erfassung dem synthetisch arbeitenden Chemiker erspart werden können, läßt sich leicht ermessen.

## 4. Das Arbeiten mit der Reaktionenkartei

### 4.1. Die Anfragemöglichkeiten

Die verschlüsselten Begriffe der willkürlich (aus W. Theilheimers „Synthetic Methods of Organic Chemistry“ Band 23) ausgewählten Reaktion (24) sind in Schema 1 wiedergegeben. Die genannten 24 Begriffe sind in Form von 24 Lochungen auf der Lochkarte vorhanden. So findet sich die unter 7 erwähnte Nitro-Gruppe (und allgemein die Nitro-Gruppen aller Ausgangsstoffe) in Spalte 17, Zeile 4 der Lochkarte.

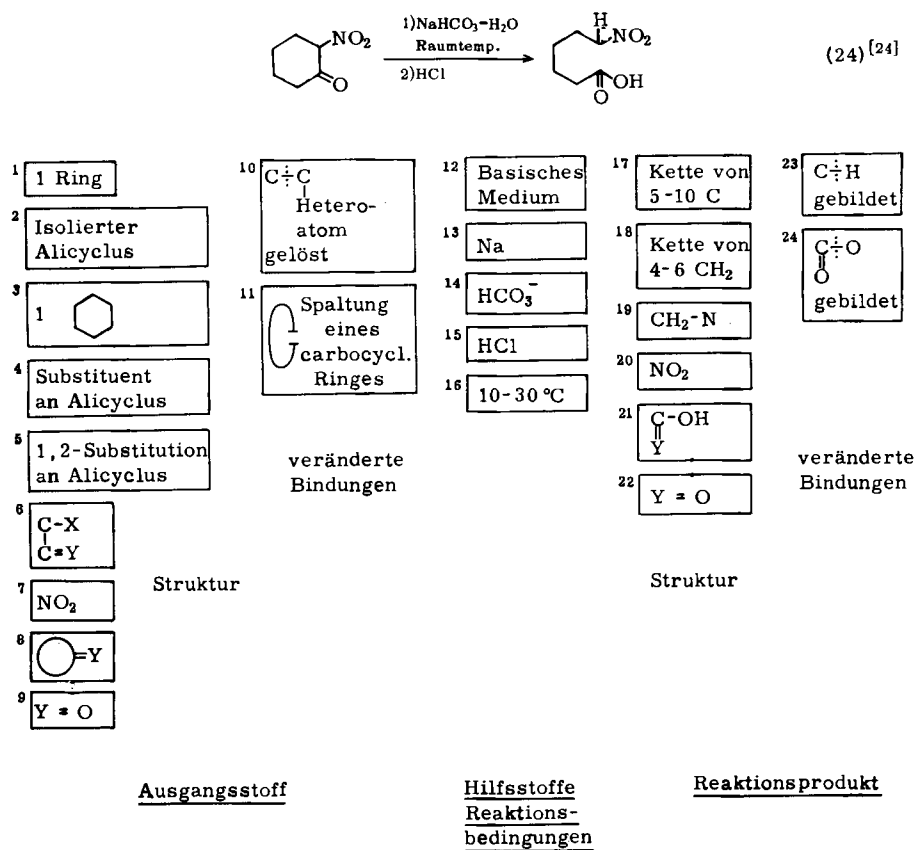
Bei der Fragestellung sind beliebige Kombinationen zwischen den Lochungen möglich („mehrdimensionale Anfrage“). Die Fragemöglichkeiten reichen von der spezifischen zur ganz allgemeinen Frage. So treffen auf die spezifische Frage nach Reaktionen von 2-Nitro-cyclohexanon die mit 1 bis 9 bezeichneten Lochungen zu. Bei der spezifischen Frage nach der Synthese von 6-Nitro-capronsäure sind es die mit 17 bis 22 bezeichneten Lochungen.

Die Verschlüsselung ist meist nicht ganz eindeutig. Mit denselben Lochpositionen 17 bis 22 werden im Code auch die nahen Verwandten 5-Nitro-n-valeriansäure und 7-Nitro-heptansäure verschlüsselt.

Neben solchen spezifischen Fragen ist eine Vielzahl allgemeiner Anfragen möglich. Die Kombination von Struktur und Reagentien, z.B. Veränderung von Oxo-Verbindungen durch Basen [Lochungen 8 + 9 + 12 bei der Reaktion (24)] gibt Auskunft, welche Reaktionen an einer Struktur eintreten können, wenn man sie mit einem bestimmten Reagens behandelt. Dies ist z.B. bei der Vorhersage von Nebenreaktionen von Bedeutung.

[22] G. Billek, Org. Syntheses 43, 49 (1963).

[23] P. A. S. Smith u. R. O. Kan, Org. Syntheses 44, 62, 91 (1964).



Schema 1.

Natürlich sind auch allgemeinere Fragen nach der Stabilität von Stoffen möglich. Die Verschlüsselung der Reaktion (24) antwortet z. B. auch auf die folgenden Anfragen: Spaltung von alicyclischen Ringen (Lochungen 3 + 11), Stabilität von Nitro-Gruppen an aktivem Methylen (6 + 7), Reaktionen von Cyclohexanon (2 + 3 + 8 + 9), Einfluß von  $\alpha$ -Substituenten auf die Stabilität von Oxo-Verbindungen (6 + 8 + 9).

Eine Anfrage wie nach Reaktionen von  $\alpha$ -Nitroketonen bei Raumtemperatur [Lochungen 6-9 und 16 bei der Reaktion (24)], die mit dem Begriff „Raumtemperatur“ verbunden ist, kann man z. B. bei Synthesen mit empfindlichen Stoffen stellen, bei denen man von vornherein erhöhte Temperaturen ausschließen muß.

Auch bei den allgemeinen Fragen, die das Reaktionsprodukt betreffen, gibt es verschiedene Möglichkeiten. Bei der Lochkarte für die Reaktion (24) antworten die genannten Lochungen auf die folgenden Anfragen: Synthese von  $\omega$ -Nitrocarbonsäuren aus anderen Nitroverbindungen (Lochungen 19-22 und 7), Synthese von Nitrocarbonsäuren (Lochungen 20-22), Synthese von Carbonsäuren (Lochungen 21 und 22).

#### 4.2. Die Qualität der Antworten

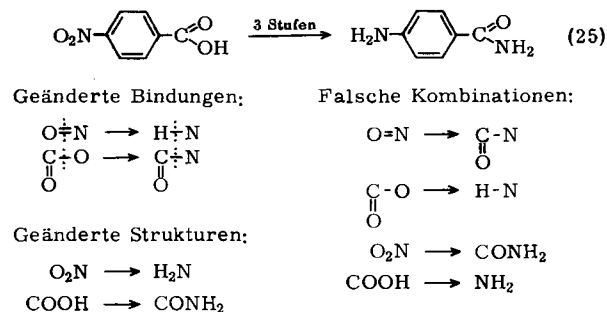
Der beschriebene Code arbeitet bewußt mit möglichst einfachen Begriffen, die jedem Chemiker geläufig sind. Die Verschlüsselungskosten sind daher relativ niedrig

[24] A. S. Matlack u. D. S. Breslow, J. org. Chemistry 32, 1959 (1967).

und liegen bei ca. 8 DM pro Referat. Auch die Fragestellung ist recht einfach. Muß dies mit einem erhöhten Anfall falscher Antworten erkaufte werden?

Ein Code, der mit Teilstrukturen arbeitet, gibt im allgemeinen keine für jede einzelne Substanz eindeutige und einmalige Verschlüsselung. Am Beispiel des Reaktionsproduktes der Reaktion (24) wurde erwähnt, daß die Verschlüsselung auch für zwei nahe verwandte Nitro-carbonsäuren zutrifft. Dies ist aber kein Nachteil, da bei einer Reaktionenkartei eine so detaillierte Erfassung meist gar nicht gewünscht wird.

Die parallelen Verschlüsselungen von Substanzen und Bindungen ergänzen sich gegenseitig und erlauben eine ausgezeichnete Selektion. Selbst bei Parallelreaktionen mehrerer Gruppen desselben Moleküls wie in (25), bei



welcher durch falsche Kombination von Lochpositionen Irrtümer möglich wären, schließt die gleichzeitige Anfrage an Struktur und Bindungen falsche Antworten im allgemeinen aus. Die Lochkarte der

Reaktion (25) antwortet daher nicht auf die genannten falschen Kombinationen.

Durch eine wesentliche Verfeinerung des Codes könnte man den sehr kleinen Anfall nicht zutreffender Karten noch weiter vermindern. Dadurch würde jedoch die Verschlüsselung langwieriger und der Code kompliziert; kaum ein Benutzer wäre dann noch in der Lage, durch eigene Verschlüsselungen die Kartei nach seinem Bedarf zu ergänzen. Überdies würde die getrennte Verschlüsselung unabhängiger Reaktionen im Fall von Parallelreaktionen wie (25) wieder die Fragen nach solchen Reaktionen [z.B. Aminocarbonsäureamide aus Nitrocarbonsäuren bei (25)] unmöglich machen.

*Die beschriebene Reaktionenkartei ist eine Gemeinschaftsarbeit von vierzehn Firmen des Dokumentationsringes der chemisch-pharmazeutischen Industrie. Unser Dank für ihre Mitarbeit richtet sich besonders an die folgenden Damen und Herren: O. Dold (Boehringer Mannheim), G. Hallermann (CIBA A.G., Basel/Schweiz), R. Payer (Chemie Grünenthal, Stolberg), H. Lebrecht und M. Scheublein (Knoll A.G., Ludwigs-hafen), K.-H. Bork (Merck, Darmstadt), U. de la Vaissiere (Roussel-Uclaf S.A., Romainville/Frankreich), H. Seidel (Schering A.G., Berlin), I. Schuler (Thomae GmbH, Biberach) und H. Lehwald (Troponwerke, Köln-Mülheim).*

Eingegangen am 13. Mai 1970 [A 772]

## CCBF – Ein System zur Computerbearbeitung chemischer und biologischer Forschungsergebnisse<sup>[\*\*]</sup>

Von Gerhard Ohnacker und Werner Kalbfleisch<sup>[\*]</sup>

*Es wird eine firmeninterne, computerorientierte Dokumentation für Ergebnisse aus der Arzneimittelforschung beschrieben. Das System speichert die zahlenmäßigen Resultate aus standardisierten biologischen Testen und verarbeitet chemische Formeln mit topologischen Methoden. Es liefert maschinell gedruckte Karteien und Listen mit chemischen und/oder biologischen Sachverhalten aus dem gespeicherten Material. Seine Programme erlauben Recherchen nach beliebigen Substrukturen und nach pharmakophoren Gruppen; damit kann die Suche nach Beziehungen zwischen chemischer Konstitution und biologischer Aktivität wirksam unterstützt werden.*

### 1. Einleitung

Die Entwicklung einer neuen, biologisch aktiven Substanz bis zu ihrer Einführung als Arzneimittel dauert heute im Mittel sechs bis acht Jahre. Dabei müssen in der Regel einige Tausend neue chemische Verbindungen synthetisiert und biologisch in vielen Richtungen geprüft werden. Der Beweisstandard, der zum Nachweis der biologischen Aktivität, der Brauchbarkeit als Heilmittel und der relativen Ungefährlichkeit gefordert wird, steigt ständig, und es gibt ernstzunehmende Prognosen<sup>[1]</sup>, nach denen die Zeit für die Entwicklung eines neuen Medikamentes, das aus einem größeren Forschungsprojekt resultieren soll, bis auf 15 oder 20 Jahre zunehmen könnte.

Während dieser langen Dauer arbeiten Wissenschaftler vieler Fachrichtungen gleichzeitig oder nacheinander an zahlreichen erfolgversprechenden Substanzgrup-

pen, und mit dem Fortschreiten des Projekts wächst die Zahl der Befunde, die beurteilt werden müssen. Dabei sind Einzelaussagen erst dann verwertbar, wenn sie durch andere, ähnliche Ergebnisse bestätigt werden, und sehr viele Resultate, die zu einer Substanz synchron oder in längeren Zeitabständen eintreffen, sind immer wieder mit anderen Einzelwerten oder auch mit den Wirkungsprofilen anderer Substanzen zu vergleichen.

Solche Wirkungsvergleiche werden um so mehr Hinweise für die Optimierung der biologischen Aktivität durch Substanzgruppen- oder Substituentenvariation liefern,

je größer das kooperierende Team ist, d.h. je mehr und je unterschiedlicheren Testen die erfaßten Einzelergebnisse entstammen,

je sicherer die biologischen Aussagen sind, d.h. je mehr „standardisierte Tests“ benutzt werden können, deren Ergebnisse über längere Zeit vergleichbar bleiben und die von äußeren Einflüssen weitgehend unabhängig sind;

je umfangreicher und genauer die physikalisch-chemischen Daten sind, die für Regressions- und Korrelationsanalysen herangezogen werden können.

[\*] Dr. G. Ohnacker und W. Kalbfleisch  
Chemische Forschung der Dr. K. Thomae GmbH  
795 Biberach an der Riß

[\*\*] CCBF = Computerbearbeitung chemischer und biologischer Forschungsergebnisse.

[1] Die Medizin um 1990 – eine technische Prognose des Office of Health Economics, London. Med. Pharm. Studienges. e.V., Frankfurt 1969.