

Qualitative Struktur-Wirkungs-Korrelation

auf der Basis von Substrukturen

an ausgewählten Beispielen

Zur Erlangung des Grades eines Diplomchemikers  
eingereicht von

Eric Fontain

Diese Arbeit entstand am Lehrstuhl 1 des Organisch-chemischen Instituts der Technischen Universität München in 8046 Garching, Lichtenbergstraße 4, in der Zeit vom 1.2.1983 bis zum 31.7.1983 unter der Leitung von

Herrn o. Prof. Dr. Ivar Ugi

Ich danke Herrn Dr. W. Schubert für die Bereitstellung des Themas und seine kompetente Hilfestellung bei der Lösung mancher Probleme.

Mein Dank gilt auch Frl. Eva Bernhofer, die mich mit Geduld in das Programmsystem eingewiesen, und einen großen Teil der Umstellungsarbeiten geleistet hat.

Dem Leiter des Instituts, Herrn o. Prof. Dr. Ivar Ugi danke ich für die stets wohlwollende und großzügige Unterstützung dieser Arbeit.

Meinen Eltern

## Inhalt

1.	Einleitung.....	1
2.	Methodik.....	3
2.1.	Struktur-Wirkungs-Korrelation.....	3
2.1.1	Substrukturen als molekulare Deskriptoren.....	5
2.2.	Arbeitsweise.....	7
2.2.1.	Generierungsphase.....	7
2.2.2.	Korrelationsphase.....	8
2.3.	Fragmentierungsalgorithmen.....	9
2.3.1.	Sphärenfragmentierung.....	9
2.3.2.	Stufenfragmentierung.....	13
2.3.3.	Sphärenfragmentierung mit anschließ- ender Stufenfragmentierung.....	16
2.3.4.	Sphärenfragmentierung unter Berück- sichtigung reaktiver Zentren.....	17
2.4.	Statistische Bewertung von Substruk- turen.....	21
3.	Anwendungen.....	25
3.1.	Bioakkumulation in Fischen.....	26
3.1.1.	Datenbasis.....	26
3.1.2.	Ergebnisse.....	26
3.1.2.1.	Starke und schwache Bioakkumulation....	27
3.1.2.2.	Starke Bioakkumulation.....	30
3.1.3.	Kombinationen.....	31
3.1.4.	Diskussion.....	31
3.2.	Aromastoffe.....	33
3.2.1.	Datenbasis.....	33
3.2.2.	Ergebnisse.....	34
3.2.2.1.	Geschmacksqualität "süß".....	34
3.2.2.2.	Geschmacksqualität "sauer".....	35
3.2.2.3.	Geschmacksqualität "karamel".....	36
3.2.3.	Diskussion.....	36

4.	Zusammenfassung.....	39
5.	Literatur und Anmerkungen.....	40
	Anhang A	
	Anhang B	

## 1. Einleitung

Etwa 60000 chemische Stoffe werden derzeit in der Bundesrepublik im gewerblichen Maßstab produziert, und alljährlich kommen 500 bis 2000 Stoffe hinzu. Doch das ist nur ein Bruchteil dessen, was tatsächlich an Stoffen mit sehr kostenintensiver Laborarbeit entwickelt und auf mögliche Wirkungen geprüft werden muß [1].

Wie sich chemische Verbindungen auf einen lebenden Organismus auswirken wird vornehmlich durch Tierversuche geklärt. Diese Tierversuche haben den Nachteil, daß ihre Ergebnisse nicht ohne weiteres auf den Menschen übertragbar sind. Es stellt sich hier die Frage, wie der Gebrauch moderner Methoden und speziell der Einsatz von Computern helfen können, diese Probleme zu lösen.

Sowohl zur reinen Speicherung der Daten über die Verbindungen, als auch zu ihrer "intelligenten" Nutzung und Auswertung kann man heute Computer heranziehen. Auf dem Gebiet der Chemie sind in den letzten 10 Jahren Programme zur Syntheseplanung [2, 3], Struktur-Wirkungs-Korrelation [4] und Reaktionssimulation [5] entwickelt worden, die dem Chemiker zumindest einen Teil der sehr aufwendigen Laborarbeit abnehmen können. Diese Programme sind in der Lage, menschliche Intelligenz zu simulieren. Es sollte möglich sein, Algorithmen zu entwickeln, die die Eigenschaften von chemischen Verbindungen vorhersagen können. Diese Programme könnten ein grobes Sieb darstellen, mit Hilfe dessen man in der Lage wäre, einen großen Teil der Synthesearbeit zu ersetzen, indem sie das Hauptinteresse auf die Verbindungen lenken, die mit größter

Wahrscheinlichkeit die gewünschte Eigenschaft zeigen. Andererseits können auf diese Weise bereits vor der Synthese einer Verbindung mögliche Nebenwirkungen oder von ihr ausgehende Gefahren für die Umwelt erkannt werden.

## 2. Methodik

### 2.1. Struktur-Wirkungs-Korrelation

Unter der Struktur eines Moleküls versteht man im allgemeinen die Informationen über chemische Zusammensetzung, Konstitution, Konfiguration und Konformation. Diese vier Einzelaussagen beschreiben ein Molekül nahezu vollständig. Somit müßten die Eigenschaften eines Moleküls letztendlich durch die aus ihrer Kombination hervorgehenden Gesamtstruktur bedingt sein. Die Art und Weise, wie diese Gesamtstruktur das Eigenschaftsprofil einer Verbindung beeinflußt, ist nur in einfachen Fällen bekannt. So hat sich das Konzept der funktionellen Gruppen in der Beschreibung von Reaktivitäten gut bewährt. Besitzt man Informationen über die Struktur-Wirkungs-Korrelation, so lassen sich zum Beispiel Verbindungen mit gewünschten Eigenschaften planmäßig konstruieren, oder es können die Eigenschaften von noch wenig erforschten Molekülen vorhergesagt werden. Im allgemeinen erfordert es jedoch einen sehr großen Aufwand an Erfahrung, Intuition und Zeit, um etwas kompliziertere Zusammenhänge zwischen Struktur und Eigenschaft zu erkennen.

Besonders komplex werden die Verhältnisse, wenn man die Wirkungen chemischer Verbindungen in biologischen Systemen verstehen will. Dies ist z.B. der Fall, wenn man neue Pharmaka auf ihre Wirkungen untersucht, oder die Gefahren von Umweltchemikalien erkennen möchte. Die Wechselwirkung von Chemikalien mit Organismen geschieht im allgemeinen nach folgendem Schema :

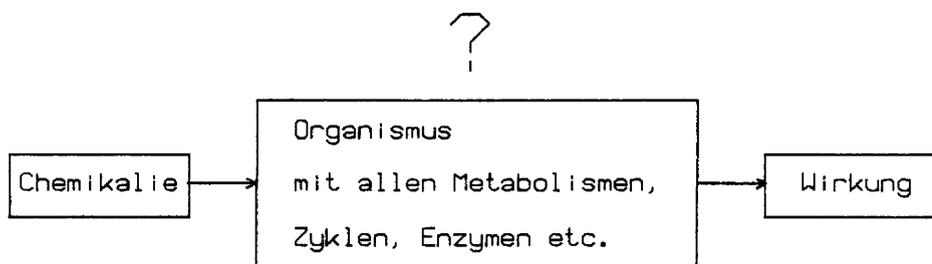


Abb. 1 : Wirkung einer Chemikalie auf einen Organismus

Zum Verständnis der Struktur-Wirkungs-Korrelation kann man auf mindestens zwei Wegen gelangen.

Man versucht zu ergründen, welche Veränderungen die zu untersuchende Substanz im Organismus auslöst, in welche Zyklen sie eingreift, welche Enzyme auf sie einwirken und mit welchen Reaktionspartnern sie reagiert. Dies erfordert eine vollständige Information über alle wesentlichen metabolischen Vorgänge, biologischen Strukturen und chemischen Verbindungen in einem Organismus, eine Information, die nur zum Teil oder überhaupt nicht verfügbar ist.

Andererseits kann man versuchen, ohne Kenntnisse der Vorgänge im Organismus, mit statistischen Verfahren Informationen darüber zu gewinnen, welche Chemikalien mit welcher Struktur welche Wirkung zeigen. Dies erfordert eine große Sammlung von Daten über die Wirksamkeit bzw. die Nicht-Wirksamkeit von vielen verschiedenen Verbindungen. Zur Verarbeitung dieser Daten bieten sich Computerprogramme aus dem Bereich der "maschinellen Intelligenz" an.

Quantitative Korrelationsverfahren dieser Art sind zum Beispiel die lineare Abhängigkeit der Aktivität von der freien Energie (Hansch [6]), Free-Wilson-Verfahren [7], DARC-PELCO-Verfahren [8], und neuere Methoden, wie Clusteranalyse [9] und Pattern Recognition [10,11]. Die neueren Verfahren zeichnen sich dadurch aus, daß die zu untersuchenden Verbindungen kein gemeinsames Grundgerüst mit variablen Substituenten besitzen müssen. Ihre Stärke besteht vielmehr darin, in sehr heterogenen Datensammlungen Gemeinsamkeiten zu finden.

Die Methoden des Pattern Recognition benötigen zur Beschreibung der Moleküle sogenannte "features" oder molekulare Deskriptoren. Diese können vorhandene Substrukturen, physikochemische Konstanten, aber auch die Geometrie beschreibende Größen sein.

### 2.1.1. Substrukturen als molekulare Deskriptoren

Wird ein Molekül in alle möglichen, allein die Konstitution beschreibenden Substrukturen zerlegt, so wird dieses Molekül in den durch die Konstitution bestimmten Eigenschaften durch die Gesamtheit dieser Teilstrukturen recht gut beschrieben. Viele Eigenschaften chemischer Verbindungen lassen sich durch die An- oder Abwesenheit bestimmter Substrukturen erklären. Dies trifft insbesondere im Bereich der chemischen Reaktivität zu. So ist zum Beispiel die Fähigkeit zur Bromaddition auf das Vorhandensein von einer oder mehreren nichtaromatischen Mehrfachbindungen angewiesen. Aber auch andere, komplexere Eigenschaften wie Farbe oder Geschmack sind zum Teil mit auftretenden Substrukturen korrelierbar.

Substrukturen in rein konstitutioneller Betrachtungsweise eignen sich zur Beschreibung von Molekülen und zur Korrelation aber nur dann, wenn die untersuchte Eigenschaft nicht eine Funktion von geometrischen Faktoren oder speziellen elektronischen Einflußgrößen ist.

Das hier angewandte qualitative Modell erlaubt es festzustellen, ob eine Eigenschaft überhaupt mit dem Auftreten von konstitutionellen Substrukturen korrelierbar ist. Ist dies der Fall, so können weiterhin die für die Eigenschaft verantwortlichen Teilstrukturen ermittelt werden. Dies geschieht, indem Substrukturen gesucht werden, die möglichst nur in Molekülen mit der untersuchten Eigenschaft vorkommen. Diese Substrukturen werden Effektoren genannt :

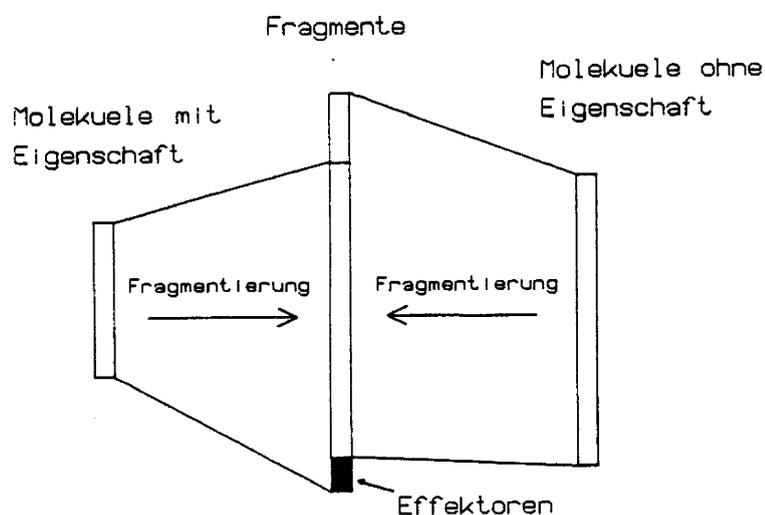


Abb. 2 : Ermittlung von Effektoren

## 2.2. Arbeitsweise

Es wurde ein bereits existierendes Programmsystem [12] angewandt, das auf einem neuen, qualitativen Modell der Struktur-Wirkungs-Korrelation beruht. Das Vorgehen teilt sich in zwei Arbeitsabschnitte. In der Generierungsphase werden aus den zu untersuchenden Molekülen Teilstrukturen erzeugt, die in der anschließenden Korrelationsphase auf ihre statistische Bedeutung hin untersucht werden.

### 2.2.1. Generierungsphase

Die in dieser Arbeit angewandte Korrelationsmethode verwendet zur Beschreibung von Molekülen Teilstrukturen, die nach verschiedenen Verfahren systematisch aus den Datenbasismolekülen erzeugt werden. Die Prozedur "GEN" [12] legt eine Liste von konstitutionell beschriebenen Strukturen an, in der sowohl die Stammoleküle, als auch alle entstehenden Fragmente eingetragen werden. Die Eintragung erfolgt in der Form von Bindungslisten mit zugeordneten Atomvektoren.

Das Substrukturspektrum ist die Gesamtheit aller Teilstrukturen, die nach Anwendung eines Fragmentierungsalgorithmus aus einem Molekül hervorgehen. Jedes Element dieser Menge ist mit seinen unmittelbaren Ausgangsmolekülen oder -fragmenten verknüpft (Vaterverweise) und enthält wiederum Verweise auf alle Fragmente, die aus ihm hervorgehen (Sohnverweise).

Werden viele oder große Moleküle einer Fragmentierung unterzogen, kommt es in steigendem Maße vor, daß Teilstrukturen entstehen, die schon erzeugt worden sind. In diesen Fällen wird das bestehende Fragment nur mit

einem Eintrag eines weiteren Vaters versehen und nicht neu in das Netzwerk eingetragen.

### 2.2.2. Korrelationsphase

Nach einer kompletten Fragmentierung werden die E-Moleküle in der Menge der Datenbasismoleküle markiert (E-Moleküle sind Moleküle mit einer Eigenschaft E). Die aus ihnen entstandenen Fragmente werden nun mit der Prozedur "CORREL" [12] daraufhin untersucht, zu wievielen E-Molekülen bzw. NE-Molekülen sie gehören (NE-Moleküle sind Moleküle, die die Eigenschaft E nicht besitzen). Dies führt zu einer ordnenden Klassifizierung der entstandenen Substrukturen in solche, die für eine Eigenschaft verantwortlich sind, und solche, die durch ihre Anwesenheit in einem Molekül keine Wirkung hervorrufen. Eine numerische Größe für den Grad dieser Korrelation bildet ein Konfidenzwert.

### 2.3. Fragmentierungsalgorithmen

An die erzeugten Fragmente werden im Rahmen einer Korrelation mehrere Forderungen gestellt :

- a) Die Fragmente sollten von chemischer Relevanz sein, bzw. möglichst großen statistischen Aussagewert haben.
- b) Die Zahl der Fragmente sollte nicht zu groß sein.
- c) Ähnliche Verbindungen sollten mindestens ein gemeinsames Fragment besitzen.

zu a) "möglichst großer statistischer Aussagewert" bedeutet, daß die Fragmente klein genug sein sollten, um in vielen Molekülen vorkommen zu können, aber nicht so klein, daß sie trivial werden (zB.: C-H, CH<sub>3</sub>, etc.).

Aus diesen Anforderungen ersieht man, daß die Wahl des geeigneten Fragmentierungsalgorithmus von entscheidender Bedeutung ist. Zur Wahl stehen in diesem Fall drei Algorithmen, wobei ein vierter in seinem Konzept vorgestellt werden soll. Die ersten drei Algorithmen arbeiten nach rein graphentheoretischen Kriterien, ohne besondere chemische Informationen, die in den Stammolekülen stecken, zu verwenden.

#### 2.3.1 Sphärenfragmentierung

Ein Sphärenfragment mit dem Radius  $r$  um ein Atom  $Z$  bezeichnet alle Atome und die sie verknüpfenden Bindungen, die nicht mehr als  $r$  Bindungen von  $Z$  entfernt sind. Es wird für jedes mindestens zweifach koordinierte Atom eine Sphäre mit wählbarem Radius erzeugt.

Algorithmus für die Sphärenfragmentierung :

- 1) Suche das nächste mindestens zweifach koordinierte Atom im Molekül.
- 2) Lege um dieses Atom eine Sphäre mit Radius  $r$  und speichere sie im Fragmentnetzwerk ab.
- 3) Wenn es noch unbehandelte Atome gibt, dann gehe nach 1, sonst "Stop".

Die Sphärenfragmentierung wird am Beispiel der 3-Methylbenzoesäure erläutert.

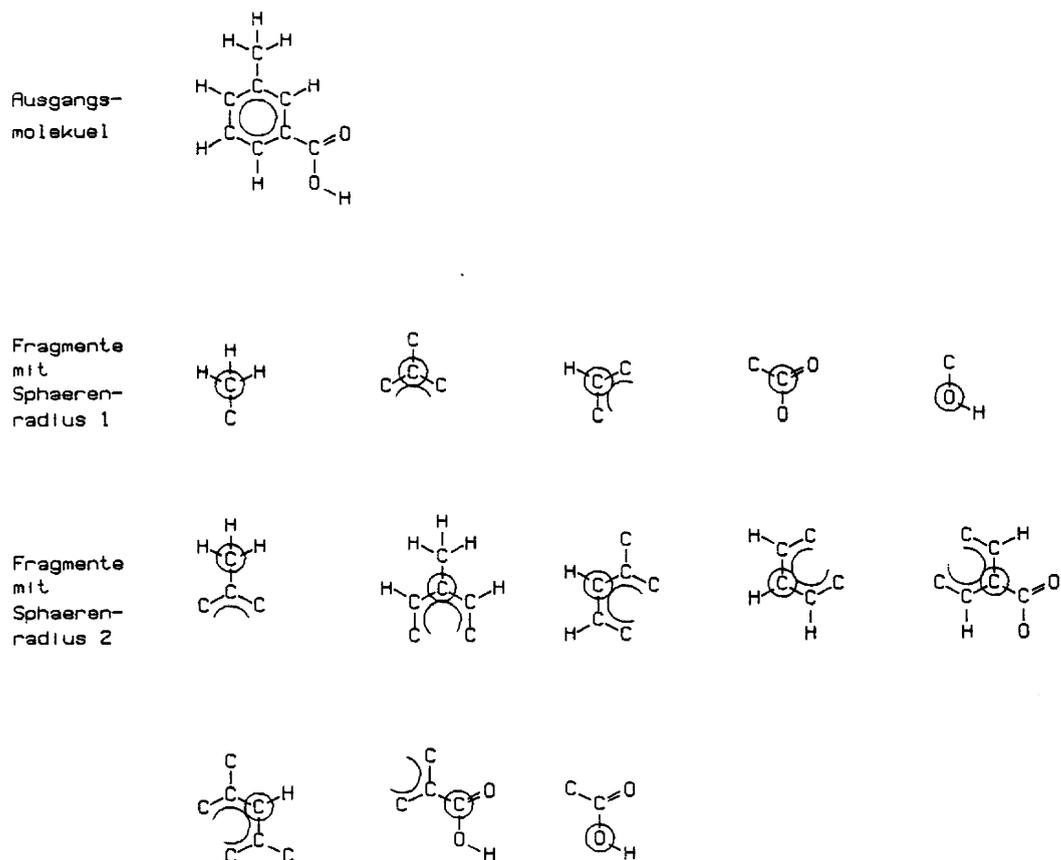


Abb. 3 : Sphärenfragmentierung von 3-Methylbenzoesäure; Zentralatom Z mit Kreis; Duplikate sind nicht aufgeführt.



Die folgende Graphik zeigt, wie sich nach der vollständigen Sphärenfragmentierung mit Radius 2 von 206 Molekülen die Zahl der Fragmente auf die verschiedenen Fragmentgrößen verteilt.

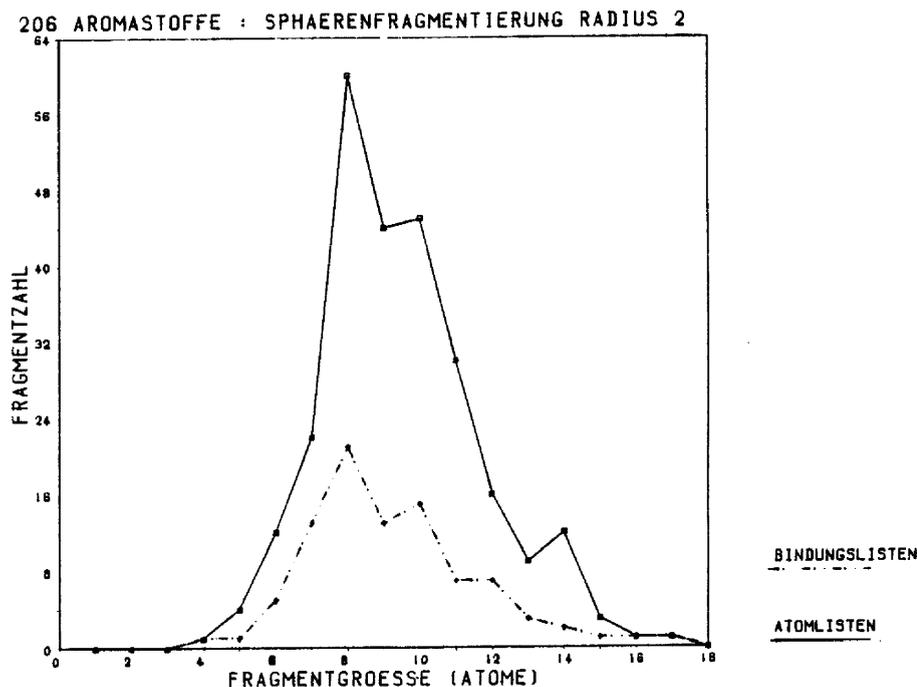


Abb. 4 : Fragmentverteilung bei einer Sphärenfragmentierung

Ein Vorteil der Sphärenfragmentierung ist, daß bei günstig gewähltem Radius Fragmente entstehen, die u.a. auch die bekannten funktionellen Gruppen beinhalten. So entstehen zum Beispiel bei der Sphärenfragmentierung von Acetaldehyd mit Radius 1 nur die Fragmente C-CHO und H<sub>3</sub>C-C, also eine Aldehydfunktion und eine Methylgruppe am C.

Ein Sphärenradius von 1 führt zu Fragmenten, wie sie als molekulare Deskriptoren in der "Quantitativen Struktur-Wirkungs-Korrelation" mit Pattern Recognition Methoden verwendet werden (augmented atom) [4].

### 2.3.2. Stufenfragmentierung

Das Konzept der Stufenfragmentierung folgt der Vermutung von Ulam [13], die, aus der Graphentheorie in die Sprache der Chemie übertragen, folgendes aussagt :

Zwei n-atomige Moleküle oder Fragmente sind identisch genau dann, wenn alle ihre (n-1)-atomigen Bruchstücke paarweise identisch sind.

Die (n-1)-atomigen Bruchstücke werden hier durch Entfernung von Randatomen erzeugt. Bei nicht identischen, aber ähnlichen Molekülen gilt, daß sich zwei Moleküle umso ähnlicher sind, je mehr paarweise identische Bruchstücke sie haben.

Die Stufenfragmentierung arbeitet nach folgendem Algorithmus :

- 1) Sämtliche Datenbasismoleküle werden entsprechend ihrer Zahl von Bindungen in ein Niveau des Fragmentnetzwerks eingetragen.
- 2) Von allen Molekülen oder Fragmenten mit n Bindungen aus einem Niveau werden nacheinander die Randatome entfernt und die entstehenden Fragmente mit n-1 Bindungen in das darunterliegende Niveau eingetragen. Eventuell vorhandene Ringe werden an allen möglichen Stellen aufgebrochen.
- 3) Schritt 2 wird für alle Niveaus, vom höchsten zum niedrigsten durchgeführt.

Die Stufenfragmentierung wird am Beispiel des Acetaldehyds erläutert :

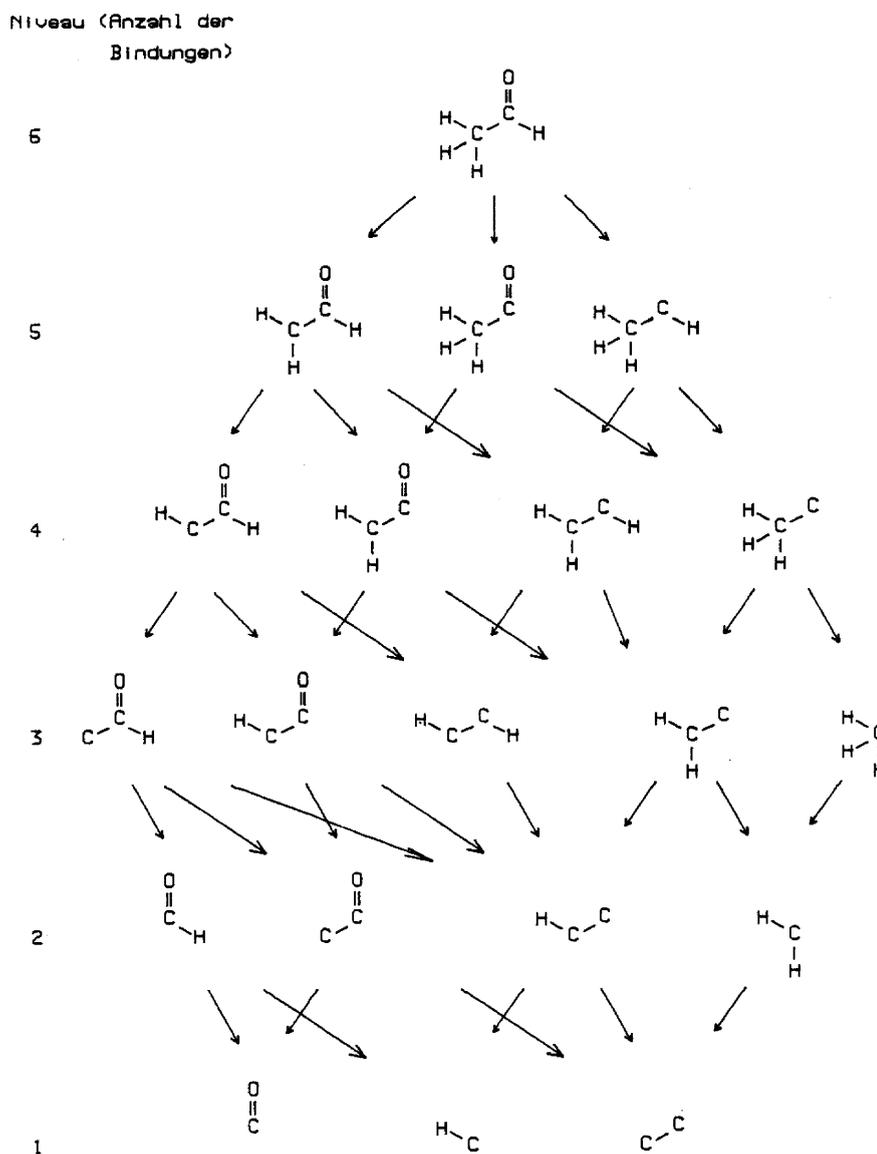


Abb. 5 : Stufenfragmentierung von Acetaldehyd

Auf diese Weise werden alle möglichen Substrukturen erzeugt, und ein Erkennen von Gemeinsamkeiten in verschiedenen, aber sehr ähnlichen Molekülen ist in jedem

Fall möglich. Erfahrungsgemäß ist jedoch die Anzahl der erzeugten Fragmente relativ groß und eine vollständige Stufenfragmentierung bei Molekülen mit mehr als 25 Atomen ist wegen des zu großen Rechenaufwandes im allgemeinen nicht mehr möglich.

Hierbei ist wichtig, daß bei einer hinreichend großen Anzahl von Molekülen die Anzahl der erzeugten Substrukturen nicht von der Zahl der fragmentierten Moleküle abhängig ist, sondern nur von der Zahl der Bindungen im größten beteiligten Molekül. Bei organischen Molekülen wird die aus graphentheoretischen Überlegungen kombinatorisch zu errechnende maximale Fragmentzahl reduziert. Das Atomsortiment beschränkt sich meist auf die Atome C, H, O, N, S und P. Nachdem etwa  $2/3$  der Bindungen gebrochen sind, werden größtenteils Duplikate erzeugt. Dies hat seinen Grund darin, daß Kohlenstoff eine maximale Koordinationszahl von 4 besitzt und nur Bindungen bis zur Bindungsordnung 3 möglich sind, was die Anzahl der unterschiedlichen Kombinationsmöglichkeiten bei kleineren Fragmenten stark einschränkt.

Nimmt man jedoch für ein  $n$ -atomiges Molekül  $(n-5)^5$  Fragmente als obere Grenze an [12], so ergibt sich für ein Molekül mit 90 Atomen eine Zahl von 4,4 Milliarden Fragmenten. Diese beanspruchen einen Speicherplatz von 6000 Gigabytes, unter der Annahme, daß jedes Fragment mit allen zugeordneten Informationen 1 Kilobyte Speicherplatz beansprucht. Dies sind Größenordnungen, wie sie heute in keinem Rechenzentrum verfügbar sind. Der größte Teil dieser Fragmente ist nur um 2 oder 3 Atome kleiner als das Stammolekül und hat, da jedes Fragment meist nur einem Stammolekül zugeordnet ist, keinen statistischen Aussagewert für die Korrelation.

### 2.3.3. Sphärenfragmentierung mit anschließender Stufenfragmentierung

Dieses Verfahren stellt eine Kombination der beiden im vorangegangenen beschriebenen Fragmentierungsalgorithmen dar. Es vermeidet die Nachteile der Stufenfragmentierung, die zu große Zahl der Fragmente und der Sphärenfragmentierung, das unvollständige Fragmentspektrum und kombiniert die Vorteile beider Fragmentierungstypen.

Es erfolgt zunächst eine Sphärenfragmentierung mit wählbarem Radius. Die entstandenen Fragmente, und nur diese, werden anschließend einer vollständigen Stufenfragmentierung unterzogen. Dadurch werden bis zu einer Maximalgröße, die durch den Radius der Sphärenfragmentierung limitiert ist, alle möglichen Substrukturen erzeugt.

Auch in diesem Fall ist die Anzahl der anfallenden Fragmente und somit Speicherplatzbedarf und benötigte Rechenzeit sehr groß. Bei 206 Molekülen mit bis zu 55 Atomen liegt der benötigte Plattenspeicherplatz für eine komplette Sphären/Stufenfragmentierung bei ca. 14 Megabytes. Diese Größenordnung liegt im Rahmen dessen, was den Benutzern moderner Rechenanlagen zur Verfügung gestellt werden kann.

Die folgende Graphik zeigt, wie sich die Fragmente auf die verschiedenen Fragmentgrößen (gemessen an der Atomzahl) verteilen (206 Moleküle, Sphärenradius 2).

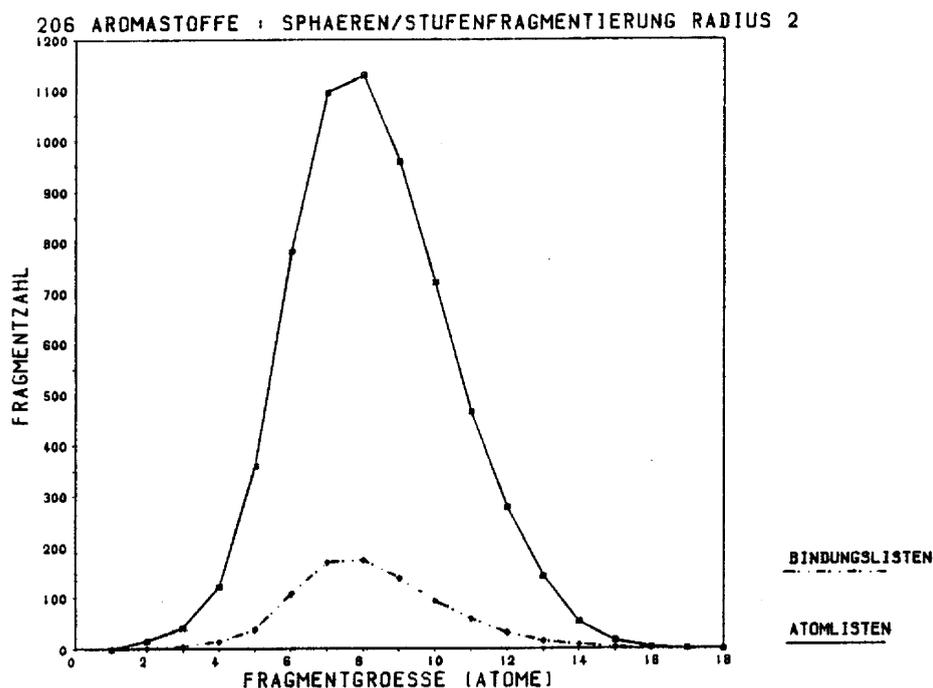


Abb. 6 : Fragmentverteilung einer Sphären/-  
Stufenfragmentierung

#### 2.3.4. Sphärenfragmentierung unter Berücksichtigung reaktiver Zentren

Die in den vorangegangenen Abschnitten beschriebenen Fragmentierungsalgorithmen berücksichtigen in den betrachteten Molekülen keine reaktiven Zentren, die durch das Auftreten von Heteroatomen oder Mehrfachbindungen gekennzeichnet sind. Von diesen Zentren und ihrer Umgebung kann angenommen werden, daß sie in besonderem Maße für die Eigenschaften eines Moleküls verantwortlich sind.

Aus diesem Grund wurde ein Algorithmus zur Fragmentierung entworfen, der diese chemischen Informationen berücksichtigt und die Erzeugung von Teilstrukturen auf diese besonders ausgezeichneten Zentren konzentriert.

Das Verfahren unterscheidet sich von der reinen Sphärenfragmentierung in folgenden Punkten:

- 1) Es werden Sphären mit unterschiedlichen Radien erzeugt.
- 2) Als Sphärenzentren werden nur Atome gewählt, die chemisch besonders ausgezeichnet sind.
- 3) Ringe oder Ringsysteme werden zusätzlich mit in das Fragmentspektrum aufgenommen.

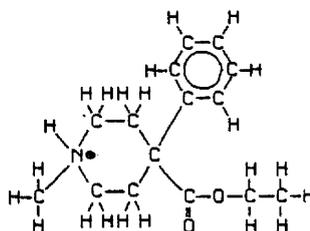
Es wird folgender Algorithmus zur Fragmentierung eines Moleküls vorgeschlagen :

- 1) Es werden alle Atome mit mindestens einer der folgenden Eigenschaften in charakteristischer Weise markiert :
  - a) Das Atom ist ein Heteroatom (verschieden von C und H)
  - b) Von dem Atom geht mindestens eine Bindung der Ordnung 2 oder 3 aus.
  - c) Das Atom ist Element eines Rings.
- 2) Um jedes nach a oder b markierte Atom werden hintereinander Sphären mit den Radien 1, 2 und 3 gelegt und in das Fragmentnetzwerk eingetragen.
- 3) Alle Atome, die nicht unmittelbar an ein Atom mit einer Marke nach c gebunden sind, werden gelöscht und die verbleibenden Ringfragmente in das Netzwerk eingetragen.
- 4) In den Ringfragmenten werden alle Atome, die nicht eine Marke nach c tragen gelöscht und die entstehenden Ringkernfragmente in das Netzwerk eingetragen.

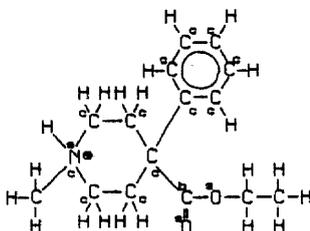
Die Vielfalt der entstehenden Teilstrukturen sollte in vielen Fällen genügen, die Moleküle hinreichend zu beschreiben. Jedoch kann auch hier, wie im Fall der reinen Sphärenfragmentierung, zur Vervollständigung des Fragmentspektrums eine Stufenfragmentierung (ohne Aufbrechen der Ringe) aller entstandenen Fragmente folgen. In diesem Fall reicht es natürlich aus, Sphären vom Radius 3 zu erzeugen, da alle kleineren Fragmente damit automatisch generiert werden.

Der vorgeschlagene Algorithmus wird am Beispiel des Kations des 4-Carboxy-1-methyl-4-phenyl-piperidinhydrochlorids (Dolantin) verdeutlicht.

Ausgangs-  
molekül



Markierungen



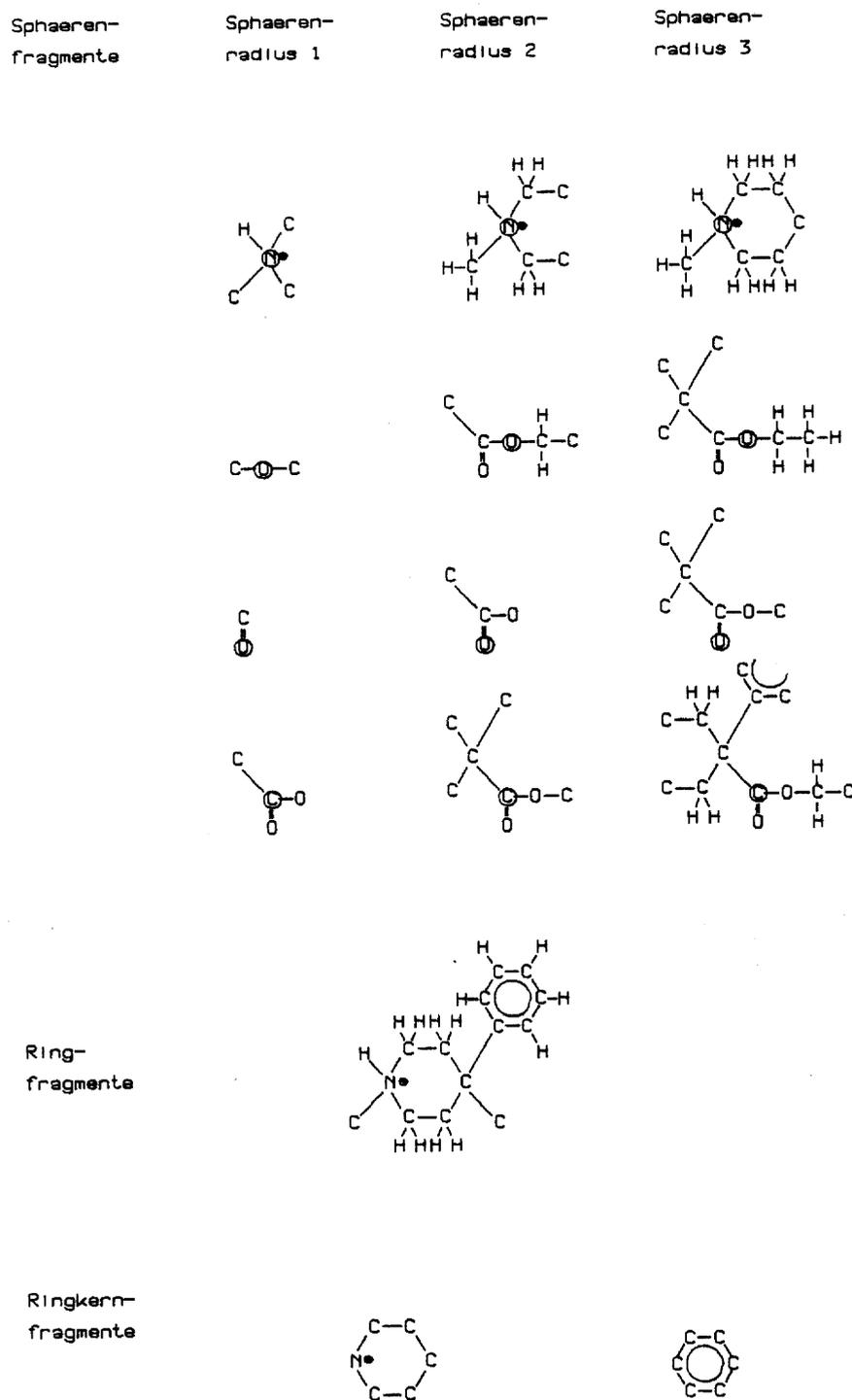


Abb. 7 : Fragmentierung von Dolantin unter Berücksichtigung reaktiver Zentren

#### 2.4. Statistische Bewertung von Substrukturen

Eine Eigenschaft  $E$  definiert in der Menge der Datenba-  
sismoleküle  $B$  eine Teilmenge  $E$  mit  $n^e$  Elementen. Es  
sind dies alle Moleküle, die die Eigenschaft besitzen;  
über eine quantitative Abstufung in eine mehr oder  
weniger starke Ausprägung dieser Eigenschaft wird  
keine Aussage gemacht.

Eine Substruktur  $S$  definiert in der Menge  $B$  eine Teil-  
menge  $S$  mit  $n_s$  Elementen. Es sind dies alle Moleküle,  
die die Substruktur enthalten. Wie oft eine Substruk-  
tur in einem Molekül vorkommt, wird nicht betrachtet.

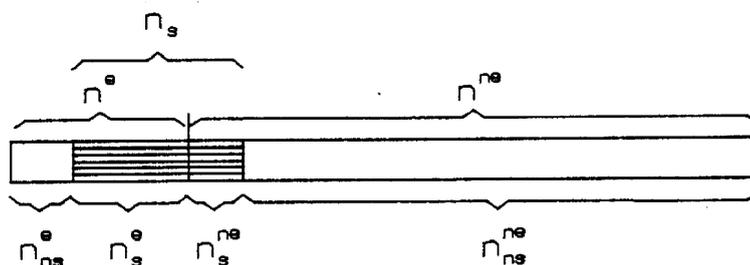
Der Durchschnitt der Mengen  $E$  und  $S$ ,  $ES = E \cap S$  mit  $n_s^e$   
Elementen sind alle Moleküle, die die Substruktur  $S$   
enthalten und die Eigenschaft  $E$  besitzen.

Eine statistische Bewertung von Substrukturen, die zum  
Ziel hat, für eine Eigenschaft verantwortliche Frag-  
mente zu finden, ordnet jeder Substruktur von  $E$  einen  
Wert zu, der einerseits den relativen Anteil von  $ES$  an  
 $E$ , sowie den relativen Anteil von  $ES$  an  $S$   
berücksichtigt. Diesen Wert, der ein Maß für die  
Wahrscheinlichkeit ist, daß eine Substruktur für eine  
Eigenschaft verantwortlich ist, nennen wir den Konfi-  
denzwert. Die Quotienten  $n_s^e/n^e$  und  $n_s^e/n_s$  sollten für  
eine Substruktur, die mit ihrem Auftreten in einem  
Molekül mit hoher Wahrscheinlichkeit eine bestimmte  
Eigenschaft hervorruft, gegen den Wert 1 streben.  
Solch eine Substruktur nennt man Effektor. Sie ist in  
vielen Molekülen von  $E$  enthalten, jedoch nur in ge-  
ringer Zahl in Molekülen, die nicht  $E$  angehören. Um  
eine ordnende Klassifizierung von Substrukturen zu er-  
reichen, wie es in der Prozedur "CORREL" durchgeführt

wird, muß für jede Substruktur der Moleküle in E ein solcher Konfidenzwert berechnet werden. Die Substrukturen mit den höchsten Konfidenzwerten werden als Effektoren ausgegeben.

Der Konfidenzwert wird folgendermaßen errechnet :

Es sei S : eine Substruktur,  
 E : eine Eigenschaft,  
 $n_s$  : Anzahl d. Mol. mit S,  
 $n^e$  : Anzahl d. Mol. mit E,  
 $n^{ne}$  : Anzahl d. Mol. ohne E,  
 $n_s^e$  : Anzahl d. Mol. mit S und E,  
 $n_s^{ne}$  : Anzahl d. Mol. mit S ohne E,  
 $n_{ns}^e$  : Anzahl d. Mol. ohne S mit E,  
 $n_{ns}^{ne}$  : Anzahl d. Mol. ohne S ohne E,  
 C : Konfidenzwert.



$$C = \frac{n_s^e}{n^e} \cdot \frac{n_s^{ne}}{n_s} \quad (0 \leq C \leq 1)$$

Abb. 8 : Konfidenzwert

Der absolute Wert von C hängt somit ab von der Zusammensetzung der Datenbasis. Dies hat zur Folge, daß

Konfidenzwertberechnungen, denen unterschiedliche Datenbasen zugrunde liegen, nicht unmittelbar miteinander verglichen werden können. C-Werte dienen in dieser Form innerhalb eines Programmlaufs nur als Beschreibung für die relative Verteilung einer Substruktur und sind ein Hilfsmittel zur Ordnung der Fragmente in Effektoren und in für eine Eigenschaft bedeutungslose Substrukturen.

Um eine Bewertung von Substrukturen zu finden, die eine absolute, von der Struktur der Datenbasis unabhängige Größe darstellt, muß man den relativen Anteil von E an B mit berücksichtigen. Dies kann geschehen, indem man versucht, an Hand des Kriteriums - Molekül enthält Effektor ? - eine Klassifizierung der Moleküle in E- und NE-Moleküle vorzunehmen.

Die gefundenen Effektoren decken jeder für sich nicht den gesamten Bereich der E-Moleküle ab. Der relative Anteil der als NE-Moleküle klassifizierten E-Moleküle bezogen auf alle als NE-Moleküle klassifizierten Moleküle ist ein Maß für die Unsicherheit der Aussage über die Klassifizierung. Um die Skala dieser Fehlerquote auf einen Bereich zwischen 0 und 1 zu normieren, wird dieses Verhältnis noch durch den relativen Anteil der E-Moleküle an der Datenbasis B dividiert.

Die niedrigsten Fehlerquoten erreicht man, wenn man die Anfrage auf mehrere Effektoren mit einer "oder"-Verknüpfung auf die Moleküle anwendet, d.h. als E-Moleküle solche klassifiziert, die entweder den einen oder einen anderen Effektor enthalten. Hierbei muß nicht immer die Kombination der Effektoren mit den höchsten Konfidenzwerten zu einer niedrigen Fehlerquote führen, denn es sollte ja der zweite Effektor

möglichst den Teil des Eigenschaftsbereichs abdecken,  
den der erste nicht erreicht hat.

Für die Fehlerquote F gilt :

$$F = \frac{n_{ns}^e}{n_{ns}^e + n_{ns}^{ne}} = \frac{n^e + n^{ne}}{n^e}$$

Abb. 9 : Fehlerquote der Klassifizierung

### 3. Anwendungen

Anhand von zwei ausgewählten Testdatensätzen wurde das im vorigen beschriebene Modell und die zugehörigen Computerprogramme auf ihre Anwendbarkeit geprüft.

Das Programmsystem besteht aus ca. 60 modular aufgebauten Prozeduren und ist in der Programmiersprache PL/I geschrieben. Es ist implementiert auf den Computern AMDAHL 470 V/6 und SIEMENS 7.880 im Rechenzentrum des Max-Planck-Instituts für Plasmaphysik in Garching.

Um die Datensätze der hier notwendigen Größenordnung (Zahl der Moleküle größer als 200, Zahl der Atome in einem Molekül größer als 90) bearbeiten zu können, mußten in den einzelnen Prozeduren Änderungen und Erweiterungen vorgenommen werden [14].

Die computergerechte Beschreibung der Datenbasismoleküle (Bindungslisten und Atomvektoren) wurde mit einem interaktiven Moleküleingabeprogramm ("MAGIC") erzeugt [15].

### 3.1. Bioakkumulation in Fischen

#### 3.1.1. Datenbasis

Im Rahmen eines gemeinsamen Projektes mit Prof. Korte von der Gesellschaft für Strahlen- und Umweltforschung (GSF) steht uns ein Datensatz von 105 Molekülen zur Verfügung, die im Anhang A zusammengestellt sind (Nr. 001 01 bis Nr. 091 01). Nach einer von der GSF durchgeführten Literaturrecherche ergibt sich, daß von diesen Molekülen 21 die Eigenschaft zeigen, von Fischen akkumuliert zu werden, und zwar Nr. 001 01 bis Nr. 006 01 schwach bis mittel und Nr. 007 01 bis Nr. 010 01 stark akkumulierend.

Diese E-Moleküle sind in der Hauptsache polychlorierte aromatische Kohlenwasserstoffe. Der Bereich der NE-Moleküle ist sehr heterogen und enthält substituierte Aromaten, Heterocyclen, Carbonylverbindungen, gesättigte Kohlenwasserstoffe und Alkohole.

#### 3.1.2. Ergebnisse

Da das Korrelationsverfahren nicht zwischen mittlerer und starker Bioakkumulation unterscheidet, wurden in zwei Programmläufen einmal die Moleküle Nr. 001 01 bis Nr. 010 01 (stark und schwach akkumulierend) und einmal die Moleküle Nr. 007 01 bis Nr. 010 01 (stark akkumulierend) als E-Moleküle betrachtet.

Um die Auswirkungen verschiedener Fragmentierungsalgorithmen auf das Ergebnis der Korrelation studieren zu können, wurden mehrere Programmläufe mit Sphärenfragmentierungen unterschiedlicher Radien, sowie eine kombinierte Sphären/Stufenfragmentierung

durchgeführt. Die Tabellen zeigen jeweils eine Auswahl der ermittelten Effektoren. Die Auswahlkriterien sind angegeben.

### 3.1.2.1. Starke und schwache Bioakkumulation

Tabelle 1 Sphärenfragmentierung Radius 1

Auswahlkriterium :  $C \geq 0.15$

$n^e = 21$   $n^{ne} = 84$

Effektor	$n_s$	$n_s^e$	C	F
1 	32	19	0.54	0.14
2 	77	18	0.20	0.54
3 	7	5	0.17	0.82

Tabelle 2 Sphärenfragmentierung Radius 2

Auswahlkriterium :  $C \geq 0.15$  und  $n_s^e/n_s \geq 0.7$   
 $n^e = 21$   $n^{ne} = 84$

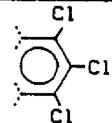
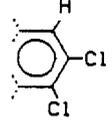
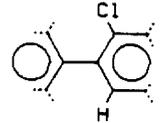
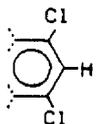
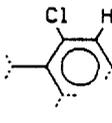
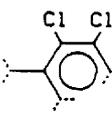
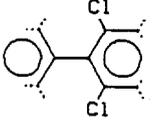
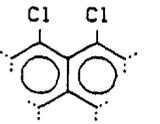
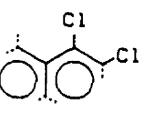
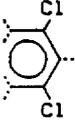
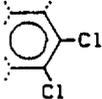
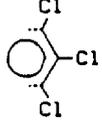
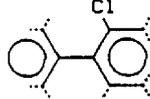
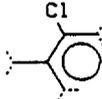
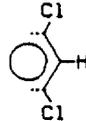
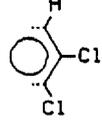
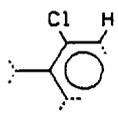
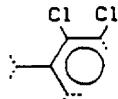
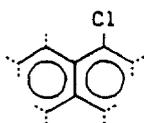
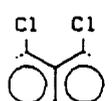
Effektor	$n_s$	$n_s^e$	C	F
1 	11	10	0.43	0.59
2 	8	7	0.29	0.72
3 	6	6	0.29	0.76
4 	6	6	0.29	0.76
5 	8	6	0.21	0.77
6 	6	5	0.20	0.81
7 	4	4	0.19	0.84
8 	4	4	0.19	0.84
9 	4	4	0.19	0.84

Tabelle 3 Sphären/Stufenfragmentierung Radius 2

Auswahlkriterium :  $C \geq 0.15$  und  $n_S^e/n_S \geq 0.7$

$n^e = 21$   $n^{ne} = 84$

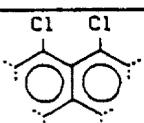
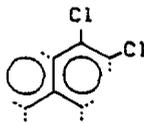
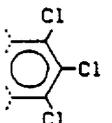
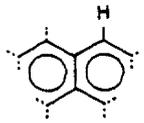
Effektor	$n_S$	$n_S^e$	C	F
1 	16	15	0.67	0.34
2 	14	12	0.49	0.49
3 	11	10	0.43	0.59
4 	8	8	0.38	0.67
5 	12	9	0.32	0.65
6 	6	6	0.29	0.76
7 	8	7	0.29	0.72
8 	8	6	0.21	0.77
9 	6	5	0.20	0.81

10		4	4	0.19	0.84
11		4	4	0.19	0.84

### 3.1.2.2. Starke Bioakkumulation

Tabelle 4 Sphären/Stufenfragmentierung Radius 2

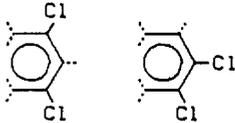
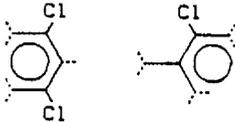
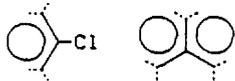
Auswahlkriterium :  $C \geq 0.15$  und  $n_s^e/n_s \geq 0.45$   
 $n^e = 10$   $n^{ne} = 95$

Effektor	$n_s$	$n_s^e$	C	F
1 	4	4	0.40	0.62
2 	4	4	0.40	0.62
3 	11	5	0.23	0.56
4 	5	3	0.18	0.74

### 3.1.3. Kombinationen

Tabelle 5 "oder"-Kombinationen von Effektoren

$$n^e = 21 \quad n^{ne} = 84$$

Kombination von Effektoren		$n_s$	$n_s^e$	C	F
1		18	16	0.68	0.29
2		21	18	0.78	0.18
3		35	20	0.54	0.07 !

### 3.1.4. Diskussion

Bei fast allen gefundenen Effektoren handelt es sich um Fragmente, die einfach oder mehrfach chlorierten Aromaten angehören. Es finden sich Substrukturen von Bi- und Terphenylen und kondensierten Aromaten. Dies entspricht in jedem Fall auch den experimentellen Befunden. Die relativ hohen C-Werte und die kleinen erreichbaren Fehlerquoten in der Klassifizierung zeigen, daß eine Korrelation zwischen Substrukturen und der Eigenschaft der Bioakkumulation besteht.

Das Fragment "Chlor an Aromat" (Tabelle 1 Nr.1) erfaßt zwar annähernd alle akkumulierenden Moleküle, ist aber auch in vielen nichtakkumulierenden Molekülen enthal-

ten. Aus diesem Grund ist sein Konfidenzwert niedriger als der von Fragmenten aus Tabelle 2 und 3, die in jeweils weniger E-Molekülen vorkommen. In der "oder"-Kombination mit dem Fragment "kondensierter Aromat" führt diese Substruktur jedoch zu einer relativ guten Klassifizierung, deren Fehlerquote von 7% hauptsächlich darin begründet ist, das einige NE-Moleküle als bioakkumulierend eingestuft werden.

Der Unterschied im C-Wert der Fragmente Nr.1 in Tabelle 2 und Nr.1 in Tabelle 3 zeigt deutlich den Einfluß der Stufenfragmentierung, die durch Freilassen von variablen Positionen zu Fragmenten führt, die allgemeinerer Natur, und so in viel mehr E-Molekülen enthalten sind.

Der Effektor mit dem höchsten Konfidenzwert ist unter diesen Voraussetzungen ein meta-dichlorsubstituierter Aromat mit einer variablen Stelle zwischen den zwei Chloratomen. Diese variable Position kann durch ein Chloratom (Nr.3 in Tabelle 3) oder ein Wasserstoffatom (Nr.6 in Tabelle 3) besetzt werden.

Die ermittelten Effektoren eignen sich durchaus zur Klassifizierung von Molekülen, die nicht an dieser Untersuchung beteiligt waren.

## 3.2. Aromastoffe

### 3.2.1. Datenbasis

Grundlage des Korrelationsversuchs bildet eine Liste der 206 gebräuchlichsten chemisch definierten Geschmacksstoffe [16]. Eine Zusammenstellung der Verbindungen findet sich im Anhang B. Jedem dieser Moleküle ist über eine spezielle Kodierung eine Reihe von Geschmacksqualitäten zugeordnet. So besitzt zum Beispiel die Verbindung Nr. 106 (Nonalid 1,4 oder  $\gamma$ -Nonalacton) die Geschmacksqualitäten "kokosartig", "cremig", "fruchtig" und "nußartig".

Es steht hiermit eine sehr umfangreiche Zusammenstellung von Molekülen mit vielen unterschiedlichen Eigenschaften zur Verfügung. Da die einzelnen Geschmacksqualitäten in verschiedener Zahl in der Datenbasis vertreten sind, bietet es sich an, zur Korrelation solche Geschmacksqualitäten auszuwählen, die relativ vielen Molekülen und solche, die nur wenigen Molekülen der Datenbasis zu eigen sind, um den Einfluß des Anteils der E-Moleküle an der Datenbasis auf das Ergebnis der Korrelation studieren zu können.

Von den 206 gebräuchlichsten Geschmacksstoffen besitzen 86 die Geschmacksqualität "süß", 14 schmecken "sauer", und nur 5 haben die Geschmackseigenschaft "karamel". Die Verbindungen mit diesen Geschmacksqualitäten, mit denen der Korrelationsversuch durchgeführt wurde, sind ebenfalls im Anhang B zusammengestellt.

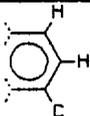
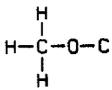
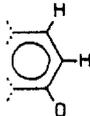
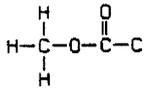
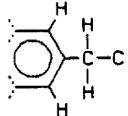
### 3.2.2. Ergebnisse

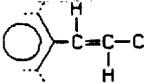
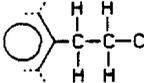
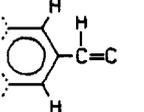
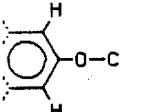
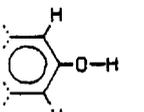
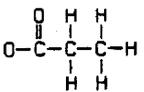
Aus technischen Gründen (Systemumstellung an der verwendeten Rechenanlage) konnte keine kombinierte Sphären/Stufenfragmentierung durchgeführt werden. In einigen Fällen wäre dadurch eine etwas differenziertere Aussage möglich gewesen. Alle ermittelten Effektoren sind Sphärenfragmente vom Radius 2.

#### 3.2.2.1. Geschmacksqualität "süß"

Tabelle 6 Sphären/Stufenfragmentierung Radius 2

Auswahlkriterium :  $C \geq 0.05$  und  $n_s^e/n_s \geq 0.7$   
 $n^e = 86$   $n^{ne} = 120$

Effektor	$n_s$	$n_s^e$	C	F
1 	52	37	0.31	0.76
2 	26	20	0.18	0.88
3 	23	17	0.15	0.90
4 	15	12	0.11	0.93
5 	16	12	0.10	0.93

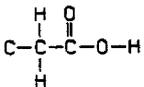
6		12	9	0.08	0.95
7		6	6	0.07	0.96
8		11	8	0.07	0.96
9		11	8	0.07	0.96
10		11	8	0.07	0.96
11		7	6	0.06	0.96

### 3.2.2.2. Geschmacksqualität "sauer"

Tabelle 7 Sphären/Stufenfragmentierung Radius 2

Auswahlkriterium :  $C \geq 0.12$

$n^e = 14$   $n^{ne} = 192$

Effektor	$n_s$	$n_s^e$	C	F
1 	15	11	0.58	0.23
2 	9	6	0.29	0.60

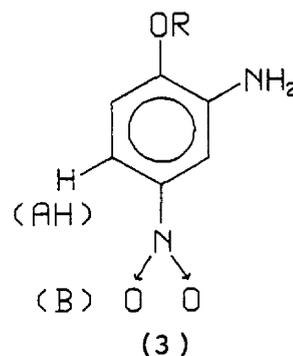
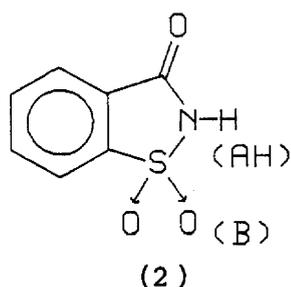


Protonen gebunden sieht [17].

Die gute Übereinstimmung zwischen dem mit dem Programm gefundenen Ergebnis und der praktischen Erfahrung zeigt, daß das Modell eine vorhandene Korrelation zwischen Eigenschaft und Struktur eindeutig erkennt und die richtigen Effektoren ermittelt.

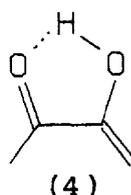
Bei den Geschmacksqualitäten "süß" und "karamel" sind die Konfidenzwerte der Effektoren niedriger und die errechneten Fehlerquoten der Klassifizierung höher als in den Fällen der Geschmacksqualität "sauer" und der Bioakkumulation. Hier ist eine eindeutige Korrelation zwischen Geschmack und Struktur nicht offensichtlich. Doch zeigen auch die Ergebnisse dieser beiden Korrelationsversuche in gewissem Sinne eine Übereinstimmung mit den mechanistischen Theorien der Struktur-Geschmacks-Korrelation.

Süßer Geschmack wird von sehr verschiedenen Verbindungen hervorgerufen. Als gemeinsame strukturelle Voraussetzung wird ein Protonendonator/-acceptor System (AH/B-System) angesehen, das mit einem komplementären Geschmacksrezeptor über Wasserstoffbrücken in Wechselwirkung treten kann [17,18]. Dieses Modell wird anhand der Beispiele Saccharin (2) und eines Nitrobenzolderivats (3) erläutert.



Die in Tabelle 5 aufgeführten Effektoren erfüllen zum Teil die Bedingungen dieser mechanistischen Theorie. Das electrophile AH-Zentrum wird in diesen Fällen gebildet durch ein Wasserstoffatom am Aromat (bei Nr.3, Nr.6, Nr.8, Nr.9 und Nr.10). Die korrespondierenden nucleophilen B-Zentren stellen bei Nr.3, Nr.9 und Nr.10 die Sauerstoffatome und bei Nr.6 und Nr.8 die  $\pi$ -Elektronen der Kohlenstoff-Kohlenstoff-Doppelbindung dar. Diese Interpretation steht durchaus im Einklang mit einer Erweiterung der AH/B-Theorie zu einem Modell, in dem das AH/B-System durch ein electrophiles/nucleophiles System ersetzt wird [19], was bedeutet, daß die Säure/Base-Eigenschaften von Wasserstoffatomen in ihrer Bedeutung gegenüber einer Betrachtung der Ladungsverteilung und der Polarisierbarkeit zurücktreten.

Für die Geschmacksqualität "karamel" sieht die Lebensmittelchemie eine Enol-Oxo-Gruppierung (4) als eigenschaftsgebendes Strukturmerkmal an [19].



Es ist auffallend, daß, obwohl die Anzahl der gemeinsamen Teilstrukturen denkbar gering ist, die gefundenen Effektoren diese Gruppierung enthalten. Bei einer nachfolgenden Stufenfragmentierung hätte es sich herausgestellt, daß diese Enol-Oxo-Gruppierung in 3 der 5 E-Moleküle als Substruktur vorkommt.

#### 4. Zusammenfassung

Ein Programm zur Struktur-Wirkungs-Korrelation mittels Teilstrukturgenerierung wurde erweitert und den Anforderungen größerer Datensätze angepaßt.

Anhand von zwei Testdatensätzen wurde die Anwendbarkeit des Programms über das bereits beschriebene hinaus [12] geprüft. Für die Bioakkumulation von chemischen Stoffen in Fischen wurde eine eindeutige Korrelation mit dem Auftreten bestimmter Strukturfragmente von polychlorierten Aromaten festgestellt. Eine automatische Klassifizierung des Testdatensatzes in bioakkumulierende und nicht bioakkumulierende Substanzen zeigt eine niedrigere Fehlerquote.

Für den Korrelationsversuch Geschmack-Struktur konnte im Fall der Geschmacksqualität "sauer" eine eindeutige Korrelation nachgewiesen werden. Die Ergebnisse bei den Geschmacksqualitäten "süß" und "karamel" erfüllen nach vorsichtiger Deutung zum Teil die Anforderungen mechanistischer Theorien über die Struktur-Geschmacks-Korrelation. Eine Klassifizierung des Testdatensatzes ist jedoch nicht möglich.

## 5. Literatur und Anmerkungen

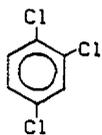
- 1) Spektrum der Wissenschaften, Nr. 6, S.17 (1983).
- 2) E.J.Corey, W.T.Wipke, Science, 166, S.178 (1969).
- 3) J.Brandt, J.Friedrich, J.Gasteiger, C.Jochum, W.Schubert, I.Ugi, Am. Chem. Soc. Symposium Series, 61, S.33 (1977).
- 4) J.K.Seydel, K.J.Schaper, "Chemische Struktur und biologische Aktivität von Wirkstoffen", Verlag Chemie (1979).
- 5) W.Schubert, Informal Commun. Math. Chem., 6, S.213 (1979).
- 6) J.W.McFarland, E.Jucker (Ed.), Fortschr. d. Arzneimittelforschung, Birkhäuser Verlag Stuttgart, 15, S.123 (1971).
- 7) S.M.Free, J.W.Wilson, J. Med. Chem., 7, S.395 (1964).
- 8) J.E.Dubois, C. R. Acad. Sci., 264, S.1019 (1966).
- 9) K.C.Chu, Anal. Chem., 46, S.1181 (1974).
- 10) B.R. Kowalsky, C.F.Bender, J. Am. Chem. Soc., 94, S.5632 (1972).
- 11) B.R. Kowalsky, C.F.Bender, J. Am. Chem. Soc., 95, S.686 (1973).

- 12) J.Friedrich, I.Ugi, J. Chem. Res. (S), S.70 (1980).
- 13) S.M.Ulam, "A Collection of Mathematical Problems", Wiley (Interscience), New York (1960).
- 14) In Zusammenarbeit mit E.Bernhofer.
- 15) J.Bauer, W.Schubert, Comp. Chem., 7, S.61 (1983).
- 16) W.Pilz, E.Ziegler (Ed.), "Die natürlichen und künstlichen Aromen", Verlag Hüthig (1982).
- 17) H.-D.Belitz, H.Wieser, Nachr. aus Chemie, Technik und Lab., 25, S.442 (1977).
- 18) R.S.Shallenberger, T.E.Acree, L.M.Beidler (Ed.), "Handbook of Sensory Physiology" 4/2., Springer Verlag, S.257 (1971).
- 19) H.-D.Belitz, W.Grosch, "Lehrbuch der Lebensmittelchemie", Springer Verlag, (1982).

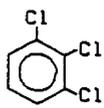
Anhang A

Datenbasismoleküle zu :

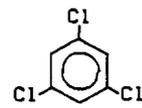
"3.1. Bioakkumulation in Fischen"



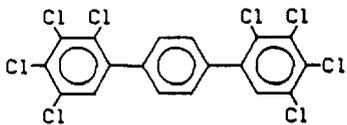
001 01



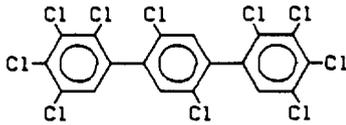
001 02



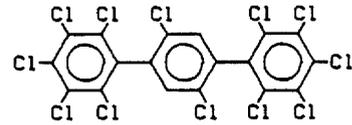
001 03



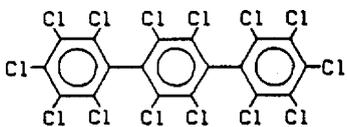
002 01



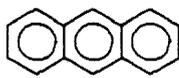
002 02



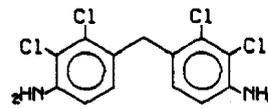
002 03



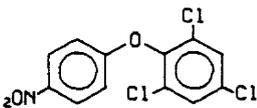
002 04



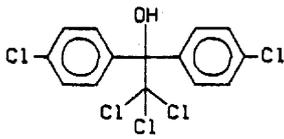
003 01



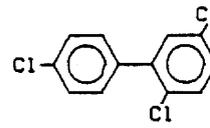
004 01



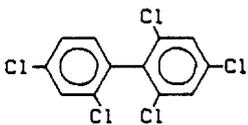
005 01



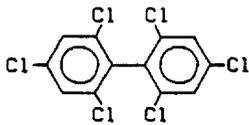
006 01



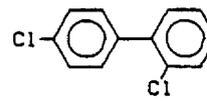
007 01



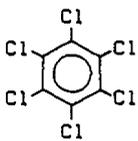
007 02



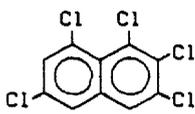
007 03



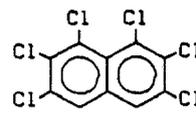
007 04



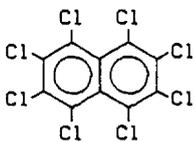
008 01



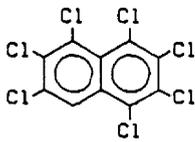
009 01



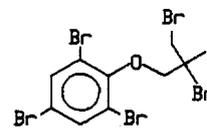
009 02



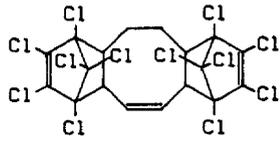
009 03



009 04



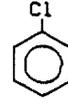
010 01



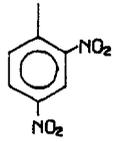
011 01



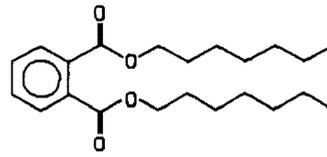
012 01



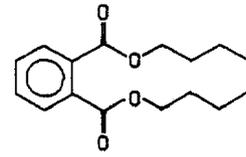
013 01



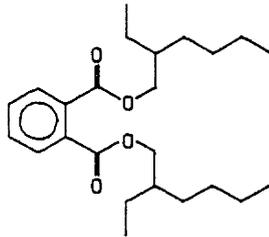
014 01



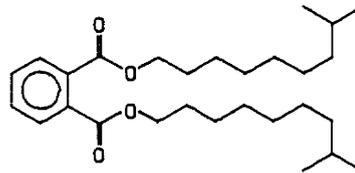
015 01



016 01



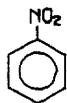
017 01



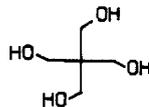
018 01



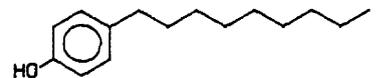
019 01



020 01



021 01



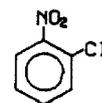
022 01



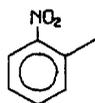
023 01



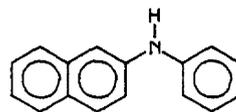
024 01



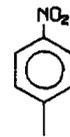
025 01



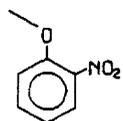
026 01



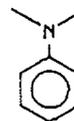
027 01



028 01



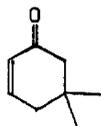
029 01



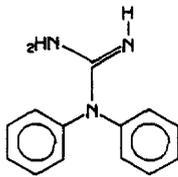
030 01



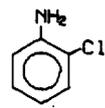
031 01



032 01



033 01



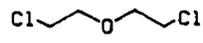
034 01



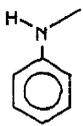
035 01



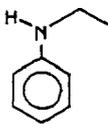
036 01



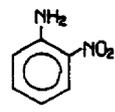
037 01



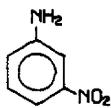
038 01



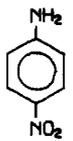
039 01



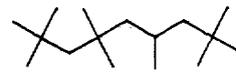
040 01



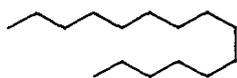
040 02



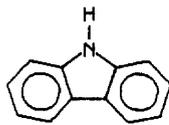
040 03



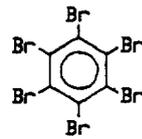
043 01



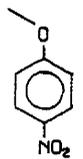
044 01



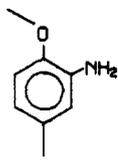
045 01



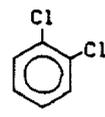
046 01



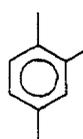
047 01



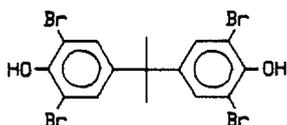
048 01



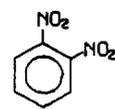
049 01



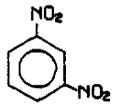
050 01



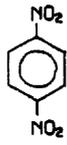
051 01



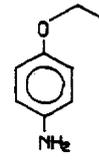
054 01



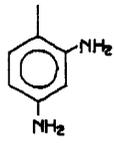
054 02



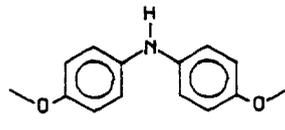
054 03



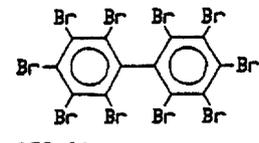
055 01



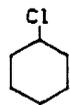
056 01



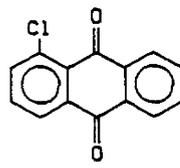
057 01



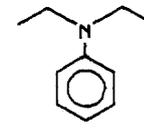
058 01



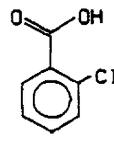
059 01



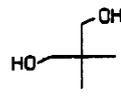
060 01



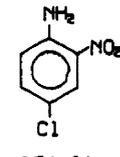
061 01



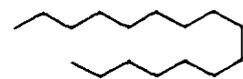
062 01



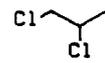
063 01



064 01



065 01



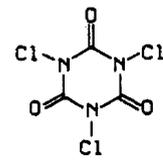
066 01



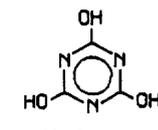
066 02



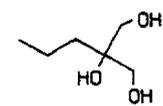
067 01



068 01



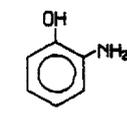
069 01



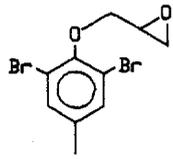
070 01



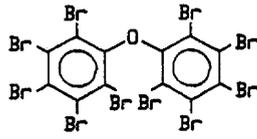
071 01



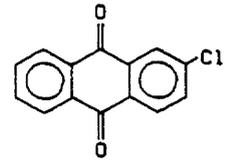
072 01



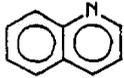
073 01



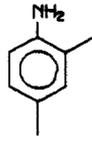
074 01



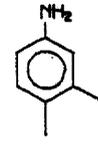
076 01



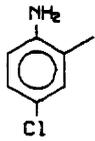
077 01



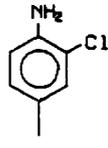
078 01



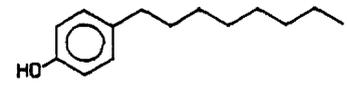
079 01



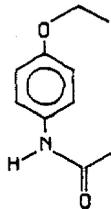
080 01



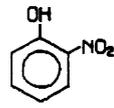
080 02



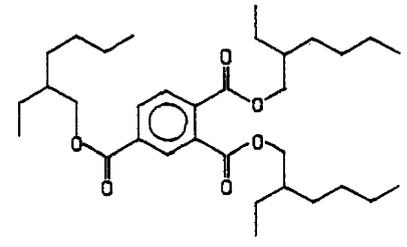
081 01



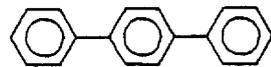
082 01



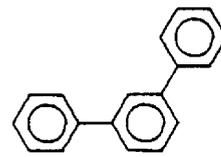
083 01



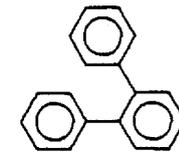
084 01



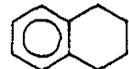
085 01



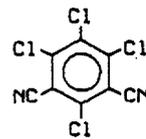
085 02



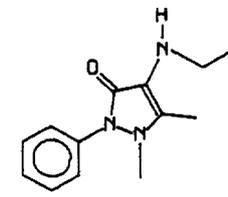
085 03



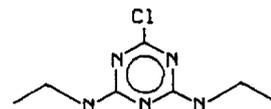
086 01



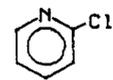
087 01



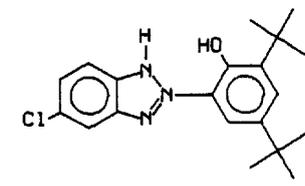
088 01



089 01



090 01



091 01

Anhang B

Datenbasismoleküle zu :

"3.2. Aromastoffe"

## Geschmacksstoffe

001	ACETALDEHYD DIETHYLACETAL	062	CITRONELLYLFORMIAT
002	ETHANAL	063	CITRONENSAEURE
003	ACETOIN	064	CUMARIN
004	ACETOPHENON	065	DECANOLID-1,4
005	2,3-PENTANDION	066	DECANOLID-1,5
006	ACONITSAEURE	067	N-DECANAL
007	ADIPINSAEURE	068	N-DECANOL
008	ETHYL-2,3-EPOXY-3-PHENYL-BUTANOAT	069	2,3-BUTANDION
009	ALLYLHEXANOAT	070	DIETHYLMALONAT
010	ALLYL-ISO-THIOCYANAT	071	DIALLYLDISULFID
011	AMEISENSAEURE	072	DIALLYLSULFID
012	1-PENTANOL	073	DIHYDROCARVEOL
013	3-METHYL-1-BUTANOL	074	DIHYDROCARVON
014	3-METHYL-1-BUTYL-ACETAT	075	DIHYDROCUMARIN
015	N-PENTYL-N-BUTANOAT	076	2-METHYLAMINO-METHYLBENZOAT
016	ISO-PENTYL-N-BUTANOAT	077	DIHYDROCITRONELLICL
017	N-PENTYL-HEXANOAT	078	DIMETHYLSULFID
018	ISO-PENTYL-N-HEXANOAT	079	5-HEPTYLBUTANOLID
019	ISO-PENTYL-3-PHENYL-PROPENOAT	080	ETHYLACETAT
020	N-PENTYLFORMIAT	081	ESSIGSAEURE
021	ISO-PENTYLFORMIAT	082	3-(4-METHOXYPHENYL)-1-PROPEN
022	ISO-PENTYL-ISO-PENTANOAT	083	ETHANOL
023	ISO-PENTYL-DCDECANOAT	084	ETHYLPROPENOAT
024	ISO-PENTYL-PROPANOAT	085	3-HEPTANON
025	ANETHOL	086	ETHYLBUTANOAT
026	ANISOL	087	ETHYL-3-PHENYL-PROPENOAT
027	4-METHOXYBENZALDEHYD	088	ETHYL-2-BUTENOAT
028	4-METHOXYBENZYLALKOHOL	089	ETHYLHEXANOAT
029	4-METHOXYBENZYLFORMIAT	090	ETHYLFORMIAT
030	APPELSAEURE	091	ETHYLHEPTANOAT
031	BENZALDEHYD	092	ETHYLDECANOAT
032	BENZYLACETAT	093	ETHYLOCTANOAT
033	BENZYLALKOHOL	094	ETHYL-ISO-PENTANOAT
034	BENZYLBENZOAT	095	ETHYLLACTAT
035	BENZYL-3-PHENYL-PROPENOAT	096	ETHYLDODECANOAT
036	BENZYLPHENYLACETAT	097	ETHYL-2-METHYL-BUTANOAT
037	BENZYLPROPANOAT	098	ETHYLPROPANOAT
038	BENZYL SALICYLAT	099	ETHYLSALICYLAT
039	N-BUTANSAEURE	100	ETHYLVANILLIN
040	ISO-BUTANSAEURE	101	1,8-CINEOL
041	N-BUTYLACETAT	102	EUGENOL
042	N-BUTANOL	103	ISO-EUGENOL
043	ISO-BUTANOL	104	EUGENOLACETAT
044	N-BUTANAL	105	FUMARSAEURE
045	ISO-BUTANAL	106	FURFUROL
046	N-BUTYL-BUTANOAT	107	2-FURFURYL-2-PROPANON
047	BUTYL-MILCHSAEURE-BUTYRAT	108	GERANIOL
048	ISO-BUTYL-BUTANOAT	109	GERANYLACETAT
049	ISO-BUTYL-PROPIONAT	110	GERANYLBUTYRAT
050	N-HEXANSAEURE	111	GERANYLFORMIAT
051	2-METHYL-5-ISOPROPYLPHENOL	112	HELIOTROPIN
052	CARVEOL	113	HEPTANAL
053	CARVON	114	HEPTANOL
054	CARVON	115	HEXANOLID-1,4
055	BETA-CARYOPHYLLEN	116	HEXANAL
056	3-PHENYL-PROPENYL-ACETAT	117	2-HEXENAL
057	CITRAL	118	2-HEXENOL
058	CITRONELLAL	119	3-HEXENOL
059	CITRONELLOL	120	HEXYLACETAT
060	CITRONELLYLACETAT	121	HEXANOL
061	CITRONELLYLBUTYRAT	122	HEXYLHEXANOAT
		123	HEXYLOCTANOAT
		124	4-(4-HYDROXYPHENYL)-BUTAN-2-ON
		125	3-PHENYL-1-PROPANOL
		126	3-PHENYL-1-PROPANAL
		127	3-PHENYL-1-PROPYLACETAT

128 3-PHENYL-1-PROPIONSAEURE  
129 INDOL  
130 ALPHA-IRON  
131 DODECANOL  
132 DODECANAL  
133 D-LIMONEN  
134 LINALOOL  
135 LINALYLACETAT  
136 LINALYLISOBUTYRAT  
137 LINALYLPROPIONAT  
138 3-HYDROXY-2-METHYL-4-PYRON  
139 L-MENTHOL  
140 2-HEPTANON  
141 METHYLACETAT  
142 P-METHYLACETOPHENON  
143 4-METHOXY-1-METHYLBENZOL  
144 METHYL-2-AMINOENZOAT  
145 METHYLBENZOAT  
146 METHYLHEXANOAT  
147 METHYLOCTANOAT  
148 METHYL-3-PHENYLPROPENOAT  
149 2-HYDROXY-3-METHYL-CYCLOPENT-2-EN-1-ON  
150 2-BUTANON  
151 METHYLEUGENOL  
153 METHYLISOBUTYRAT  
154 METHYL-4-METHYLPENTANOAT  
155 METHYL-3-METHYLTHIOPROPANOAT  
156 METHYL-2-METHYLBUTYRAT  
157 2-UNDECANON  
158 2-PENTANON  
159 METHYLPROPANOAT  
160 METHYLSALICYLAT  
161 METHYL-3-METHYL-BUTANOAT  
162 METHYLPENTANOAT  
163 MILCHSAEURE  
164 NEROLIDOL  
165 2,6-NONADIEN-1-OL  
166 NONALID-1,4  
167 N-NONANAL  
168 N-NONYLACETAT  
169 N-NONANOL-1  
170 N-NONYLOCTANOAT  
171 OCTANOLID-1,4  
172 N-OCTANAL  
173 N-OCTANOL  
174 N-OCTYLACETAT  
175 N-OCTYL-3-METHYLBUTANOAT  
176 N-OCTYLOCTANOAT  
177 NONANSAEURE  
178 ALPHA-PHELLANDREN  
179 PHENOXYESSIGSAEURE  
180 1-OXO-2-PHENYLETHAN  
181 2-PHENYLETHYLACETAT  
182 2-PHENYLETHANOL  
183 2-PHENYLETHYL-2-METHYLPROPANOAT  
184 2-PHENYLETHYL-2-METHYLBUTANOAT  
185 PHENYLESSIGSAEURE  
186 PIPERIDIN  
187 PIPERIN  
188 5-PROPENYL-2-ETHOXYPHENOL  
189 PROPANAL  
190 1-PROPANOL  
191 2-PROPANOL  
192 ISO-PROPYL-N-BUTANOAT  
193 ISO-PULEGOL  
194 RHODINOL (LAevo-CITRONELLOL)  
195 TERPINEOL  
196 TERPINYLACETAT  
197 TERPINYLPROPIONAT  
198 GLYCERYLTRIACETAT  
199 UNDECANOLID-1,4  
200 UNDECANAL  
201 PENTANSAEURE  
202 3-METHYLBUTANSAEURE  
203 VANILLIN  
204 3-PHENYLPROPENAL  
205 3-PHENYLPROPENOL  
206 3-PHENYLPROPENSAEURE

Geschmacksqualität "süß"

004 ACETOPHENON  
 005 2, 3-PENTANDION  
 008 ETHYL-2,3-EPOXY-3-PHENYL-BUTANOAT  
 016 ISO-PENTYL-N-BUTANCAI  
 019 ISO-PENTYL-3-PHENYL-PROPENOAT  
 021 ISO-PENTYLFORMIAT  
 024 ISO-PENTYL-PRCPANOAT  
 025 ANETHOL  
 026 ANISOL  
 027 4-METHOXYBENZALDEHYD  
 028 4-METHOXYBENZYLALKOHCL  
 029 4-METHOXYBENZYLFORMIAT  
 035 BENZYL-3-PHENYL-PROPENOAT  
 036 BENZYLPHENYLACETAT  
 037 BENZYLPROPANOAT  
 043 ISO-BUTANOL  
 046 N-BUTYL-BUTANOAT  
 049 ISO-BUTYL-PRCPIONAT  
 053 CARVON  
 054 CARVON  
 056 3-PHENYL-PROPENYL-ACETAT  
 061 CITRONELLYLBUTYRAT  
 064 CUMARIN  
 066 DECANOLID-1,5  
 067 N-DECANAL  
 070 DIETHYLMALONAT  
 075 DIHYDROCUMARIN  
 076 2-METHYLAMINO-METHYLBENZOAT  
 080 ETHYLACETAT  
 082 3-(4-METHOXYPHENYL)-1-PROPEN  
 086 ETHYLBUTANOAT  
 087 ETHYL-3-PHENYL-PROPENOAT  
 090 ETHYLFORMIAT  
 092 ETHYLDECANOAT  
 093 ETHYLOCTANOAT  
 094 ETHYL-ISO-PENTANOAT  
 099 ETHYLSALICYLAT  
 100 ETHYLVANILLIN  
 103 ISO-EUGENOL  
 104 EUGENOLACETAT  
 109 GERANYLACETAT  
 110 GERANYLBUTYRAT  
 112 HELIOTROPIN  
 114 HEPTANOL  
 115 HEXANOLID-1,4  
 120 HEXYLACETAT  
 124 4-(4-HYDROXYPHENYL)-BUTAN-2-ON  
 125 3-PHENYL-1-PRCPANCL  
 126 3-PHENYL-1-PROPANAL  
 127 3-PHENYL-1-PRCPYLACETAT  
 129 3-PHENYL-1-PROPIONSÄURE  
 130 ALPHA-IRON  
 135 LINALYLACETAT  
 136 LINALYLISOBUTYRAT  
 137 LINALYLPROPIONAT  
 141 METHYLACETAT  
 142 P-METHYLACETOPHENON

144 METHYL-2-AMINOBENZOAT  
 145 METHYLBENZOAT  
 146 METHYLHEXANOAT  
 149 2-HYDROXY-3-METHYL-CYCLOPENT-2-EN-1-ON  
 153 METHYLISOBUTYRAT  
 154 METHYL-4-METHYLPENTANOAT  
 155 METHYL-3-METHYLTHIOPROPANOAT  
 156 METHYL-2-METHYLBUTYRAT  
 159 METHYLPROPANOAT  
 160 METHYLSALICYLAT  
 162 METHYLPENTANOAT  
 171 OCTANOLID-1,4  
 172 N-OCTANAL  
 173 N-OCTANOL  
 176 N-OCTYLOCTANOAT  
 181 2-PHENYLETHYLACETAT  
 182 2-PHENYLETHANOL  
 185 PHENYLESSIGSÄURE  
 186 PIPERIDIN  
 188 5-PROPENYL-2-ETHOXYPHENOL  
 192 ISO-PROPYL-N-BUTANOAT  
 194 RHODINOL (LAEVO-CITRONELLOL)  
 195 TERPINEOL  
 197 TERPINYLPROPIONAT  
 198 GLYCERYLTRIACETAT  
 199 UNDECANOLID-1,4  
 203 VANILLIN  
 204 3-PHENYLPROPENAL  
 205 3-PHENYLPROPENOL

Geschmacksqualität "sauer"

006 ACONITSAURE  
 007 ADIPINSÄURE  
 030 AEPFELSÄURE  
 039 N-BUTANSÄURE  
 040 ISO-BUTANSÄURE  
 047 BUTYL-MILCHSÄURE-BUTYRAT  
 063 CITRONENSÄURE  
 081 ESSIGSÄURE  
 088 ETHYL-2-BUTENAT  
 105 FUMARSÄURE  
 163 MILCHSÄURE  
 179 PHENOXYESSIGSÄURE  
 191 2-PROPANOL  
 202 3-METHYLBUTANSÄURE

Geschmacksqualität "karamel"

030 AEPFELSÄURE  
 088 ETHYL-2-BUTENAT  
 106 FURFUBOL  
 139 3-HYDROXY-3-METHYL-4-PYRON  
 149 2-HYDROXY-3-METHYL-CYCLOPENT-2-EN-1-ON

Alle graphischen Darstellungen in dieser Arbeit  
wurden mit Hilfe des Systems "GRECO"  
entworfen und gezeichnet.

