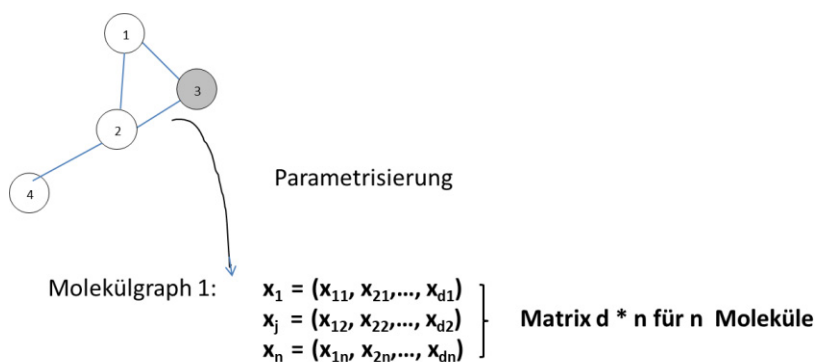


Das Buch „Chemische Karzinogenese von A bis Z“ [1] beschreibt die Simulationstechniken sowie die Prognoseergebnisse bei der Klassifizierung mutagen wirkender Verbindungen. Basismaterial bildeten die Monograph-Publikationen der **IARC**)<sup>1</sup>. Bei einer Klasse wichtiger organischer Zwischenprodukte, den synthetischen Epoxiden, gelang jedoch eine Einteilung in mutagen und nicht mutagen wirkende Strukturen nicht. Der Datensatz ließ sich nicht in Klassen partitionieren. D. h. alle damals produzierten primären Epoxide (Abb. 2-1, Teil A) zeigten ein mutagenes Eigenschaftsbild. Doch wie erfolgt eine solche computergestützte Klassifizierung, die eine Menge vorgegebener Valenzstrichformeln in wirkende und nicht wirkende Eigenschaften einteilt?

Zur Simulationen von Struktur-Eigenschaftsbeziehungen sind üblicherweise zwei mathematische Operationen erforderlich. Im ersten Schritt werden die Valenzstrichformeln parametrisiert. Man erhält für jede vorgegebene Struktur in Form ihres Molekülgraphen einen d-dimensionalen **Objektvektor** aus elektronischen und/oder topologischen Parametern. Die Mächtigkeit von d ist zunächst frei wählbar, sollte im Verlauf der Simulationen aber bewusst auf ein Minimum reduziert werden. Zur **Parametrisierung** dienen häufig Verfahren, auf deren Basiswerte man problemlos zugreifen kann. Hierzu gehört z. B. die Bildung elektrotopologischer Indices gemäß Tab. 2-1 nach Kier und Hall [11]. Andere typische Parameter wären die **Graphenmetrik** [19, 71], die Trägheitsmomente eines Moleküls oder spektroskopische Parameter [18]. Alle n zu untersuchenden Strukturen zusammen bilden dann eine d \* n große Wertematrix (Abb. 2-1). Sie wird im zweiten Simulationsschritt partitioniert, für die Mutagenitätsanalyse in die Klasse der mutagenen und in die Klasse der nicht mutagenen Strukturen.

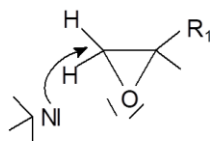


**Abb. 2-1: Molekülgraph, Bildung der Objektvektoren und die Vereinigung der Vektoren zur Matrix**

<sup>1</sup> Die Ersterwähnung unbekannter Begriffe oder Abkürzungen wird im Text fett gedruckt angezeigt. Die Begriffserklärungen befinden sich im Glossar Kap. 13.

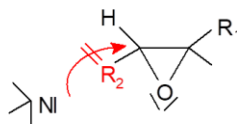
## 2.1 Parametrisierung der Molekülabbildung

Prinzipiell könnte jedes Skelettatom eines Molekülgraphen ein Merkmal im Vektor darstellen. Doch es ist keinesfalls sinnvoll, mit so vielen Merkmalen zu arbeiten. Man sucht vielmehr gerade beim Simulieren mit **Neuronalen Netzen** stets nach jener Abbildung, die mit minimaler Zahl von Merkmalen und in kurzer Zeit Ergebnisse fehlerfrei generieren kann [53].



primäres Epoxid

mit nucleophilem Angriff auf DNA



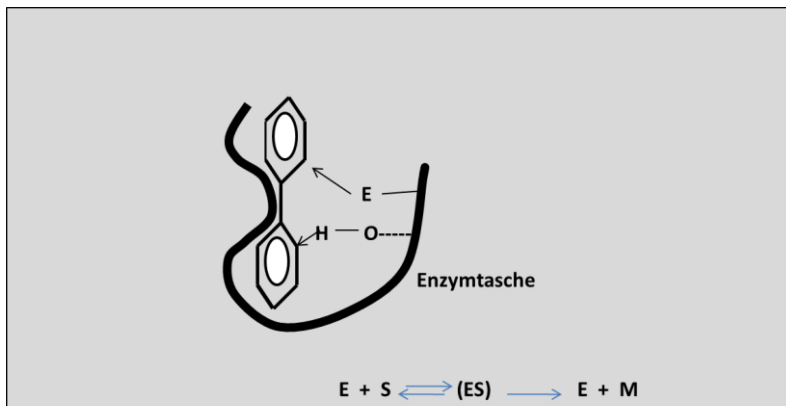
sekundäres Epoxid

kein nucleophiler Andockvorgang  
sterisch an DNA möglich

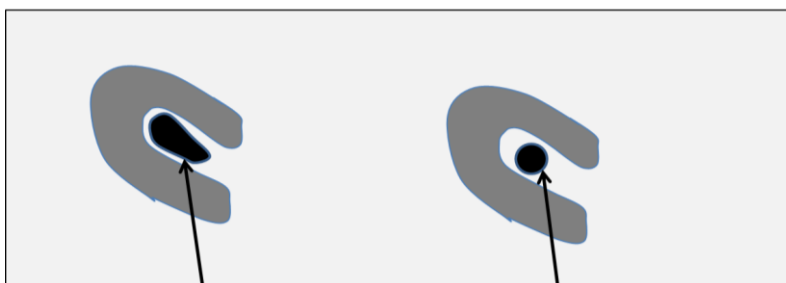
**Abb. 2-2: Alkylierungsmodell für Epoxide, links: primäre Epoxidstrukturen mit erlaubtem Andockvorgang, rechts: sekundäres Epoxid mit verbotener Andockung**

Einer solchen iterativen Merkmalssuche bedarf es im speziellen Falle der synthetischen Epoxide jedoch nicht. Aus der Organischen Chemie kennt man die Reaktionsorte der Epoxide. Es sind die dem O-Atom im Oxyranring benachbarten C-Atome [12]. Sie sollten auch das Reaktionszentrum für den Andockvorgang an eine DNA- oder RNA-Base in der lebenden Zelle für den Initialschritt der Mutagenese bilden (Abb. 2-2). Natürlich besteht zwischen In-vitro- und In-vivo-Reaktionen ein gravierender Unterschied. Beim Eindringen in die Zelle müssen die Noxen u. a. geometrische Sperren überwinden oder laufen Gefahr, durch Metabolisierungsreaktionen detoxifiziert zu werden. Das sind Hindernisse, die es in der klassischen Chemie so nicht gibt. Diese, dem eigentlichen Andockvorgang vorgelagerten Reaktionen lassen sich u. a. durch Parameter der Graphenmetrik oder durch Trägheitsmomente abbilden.

Die Wirksamkeit der Trägheitsmomente auf die Klassifizierung ergibt sich u. a. aus der Besonderheit der Enzymkatalyse (Abb. 2-3, oberer Teil). Sie erfolgt in sogenannten Enzymtaschen und kommt nur dann zustande, wenn die Geometrie des Targets mit der Geometrie der Tasche übereinstimmen. Und diese Ähnlichkeiten in den Geometrien von Tasche und Target lassen sich durch die Trägheitsmomente der Targetmoleküle abbilden (Abb. 2-3, unterer Teil). So kann z. B. ein kugelförmiges Molekül wesentlich schlechter in der Enzymtasche fixiert werden als ein ellipsoidförmiges. Bei kugelförmiger Geometrie beträgt das Verhältnis von maximalem zu minimalem Trägheitsmoment etwa 1, bei Ellipsoiden dagegen  $> 1$ .



E Enzym S Substrat M Metabolit



Trägheitsmomente

$I_{\max} / I_{\min} \gg 1$

$I_{\max} / I_{\min} = 1$

**Abb. 2-3: Modell einer Enzymkatalyse (oben) und Form seiner mathematischen Abbildung durch Trägheitsmomente (unten)**

Im Falle der Epoxide handelt es sich z. T. um sehr kleine Moleküle, sodass sterische Hindernisse kaum von Relevanz sind. Deshalb werden zunächst aus didaktischen Gründen zur Abbildung der Epoxide nur die beiden, dem Oxyransauerstoff benachbarten C-Atome als Merkmale vorgegeben und durch elektrotopologische Indices gemäß Tab. 2-1 parametrisiert. Bei den elektrotopologischen Indices handelt es sich um sogenannte **Konnektivitäten**, die durch elektronische Parameter korrigiert sind. Diese Grundwerte der Skelettatome werden durch eine weitere Korrektur ergänzt, die das topologische Umfeld des jeweiligen Atoms abbildet. Damit berechnen sich die elektrotopologischen Indices zu:

$$S_i = I_i + \Delta I_i$$

Gl. 2-1

Die inhärenten Werte sind für die in organischen Verbindungen typischen Skelettatome in Tab. 2-1 aufgelistet. Das topologische Umfeld ergibt sich zu:

$$\Delta I_i = \sum_{j=1}^n (I_i - I_j) / r_{ij}^2$$

Gl. 2-2

mit: n = Gesamtzahl der Skelettatome j im Umfeld des Knotens i,  
I<sub>i</sub> , I<sub>j</sub> inhärente Werte der Skelettatome gemäß Tab. 2-1,  
r<sub>ij</sub> ihre topologische **Distanz** zueinander.

Tab. 2-1: Elektrotopologische Indices [11]

Atom/Gruppe	I-Wert	Atom/Gruppe	I-Wert	Atom/Gruppe	I-Wert
>C<	1,25	=CH <sub>2</sub> , =N-	3	-F	8
>CH-	1,33	-O-	3,5	-Cl	4,11
-CH <sub>2</sub> -	1,5	≡CH, NH <sub>2</sub>	4	-Br	2,75
>C=	1,667	=NH	5	-I	2,12
-CH <sub>3</sub> , =CH-, >N-	2	≡N, -OH	6	=S	3,667
≡C-, -NH-	2,5	=O	7	-SH	3,222
				-S-	1,833

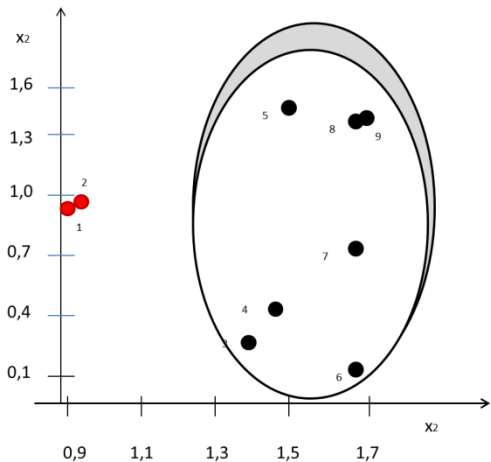


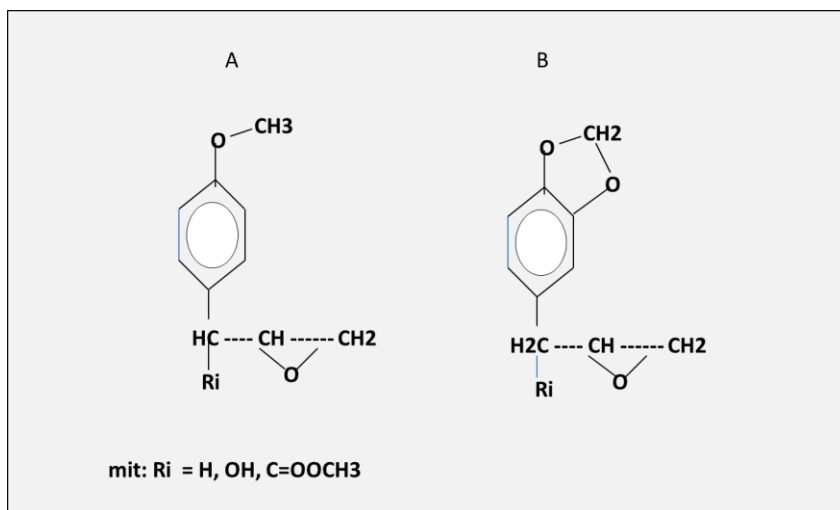
Abb. 2-4: 2-dimensionale Darstellung der Epoxide (Notation: 1 Cyclohexenoxid, 2 C 18:1 Fettsäureepoxid, 3 PO, 4 Styrenoxid, 5 EO, 6 Epichlorhydrin, 7 1,2-Butenoxid, 8 Saftrolderivat, 9 Estragolderivat)

In Abb. 2-4 sind nun die elektrotopologischen Parameter für die beiden C-Atome gegeneinander aufgetragen. Es fällt sofort auf, dass ein C-Atom primärer Epoxide, also die CH<sub>2</sub>-Struktur, in allen mutagen wirkenden Epoxidstrukturen eine fast gleiche Abbildung erfährt und nur das C-Atom der R-CH-Gruppe in Abhängigkeit des Umfeldes stärker veränderliche Parameter aufweist. Diese Aussage trifft sowohl für die mutagen wirken-

den synthetischen Epoxide, also z. B. EO, PO oder Epichlorhydrin (Abb. 2-4, Notation 4–6) als auch für die im Nagetierexperiment hepatokarzinogen [14] wirkenden nativen **Safrol-** (8) bzw. **Estragolderivate** (9) zu.

## 2.2 Klassifizierungsverfahren

Die eingangs erwähnte Wertematrix unterwirft man bei Merkmalssätzen mit  $d \geq 3$  Merkmalen im zweiten Simulationsschritt einem mathematischen Klassifizierungsverfahren. Das kann u. a. eine **Hauptkomponentenanalyse**, für Mehrklassenprobleme ein **Clusterverfahren** oder eine Simulation mit einem N N sein. Eine Klassifizierung mit einem N N soll modellhaft am Beispiel nativer Epoxide demonstriert werden. Als Ausgangsdaten dienten sowohl die in Abb. 2-5 aufgeführten Estragol- (A) als auch Safrol- (B). In beiden Substanzklassen treten Vertreter mit primären Oxyranstrukturen auf (Abb. 2-5). Die nativen Epoxide der genannten Substanzklassen weisen im Nagetierexperiment ein hohes hepatokarzinogenes Potenzial auf [1].



**Abb. 2-5: Strukturen nativer, mutagen wirkender Epoxide [14], links: Estragolderivate, rechts: Safraolderivate**

Das Ergebnis der Klassifizierung des oben genannten Datensatzes mit einem N N ist in Tab. 2-2 dargestellt. Das Netz wurde mit elektrotopologischen Indizes (Spalte 2), mit empirischen Parametern bestehend aus dem Trägheitsmoment, dem **DHMO-Wert** und Konnektivitäten (Spalte 3) sowie mit Parametern der Graphenmetrik (Spalte 4) parametrisiert.

Das N N liefert für die ersten beiden Datensätze fehlerfreie Klassifizierungen in kurzen Simulationszeiten (Zeilen 3 und 4). Der Netzaufbau des N N besteht dabei aus einer variablen Inputschicht entsprechend der Anzahl der vorgegebenen Merkmale, 8 bzw. 10 Knoten in der Zwischenschicht und einem Outputknoten für die mutagenen Strukturen.

Der Datensatz mit Werten der Graphenmetrik weist dagegen einen Fehler auf (Zeile 4). Allein mit Molekülgraphen ist der Wirkunterschied von mutagen wirkendem Estragol und nicht mutagenem Anethol nicht abzubilden. Beide Strukturen sind Strukturisomere und besitzen die gleiche Graphenmetrik.

**Tab. 2-2: Mutagenitätsprognose für epoxidierte und nicht epoxidierte Safrolerivate durch Simulationen mit einem N N**

Parameter 1	elektrotopologische Indizes [11] 2	Trägheitsmoment DHMO, Konnektivitäten [1] 3	Metrik von Molekülgraphen [19] 4
<sup>1</sup> Zahl der Merkmale	3	5	5
<sup>2</sup> Netzaufbau) <sup>1</sup>	3/8/1	5/8/1	5/10/1
<sup>3</sup> Simulationszeit in min	8	> 1	2
<sup>4</sup> Fehler	0	0	1

)<sup>1</sup> Knotenzahl in der Input-, Hidden- und Outputsicht

Allgemein erwartet man, dass nach Ausführung des Klassifizierungsverfahrens sich Cluster mit unterschiedlicher Wirkung bilden, speziell bei der Mutagenitätsprognose zwei Klassen: die der mutagen wirkenden und die der nicht mutagen wirkenden Strukturen (Abb. 2-6). Mitunter überschneiden sich jedoch beide Klassen, d. h. für die Objekte der Schnittmenge ist dann leider keine Aussage über ihre Wirkung zu treffen. Im Falle der primären Epoxide ergab sich, wie bereits angedeutet, jedoch ein völlig anderes Bild. Alle Objekte gehörten nur einer einzigen Klasse an.

**Tab. 2-3: Verfahren zur Klassifizierung von Daten**

Verfahren	Zahl der Merkmale	Anzahl der Klassen
<sup>1</sup> visuell	2 maximal 3	2
<sup>2</sup> Neuronales Netz [22]	2 bis 99	≥ 2
<sup>3</sup> Hauptkomponentenanalyse [1, 21]	> 3	≥ 2
<sup>4</sup> Clusteranalyse [20]	beliebig	beliebig

In einer gemeinsamen 2-dimensionalen Abbildung aus elektrotopologischen Indices gemäß (Gl. 2-2) unterscheiden sich die Epoxidstrukturen von primären und sekundären Epoxiden. Sekundäre Epoxide, z. B. von nativen Fettsäureestern, aber auch ein Derivat vom Cyclohexenepoxid, heben sich deutlich vom Bereich der primären Epoxide ab. Ein Syntheseprodukt aus zwei Cyclohexenoxidstrukturen war das erste synthetisch hergestellte, nicht mutagen wirkende Epoxid, das mit der Handelsbezeichnung „Epoxy-221“ [13] im Jahre 1996 auf den Markt kam. Aus Abb. 2-6 ist ferner erkenntlich, dass für die 2-dimensionale Darstellung des Epoxid-Datensatzes das menschliche Auge als Klassifikator ausreichend ist. Mathematische Darstellungen zur Klassifizierung können damit entfallen.



**Abb. 2-6: Klassifizierung von Strukturen zur Eigenschaftsprognose (2-Klassenproblem, Abbildung der Objekte durch zwei Merkmale  $X_1$  und  $X_2$ )**

Die Suche nach nicht mutagen wirkenden nativen Epoxiden führte auf die Struktur der Vernolsäure (Tab. 13-4). Sie ist bis zu 80 % im Vernoniaöl enthalten, was u. a. zu 42 % Bestandteil der Ölpflanze *Vernonia galamensis* ist. Dieses Vernoniaöl wird u. a. seit mehreren Jahrzehnten in der US-Kosmetikindustrie als Rohstoff eingesetzt. Negative Nebenwirkungen sind bisher nicht bekannt geworden. Allein – es war bis Ende des 20. Jahrhunderts nicht gelungen, die wildwachsende Steppenpflanze für eine maschinelle Ernte zu kultivieren; der Ertrag der Wildform für chemische Großsynthesen jedoch viel zu gering. Aber mit der Kenntnis einer nicht mutagen wirkenden, nativen Epoxidstruktur war das Bauprinzip zu ihrer Herstellung gefunden. Man musste ungesättigte Fettsäuren epoxidieren!

Native Epoxide und Epoxidharze - ein Beitrag zur  
ökologischen Chemie

Adler, B.

2017, VII, 147 S. 49 Abb. Book + eBook., Hardcover

ISBN: 978-3-662-55613-9