

QSAR-Ansätze für die toxikologische Bewertung von Industriechemikalien

E. Jacob (elard.jacob@basf-ag.de), Ludwigshafen/Rhein

Historie

Die hier betrachteten (Q)SAR Modelle und ihr Einsatz in der toxikologischen Bewertung von Industriechemikalien unterscheiden sich wesentlich von früheren Entwicklungen dieses Namens, die einem anderen Zweck dienen. Sie waren auf die Ableitung pharmakologischer und physikalisch-chemischer Eigenschaften von Entwicklungssubstanzen in der pharmazeutischen Forschung und im Pflanzenschutz ausgerichtet mit dem Ziel der Vorhersage erhöhter Wirksamkeit strukturverwandter Substanzen mit gleichem zellulären Wirkort.

Insofern fokussierten sie sich fast ausschließlich auf "Congeneren"-Betrachtungen, bei denen in einer Familie strukturverwandter chemischer Substanzen der Einfluss verschiedener "Themen"-Variationen auf eine höhere Wirksamkeit, ggf. in Verbindung mit reduzierten Nebenwirkungen und für die geplante Anwendung günstigeren physikalisch-chemischen Eigenschaften betrachtet bzw. vorausgesagt wurde.

Insofern sind derartige (Q)SAR-Methoden, die teilweise auch den Charakter von "Molecular Modelling" (also lokale Variationen von Substituenten und homologen strukturellen Änderungen im Rückgrat des Moleküls) hatten, in der pharmazeutischen und chemischen Industrie bewährte und gut etablierte Werkzeuge.

Dies gilt insbesondere für die zur direkten Messung alternative Methode der Berechnung physikalisch-chemischer Eigenschaften, die zur Ableitung ökotoxikologischen Verhaltens (u.a. Verteilung in unterschiedlichen Umweltkompartimenten) heran gezogen wird.

Bei den neuen Bewertungsmodellen, die als (Q)SAR-Programme in der Toxikologie eingesetzt werden, gibt es jedoch bedeutende Unterschiede zu den früheren Modellen. Sie liegen vor allem in der sehr viel größeren Strukturvielfalt der Trainingsdaten und der zu bewertenden Substanzen einschließlich der Bewertung völlig unbekannter Moleküle. Der Congenerenansatz ist hier weitgehend in den Hintergrund getreten.

Begriffsdefinitionen

Es erscheint hilfreich, an dieser Stelle vor einer weiteren Diskussion der verschiedenen Entwicklungen und Einsatzmöglichkeiten, eine kurze Begriffsdefinition einzuführen.

Danach soll mit "**QSAR**" eine quantitative Beziehung zwischen einer pharmakologischen, chemischen, biologischen, physikalischen (z.B. Siedepunkt) Wirkung eines Moleküls mit seiner chemischen Struktur verstanden werden. Teilweise findet man in der Literatur auch die Bezeichnung QSPR, als Akronym für **Quantitative Structure Property Relationship**, hier beschränkt man sich darauf, die Beziehung zwischen den physikalischen und chemischen Eigenschaften eines Moleküls und seiner Struktur darzustellen.

Erste Arbeiten in diesem Gebiet sind vermutlich bereits im Jahre 1842 durch Kopp /1/ entstanden, welcher eine lineare Beziehung zwischen den Siedepunkten von Alkanen mit deren Kettenlänge herstellte.

Mit "**SAR**" (Struktur-Wirkungsbeziehung) ist die quantitative Beziehung in den Hintergrund getreten. Mit diesem Akronym sollen die qualitativen Beziehungen zwischen einer chemischen Struktur und dem Potential für eine bestimmte chemische oder biologische Eigenschaft bezeichnet werden. Beispiele im Bereich Säugertoxizität könnten dafür sein: Mutagenität, Sensibilisierung und Reizwirkung.

SARs geben deshalb i.A. eine Ja/Nein Aussage, während die Skalierung des Effekts häufig nicht vertretbar ist, da ein Vergleich zwischen Wirksubstanzen unterschiedlicher Potenz nicht massenproportional abgebildet werden kann. Exemplarisch für solche Effekte soll hier mutagene Potenz genannt werden, jedoch ist dieses Charakteristikum auch für andere Endpunkte gültig. Eine besondere Stellung zwischen diesen beiden Polen kann eine Bewertung der quantitativen Verteilung von aktiven Metaboliten einnehmen, wie sie z.B. in dem Programm TIMES von LMC implementiert ist (siehe unten).

Da auch in diesem Beitrag häufig ohne direkte Diskriminierung zwischen diesen beiden Vorhersagemethoden gesprochen wird, soll als gemeinsamer Begriff **(Q)SAR** verwendet werden, der sowohl den qualitativen wie den quantitativen Aspekt berücksichtigt soll.

Schließlich soll der Begriff des Expertensystems kurz angesprochen werden.

Unter "**Expertensystem**" wird ein computer-gestütztes wissensbasiertes System zur Vorhersage der stofflichen Eigenschaften aus der chemischen Struktur verstanden, das auf einem empirischen Regelwerk ("Rules") aufbaut. Ausgehend von diesen Regeln wird die Information zu einer spezifischen Problemklasse analysiert und auf Grundlage der mathematischen Analyse eine Reaktion (hier eine Bewertung) empfohlen. Damit stellt ein solches Regelwerk u.a. eine strukturell effiziente Organisation bekannten (öko-)toxikologischen Wissens dar, das als die Grundlage zu einem "Expert Judgement" dienen kann.

Die Berechnung physikalisch-chemischer und struktureller Parameter als Beitrag zur Einschätzung der (biologischen) Wirksamkeit tritt dabei eher in den Hintergrund.

(Q)SAR-Methoden und Expertensysteme in der Öko- und Säuger-Toxikologie

In den letzten fünfzehn Jahren wurden (Q)SAR und Expertensystem vermehrt auf (öko-)toxikologische Endpunkte wie Bioabbau, Bioakkumulation, Akute Fischtoxizität, Gentoxizität, Hautsensibilisierung und Kanzerogenität angewendet.

Insbesondere die Vorhersagen zu Mutagenität im Ames-Test mit der verfügbaren breiten Datenbasis haben bei der Entwicklung von (Q)SAR-Systemen eine wichtige Rolle gespielt, wobei in den neueren Entwicklungen zunehmend auch die Rolle metabolischer (De-)Aktivierung in Form der *in-vitro* Simulation organspezifischer Stoffwechselleistungen berücksichtigt wurde.

Im Gegensatz zur Mutagenität erscheinen (Q)SAR-Vorhersagen zu anderen, komplexeren Endpunkten eher unterentwickelt, wobei häufig eine unzureichende Datenlage aus experimentellen Prüfungen der begrenzende Faktor ist, jedoch auch mangelnde Erkenntnisse zu den pharmakokinetischen "ADME"-Eigenschaften (ADME: Absorption, Distribution, Metabolism, Excretion).

Einschränkungen erfahren diese Systeme weiterhin durch ihre fehlende Prädiktivität bei wirtschaftlich wichtigen Stoffklassen, z.B. Polymeren, anorganischen Verbindungen, Metallkomplex-Verbindungen und Fasern.

Für diese Klassen sind Analogieschlüsse in Verbindung mit "Expert-Judgement", ggf. in Verbindung mit strukturierten, behördlich anerkannten Modellen wie OncoLogic (US-EPA) und die Prüfung in *in-vitro* Surrogat Testsystemen (z.B. Reizwirkung, kanzerogenes Potential im Syrian Hamster Embryo (SHE) Zelltransformations-Assay) eine Alternative zur Prüfung im Tierversuch.

In der chemischen Industrie kommen den folgenden Eigenschaften eines (Q)SAR Systems wesentliche Bedeutung zu:

- Mechanistisches Verständnis des vorhergesagten Effekts
- Berücksichtigung des relevanten Metabolismus (z.B. Leber oder Haut)
- Akzeptable Prädiktivität auch bei wenig datenreichen Endpunkten
- Integrierbarkeit von eigenen Prüfdaten mit proprietärer Chemie
- Transparenz der Aussage zur Vorhersage-Domäne
- Vollständigkeit des vorhergesagten Endpunkts (z.B. bei Gentoxizität)

Neben der internen Nutzung als Priorisierungswerkzeug wurde der Umgang mit solchen Software-Tools auch deshalb relevant, um bei Einsatz durch Zulassungsbehörden eigene Erfahrungen und Detailkenntnisse bzgl. ihrer Aussagekraft u. -grenzen zu haben.

Im Rahmen der regulatorischen Akzeptanz von (Q)SAR-Methoden wurden 2002 vor dem ICCA-Setubal Workshop in Portugal von einer ECETOC /2/ Gruppe die zu der Zeit kommerziell verfügbaren Programme bewertet. Die Bewertung /3/ stand für den Workshop als Basisinformation zur Verfügung und wurde später als "Technical Report" veröffentlicht. Die weiteren Aktivitäten in der EU in Zusammenarbeit mit der OECD bauen auf den Ergebnissen dieses Workshops auf.

Die chemische Industrie verfolgt diese Entwicklungen mit großer Aufmerksamkeit, da die behördliche Akzeptanz, vor

allem wenn auch negative Vorhersagen als valide betrachtet werden, im Rahmen von REACH signifikante Erleichterungen bzgl. der Priorisierung experimenteller Prüfungen oder als Ersatz von ihnen bedeuten kann.

Verfügbare (Q)SAR-Modelle

Vorläufige Bewertung aus industrieller Sicht

Die im Rahmen einer ECETOC /2/ Experten Gruppe 2003 erstellte Bewertung /3/ soll hier aus der gegenwärtigen Sicht einer toxikologischen Fachabteilung differenziert dargestellt werden (Tabelle 1).

Physiko-chemische Eigenschaften		Umweltverhalten/ Ökotoxizität		Säugertoxizität	
Schmelzpunkt	gut	Photoabbau	eingeschränkt	Akute orale Toxizität	eingeschränkt
Siedepunkt	gut	Hydrolyse	eingeschränkt	Hautreizung	eingeschränkt
Dampfdruck	sehr gut	Bioabbau	gut	Augenreizung	eingeschränkt
Wasserlöslichkeit	sehr gut	Bioakkumulation	gut	Hautsensibilisierung	eingeschränkt
Verteilungskoeff.	sehr gut	Akute Toxizität für Fische	gut	Chronische Toxizität	sehr eingeschränkt
		Akute Toxizität für Invertebraten	eingeschränkt	Mutagenität <i>in vitro</i>	eingeschränkt bis gut
		Akute Toxizität für Algen	eingeschränkt	Karzinogenität	sehr eingeschränkt
		Chronische Toxizität	sehr eingeschränkt	Repro-Tox Entw.-Tox	sehr eingeschränkt

Tab. 1: Vorläufige Bewertung verfügbarer (Q)SARs bzgl. verschiedener Endpunkte

Potenzieller Nutzen von (Q)SARs aus Industrieperspektive

Die prinzipiellen Vorteile werden zunächst im Bereich der Vorauswahl von Stoffen im Prozess eines "Screenings" und der Identifizierung von möglichen Problemsubstanzen gesehen. Diese gezielte Auswahl kann als Hilfsinstrument bei den wachsenden Anforderungen der Gesetzgebung genutzt werden.

Insbesondere wird dadurch eine effektivere Nutzung der begrenzten experimentellen und Bewertungs-Ressourcen erwartet. Vor allem im Bereich von Tierversuchen wird eine

Reduzierung erhofft. Ob ein zumindest teilweiser Ersatz möglich ist, wird von der regulatorischen Akzeptanz der z. Zt. in Entwicklung befindlichen (Q)SAR-Werkzeuge abhängen.

Damit einher gehend sollte eine Beschleunigung von Priorisierungs- und Bewertungsverfahren sowie mehr Transparenz in Stoff- und Produktbewertungen möglich sein.

Jedoch werden die gegenwärtig verfügbaren (Q)SAR-Werkzeuge – speziell für Vorhersagen im Säugertoxizitäts-Bereich - auch kritisch gesehen. Während für die Vorhersage physiko-chemischer Endpunkte und akuter Effekte in der Ökotoxikologie eine gute Verlässlichkeit erkannt wird, erscheinen (Q)SARs im Säugertox-Bereich zwar relativ gut für Endpunkte mit validierten *in-vitro* Assays, können aber nicht die aufwendigen und deshalb teureren *in-vivo* Studien ersetzen.

Gründe dafür sind, dass für *in-vivo* Studien eine kritische Masse zum Aufbau eines verlässlichen Lerndatensatzes fehlt, der Endpunkt häufig stark durch ADME-Effekte beeinflusst wird und der zugrunde liegende Wirkmechanismus oft nur unzureichend verstanden ist.

Unter den derzeit laufenden Aktivitäten in Kooperation zwischen OECD und EU unter Leitung des ECB (European Chemical Bureau) wird die Weiterentwicklung der (Q)SARs auf Basis allgemein akzeptierter Kriterien ("Solna" Prinzipien in der Modifikation gemäß den Setubal-Vereinbarungen) vorangetrieben. Wesentliche Punkte dabei sind die Transparenz des Vorhersage-Algorithmus und des Lerndatensatzes ("Applicability Domain"). Die validierten Modelle müssen allgemein verfügbar sein, und es wird eine Koppelung der Aussage mehrerer unabhängiger Modelle im Sinn eines "Weight of Evidence" angestrebt.

Aus Sicht der Industrie ist für die komplexen *in-vivo* Endpunkte die Erstellung großer gemeinsamer Datenbasen unabdingbar, die allerdings nur mit einem ausreichenden Datenschutz realistisch erscheint.

Mögliche Anwendungsgebiete für (Q)SAR's

Sie werden auf folgenden Ebenen gesehen:

- Risikobasierte Priorisierung von Stoffen / Produkten
- Risikobewertung (besonders für Stoffe mit niedrigen Produktionsvolumina)
- Einstufung und Kennzeichnung von Stoffen auf Basis der intrinsischen Eigenschaften einschließlich Plausibilitätsüberprüfung (Verunreinigungsproblematik?)
- Entscheidungshilfe bei der Entwicklung von Teststrategien

Schlussfolgerungen für den Einsatz von (Q)SAR Methoden

Zusammenfassend ergeben sich aus Industriesicht daraus die folgenden Schlussfolgerungen für den Einsatz von (Q)SAR Methoden:

- Ein Ersatz von experimentellen Prüfungen im Bereich Ökotoxikologie und Bioabbau ist möglich.
- Bei Endpunkten der Säugertoxizität ist für die Erstellung von Teststrategien eine Priorisierung von Substanzen

möglich, falls ausreichende Informationen zu den Stoffeigenschaften über ADME-Einflüsse bekannt sind.

Hier könnte - als Denkmöglichkeit - die begrenzte Prüfung (im Sinn eines "Screening"-Verfahrens) von plasmakinetischen Daten durchgeführt werden. Zusammen mit der Vorhersage z.B. leberspezifischer Metaboliten wie es im (Q)SAR-Modell TIMES von LMC möglich ist, könnten daraus Eckwerte für Plasmahalbwertszeiten und potentielle Metaboliten erstellt werden, die als Surrogat-Daten für eine volle ADME-Prüfung herangezogen werden können.

Für den Endpunkt ADME zeichnet sich ein möglicher Einsatz von spezifisch dafür ausgelegten (Q)SAR-Programmen aus dem Pharmabereich ab. Ihre Validierung für den Einsatz an Industrie-Chemikalien mit deren ganz andersartigen proprietären Chemie ist allerdings noch offen und müsste ggf. in Angriff genommen werden.

Bei Congeneren erscheint ein Ersatz von tier-experimentellen Untersuchungen durch (Q)SAR-Betrachtungen möglich, vor allem wenn eine positive Aussage im Sinn einer Belastung der Substanz und ein Kennzeichnungs- und Einstufungsvorschlag daraus abgeleitet werden.

Die regulatorische Akzeptanz von negativen Vorhersagen einer Wirkpotenz ("Entlastung" der Substanz) wird z.Zt. allerdings noch als problematisch gesehen.

Voraussetzungen und Einschränkungen für den gegenwärtigen Einsatz von (Q)SAR- Vorhersagen in der (Öko-)Toxikologie

Die Nutzung von (Q)SAR Vorhersagen unterliegt z. Zt. einer Reihe stringenter Voraussetzungen:

- keine Treffer bei Literatur-Recherchen in in-/externen Datenbanken
- Analogieschluss scheidet aus
- Expert Judgement nicht möglich
- Ausschlusskriterien für die Anwendung von (Q)SAR liegen nicht vor (z.B. Polymere, Anorganika, instabile Substanzen, Mischungen, Salze (einige (Q)SAR-Modelle), Metalle, Metallkomplex-Verbindungen, Fasern)
- Zuverlässigkeit des (Q)SAR-Modells für Stoffklasse und Endpunkt ist bekannt

Hinzu kommen einige technische Einschränkungen, unter denen die Erstellung von (Q)SAR-Vorhersagen nicht sinnvoll erscheint:

So werden für F&E (Forschungs- und Entwicklungs)-Produkte, die z.B. als Muster in kleinen Mengen versandt werden (relativ häufig) keine derartigen Beurteilungen erstellt.

Dies gilt ebenfalls für Zwischenprodukte, wenn sie stofflich nur in kleinen Mengen akkumulieren bzw. eine ausschließliche Handhabung im geschlossenen Systemen erfolgt sowie für Zwischenprodukte, deren Umsetzung innerhalb des Firmengeländes erfolgt (sehr selten).

Ausnahmen stellen Handels- und Verkaufsprodukte dar, die einer regulären Prüfung unterzogen werden.

(Q)SAR Bewertungen werden i. A. nicht durchgeführt, wenn Alternativen durch *in-vitro* Testsysteme verfügbar sind. Dazu gehören beispielhaft:

- Reizwirkung am Auge und an der Haut: HetCam (Auge); EpiDerm (Haut)
- *in-vitro* Hautpenetration (Ratten- u. Humanhaut)
- Kanzerogenität SHE (Syrian Hamster Embryo)-Zelltransformationstest
- Entwicklungstoxizität im. Limb-Bud Assay bzw. am Fischei (mit der Einschränkung für parental aktive Substanzen)

Eigene Bewertungen und Erfahrungen

Eigene Erfahrungen wurden bei der Prädiktion akuter Säugertoxizität, speziell bei oraler Zufuhr, mit dem Einsatz neuronaler Netze gemacht.

Bezüglich Reizwirkung wurde (zeitlich begrenzt) bei dem BfR SAR Expertensystem (Skin Irritation Corrosion Rules Estimation Tool, SICRET; /4/) mitgearbeitet.

Zur Mutagenität als Endpunkt mit der umfangreichsten Datenlage wurden verschiedene (Q)SAR Modelle und Expertensysteme bewertet (DEREK, MultiCase und TopKat) und im Rahmen der Erstellung eigener Trainingsdaten insbesondere mit MultiCase bei der Mutagenitätsvorhersage detaillierte Erfahrungen gesammelt. Eine wesentliche Erfahrung war die Bedeutung, bei dem Einsatz kommerzielle (Q)SAR Modelle eigene Trainingsdaten einzubringen. Da sich die Strukturdomänen kommerzieller Software, die zumeist auf öffentlich zugänglichen Trainingsdaten basieren, von Industriechemikalien mit proprietärer Chemie signifikant unterscheiden, entstehen sonst zwangsweise Defizite in der Zuverlässigkeit der Vorhersage. Ein weiteres wichtiges Argument ist die Möglichkeit, die Daten bezüglich des verwendeten Testprotokolls rigoros zu definieren und Unsicherheiten auf Grund aktiver NebenkompONENTEN zu erfassen bzw. zu eliminieren.

In einer neueren Entwicklung wurde an der Modellierung von TIMES (LMC Bourgas, Bulgarien; /5/) mit eigenen Trainingsdaten zur Mutagenität und zur Erweiterung auf den Endpunkt Chromosomenaberration mitgearbeitet.

TIMES (Tissue **ME**tabolism **S**imulator) ist ein heuristischer Algorithmus, mit dem plausible Metaboliten von einer umfassenden Bibliothek bekannter Biotransformationen und abiotischer Reaktionen abgeleitet und semiquantitativ vorhergesagt werden.

Durch die Integration dieses Moduls gestattet TIMES im Bereich Säugertoxizität neben der Aktivitätsvorhersage für die Parentalstrukturen auch die der wahrscheinlichen Metabolite im Bereich Gentoxizität, Hautsensibilisierung (in Entwicklung) und Estrogen/ Androgen-Rezeptor Bindungsaffinitäten.

Ein weiteres Modul (CATABOL) zur Vorhersage der Kinetik des Bioabbaus wurde ebenfalls in seiner Modellierung mit einer umfassenden Sammlung eigener Prüfdaten unterstützt.

Zusätzliche Motivation

Abgesehen von der Nutzung als internes Bewertungs- und Priorisierungswerkzeug hat die Beschäftigung mit diesen Modellen ein besseres mechanistisches Verständnis des vorhergesagten Effekts zur Folge gehabt. Dabei spielt auch die Berücksichtigung des spezifischen Metabolismus (z.B. Leber oder Haut) eine wesentliche Rolle.

Insbesondere wurde der Wert von Expertensystemen für die Organisation des eigenen Fachwissens im Bereich "Expert Judgement" deutlich, dem auch bei den Einstufungsbegründungen des GHS (Global Harmonised System) eine zunehmend bedeutende Rolle zugeordnet wird.

Nutzen der eigenen Erfahrungen

Die eigenen Erfahrungen im Umgang mit den z.Zt. verfügbaren (Q)SAR Programmen waren eine wesentliche Voraussetzung dafür, in den entsprechenden Expertengremien des Europäischen Chemieverbandes und der OECD eigene Bewertungen der Einsatztauglichkeit und notwendiger Weiterentwicklung – auch im Hinblick auf Akzeptanzkriterien - aus Sicht der Industrie einzubringen. Sie haben sich auch in Behördengesprächen bewährt, wenn Analogien zur Bewertung kritischer Substanzen herangezogen wurden und die eigenen Erfahrungen für eine kritische Wertung dieser Schlussfolgerungen entgegen gestellt werden konnten.

Literatur:

- /1/ D. Bonchev, D.H. Rouvray: Chemical Graph Theory: Introduction and Fundamentals. Gordon and Breach Science Publishers, 1990, ISBN 0-85626-454-7
- /2/ www.ecetoc.org
- /3/ (Q)SARs: Evaluation of the commercially available software for human health and environmental endpoints with respect to chemical management applications (2003), ECETOC Technical Report No. 89
- /4/ www.bfr.bund.de/cm/232/das_neue_europaeische_chemikalien_recht_und_der_verbraucherschutz_langfassung.pdf
- /5/ <http://www.oasis-lmc.org/>

Korrespondenzadresse

Dr. E. Jacob

Abt. Produktsicherheit, Chemikalienrecht, Toxikologie und Ökologie,
GUP/CB - Z470, BASF AG
D-67056 Ludwigshafen/Rhein