

# Organische Reaktionen mit Hilfe neuronaler Netze klassifiziert: Michael-Additionen, Friedel-Crafts-Alkylierungen durch Alkene und verwandte Reaktionen\*\*

Lingran Chen und Johann Gasteiger\*

Organische Reaktionen werden durch viele Faktoren beeinflusst: durch die Struktur der Ausgangsmaterialien, die Reagentien und Katalysatoren sowie die Reaktionsbedingungen, z. B. Temperatur, Lösungsmittel, Druck und Licht. Jeder dieser Faktoren kann als eine Koordinate betrachtet werden, die gemeinsam einen mehrdimensionalen Raum aufspannen, und in diesem Sinne ist eine chemische Reaktion ein Ereignis in einem solchen mehrdimensionalen Raum. Chemiker haben ihr Wissen über organische Reaktionen zum großen Teil aus Beobachtungen an Serien von Reaktionen gewonnen, indem sie nach Gemeinsamkeiten und Unterschieden suchten und so Reaktionen in Klassen einteilten, die meist nach ihren Entdeckern benannt wurden, wie Wittig-Reaktion, Michael-Addition oder Beckmann-Umlagerung. Natürlich kann solch ein eindimensionales Klassifizierungsschema nur ungenügend den vielfältigen Beobachtungen bei chemischen Reaktionen Rechnung tragen.

Wir zeigen nun, daß ein Satz chemischer Reaktionen durch ein selbstorganisierendes neuronales Netz in eine zweidimensionale Karte projiziert werden kann. Eine Reaktion entspricht dann einem Punkt in einer solchen Karte, und der Abstand zweier Punkte spiegelt den Grad der Ähnlichkeit der zugehörigen Reaktionen wider. Je nachdem in welcher Richtung in eine solche zweidimensionale Karte projiziert wird, lassen sich unterschiedliche Arten von Ähnlichkeit darstellen.

Das menschliche Gehirn erzeugt zweidimensionale Karten im visuellen, auditiven oder somatosensorischen Kortex aus sensorischen Informationen der Umgebung. Das selbstorganisierende neuronale Netz nach Kohonen<sup>[1]</sup> nimmt diese Fähigkeit des menschlichen Gehirns als Vorbild<sup>[2]</sup>. In ihm sind die Grundelemente, die künstlichen Neuronen, in zwei Dimensionen angeordnet, so daß „Landkarten“ der analysierten Daten erhalten werden.

Organische Reaktionen überstreichen einen gewaltigen Bereich an Variationen, und eine Karte, die dieses breite Spektrum an Möglichkeiten widerspiegelt, müßte sehr groß sein. Um die Leistungsfähigkeit unserer Methode zu veranschaulichen, beschränken wir uns deshalb auf einen Satz an Reaktionen, die ein wichtiges Charakteristikum gemeinsam haben: das Reaktionszentrum, d. h. die Atome und Bindungen, die direkt am Bindungsumordnungsprozeß teilnehmen. Wir entschieden uns für die Addition einer C-H-Bindung an eine C-C-Doppelbindung. Damit sind so bedeutende Reaktionstypen wie Michael-Additionen, Friedel-Crafts-Alkylierungen durch Alkene und die Addition freier Radikale an Alkene erfaßt. In Tabelle 1 ist in der Kopfzeile das Reaktionszentrum dargestellt.

120 Reaktionen wurden bei einer Suche mit dem Reaktions-Retrieval-System ISIS Host<sup>[3]</sup> in der Version von 1991 der Reaktionsdatenbank ChemInform-RX erhalten<sup>[4]</sup>. Da Änderun-

Tabelle 1. Elektronische Variablen zur Charakterisierung der Reaktionszentren der diskutierten Reaktionen [a].

	$\begin{array}{ccccccc} \text{C}=\text{C} & + & \text{H}-\text{C} & \rightarrow & \text{H}-\text{C}-\text{C}-\text{C} \\ 1 & 2 & 4 & 3 & 4 & 1 & 2 & 3 \end{array}$							
$\alpha_i$								
$q_{\text{tot}}$								
$\chi_\sigma$								
$\chi_\pi$								

[a] Für die Bedeutung der Symbole siehe Text.

gen in der Struktur des Ausgangsmaterials von ausschlaggebendem Einfluß auf eine chemische Reaktion sind, haben wir uns auf die Untersuchung des strukturellen Einflusses konzentriert. Die Frage war, wie sollten die Strukturen kodiert werden? Offensichtlich kann eine Liste der funktionellen Gruppen um das Reaktionszentrum nicht die Methode der Wahl sein, da diese Liste sehr umfassend sein müßte. Die Chemiker haben das Problem, unterschiedliche funktionelle Gruppen zu vergleichen, bereits angegangen und deren Einfluß mit Konzepten wie Partialladung, induktiver/Feld-, Resonanz- oder Polarisierbarkeits-effekt verallgemeinert. Methoden zur empirischen Berechnung dieser Effekte sind in unserer Gruppe entwickelt worden<sup>[5-8]</sup> und wurden eingesetzt, um den Einfluß funktioneller Gruppen auf die Atome des Reaktionszentrums zu berechnen. Dabei wurden bei den Atomen C-1 und C-3 die  $\sigma$ - und  $\pi$ -Elektronenaktivitäten  $\chi_\sigma$  bzw.  $\chi_\pi$  betrachtet, bei den Atomen C-2 und C-3 die Gesamtladungen  $q_{\text{tot}}$  und beim Atom C-3 die effektive Polarisierbarkeit  $\alpha_i$  (siehe Tabelle 1).

Diese sieben Variablen wurden zur Beschreibung jeder einzelnen Reaktion des Datensatzes und somit als Eingabe in das neuronale Kohonen-Netz verwendet. Dieses projizierte dann den sieben-dimensionalen Raum in eine Karte aus  $12 \times 12$  Neuronen.

Jedes Neuron hat so viele Gewichte, wie es Eingabevariablen für das beschriebene Objekt gibt, in unserem Fall also sieben. Eine Reaktion  $s$  wird in dasjenige Neuron  $c$  projiziert, dessen Gewichte  $w_{ji}$  ihren Eingabedaten  $x_{si}$  am ähnlichsten sind [Gl. (a)]. Nach jeder Eingabe einer Reaktion wurden die Ge-

$$c \leftarrow \min \left\{ \sum_{i=1}^7 (w_{ji} - x_{si})^2 \right\} \quad j = 1-144 \quad (\text{a})$$

wichte aller Neuronen den Eingabewerten dieser Reaktion angeglichen. Die Änderungen sind dabei für das Gewinnerneuron  $c$  am größten und nehmen mit zunehmendem Abstand von diesem ab. Reaktionen mit ähnlichen elektronischen Faktoren werden daher in das gleiche oder in nah benachbarte Neuronen abgebildet. In unserem Fall erhielt eine Reihe von Neuronen mehrere Reaktionen (bis zu 5), was das hohe Maß an Ähnlichkeit bei diesen Reaktionen aufzeigt. Einer erheblichen Zahl an Neuronen (78) wurde überhaupt keine Reaktion zugewiesen.

Um die Ergebnisse dieser Selbstorganisation von 120 Reaktionen in einem Kohonen-Netz zu veranschaulichen, wurden die Reaktionen von Chemikern den Reaktionstypen zugeordnet und diese durch Symbole charakterisiert (Tabelle 2). Die Neuronen wurden dann mit diesen Symbolen markiert, um zu zeigen, welche Reaktionen in welche Neuronen projiziert wurden. Die Ergebnisse sind in Abbildung 1 dargestellt.

Ziemlich viele Neuronen enthalten mehrere Reaktionen, die immer vom gleichen Typ sind. Dies unterstreicht die Leistungsfähigkeit des Kohonen-Netzes, die Ähnlichkeit von Reaktionen zu erkennen. Nur ein einziges Neuron, das Neuron (12,1), enthält einen Widerspruch, da ihm sowohl eine Kondensations-

[\*] Prof. Dr. J. Gasteiger, Dr. L. Chen  
Computer-Chemie-Centrum  
Institut für Organische Chemie der Universität Erlangen-Nürnberg  
Nägelbachstraße 25, D-91052 Erlangen  
Telefax: Int. + 9131/85-6566  
E-mail: gasteiger@eros.ccc.uni-erlangen.de

[\*\*] Diese Arbeit wurde von der Alexander-von-Humboldt-Stiftung unterstützt. Wir danken MDL Information Systems für die Überlassung von ISIS Host und, zusammen mit dem Fachinformationszentrum Chemie, Berlin, für den Zugang zur Reaktionsdatenbank ChemInform-RX.

Tabelle 2. Identifizierte Reaktionstypen mit den in Abb. 1 verwendeten Symbolen und der Zahl an Reaktionsbeispielen.

	Reaktionstyp	Zahl der Reaktionen	Reaktionsindex
A	Michael-Addition	75	1–14, 24–84
B	fälsch bezeichnetes Reaktionszentrum	1	85
C	Friedel-Crafts-Alkylierung durch Alkene	18	15–23, 86–92, 115, 116
D	Elektronentransferreaktion	1	93
E	photochemische Addition eines Acylradikals an ein elektronenarmes Olefin	1	94
F	photolytisch induzierte Alkylierung	4	95–97, 101
G	H-Radikal-Transferreaktion	3	98–100
H	Reaktion mit speziellem Mechanismus	1	102
I	Hydridabstraktionsreaktion	3	103–105
J	Nazarov-Reaktion	5	106–110
K	Allylierung eines Oxoesters mit anschließender Anti-Markovnikov-Addition von HBr[a]	1	111
L	palladiumkatalysierte Hexatrienolat-Cyclisierung[a]	3	112–114
M	Kondensationsreaktion	1	117
N	photoinduzierte Radikaladdition an einen Kronenether	3	118–120

[a] Zwei Reaktionsschritte.

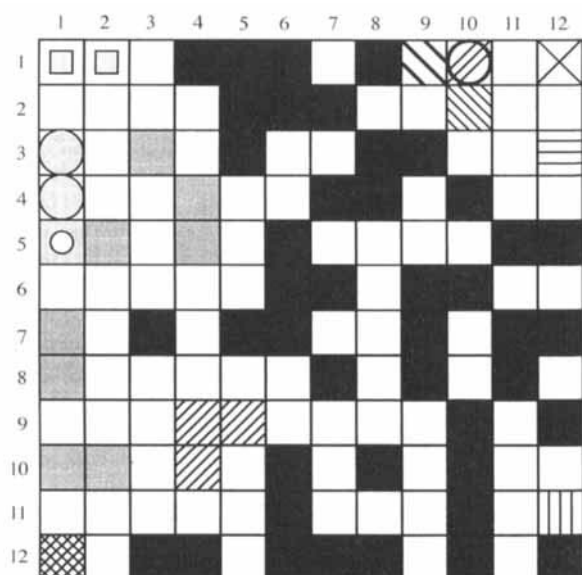


Abb. 1. Kohonen-Karte zur Klassifizierung von 120 Reaktionen. Leere Kästchen stehen für leere Neuronen, durchgekreuzte für Konfliktnuronen. Die Bedeutung der übrigen Symbole ist in Tabelle 2 erklärt.

reaktion zugeordnet wurde als auch eine Reaktion, die in der Datenbank mit einem falschen Reaktionszentrum bezeichnet wurde.

Bei den meisten Beispielreaktionen handelt es sich um Michael-Additionen, so daß das Kohonen-Netz logischerweise für diesen Reaktionstyp die größte Fläche vorbehält. Die Spezialfälle und die Reaktionstypen mit nur wenigen Beispielen wurden stärker an den Rand der Karte geschoben. Es ist beachtlich, daß die von einem Chemiker zugeteilten Reaktionstypen auch von einem Kohonen-Netz erkannt werden, das als Zuordnungskriterium

physikochemische Variablen der Reaktionszentren der einzelnen Reaktionen nutzt.

Innerhalb des Gebietes eines Reaktionstyps enthält die genaue Stelle, an der sich eine Reaktion befindet, noch eine Menge an weiterer chemischer Information. Abbildung 2 zeigt den Teil

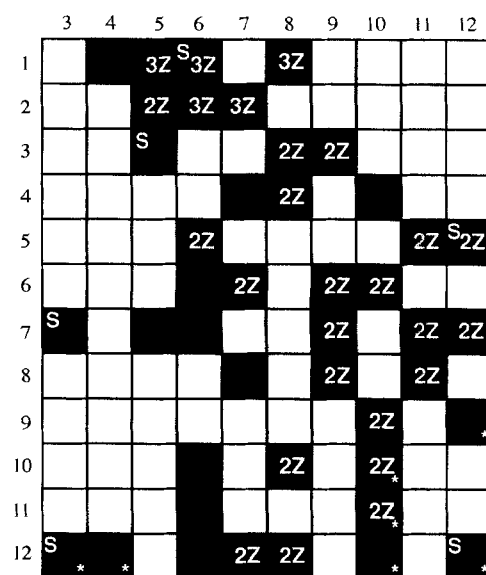
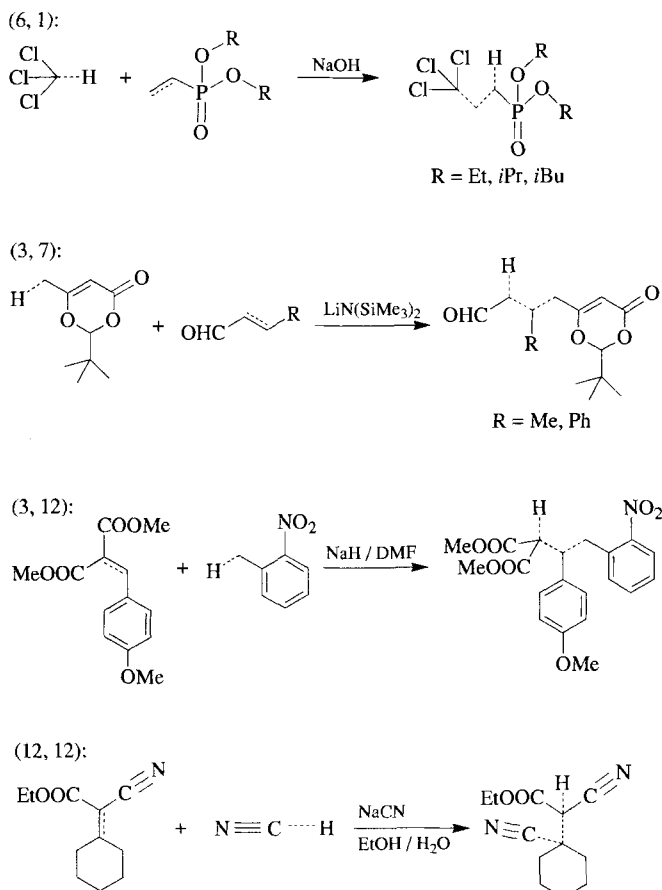


Abb. 2. Detaillierte Analyse der Gruppe der Michael-Additionen in der Kohonen-Karte von Abb. 1. Die mit 2Z und 3Z gekennzeichneten Neuronen enthalten Michael-Additionen, bei denen die reagierenden C-H-Bindungen durch zwei bzw. drei stark elektronenziehende Gruppen aktiviert werden. Die mit \* markierten Neuronen weisen auf Michael-Additionen hin, bei denen die reagierende C=C-Bindung durch zwei stark elektronenziehende Gruppen aktiviert wird. Die mit S gekennzeichneten Neuronen enthalten spezielle Michael-Additionen.

der Karte von Abbildung 1, in dem sich die Michael-Additionen befinden. Zusätzliche Symbole geben hier weitere Details zu den Reaktionen an. Reaktionen, die eine (keine zusätzliche Bezeichnung), zwei (2Z) und drei (3Z) elektronenziehende Gruppen an der reagierenden C-H-Bindung enthalten, sind recht gut voneinander getrennt. Genauso werden die Reaktionen, bei denen die C=C-Bindung zwei stark elektronenziehende Gruppen trägt (\*), von denen räumlich getrennt, die nur eine derartige Gruppe enthalten (kein \*). Eine zweidimensionale Karte bietet weiterhin die Möglichkeit, Strukturänderungen an beiden reagierenden Bindungen, der C-H- und der C=C-Bindung, gleichzeitig Rechnung zu tragen. Außerdem findet man innerhalb eines Reaktionstyps wie der Michael-Addition die spezielleren Beispiele (S in Abb. 2) an den Randgebieten der Karte.

Einige der spezielleren Michael-Additionen werden im folgenden näher vorgestellt. Schema 1 zeigt die Reaktionen, die den Neuronen (6,1), (3,7), (3,12) und (12,12) zugeordnet wurden. Während die meisten Michael-Additionen von einer C-H-Bindung ausgehen, die durch Gruppen mit einem –M-Effekt aktiviert wird, wird bei den drei Reaktionen, die in das Neuron (6,1) projiziert wurden, das Carbanion durch drei Substituenten mit einem –I-Effekt stabilisiert<sup>[9]</sup>. Die zwei Reaktionen im Neuron (3,7) haben die Besonderheit, daß eine CH<sub>3</sub>-Gruppe reagiert, die durch eine Esterfunktion aktiviert wird, die ihren Einfluß durch Konjugation über eine Doppelbindung ausübt<sup>[10]</sup>. Bei der Reaktion, die das Neuron (3,12) enthält, wird ebenfalls eine CH<sub>3</sub>-Gruppe durch Konjugation aktiviert, hier allerdings durch eine Nitrogruppe über einen Benzolring<sup>[11]</sup>. Die Reaktionen in den Neuronen (3,7) und (3,12) sind beide durch eine CH<sub>3</sub>-Gruppe charakterisiert, die durch konjugativ



Schema 1. Die Michael-Additionen, die in die Neuronen (6,1), (3,7), (3,12) und (12,12) projiziert wurden; die gestrichelten Bindungen geben die in der Reaktion gebrochenen und geknüpften Bindungen wieder.

elektronenziehende Gruppen aktiviert wird, jedoch wird dieser Effekt über unterschiedliche  $\pi$ -Elektronensysteme übertragen. Es befriedigt, daß diese beiden speziellen Michael-Additionen in ähnliche Regionen der Karte gelangen.

Die Reaktion im Neuron (12,12)<sup>[12]</sup> ist die einzige Michael-Addition in diesem Satz an Reaktionen, bei der eine  $C_{sp^3}$ -H-Bindung reagiert; in allen anderen Fällen ist eine  $C_{sp^2}$ -H-Bindung beteiligt. Diese Reaktion erweitert also den Bereich der Michael-Additionen und wird folgerichtig im Randbereich des Reaktionstyps abgelegt.

Eine der aufregendsten Eigenschaften des Kohonen-Netzes ist, daß es nach der Reaktionsklassifizierung automatisch chemisches Wissen aus diesen Reaktionsbeispielen erworben hat. Folglich kann das „trainierte“ Netz auch eingesetzt werden, um den Reaktionstyp einer unbekannten Reaktion vorherzusagen. Um dies zu verdeutlichen, wurden die 120 Beispielreaktionen in zwei Gruppen aufgeteilt: Die 60 Reaktionen mit ungeraden Indices (siehe Tabelle 2) wurden als Trainingssatz verwendet, die übrigen 60 als Testsatz. Das Ergebnis ist recht eindrucksvoll: Die Reaktionstypen von 95% der Reaktionen aus dem Testsatz (57 Beispielen) wurden korrekt vorhergesagt. Der Grund für zwei nicht eindeutig und eine falsch zugeordnete Reaktion ist sehr einfach: Es gibt kein Beispiel für diese drei speziellen Reaktionen im Trainingssatz. Sie wurden folgerichtig in leere Neuronen am Rand abgelegt.

Zweidimensionale Kohonen-Karten können die Beziehungen zwischen den untersuchten chemischen Reaktionen aufzeigen, die wichtigen Reaktionstypen herausheben, deren Geltungsbereiche darstellen und außerdem auf ungewöhnliche Reaktionen

hinweisen. Sie ermöglichen es dem Chemiker, seine Beobachtungen über chemische Reaktionen in einer intuitiv ansprechenden und chemisch richtigen Art und Weise zu ordnen.

Eingegangen am 6. Oktober 1995 [Z8450]

**Stichworte:** Kohonen-Karte · Michael-Additionen · Neuronale Netze · Reaktionsklassifizierung

- [1] a) T. Kohonen, *Biol. Cybern.* **1982**, 43, 59–69; b) *Self-Organization and Associative Memory*, 3. Aufl., Springer, Berlin, **1989**.
- [2] J. Zupan, J. Gasteiger, *Neural Networks for Chemists: An Introduction*, VCH, Weinheim, **1993**; J. Gasteiger, J. Zupan, *Angew. Chem.* **1993**, 105, 510–536; *Angew. Chem. Int. Ed. Engl.* **1993**, 32, 503–527.
- [3] ISIS Host wird von MDL Information Systems, San Leandro, CA, USA, vertrieben.
- [4] a) A. Parlow, C. Weiske, J. Gasteiger, *J. Chem. Inf. Comput. Sci.* **1990**, 30, 400–402; b) ChemInform-RX: produziert vom Fachinformationszentrum Chemie, Berlin, und vertrieben von MDL Information Systems, San Leandro, CA, USA.
- [5] J. Gasteiger, M. Marsili, *Tetrahedron* **1980**, 36, 3219–3228.
- [6] J. Gasteiger, H. Saller, *Angew. Chem.* **1985**, 97, 699–701; *Angew. Chem. Int. Ed. Engl.* **1985**, 24, 687–689.
- [7] M. G. Hutchings, J. Gasteiger, *Tetrahedron Lett.* **1983**, 24, 2541–2544.
- [8] J. Gasteiger, M. G. Hutchings, *J. Chem. Soc. Perkin Trans. 2* **1984**, 559–564.
- [9] O. A. Kolyamshin, V. V. Kormachev, Y. N. Mitrasov, *Zh. Obshch. Khim.* **1990**, 60, 2170–2171.
- [10] D. Seebach, U. Mißlitz, P. Uhlmann, *Chem. Ber.* **1991**, 124, 1845–1852.
- [11] D. M. Floyd, R. V. Moquin, K. S. Atwal, S. Z. Ahmed, S. H. Spergel, J. Z. Gougoutas, M. F. Malley, *J. Org. Chem.* **1990**, 55, 5572–5579.
- [12] G. Griffiths, H. Mettler, L. S. Mills, F. Previdoli, *Helv. Chim. Acta* **1991**, 74, 309–314.

## Elektronenstruktur von Nicht-Häm-Oxo-eisen-komplexen in hohen Oxidationsstufen mit der bislang unbekannten $[\text{Fe}_2(\mu\text{-O})_2]^{3+}$ -Einheit\*\*

Abhik Ghosh, Jan Almlöf† und Lawrence Que, Jr.\*

Die Chemie der Disauerstoffaktivierung an Nicht-Häm-Dieisenzentren in Enzymen wie Ribonucleotid-Reduktase und Methan-Monooxygenase hat in letzter Zeit viel Aufmerksamkeit erregt<sup>[1,2]</sup>. Für das letztere Enzym ist eine instabile Zwischenstufe, Verbindung Q genannt, Mößbauer-spektroskopisch charakterisiert worden<sup>[2]</sup>. Die gekoppelte Dieisen(IV)-Spezies ist in der Lage, Methan zu hydroxylieren. Unsere Anstrengungen zur Synthese von Modellverbindungen solcher Zwischenstufen mit hohen Oxidationsstufen führten zu zwei Komplexen der allgemeinen Formel  $[\text{LFe}(\mu\text{-O})_2\text{FeL}](\text{ClO}_4)_3$  [1: L = Tris(2-pyridylmethyl)amin (tpa); 2: L = Tris(5-methyl-2-pyridylmethyl)amin (tpa')], die die bislang unbekannte hochvalente  $[\text{Fe}_2(\mu\text{-O})_2]^{3+}$ -Einheit enthalten<sup>[3,4]</sup>. Die ESR-Spektren von 1 und 2 deuten auf einen Quartett-Spinzustand ( $S = 3/2$ ) hin, und die Parameter der Mößbauer-Spektren ähneln denen der Verbindung Q<sup>[4]</sup>. Obwohl formal die Oxidationsstufen der Eisenzentren in der  $[\text{Fe}_2(\mu\text{-O})_2]^{3+}$ -Einheit einfach mit III und IV ange-

[\*] Prof. Dr. L. Que, Jr., Dr. A. Ghosh<sup>[†]</sup>, Prof. Dr. J. Almlöf<sup>†</sup>  
Department of Chemistry  
University of Minnesota  
Minneapolis, MN 55455 (USA)  
Telefax: Int. + 612/624-7029

[†] Neue Adresse: Department of Chemistry  
Institute of Mathematical and Physical Sciences (IMR)  
University of Tromsø  
Breivika, N-9037 Tromsø (Norwegen)

[\*\*] Diese Arbeit wurde vom Minnesota Supercomputer Institute (Rechenzeit auf einem Cray-C90-Computer), von den National Institutes of Health (GM-38767(LQ)), der National Science Foundation (CHE-9223782(JA)) und von Biosym Technologies, Inc. (San Diego, CA) unterstützt.