

eder sind vermutlich infolge der sterischen Anforderungen der Phenylringe verzerrt tetraedrisch (Bereich der  $\mu$ -P-Cd-P-terminal-Winkel 105.20(9)–126.24(8) $^\circ$ , Bereich der  $\mu$ -P-Cd- $\mu$ -P-Winkel 99.11(8)–109.43(9) $^\circ$ ). Dagegen unterscheiden sich die Cd-P-Cd-Bindungswinkel innerhalb des Käfigs nicht erheblich (Bereich von 115.97(9)–121.62(9) $^\circ$ ).

Adamantangerüste wurden auch in einigen Strukturen von Cadmiumsulfiden beobachtet, und zwar in  $(\text{NMe}_4)_2[(\mu\text{-SC}_6\text{H}_5)_6\text{Cd}_4\text{Br}_4]$ <sup>[14]</sup>,  $(\text{Et}_3\text{NH})_2[(\mu\text{-SC}_6\text{H}_5)_6\text{Cd}_4(\text{SC}_6\text{H}_5)_4]$ <sup>[15]</sup>,  $(\text{NEt}_4)_2[(\mu\text{-SiPr})_6\text{Cd}_4\text{Br}_4]$ <sup>[16]</sup> und  $[(\mu\text{-SiPr})_6(\text{CdPPPh}_3)_2(\text{CdOClO}_3)_2] \cdot \text{EtOH}$ <sup>[17]</sup>. Obwohl  $[\text{Cd}_4\text{Cl}_4(\text{PPh}_2)_4(\text{PnPr}_3)_2]$  ein neutrales Adamantangerüst aufweist, enthält 1 das erste strukturell charakterisierte, homoleptische Cadmiumphosphid-Adamantanfragment. Der Ausschluß von Halogeniden im Anion von 1 ist für seine mögliche Anwendung als Einkomponenten-Precursor für halbleitende Cd/P-Phasen von Bedeutung.

Das temperaturabhängige  $^{31}\text{P}$ -NMR-Spektrum von 1 in Lösung zeigt bei Raumtemperatur ein einziges Resonanzsignal ( $\delta = -41.5$ ). Dieses spaltet bei sinkender Temperatur auf (<213 K). Die für diese Auflösung erforderliche tiefe Temperatur korreliert mit einem Vorgang mit geringer Aktivierungsenergie, wie einem intramolekularen Austausch der terminalen und verbrückenden PPh<sub>2</sub>-Liganden des Anions. Außerdem beobachtet man bei tieferen Temperaturen Satelliten, die einer  $^{31}\text{P}$ - $^{111/113}\text{Cd}$ -Kopplung von etwa 100 Hz entsprechen.

Der Cadmiumkomplex 1 ist ein seltenes Beispiel für ein strukturell charakterisiertes, metallorganisches Cadmiumphosphid, das abgesehen von seinem kristallographischen Interesse ein potentieller Einkomponenten-Precursor für die Bildung einer Cd/P-Halbleiterphase ist.

### Experimentelles

Synthese von 1: 3.13 mL BuLi (Aldrich, 5.0 mmol, 1.6 M Lösung in Hexan) wurden zu einer Lösung von 0.87 mL Ph<sub>2</sub>PH (Strem, 5.0 mmol) in 10 mL THF gegeben. Zu der gebildeten roten Lösung wurden 1.0 mL  $[\text{Cd}(\text{N}(\text{SiMe}_3)_2)_2]$ <sup>[18]</sup> (2.5 mmol) hinzugefügt. Nach 5 min Röhren bei Raumtemperatur erhielt man einen gelben Niederschlag, der durch Zugabe von weiteren 30 mL THF und Erhitzen bis zum Rückfluß aufgelöst wurde. Man ließ die gelbe Lösung 12 h bei Raumtemperatur stehen und erhielt gelbe, luftempfindliche Kristalle von 1 in 70% Ausbeute (1.00 g, erste Fraktion). Bei der Isolierung wurde gittergebundenes THF durch Behandlung im Vakuum ( $10^{-2}$  atm, 10 min) entfernt, was ein amorphes gelbes Pulver ergab, das bei den folgenden Analysen verwendet wurde: Zersetzung zu einem schwarzen Feststoff oberhalb 200 °C;  $^1\text{H}$ -NMR (+25 °C, 250 MHz, 0.2 mol L<sup>-1</sup>, [D<sub>6</sub>]THF):  $\delta = 7.41$ –7.26 (überlappende Multipletts, 100 H, Ph), 3.60 (m, 32 H, THF), 1.75 (m, 32 H, THF);  $^{31}\text{P}$ -NMR (+25 °C, 101.2 MHz, 0.2 mol L<sup>-1</sup>, [D<sub>6</sub>]THF, Standard 85% H<sub>3</sub>PO<sub>4</sub> bei 0):  $\delta = -41.5$ ; Elementaranalyse: ber. C 63.0, H 5.7, P 10.7; gef. C 62.4, H 5.7, P 10.4.

Synthese von 2: 3.13 mL BuLi (Aldrich, 5.0 mmol, 1.6 M Lösung in Hexan) wurden zu einer Lösung von 1.1 mL cy<sub>2</sub>PH (Strem, 5.0 mmol) in 10 mL THF gegeben. Zu der gebildeten gelben Lösung wurden 1.0 mL  $[\text{Cd}(\text{N}(\text{SiMe}_3)_2)_2]$ <sup>[18]</sup> (2.5 mmol) hinzugefügt, und das Gemisch wurde 5 min zum Rückfluß gebracht. Das Reaktionsgemisch wurde filtriert (Celite, Porosität 3), was eine gelbe Lösung ergab. Die Lösung wurde im Vakuum auf 7 mL eingeengt und 2 mL Hexan hinzugefügt. Man ließ die Lösung anschließend 24 h bei Raumtemperatur stehen und erhielt farblose, luftempfindliche Kristalle von 2 in 35% Ausbeute (0.35 g, erste Fraktion): Die physikalischen Eigenschaften von 2 stehen mit den Literaturdaten in Einklang [11, 19]; Fp. 173 °C;  $^1\text{H}$ -NMR (+25 °C, 250 MHz, 0.2 mol L<sup>-1</sup>, [D<sub>6</sub>]THF):  $\delta = 2.04$ –1.22 (überlappende Multipletts, 44 H, Cyclohexyl);  $^{31}\text{P}$ -NMR (+25 °C, 101.2 MHz, 0.2 mol L<sup>-1</sup>, [D<sub>6</sub>]THF, Standard 85% H<sub>3</sub>PO<sub>4</sub> bei 0):  $\delta = -22.2$  (s); Elementaranalyse: ber. C 73.0, H 11.2, P 15.7; gef. C 72.4, H 11.0, P 15.0.

Eingegangen am 2. August 1995 [Z 8273]

**Stichworte:** Adamantanartige Verbindungen · Cadmiumverbindungen · Komplexe mit Phosphorliganden

- [1] A. H. Cowley, R. A. Jones, *Angew. Chem.* **1989**, *101*, 1235; *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 1208.
- [2] R. L. Wells, *Coord. Chem. Rev.* **1992**, 273.
- [3] M. A. Paver, C. A. Russell, D. S. Wright, *Comprehensive Organometallic Chemistry, Vol. II* (Hrsg.: F. G. A. Stone, G. Wilkinson, E. W. Abel), Pergamon, Oxford, 1995.
- [4] M. Bochmann, K. Webb, M. Harman, M. B. Hursthouse, *Angew. Chem.* **1990**, *102*, 703; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 638.
- [5] S. C. Goel, M. Y. Chiang, D. J. Rauscher, W. E. Buhro, *J. Am. Chem. Soc.* **1990**, *112*, 5636.
- [6] M. G. Davidson, A. J. Edwards, M. A. Paver, P. R. Raithby, C. A. Russell, A. Steiner, K. L. Verhorevoort, D. S. Wright, *J. Chem. Soc. Chem. Commun.* **1995**, 1989.
- [7] M. A. Matchett, M. Y. Chiang, W. E. Buhro, *Inorg. Chem.* **1994**, *33*, 1109.
- [8] S. C. Goel, M. Y. Chiang, D. J. Rauscher, W. E. Buhro, *J. Am. Chem. Soc.* **1993**, *115*, 160.
- [9] B. L. Benac, A. H. Cowley, R. A. Jones, C. M. Nunn, T. C. Wright, *J. Am. Chem. Soc.* **1989**, *111*, 4986.
- [10] A. Eichöfer, J. Eisenmann, D. Fenske, F. Simon, *Z. Anorg. Allg. Chem.* **1993**, *619*, 1360.
- [11] K. Issleib, G. Döll, *Chem. Ber.* **1963**, *96*, 1544.
- [12] Kristallstruktur von 1:  $C_{152}\text{H}_{164}\text{Cd}_4\text{Li}_2\text{O}_8\text{P}_{10} \cdot 5.5\text{ THF}$ ,  $M = 3288.58$ , triklin, Raumgruppe  $P\bar{1}$ ,  $a = 17.261(6)$ ,  $b = 17.996(6)$ ,  $c = 29.470(13)$  Å,  $\alpha = 88.89(3)$ ,  $\beta = 89.14(2)$ ,  $\gamma = 61.85(2)$ °,  $V = 8070(5)$  Å<sup>3</sup>,  $Z = 2$ ,  $\rho_{\text{ber.}} = 1.353$  Mg m<sup>-3</sup>,  $\lambda = 0.71073$  Å,  $T = 153$  K,  $\mu(\text{MoK}\alpha) = 0.679$  mm<sup>-1</sup>. Die Daten wurden auf einem Siemens-Stoe-AED-Diffraktometer gesammelt, wobei ein mit Öl überzogener, rasch abgekühlter Kristall mit den Abmessungen  $0.4 \times 0.3 \times 0.3$  mm verwendet wurde, den man direkt aus der Lösung montierte [13].  $0/\omega$ -Scan ( $5^\circ \leq 2\theta \leq 45^\circ$ ). Die Absorptionskorrektur erfolgte nach einem semiempirischen Verfahren auf der Basis von  $\Psi$ -Scans mit einer maximalen und minimalen Transmission von 0.908 bzw. 0.775. 21 897 Reflexe gemessen, 21 131 unabhängige Reflexe. Die Struktur wurde mit Direkten Methoden (SHELXTL PLUS) gelöst und nach dem Kleinst-Fehlerquadrat-Verfahren auf der Basis von  $F^2$  mit voller Matrix verfeinert;  $R_1$  [ $F > 4\sigma(F)$ ] = 0.0729,  $wR_2$  (alle Daten) = 0.1656 (SHELXL-93, Göttingen, 1993). Alle Nichtwasserstoffatome im Anion von 1 wurden mit anisotropen Auslenkungsparametern verfeinert. Eines der  $[\text{Li}(\text{thf})_4]^+$ -Kationen und 3.5 der 5.5 unabhängigen nichtkoordinierenden THF-Moleküle sind auf zwei Lagen fehlgeordnet, das halbe auf dem Inversionszentrum. Alle fünf Nichtwasserstoffatome jedes THF-Moleküls in der Zelle wurden mit denselben isotropen Parametern und mit festgelegten Abständen verfeinert. Die Wasserstoffatome wurden an geometrisch sinnvollen Positionen eingesetzt. Die maximale und minimale Restelektronendichte in der letzten Differenz-Fourier-Karte betrugen 2.018 bzw.  $-1.103$  e Å<sup>-3</sup>. Weitere Einzelheiten zur Kristallstrukturuntersuchung können beim Direktor des Cambridge Crystallographic Data Centre, 12 Union Road, GB-Cambridge CB2 1EZ, unter Angabe des vollständigen Literaturzitats angefordert werden.
- [13] T. Kotke, D. Stalke, *J. Appl. Crystallogr.* **1993**, *26*, 615.
- [14] P. A. W. Dean, J. J. Vittal, N. C. Payne, *Inorg. Chem.* **1987**, *26*, 1683.
- [15] J. L. Hencher, M. A. Khan, F. F. Said, D. G. Tuck, *Polyhedron* **1985**, *4*, 1263.
- [16] P. A. W. Dean, J. J. Vittal, Y. Wu, *Can. J. Chem.* **1992**, *70*, 779.
- [17] P. A. W. Dean, N. C. Payne, J. J. Vittal, Y. Wu, *Inorg. Chem.* **1993**, *32*, 4632.
- [18] H. Bürger, W. Sawdony, U. Wannagat, *J. Organomet. Chem.* **1965**, *3*, 113.
- [19] R. Richter, J. Kaiser, J. Sieler, H. Hartung, C. Peter, *Acta Crystallogr. Sect. B* **1977**, *33*, 1887.

### Bewertung der Ähnlichkeit und Vielfalt von Verbindungsbibliotheken mit räumlichen Autokorrelationsfunktionen und neuronalen Netzen\*\*

Jens Sadowski, Markus Wagener und Johann Gasteiger\*

Professor Ivar Ugi zum 65. Geburtstag gewidmet

Seit kurzem stehen Hochleistungsmethoden des Massen-Screenings zur Identifizierung neuer Leitstrukturen zur Verfügung. Damit ist der Bedarf an einer lückenlosen Versorgung mit neuartigen Verbindungen gewachsen. Dieser Bedarf kann durch die Herstellung von chemischen Bibliotheken mit Hilfe der

[\*] Prof. Dr. J. Gasteiger, Dr. J. Sadowski, Dr. M. Wagener  
Computer-Chemie-Centrum

Institut für Organische Chemie der Universität Erlangen-Nürnberg  
Nägelsbachstraße 25, D-91052 Erlangen  
Telefax: Int. + 9131/85-6566  
E-mail: Gasteiger@EROS.CCC.Uni-Erlangen.DE

[\*\*] Diese Arbeit wurde vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) gefördert.

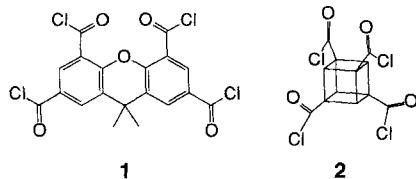
kombinatorischen Synthese<sup>[1]</sup> gedeckt werden: Die wiederholte und systematische Anwendung eines Syntheseprinzips auf eine Anzahl chemischer Bausteine ergibt große Populationen von unterschiedlichen Strukturen. So kann eine Bibliothek aus  $20^4 = 160\,000$  Verbindungen aufgebaut werden, indem 20 Bausteine in vier Reaktionsschritten systematisch miteinander verknüpft werden. Durch Änderung einiger der verwendeten Bausteine oder Reaktionen lässt sich eine fast unbegrenzte Zahl unterschiedlicher Bibliotheken verwirklichen. Daher werden effiziente Methoden benötigt, Verbindungsbibliotheken zu vergleichen und die Diversität der Verbindungen abzuschätzen. Wir stellen hier einen neuen Weg vor, Verbindungsbibliotheken zu analysieren, der auf räumlichen Autokorrelationsfunktionen und neuronalen Netzen beruht. Dabei werden mehrere Bibliotheken verglichen und ihre Ähnlichkeit und Vielfalt abgeschätzt. Ein Resultat dieser Analyse sind Empfehlungen zur geeigneten Auswahl von Bibliotheken für weitere Untersuchungen.

Vor kurzem haben wir Autokorrelationsvektoren als 3D-QSAR-Deskriptoren eingeführt. Sie kombinieren die Lage von Punkten auf einer Moleküloberfläche mit Eigenschaftswerten (z.B. dem elektrostatischen Potential) an diesen Punkten<sup>[2]</sup>. Um diesen Deskriptor zu bestimmen, werden die Abstände zwischen zufällig ausgewählten Punkten auf der Oberfläche in Intervalle mit festen Unter- und Obergrenzen,  $d_u$  bzw.  $d_o$ , aufgeteilt. Der Autokorrelationskoeffizient  $A(d_u, d_o)$  lässt sich dann gemäß Gleichung (a) ermitteln. Dabei sind  $p_i$  und  $p_j$  die Eigenschafts-

$$A(d_u, d_o) = 1/L \sum_{ij} p_i p_j (d_u \leq d_{ij} < d_o) \quad (\text{a})$$

werte an den Oberflächenpunkten  $i$  bzw.  $j$ , und  $L$  ist die Zahl der Entfernungen im Intervall  $[d_u, d_o]$ . Für eine Reihe von Distanzintervallen mit unterschiedlichen Unter- und Obergrenzen erhält man einen niederdimensionalen Vektor aus Autokorrelationskoeffizienten. Er ist eine verdichtete Darstellung – ein Fingerabdruck – der Verteilung der Eigenschaft  $p$  auf der Moleküloberfläche. Da in Gleichung (a) nur Abstände verwendet werden, müssen die 3D-Strukturen nicht in vergleichbare Positionen im dreidimensionalen Raum gebracht werden. Dieser Deskriptor ist mit Erfolg eingesetzt worden, um die biologische Aktivität zweier unterschiedlicher Klassen von Verbindungen zu modellieren<sup>[2]</sup>. Er kann so effizient bestimmt werden, daß sich auch große Zahlen an Verbindungen, aus denen kombinatorische Bibliotheken bestehen, problemlos verarbeiten lassen.

Um unsere Vorgehensweise und ihre Vorteile beim Handhaben kombinatorischer Bibliotheken zu illustrieren, haben wir ein Beispiel aus der Literatur gewählt. Rebek et al. berichteten kürzlich über die Synthese von Bibliotheken halbstarrer Verbindungen in Lösung und das anschließende Screening zur Isolierung wirksamer Trypsin-Inhibitoren<sup>[3]</sup>. Die Bibliotheken wurden aus einem starren Zentralmolekül wie dem Xanthen-derivat **1** oder dem Cubanerivat **2** und einer Mischung aus 19 L-Aminosäuren hergestellt<sup>[4]</sup>.



Die Kombination von **1** mit den 19 Bausteinen ergibt theoretisch 65 341 verschiedene Strukturen. Wegen seiner höheren Symmetrie kann **2** mit dem gleichen Satz an Synthesebausteinen nur 11 191 Verbindungen bilden. Die so gewonnenen Bibliotheken wurden mit einem Trypsinassay untersucht und anschlie-

ßend sukzessive in Teilbibliotheken aufgespalten. Dieses Screening führte schließlich zur Identifizierung einer neuen wirksamen Leitstruktur<sup>[3b]</sup>.

Verbindungen, die vom Xanthen **1** abgeleitet sind, sollten eine scheibenförmige Gestalt haben, während Derivate des Cubans **2** eher einer Kugel ähneln sollten. Daher sollte ein 3D-Deskriptor für die Behandlung derartiger kombinatorischer Bibliotheken wesentlich sein.

Um die Analyse der kombinatorischen Bibliotheken vorzubereiten, erzeugten wir zunächst für alle Verbindungen in den beiden Bibliotheken 3D-Autokorrelationsvektoren als Deskriptoren. Dazu wurden die folgenden fünf Schritte durchgeführt: 1) Der 3D-Strukturgenerator Corina<sup>[5]</sup> wurde verwendet, um dreidimensionale Strukturmodelle der beiden Zentralmoleküle **1** und **2** und der 19 L-Aminosäuren zu gewinnen. 2) Ein weiterer Strukturgenerator erzeugte die 3D-Strukturen aller möglichen Kombinationen dieser Fragmente. 3) Mit schnellen, empirischen Methoden<sup>[6]</sup> wurden Atomladungen berechnet. 4) Das elektrostatische Potential wurde nach einem klassischen Coulomb-Ansatz an Punkten ermittelt, die mit einer Dichte von 1 Punkt pro Å<sup>2</sup> zufällig auf der van-der-Waals-Oberfläche verteilt waren. 5) Die Abstände zwischen diesen Punkten wurden in zwölf gleich große Bereiche zwischen 2 und 20 Å aufgeteilt. Daraufhin wurden aus den Werten des elektrostatischen Potentials Autokorrelationskoeffizienten nach Gleichung (a) berechnet. Somit ergaben sich jeweils zwölf Werte zur Beschreibung einer Bibliotheksverbindung. Auf einem Sun-Sparc-10/512-Rechner dauerte die gesamte Rechnung ungefähr 49 Stunden für die 65 341 Verbindungen der Xanthenbibliothek und benötigte ca. 4 MB Plattspeicher für die Deskriptoren – ziemlich geringe Anforderungen an die Rechnerausstattung.

Die Frage ist nun: Wie kann man aus dieser Unmenge von Zahlen ( $65\,341 \times 12$  Deskriptoren für die Xanthenbibliothek allein) nutzbringende Einsichten in die Natur der Xanthen- und Cubanbibliothek gewinnen? Wir verwendeten ein Kohonen-Netz<sup>[7]</sup>, um die Diversität der Bibliotheksverbindungen darzustellen. Das Kohonen-Verfahren, ein nichtüberwachtes Lernverfahren, ist in der Lage, einen hochdimensionalen Deskriptorraum auf eine zweidimensionale Karte zu projizieren. Dabei bleibt die Topologie der Information erhalten, so daß Punkte, die im Deskriptorraum benachbart sind, auch benachbarte Neuronen in der zweidimensionalen Karte belegen<sup>[8]</sup>. Abbildung 1 stellt diesen Prozeß dar.

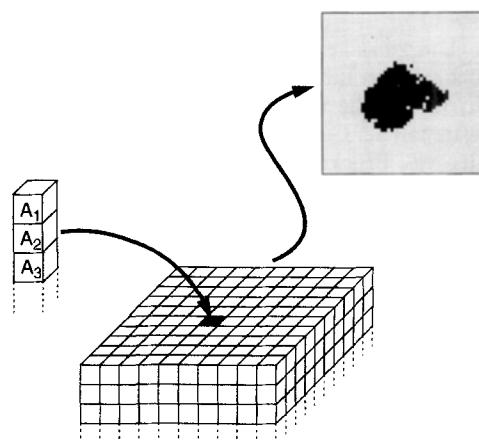


Abb. 1. Projektion der Xanthen- und Cubanbibliotheken auf eine Kohonen-Karte. Das Kohonen-Netz wird mit zwölfdimensionalen Eingabevektoren  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_{12}$  trainiert. Nach dem Training können die Neuronen dem Zentralmolekül entsprechend eingefärbt werden: Xanthen grau, Cuban schwarz.

Mit den zwölfdimensionalen Autokorrelationsvektoren A, die die vom Strukturgenerator (Schritt 2) erzeugten Verbindungen in der Xanthen- bzw. Cubanbibliothek beschreiben, wurde ein Netz mit  $50 \times 50$  Neuronen trainiert. Dies benötigte ca. 12 Minuten auf einem Sun-Sparc-10/512-Rechner. Nachdem das Training beendet war, war jede Verbindung einem bestimmten Neuron zugeordnet. Das Zentralmolekül der Bibliotheksverbindungen wurde verwendet, um die Xanthene (grau) und die Cubane (schwarz) in einer zweidimensionalen Karte anzuzeigen (Abb. 2a). Der gesamte von den beiden Bibliotheken belegte chemische Raum entspricht der  $50 \times 50$ -Karte. Alle 2500 Neuronen sind belegt. Die Cubanderivate (schwarz) bilden einen abgetrennten Cluster, der von den Verbindungen der Xanthenbibliothek umgeben ist. Nur eine geringe Zahl von Xanthen- und Cubanderivaten belegen ein und dasselbe Neuron ( $2649$  Verbindungen oder 3%). Also konnte das Netz die beiden Bibliotheken vollständig trennen. Dies legt nahe, daß es lohnend ist, beide Bibliotheken herzustellen und auf ihre Aktivität hin zu untersuchen, da sie unterschiedliche Bereiche des chemischen Raumes belegen.

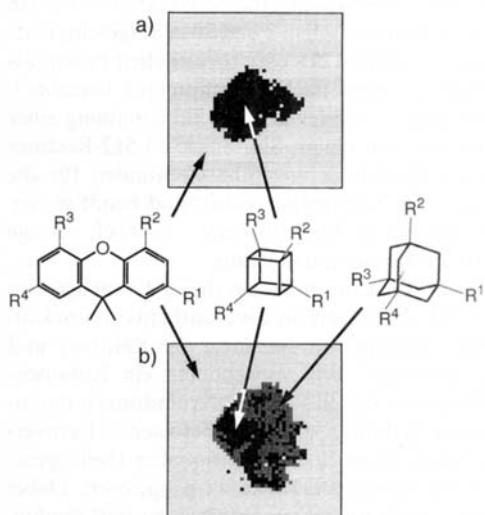


Abb. 2. a) Kohonen-Projektion der Xanthen- und der Cubanbibliothek auf eine zweidimensionale Karte. b) Kohonen-Projektion der drei Bibliotheken – Xanthen-, Cuban- und Adamantinbibliothek – auf eine zweidimensionale Karte. In beiden Fällen wird dieselbe Einfärbung verwendet: Xanthen hellgrau, Adamantan dunkelgrau, Cuban schwarz.

Um zu zeigen, daß sich mit unserer Methode auch die Ähnlichkeit von Bibliotheken bestimmen läßt, haben wir eine dritte, hypothetische Bibliothek definiert, von der wir annahmen, sie wäre mit der Cubanbibliothek eng verwandt. Wir wählten ein an den vier Brückenkopfatomen funktionalisiertes Adamantangrüss als zentralen Baustein, da damit die Aminosäuren eine räumliche Ausrichtung annehmen, die der in den Cubanderivaten ähnlich ist. Wie im Fall der Cubanbibliothek wurden 11191 Verbindungen erzeugt. Abbildung 2b zeigt die Kohonen-Karte, die mit den Daten der drei Bibliotheken trainiert wurde. Die Neuronen wurden entsprechend dem am häufigsten auftretenden Zentralmolekül eingefärbt. Wiederum läßt sich eine scharfe Grenze zwischen den Xanthenderivaten und den anderen Verbindungen beobachten. Die Cuban- und die Adamantanbibliothek jedoch konnte das Netzwerk nicht trennen; sie bilden einen einzigen Cluster in Abbildung 2b. Da 92% der Cuban- und Adamantanderivate in diesem Cluster gemeinsame Neuronen belegen, ist er noch einheitlicher, als es Abbildung 2b erscheinen läßt.

Um diesen bildlichen Eindruck quantitativ zu bestätigen, betrachteten wir Verbindungen, die in den drei Bibliotheken äquivalent sind, d. h. dasselbe Substitutionsmuster aufweisen (z.B. Ala, Gly, Leu, Pro). Ihre Entfernung in der Kohonen-Karte, die als der Median der Neuronen zwischen ihnen bestimmt wurde, sollte ein quantitatives Maß für die Ähnlichkeit der Bibliotheken sein: Xanthen – Cuban: 19 Neuronen; Xanthen – Adamantan: 19 Neuronen; Cuban – Adamantan: 4 Neuronen. Somit liegen die Cuban- und die Adamantanbibliothek dicht beieinander und belegen nahezu denselben Teil des chemischen Raums. Die eine Bibliothek sollte durch die andere ersetztbar sein, und umgekehrt. Daraus folgt, daß es nur der Mühe wert ist, die Xanthenbibliothek und *eine* der beiden anderen Bibliotheken herzustellen.

Rebek et al.<sup>[3]</sup> fanden, daß nur die Xanthenbibliothek die Trypsinaktivität signifikant verminderte. So entschlossen sie sich, diese Bibliothek für weitere Untersuchungen in Teilbibliotheken zu zerlegen. Aus dem ursprünglichen Satz von 19 Aminosäuren<sup>[4]</sup> wurde Cystein weggelassen. Die restlichen 18 Aminosäuren wurden in sechs Gruppen aufgeteilt, die jeweils drei Aminosäuren mit ähnlichen Seitenketten enthielten. Die Gruppen 1 und 2 enthielten Alkylaminosäuren, Gruppe 3 basische Aminosäuren, Gruppe 4 Hydroxy- und Methylthiofunktionalisierte Aminosäuren, Gruppe 5 aromatische Aminosäuren und Gruppe 6 Aminosäuren mit einer zweiten Säure- oder einer Amidfunktion<sup>[9]</sup>. Dann wurden Teilbibliotheken gebildet, indem jeweils die drei Aminosäuren einer Gruppe vom vollständigen Satz weggelassen wurden. Aus dem Xanthen 1 und jeweils 15 Aminosäuren ergaben sich auf diese Weise sechs Teilbibliotheken, die pro Teilbibliothek aus 25425 Verbindungen bestanden. Die Überlappung der Teilbibliotheken ist relativ hoch: Eine Teilbibliothek hat 10440 Verbindungen (41%) gemeinsam mit jeder der anderen fünf Teilbibliotheken, und jede Verbindung ist in mindestens zwei Teilbibliotheken enthalten.

Für die Untersuchung dieser Aufteilung mit dem Rechner verwendeten wir wiederum eine Kohonen-Karte. Zunächst wurde das Netz mit den Daten der 65341 Verbindungen der vollständigen Xanthenbibliothek trainiert. Dies ergab eine Karte, die dem gesamten chemischen Raum entspricht, den die Xanthenbibliothek einnimmt. Anschließend wurden im trainierten Netz diejenigen Neuronen angezeigt, die von den einzelnen Teilbibliotheken belegt werden. Abbildung 3 zeigt diese Karten und die Zahl der nicht belegten Neuronen.

Die Betrachtung der weißen Bereiche in den Karten, d. h. derjenigen Neuronen, auf die keine Verbindung projiziert wurde, vermittelt einen Eindruck von der Vielfalt der einzelnen Teilbibliotheken, und zwar gilt: Je höher die Zahl der nicht belegten Neuronen ist, um so geringer ist die Diversität der Teilbibliotheken. Trotz der hohen Redundanz der Teilbibliotheken (siehe oben) unterscheiden sich die Karten merklich. Die Teilbibliotheken 2 und 4 (Abb. 3b, d) belegen fast die gesamte Karte. Daher scheint die Diversität nicht abzunehmen, wenn die größeren Alkylaminosäuren (Gruppe 2) oder die Hydroxy- und Methylthiosubstituierten Aminosäuren (Gruppe 4) weggelassen werden. Die anderen Teilbibliotheken weisen große weiße Bereiche in ihren Karten auf – Bereiche des chemischen Raums, den sie nicht abdecken. Die meisten nicht belegten Neuronen finden sich bei der Teilbibliothek 3 (Abb. 3c). Also scheinen die basischen Aminosäuren (Gruppe 3), die in dieser Teilbibliothek fehlen, Eigenschaften zu haben, die keiner der anderen Bausteine aufweist. Den Inhalt eines willkürlich aus der Karte aller Xanthene ausgewählten Neurons, d. h. einen Satz von Verbindungen, die vom Kohonen-Netz als ähnlich angesehen wurden, zeigt Schema 1. Darunter gibt es vier Strukturpaare, die sich nur durch eine einzige Aminosäure unterscheiden (3 und 4, 5 und 6,

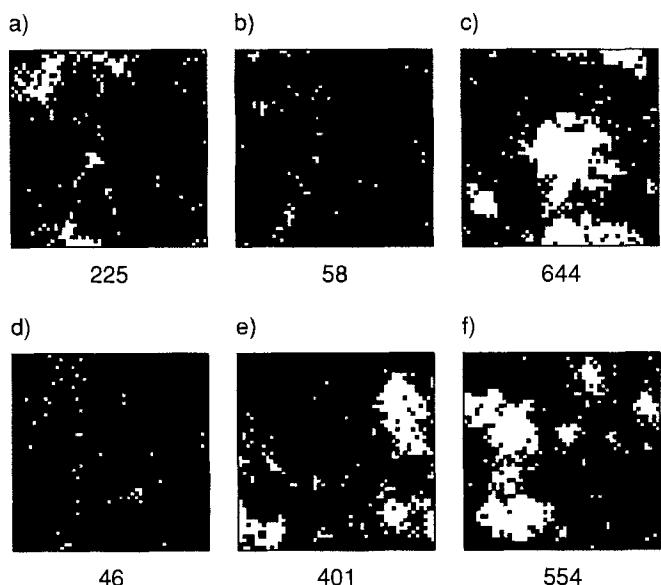
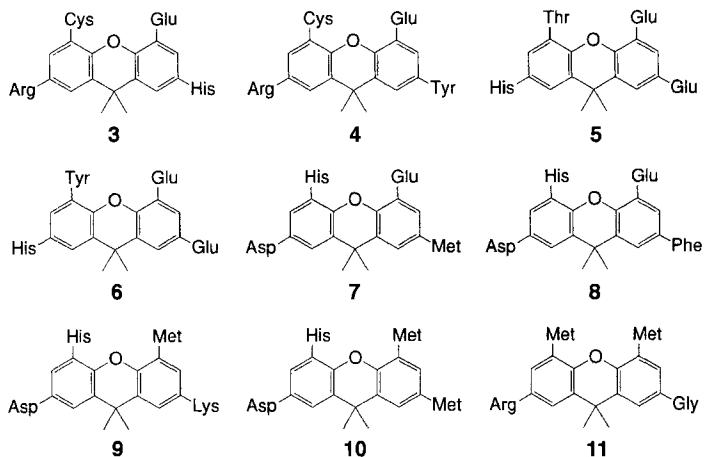


Abb. 3. Darstellung der sechs Xanthen-Teilbibliotheken in einer Karte, die mit der vollständigen Xanthenbibliothek trainiert wurde. Die Zahl der nicht belegten Neuronen ist unter den Karten angegeben. Bei a) und b) sind die Alkylaminosäuren weggelassen, bei c) die basischen Aminosäuren, bei d) die Hydroxy- und Methylthio-funktionalisierten Aminosäuren und bei f) die Aminosäuren mit einer zweiten Säure- oder einer Amidfunktion [9].



Schema 1. Neun Xanthenderivate, die das Neuron (6,13) der Kohonen-Karte, die mit der vollständigen Xanthenbibliothek trainiert wurde, belegen.

7 und 8 sowie 9 und 10 – auch ein Chemiker würde sie zusammengruppieren. Zwischen diesen vier Paaren jedoch sind die Ähnlichkeiten nicht so offensichtlich. Erst die Analyse mit dem Kohonen-Netz zeigte, daß sie sich in der Verteilung des elektrostatischen Potentials auf der Moleküloberfläche gleichen – eine Entdeckung, die sich nicht auf einfache Weise aus den Strukturformeln ableiten läßt.

Es wäre interessant, das Ergebnis dieser Art von Ähnlichkeitsanalyse mit Aktivitätsdaten zu vergleichen. Leider ist nur von zwei der Verbindungen aus der Xanthenbibliothek<sup>[3b]</sup> die genaue Aktivität bekannt. Wir fanden, daß zusammen mit dem aktivsten Inhibitor weitere 41 Verbindungen einem Neuron des Kohonen-Netzes zugeordnet wurden. Unserer Meinung nach sind diese Verbindungen ein lohnendes Ziel für weitere Untersuchungen.

Eingegangen am 17. Juli 1995 [Z 8217]

**Stichworte:** Autokorrelation · Computerchemie · Neuronale Netze · Verbindungsbibliotheken

- [1] M. A. Gallop, R. W. Barrett, W. J. Dower, S. P. A. Fodor, E. M. Gordon, *J. Med. Chem.* **1994**, *37*, 1233–1251; E. M. Gordon, R. W. Barrett, W. J. Dower, S. P. A. Fodor, M. A. Gallop, *ibid.* **1994**, *37*, 1385–1401.
- [2] M. Wagener, J. Sadowski, J. Gasteiger, *J. Am. Chem. Soc.* **1995**, *117*, 7769–7775.
- [3] a) T. Carell, E. A. Wintner, A. Bashir-Hashemi, J. Rebek, Jr., *Angew. Chem. 1994*, *106*, 2159–2162; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2059; b) T. Carell, E. A. Wintner, J. Rebek, Jr., *ibid.* **1994**, *106*, 2162–2165 bzw. **1994**, *33*, 2062; T. Carell, E. A. Wintner, A. J. Sutherland, J. Rebek, Jr., Y. M. Dunayevskiy, P. Vouros, *Chem. Biol.* **1995**, *2*, 171–183.
- [4] Dreibuchstabige Abkürzungen der 19 L-Aminosäuren, die als Bausteine eingesetzt wurden: Gly, Ala, Val, Leu, Ile, Pro, Arg, Lys, His, Ser, Thr, Met, Phe, Tyr, Trp, Glu, Asp, Asn, Cys.
- [5] J. Gasteiger, C. Rudolph, J. Sadowski, *Tetrahedron Comput. Methodol.* **1990**, *3*, 537–547; J. Sadowski, J. Gasteiger, *Chem. Rev.* **1993**, *93*, 2567–2581; J. Sadowski, J. Gasteiger, G. Klebe, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1000–1008.
- [6] J. Gasteiger, M. Marsili, *Tetrahedron* **1980**, *36*, 3219–3228; J. Gasteiger, H. Saller, *Angew. Chem.* **1985**, *97*, 699–701; *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 687–689.
- [7] T. Kohonen, *Self-Organization and Associative Memory*, 3. Aufl., Springer, Berlin, **1989**.
- [8] J. Zupan, J. Gasteiger, *Neural Networks for Chemists – An Introduction*, VCH, Weinheim, **1993**.
- [9] Die sechs Gruppen, in die die 18 L-Aminosäuren eingeteilt wurden: 1) Gly, Ala, Val, 2) Leu, Ile, Pro, 3) Arg, Lys, 4) Ser, Thr, Met, 5) Phe, Tyr, Trp, 6) Glu, Asp, Asn.

## Strukturbestimmende C–H···O–Si-Wasserstoffbrückenbindungen im Cobaltocenium-Fluorid-Nonasil\*\*

Peter Behrens\*, Gianpietro van de Goor und Clemens C. Freyhardt

Die Bildung von zeolithartigen Festkörpern durch hydrothermale Behandlung von Silicat-Gelen oder -Lösungen in Gegenwart von organischen Molekülen wird derzeit intensiv untersucht<sup>[1]</sup>. Die organischen Moleküle, beispielsweise Amine, Alkylammonium-Ionen oder cyclische Ether, werden in die Hohlräume des kristallisierenden Gerüsts eingeschlossen und beeinflussen die entstehende Struktur. Dabei bestehen allerdings kaum eindeutige Wirt-Gast-Beziehungen im Sinne eines echten Templat-Effekts. Im Einklang damit liegen die Gastmoleküle in den meisten Verbindungen bei Raumtemperatur orientierungs- oder rotationsfehlgeordnet vor<sup>[1]</sup>. Dennoch ist zumindest in eingeschränktem Maße eine Kontrolle über das Reaktionsprodukt möglich, die zum planbaren Design von Festkörperstrukturen<sup>[2]</sup> führen kann (strukturdiregierte Festkörpersynthese). Als wesentliche Parameter wurden die Größe und die Gestalt der strukturdireigend wirkenden Moleküle erkannt, die die Größe und Form der entstehenden Hohlräume im Tetraedernetzwerk der Wirtsubstanz bestimmen<sup>[3]</sup>. Einzelheiten der Wechselwirkung zwischen den organischen Gastmolekülen und dem Wirtgerüst sind aber nur in wenigen Fällen bekannt.

[\*] Prof. Dr. P. Behrens, Dipl.-Chem. G. van de Goor  
Institut für Anorganische Chemie der Universität  
Meiserstraße 1, D-80333 München  
Telefax: Int. + 89/5902-578  
E-mail: pbe@anorg.chemie.uni-muenchen.de  
Dr. C. C. Freyhardt  
Fakultät für Chemie der Universität Konstanz

[\*\*] Diese Arbeiten wurden im Arbeitskreis von Prof. J. Felsche an der Universität Konstanz durchgeführt. Wir danken ihm, der Deutschen Forschungsgemeinschaft (Be1664/1-2) und dem Fonds der Chemischen Industrie für die gewährte Unterstützung.