

organisieren, sondern die Bildung der Überstruktur dient zur Organisation der funktionellen Gruppen.

Eingegangen am 19. März 2001 [Z16805]

- [1] a) G. M. Whitesides, J. P. Mathias, C. T. Seto, *Science* **1991**, 254, 1312; b) J.-M. Lehn, *Supramolecular Chemistry*, VCH, Weinheim, **1995**.
- [2] Beispiele komplexer Überstrukturen von Stab-Knäuel-Blockcopolymeren: a) E. R. Zubarev, M. U. Pralle, L. Li, S. I. Stupp, *Science* **1999**, 283, 523; b) J. T. Chen, E. L. Thomas, C. K. Ober, G.-p. Mao, *Science* **1996**, 273, 343; c) S. A. Jenekhe, X. L. Chen, *Science* **1999**, 283, 372; d) H.-A. Klok, J. F. Langenwalter, S. Lecommandoux, *Macromolecules* **2000**, 33, 7819; e) H. Engelkamp, S. Middelbeek, R. J. M. Nolte, *Science* **1999**, 284, 785.
- [3] A. Klug, *Angew. Chem.* **1983**, 95, 579; *Angew. Chem. Int. Ed. Engl.* **1983**, 22, 565.
- [4] a) V. Percec, J. Heck, G. Johansson, D. Thomas, G. Ungar, *Macromol. Symp.* **1994**, 77, 237; b) S. D. Hudson, H.-T. Jung, V. Percec, W.-D. Cho, G. Johansson, G. Ungar, V. S. K. Balagurusamy, *Science* **1997**, 278, 449.
- [5] a) M. R. Ghadiri, J. R. Granja, R. A. Milligan, D. E. McRee, N. Khazanovich, *Nature* **1993**, 366, 324; b) G. Gattuso, S. Menzer, S. A. Nepogodiev, J. F. Stoddart, D. J. Williams, *Angew. Chem.* **1997**, 109, 1617; *Angew. Chem. Int. Ed. Engl.* **1997**, 36, 1451.
- [6] S. Höger, K. Bonrad, A. Mourran, U. Beginn, M. Möller, *J. Am. Chem. Soc.* **2001**, 123, 5651.
- [7] a) J. Zhang, J. S. Moore, *J. Am. Chem. Soc.* **1992**, 114, 9701; b) A. S. Shetty, J. Zhang, J. S. Moore, *J. Am. Chem. Soc.* **1996**, 118, 1019; c) Y. Tobe, N. Utsumi, K. Kawabata, K. Naemura, *Tetrahedron Lett.* **1996**, 37, 9325; d) Y. Tobe, N. Utsumi, A. Nagano, K. Naemura, *Angew. Chem.* **1998**, 110, 1347; *Angew. Chem. Int. Ed.* **1998**, 37, 1285; e) S. Tanikawa, J. S. Moore, *Polym. Prepr. (Am. Chem. Soc. Div. Polym. Chem.)* **1999**, 40(2), 169; f) S. Lahiri, J. L. Thompson, J. S. Moore, *J. Am. Chem. Soc.* **2000**, 122, 11315.
- [8] a) S. Höger, *J. Polym. Sci. Part A: Polym. Chem.* **1999**, 37, 2685; b) M. M. Haley, J. J. Pak, S. C. Brand, *Top. Curr. Chem.* **1999**, 201, 81; c) P. Siemsen, R. C. Livingston, F. Diederich, *Angew. Chem.* **2000**, 112, 2740; *Angew. Chem. Int. Ed.* **2000**, 39, 2632.
- [9] J. C. M. van Hest, D. A. P. Delnoye, M. W. P. L. Baars, C. Elissen-Román, M. H. P. van Genderen, E. W. Meijer, *Chem. Eur. J.* **1996**, 12, 1616.
- [10] J. S. Moore, S. I. Stupp, *Macromolecules* **1990**, 23, 65.
- [11] a) H. Kukula, U. Ziener, M. Schöps, A. Godt, *Macromolecules* **1998**, 31, 5160; b) M. Lee, B.-K. Cho, H. Kim, W.-C. Zin, *Angew. Chem.* **1998**, 110, 661; *Angew. Chem. Int. Ed.* **1998**, 37, 638; c) L. H. Radzilowski, S. I. Stupp, *Macromolecules* **1994**, 27, 7747; d) D. Marsitzky, T. Brand, Y. Geerts, M. Klapper, K. Müllen, *Macromol. Rapid Commun.* **1998**, 19, 385; e) M. A. Hempenius, B. M. W. Langeveld-Voss, J. A. E. H. van Haare, R. A. J. Janssen, S. S. Sheiko, J. P. Spatz, M. Möller, E. W. Meijer, *J. Am. Chem. Soc.* **1998**, 120, 2798.
- [12] **1a** bildet unter den gleichen Bedingungen eine Suspension.
- [13] Alle hier vorgestellten Systeme sind nicht thermotrop-flüssigkristallin; die entsprechenden Polysiloxan-substituierten Verbindungen dagegen schon: S. Höger, S. Rosselli, noch unveröffentlichte Ergebnisse.
- [14] a) H. Hoffmann, G. Hebert, *Angew. Chem.* **1988**, 100, 933; *Angew. Chem. Int. Ed. Engl.* **1988**, 27, 902; b) D. J. Abdallah, R. G. Weiss, *Adv. Mater.* **2000**, 12, 1237.
- [15] CONTIN ist eine universell einsetzbare Regularisierungsroutine mit Randbedingungen zur Invertierung von verrauchten linearen algebraischen Gleichungen und Integralen: S. W. Provencher, *Comput. Phys. Commun.* **1982**, 27, 229; S. W. Provencher, *Makromol. Chem.* **1979**, 180, 201.
- [16] Die Konzentration der untersuchten Proben betragen 0.05, 0.093 und 0.113 Gew.-%.
- [17] Die Persistenzlänge l ist ein Maß für die Rigidität eines Polymers und ist definiert durch $\langle \cos \theta(s) \rangle = \exp(-s/l)$, wobei s die Konturlänge ist. Im Falle eines supramolekularen Aggregats führt die Auswertung nach dieser Gleichung zur „virtuellen“ Persistenzlänge. Zum Vergleich: Gebräuchliche synthetische Polymere weisen eine Persistenzlänge von <1 nm auf, die Persistenzlänge von DNA liegt bei ca.

100 nm. A. Y. Grosberg, A. R. Khokhlov, *Giant Molecules*, Academic Press, London, **1997**.

- [18] Die Stablänge L wurde aus R_h mit der Gleichung $R_g/R_h = (1/\sqrt{3})(\ln(L/d) + 0.38)$ (gültig für Zylinder der Länge L und des Durchmessers d ; R_g = Gyrationradius) erhalten. In unserem Fall nehmen wir aufgrund der großen R_h -Werte $d \ll L$ an. Mit $L/d \approx 40$ ergibt sich daraus $R_g/R_h = 2.34$ (≈ 2). Mit $R_g^2 = L^2/12$ folgt $L \approx [12(2R_h)^2]^{1/2}$. a) M. Schmidt, W. H. Stockmayer, *Macromolecules* **1984**, 17, 509; b) H.-G. Elias, *Makromoleküle, Bd. 1*, Hüthig & Wepf, Basel, **1990**.
- [19] Alle Formfaktoren für Zylinder mit monodispersen Durchmesser weisen ausgeprägte Oszillationen im betrachteten q -Bereich auf.
- [20] Die verwendete Fit-Funktion ist durch den Ausdruck (1) gegeben.

$$\int_0^{\frac{\pi}{2}} \int_0^{\frac{R_a + R_i}{2}} \int_{\frac{R_a + R_i}{2}}^{\infty} \{4r_a j_1(qr_a \sin(\alpha)) - 4r_i j_1(qr_i \sin(\alpha))\} \sin(ql \cos(\alpha/2)) / \{q \sin(\alpha) q \cos(\alpha) (r_a^2 - r_i^2)\}^2 \sin(\alpha) \exp[-(r_a - R_a)^2 / (2\Delta r_a^2) - (r_i - R_i)^2 / (2\Delta r_i^2)] dr_a dr_i d\alpha \quad (1)$$

Eine noch bessere Datenanpassung wird durch ein verfeinertes Elektronendichteprofil im Bereich $2 \text{ nm} < d < 4 \text{ nm}$ erhalten. Allerdings erfordert eine detailliertere Beschreibung der Elektronendichte akkuratere Streudaten.

- [21] Vorläufige Untersuchungen ergaben, dass durch sehr langsames Eintrocknen der Polymerfilme („Gleichgewichtsbedingungen“) im festen Zustand eine andere Überstruktur erhalten wird. Daher können die hier vorgestellten Untersuchungen in Lösung nicht auf den festen Zustand übertragen werden.
- [22] Gleichzeitig bildet ein Teil der Blockcopolymere eine gewellte Untergrundstruktur auf dem Glimmer.
- [23] a) S. Höger, V. Enkelmann, *Angew. Chem.* **1995**, 107, 2917; *Angew. Chem. Int. Ed. Engl.* **1995**, 34, 2713; b) S. Höger, V. Enkelmann, K. Bonrad, C. Tschierske, *Angew. Chem.* **2000**, 112, 2355; *Angew. Chem. Int. Ed.* **2000**, 39, 2268.
- [24] a) P. Dziezok, S. S. Sheiko, K. Fischer, M. Schmidt, M. Möller, *Angew. Chem.* **1997**, 109, 2894; *Angew. Chem. Int. Ed. Engl.* **1997**, 36, 2812; b) A. D. Schlüter, J. P. Rabe, *Angew. Chem.* **2000**, 112, 860; *Angew. Chem. Int. Ed.* **2000**, 39, 864.

Von der Struktur zur Funktion: ein neuer Ansatz zum Erkennen von funktionellen Ähnlichkeiten zwischen Proteinen unabhängig von Sequenz- und Faltungshomologie

Stefan Schmitt, Manfred Hendlich und Gerhard Klebe*

Das rasante Anwachsen von Informationen über Sequenz und Struktur von Proteinen erfordert neue Computermethoden, die es erlauben, von der Struktur eines Proteins auf dessen Funktion zu schließen. Die molekulare Funktion eines Proteins ist in aller Regel direkt an die spezifische Erkennung von Substraten oder endogenen Liganden in einer wohldefinierten Proteinbindetasche gekoppelt. Proteine mit verwandten Funktionen sollten daher auch ähnliche Bindeta-

[*] Prof. Dr. G. Klebe, S. Schmitt, M. Hendlich
Philipps-Universität Marburg
Institute für Pharmazeutische Chemie
Marbacher Weg 6, 35032 Marburg (Deutschland)
Fax: (+49) 6421-28-2-8994
E-mail: Klebe@mail.uni-marburg.de

schen aufweisen. Wir haben einen neuen Ansatz entwickelt, der Aminosäuren, die eine Bindetasche begrenzen, durch einfache physikochemische Deskriptoren repräsentiert. Die abgeleiteten Eigenschaftsdeskriptoren können anschließend benutzt werden, in den Taschen nach gemeinsamen Substrukturen zu suchen und so Ähnlichkeiten zu finden. Die erhaltenen Lösungen werden nach dem Grad ihrer Ähnlichkeit bewertet und erlauben es, funktionelle Verwandtschaften zwischen Proteinen aufzufinden - unabhängig von einer möglicherweise gegebenen Sequenz- oder Faltungshomologie.

Die Entschlüsselung des menschlichen Genoms ist nur der erste Schritt zum Verständnis des funktionellen und strukturellen Zusammenspiels von Proteinen in biologischen Systemen. Durch methodische Entwicklungen auf dem Gebiet der „structural genomics“, wie die Hoch-Durchsatz-Röntgenkristallographie,^[1] werden wir in Zukunft vermehrt mit neuen 3D-Proteinstrukturen konfrontiert, für die zunehmend ihre eigentliche biochemische Funktion bei der Strukturbestimmung noch nicht bekannt ist. Darüber hinaus werden methodische Verbesserungen der Strukturvorhersage und Homologie-Modellierung dazu führen, dass es möglich wird, die grobe Proteingeometrie korrekt vorherzusagen.^[2] Methoden, die die Proteinfunktion aus ihrer 3D-Struktur ablesen können, werden daher dringend benötigt. Da die Funktion eines Proteins nicht notwendigerweise durch ein bestimmtes Faltungsmuster festgelegt wird und auch ein gleiches Faltungsmuster nicht für dieselbe Funktion stehen muss,^[3] sind diese Vorhersagen keinesfalls trivial.

Die Proteinfunktion ist oft eng an die Erkennung und/oder die chemische Modifikation eines endogenen Liganden oder Substrats gekoppelt.^[4, 5] Der Prozess der molekularen Erkennung findet üblicherweise in gut charakterisierten Spalten oder Vertiefungen auf der Proteinoberfläche statt.^[6] Eine enzymatische Reaktion benötigt eine definierte räumliche Anordnung der Reaktionspartner zueinander. Dies verlangt umgekehrt auf der Proteinseite eine hochkonservierte räumliche Anordnung der Stellen, an denen der für die Reaktion

erforderliche Prozess der molekularen Erkennung stattfindet. Die Vorstellung, dass ein molekulares Erkennungsmuster möglicherweise zwischen verschiedenen Proteinen ähnlicher Funktion konserviert sein könnte, hat uns veranlasst, einen neuen Ansatz zu entwickeln, um Proteinstrukturen miteinander zu vergleichen. Anstatt wie allgemein üblich gesamte Proteine anhand ihrer Sequenzen zu vergleichen, suchen wir nach Ähnlichkeiten physikochemischer Eigenschaften, die auf der Oberfläche der Bindetaschen präsentiert werden.

In der Literatur sind verschiedene Methoden beschrieben, die Vertiefungen auf der Proteinoberfläche als potentielle Bindestellen erkennen.^[7–11] Wir haben zur Identifizierung von Bindestellen das Programm LIGSITE^[12] verwendet und dessen Algorithmus in die Protein-Ligand Datenbank RELIBASE^[13, 14] implementiert. Zunächst wird jedes Protein in ein Gitter vorgegebener Maschenweite eingebettet. Jeder Gitterpunkt, der mit einem Proteinatom zusammenfällt, also innerhalb dessen van-der-Waals-Radius liegt, bleibt unberücksichtigt. Die verbleibenden Gitterpunkte werden nach dem Grad ihrer „Vergrabenheit“ im Rezeptor bewertet. Benachbarte Gitterpunkte mit hoher Vergrabenheit werden zu zusammenhängenden Bindetaschen verschmolzen und in der objektorientierten Datenbank CAVBASE abgelegt. Die Randpunkte einer solchen Vereinigungsmenge von Gitterpunkten beschreiben angenähert die Oberfläche einer Tasche. Die Atomkoordinaten der Aminosäuren, die die Bindetasche begrenzen, werden auf einen Satz repräsentativer Pseudozentren reduziert. Diese Pseudozentren werden nach fünf Eigenschaften klassifiziert, die essentiell für die molekulare Erkennung sind: H-Brückendonor-, -acceptor-, ambivalente Donor/Acceptor- (Hydroxygruppen oder Stickstoffatome der Histidinseitenkette), hydrophobe aliphatische und aromatische Eigenschaften.^[15] Die Pseudozentren werden außerdem noch dahingehend untersucht, ob sie ihre Eigenschaften auf die Bindetaschenoberfläche ausrichten. Dieses Verhalten wird dann den benachbarten Gitter-Oberflächenpunkten zugewiesen (Abbildung 1). Die Form der Bindetasche, die zugeordneten Deskriptoren, deren Oberflächenzugänglich-

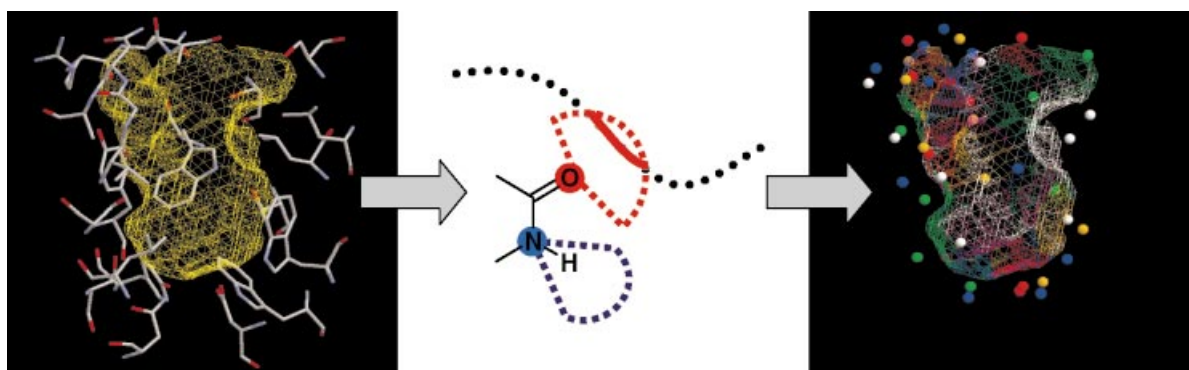


Abbildung 1. Eine Bindetasche, deren Form und Eigenschaften zunächst durch einen Satz von Aminosäuren bestimmt werden (links), wird in einen Satz von Pseudozentren mit H-Brückendonor- (blau), -acceptor- (rot), ambivalente Donor/Acceptor- (grün), hydrophobe aliphatische (weiß) und aromatische Eigenschaften (orange) übersetzt (rechts, farbige Kugeln). Jedes Pseudozentrum kann seine Eigenschaft in einen bestimmten Raumbereich exponieren und somit auf die Bindetaschenoberfläche projizieren. Es werden aber nur solche Pseudozentren berücksichtigt, die ihre Eigenschaft auf die Bindetaschenoberfläche ausrichten (in der Mitte gezeigt für die Acceptor- und die Donoreigenschaft einer Peptidbindung, wobei hier die Donorgruppe nicht als Pseudozentrum berücksichtigt wird, da sie nicht auf die Oberfläche ausgerichtet ist). Die jeweiligen Oberflächenbereiche beschreiben die Zugänglichkeit der einzelnen Pseudozentren. Sie werden ebenfalls abgespeichert und dienen später dazu, die Ähnlichkeit zwischen Bindetaschen zu bewerten. Auf diese Weise wird die gesamte Tasche in Pseudozentren mit zugewiesenen Eigenschaften übersetzt und mit den dazugehörigen Oberflächenbereichen verknüpft (für die Pseudozentren (Kugeln) und die Oberflächenanteile (Netzstruktur) wurde die gleiche Farbkodierung verwendet).

keit und die zugewiesenen Oberflächenbereiche werden zusammen mit Informationen über den ggf. gebundenen Liganden in CAVBASE abgespeichert. Die Protein- und Ligandinformationen in CAVBASE sind mit denen in RELIBASE so verknüpft, dass auf die korrespondierenden Objekte in RELIBASE zugegriffen werden kann. Ein Datensatz von 31441 Taschen wurde aus 8308 Proteinstrukturen ausgeschnitten. Die neue Datenbank wurde mit Werkzeugen zur interaktiven Visualisierung einzelner Bindetaschen sowie der Ergebnisse von Bindetaschenvergleichen versehen.

Die gespeicherten Informationen erlauben einen schnellen und effizienten Vergleich von Bindetaschen in einer großen Datenmenge. Ein Clique-Algorithmus wird angewendet, um gemeinsame Subgraphen zu finden. Diese Subgraphen werden aus den Knoten korrespondierender Pseudozentrenpaare mit gleichen Eigenschaften und ähnlichem Abstandsmuster erzeugt.^[16] Geeignete Toleranzen wurden angewandt, um der konformativen Flexibilität von Proteinen (z.B. Induced-fit-Konformationsänderungen, Bewegung von Domänen) Rechnung zu tragen und die prinzipiellen Beschränkungen der Genauigkeit von Proteinstrukturen zu berücksichtigen. Mehrere Lösungen werden generiert und dahingehend bewertet, inwieweit Oberflächenbereiche zweier Taschen so überlagert werden können, dass größtmögliche Flächenbereiche mit gleichen physikochemischen Eigenschaften überlappen.

Auf einem üblichen Einzelprozessorrechner benötigt der direkte Vergleich zweier mittelgroßer Bindetaschen (ca. 800 Å³) ungefähr 100 CPU-Sekunden.^[17] So ist unsere Methode in der Lage, die Bindetasche der Suchanfrage gegen eine Auswahl von mehreren tausend Bindetaschen – parallel auf mehreren Rechnern – zu vergleichen. Um den ursprünglichen Datensatz der CAVBASE auf eine handhabbare Größe von 5000–6000 Einträgen zu reduzieren, wurden die Daten nach Kriterien wie mittlere Taschengröße und Auflösung der Proteine gefiltert. In den nachfolgenden Suchbeispielen wurden zusätzlich Bindetaschen von Proteinen mit einer hohen Sequenzidentität zum Protein der Suchanfrage verworfen, um triviale Lösungen zu vermeiden. Die am besten bewerteten Lösungen weisen sowohl eine hohe lokale Oberflächenähnlichkeit als auch ein gemeinsames Bindungsmotiv mit der Tasche der Suchanfrage auf.

Da dieser Ansatz allein auf physikochemischen Eigenschaften beruht, die zur Oberfläche der Bindetasche hin exponiert werden, ist es möglich, Verwandtschaften unabhängig von einer gegebenen Sequenz- oder Faltungshomologie zu finden. Dies unterscheidet unseren Ansatz von anderen,^[18–23] die mehr oder weniger stark auf Informationen aus dem Faltungsmuster zurückgreifen. Kürzlich berichteten Stahl et al.^[24] über die automatische Gruppierung von 176 Zinkmetalloproteinasen mit Hilfe eines selbstorganisierenden Neuronalen Netzes. Dazu wurden Oberflächen-zugängliche Muster physikochemischer Eigenschaften in Bindetaschen ausgewertet. Der Ansatz separiert die katalytischen Zentren dieser Enzyme klar von anderen Cavitäten auf deren Oberfläche.

In unserem ersten Beispiel haben wir die Bindetaschen zweier Chorismat-Mutasen aus *Saccharomyces cerevisiae* und *Escherichia coli* verglichen.^[25, 26] Dieses Beispiel wurde auch von Rosen et al.^[19] nach einem anderen Verfahren zur Beschreibung einer Bindetasche untersucht, bei dem so genannte „sparse critical points“^[27] als Bindetaschen-Deskriptoren verwendet werden. Der gebundene Ligand – ein bicyclischer Inhibitor, der dem Übergangszustand analog ist – wird größtenteils über Seitenkettenwechselwirkungen erkannt. Ein Beispieldatensatz aus 5445 Bindetaschen, der die Chorismat-Mutase von *E. coli* enthält, wurde mit der Bindetasche von *S. cerevisiae* verglichen. Die *E. coli*-Bindetasche wurde dabei als beste Lösung gefunden. Abbildung 2 zeigt die überlagerten Pseudozentren und ihre korrespondierenden Oberflächenbereiche. Obwohl keine Informationen über die gebundenen Liganden verwendet wurden, werden die Liganden durch die Transformation auf der Basis der Oberflächenanteile nahezu perfekt einander überlagert.

In einem zweiten Beispiel wurde eine Bindetasche von Trypsin mit einem Datensatz aus 5248 Taschen verglichen, die zu Trypsin eine Sequenzidentität von weniger als 35 % aufweisen. Die als beste Lösungen bewerteten Bindetaschen repräsentieren andere Mitglieder der Trypsin-Familie und zeigen eine stetig abnehmende Homologie zu Trypsin. Auf Position 114 – also unter den ersten 3 % der betrachteten Vergleichsdaten – wurde die Bindetasche eines Enzyms aus der strukturell nicht verwandten Subtilisin-Familie gefunden. Auf den nächsten Rängen folgten weitere Bindetaschen der Subtilisin- und der Trypsin-Familie. Die Trypsin- und Subtilisin-Superfa-

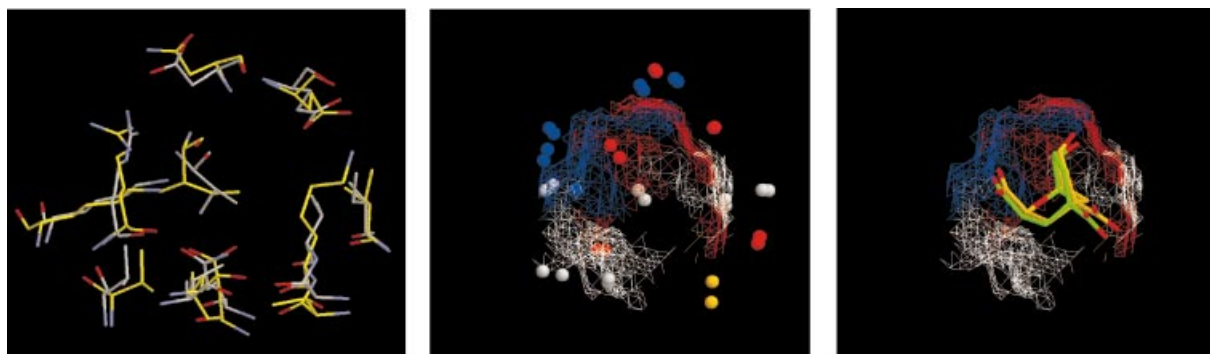


Abbildung 2. Darstellung der Bindetaschen der Chorismat-Mutasen aus *S. cerevisiae* (4csn) und *E. coli* (1ecm), die Proteine weisen untereinander eine Sequenzhomologie von nur 20 % auf. Die gebundenen Liganden werden fast ausschließlich durch Wechselwirkungen mit den Seitenkettenresten erkannt (links). Die Tasche von *E. coli* wurde mit einem Datensatz aus 5445 verschiedenen Taschen unter Verwendung der Pseudozentrenbeschreibung verglichen (Mitte). Rechts sind die überlagerten Bindetaschen zusammen mit den überlappenden gemeinsamen Oberflächenbereichen und den gebundenen bicyclischen Übergangszustands-analogen Inhibitoren abgebildet.

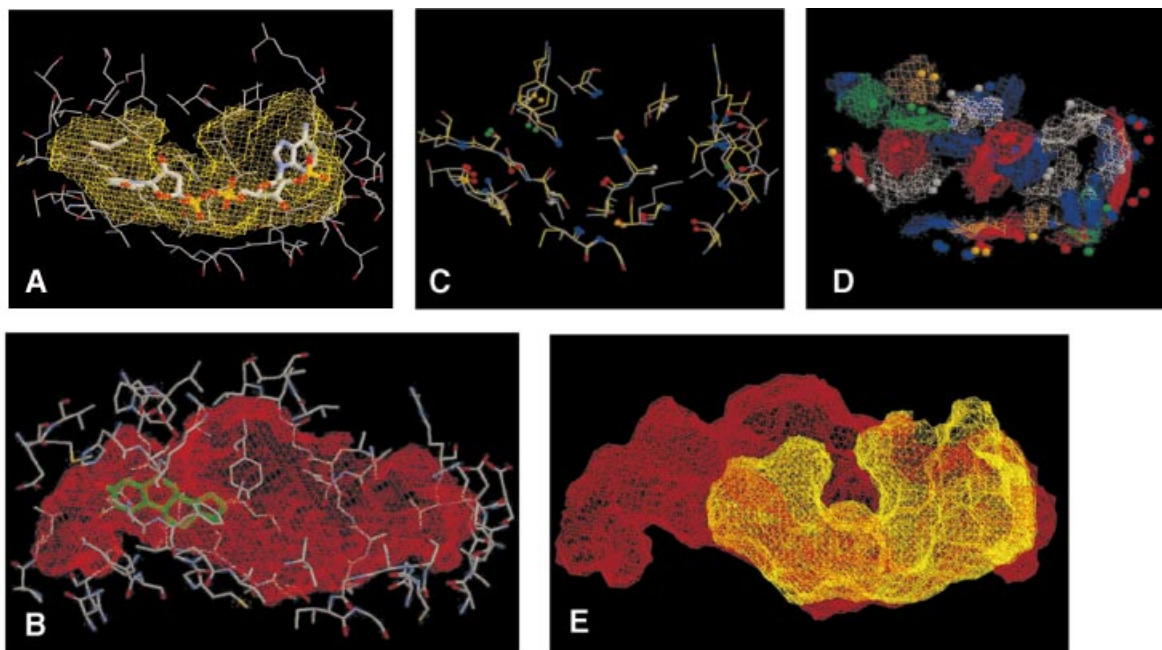


Abbildung 3. Die NADPH-Cofaktor-Bindestelle der Carbonyl-Reduktase (1cyd; A) wurde mit einem Datensatz aus 5377 Taschen verglichen. Der Algorithmus entdeckt auf Position 41 eine ähnliche Bindetasche aus einer NADPH-abhängigen Steroid-Dehydrogenase (1dfs; B). C und D zeigen die zueinander äquivalenten Aminosäuren, Pseudozentren bzw. Oberflächenanteile. Die Kristallstruktur der Steroid-Dehydrogenase wurde ohne den gebundenen Cofaktor bestimmt, jedoch lässt sich die Cofaktor-Bindestelle der Carbonyl-Reduktase (A, gelb) gut mit der unbesetzten Tasche der Steroid-Dehydrogenase (B, rot) überlagern (E).

milien gehören zur großen Klasse der Serinproteasen. Sie sind untereinander zwar in funktioneller Hinsicht ähnlich, jedoch fehlt ihnen eine Sequenz- und Faltungshomologie.

Die gesamte NADPH-Cofaktor-Bindestelle einer Carbonyl-Reduktase^[28] wurde als Suchanfrage benutzt und mit einem Datensatz aus 5377 Bindetaschen verglichen. Proteine dieses Datensatzes weisen keine signifikante Sequenzhomologie mit dem Referenzprotein auf. Die Bindetaschen der besten Lösungen haben ebenfalls NADPH als Ligand gebunden. Auf den weiteren Rängen folgten Bindetaschen mit Liganden, die sich von NADPH zunehmend unterscheiden, aber noch Teile des NADPH-Gerüsts enthalten. Auf Position 41 wurde die große Bindetasche der NADPH-abhängigen Steroid-Dehydrogenase^[29] (Abbildung 3) entdeckt, obwohl die Kristallstruktur dieses Enzyms ohne den gebundenen Cofaktor bestimmt wurde. Auf Position 116 folgte eine Phenolhydrolase, die FAD als Cofaktor gebunden hat. In diesem Fall weist die Adenin-Erkennungstasche Ähnlichkeiten mit der NADPH Bindetasche des Referenzproteins auf.

Als überraschendes Ergebnis entdeckten wir die Oberflächenähnlichkeit einer Adenin-Bindestelle mit einer unbesetzten Bindestelle in einer HIV-Protease-Struktur. In dieser Struktur ist ein makrocyclisches Peptidmimetikum im aktiven Zentrum gebunden.^[30] Der in beiden Bindetaschen als gemeinsam entdeckte Bereich bleibt jedoch in der Protease unbesetzt. In der Proteinkinase-A-Struktur bindet dort allerdings der Adeninrest des Adenyliminodiphosphats.^[31] Dieser Hinweis kann für das Drug Design sehr wertvoll sein, da der Adenylteil als Ergänzung für den makrocyclischen Inhibitor im unbesetzten Teil der HIV-Protease-Bindestelle dienen könnte, um so einen wirksameren Inhibitor zu entwickeln. Möglicherweise lassen sich durch Suchen in einer großen Datenbank von Bindetaschen interessante Vorschläge

für neue Molekülbausteine auffinden, die sich als potentielle Bioisostere für neuartige Leitstrukturen eignen.

Eingegangen am 19. März 2001 [Z16803]

- [1] V. S. Lamzin, A. Perrakis, *Nat. Struct. Biol.* **2000**, 7, 978–981.
- [2] S. Roberto, U. Pieper, F. Melo, N. Eswar, M. A. Marti-Renom, M. S. Madhusudh, *Nat. Struct. Biol.* **2000**, 7, 986–990.
- [3] J. M. Thornton, A. E. Tod, D. Milburn, N. Borkakoti, C. A. Orengo, *Nat. Struct. Biol.* **2000**, 7, 991–994.
- [4] S. L. Moodie, J. B. Mitchell, J. M. Thornton, *J. Mol. Biol.* **1996**, 263, 486–500.
- [5] R. E. Babine, S. L. Bender, *Chem. Rev.* **1997**, 97, 1359–1472.
- [6] F. K. Pettit, J. U. Bowie, *J. Mol. Biol.* **1999**, 285, 1377–1382.
- [7] I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, T. E. Ferrin, *J. Mol. Biol.* **1982**, 161, 269–288.
- [8] R. A. Laskowski, *J. Mol. Graph.* **1995**, 13, 2735–2748.
- [9] G. L. Levitt, L. J. Banaszak, *J. Mol. Graph.* **1992**, 10, 229–234.
- [10] J. Liang, H. Edelsbrunner, C. Woodward, *Protein Sci.* **1998**, 7, 1884–1897.
- [11] K. P. Peters, J. Fauck, C. Frommel, *J. Mol. Biol.* **1996**, 256, 201–213.
- [12] M. Hendlich, F. Rippmann, G. Barnickel, *J. Mol. Graph.* **1997**, 15, 359–363.
- [13] K. Hemm, K. Aberer, M. Hendlich, *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1995**, 3, 170–178.
- [14] M. Hendlich, *Acta Crystallogr.* **1998**, 54, 1178–1182.
- [15] Die Pseudozentren werden so platziert, dass sie möglichst gut die Wechselwirkungseigenschaften einer zur Oberfläche der Bindetasche exponierten funktionellen Gruppe wiedergeben. Das Pseudozentrum für einen H-Brückenacceptor wie eine Carbonylgruppe wird durch das Carbonylsauerstoffatom repräsentiert, die aromatischen Eigenschaften eines Phenylalaninrestes durch das geometrische Mittel der Koordinaten der sechs Phenylkohlenstoffatome. Weiterhin werden Zentren nur berücksichtigt, wenn sie so zur Oberfläche exponiert werden, dass sie eine funktionelle Gruppe des Liganden in einem bestimmten Winkelsegment erreichen können. Diese Segmente wurden mit dem Programm IsoStar anhand von Kristallstrukturen kalibriert (I. J. Bruno, J. Cole, J. Lommerse, R. S. Rowland, R. Taylor, M. L. Verdonk, *J. Comput. Aided Mol. Des.* **1997**, 11, 525–537).

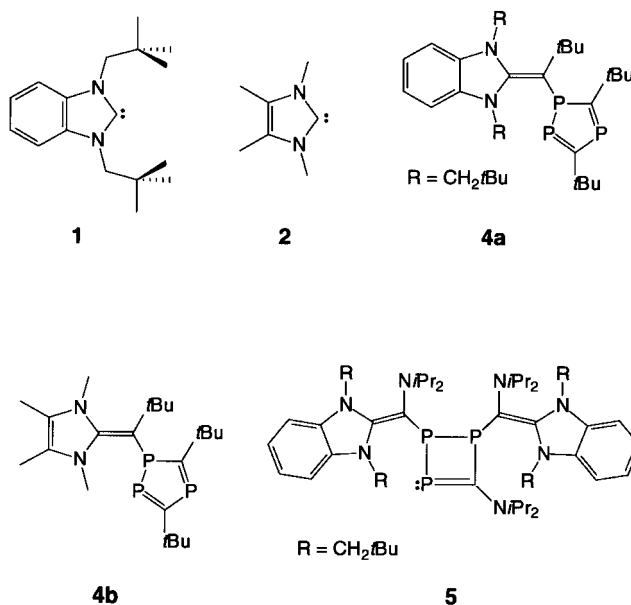
- [16] C. Bron, J. Kerbosch, *Commun. ACM* **1973**, *16*, 575–577.
 [17] Verwendet wurde ein Pentium-III-Prozessor (650 MHz) und das Betriebssystem Linux; CPU = central processing unit.
 [18] D. Fischer, H. Wolfson, S. L. Lin, R. Nussinov, *Protein Sci.* **1994**, *3*, 769–778.
 [19] M. Rosen, S. L. Liang, H. Wolfson, R. Nussinov, *J. Mol. Biol.* **1998**, *11*, 263–277.
 [20] D. Fischer, R. Norel, H. Wolfson, R. Nussinov, *Proteins* **1993**, *16*, 278–292.
 [21] A. C. Wallace, N. Borkakoti, J. M. Thornton, *Protein Sci.* **1997**, *6*, 2308–2323.
 [22] P. J. Artymiuk, A. R. Poirrette, H. M. Grindley, D. W. Rice, P. Willett, *J. Mol. Biol.* **1994**, *243*, 327–344.
 [23] R. B. Russell, *J. Mol. Biol.* **1998**, *279*, 1211–1227.
 [24] M. Stahl, C. Taroni, G. Schneider, *Protein Eng.* **2000**, *13*, 83–88.
 [25] A. Y. Lee, P. A. Karplus, B. Ganem, J. Clardy, *J. Am. Chem. Soc.* **1995**, *117*, 3627–3628.
 [26] N. Strater, G. Schnappauf, G. Braus, W. N. Lipscomb, *Structure* **1997**, *5*, 1437–1452.
 [27] S. L. Lin, R. Nussinov, D. Fischer, H. J. Wolfson, *Proteins* **1994**, *18*, 94–101.
 [28] N. Tanaka, T. Nonaka, M. Nakanishi, Y. Deyashiki, A. Hara, Y. Mitsui, *Structure* **1996**, *4*, 33–45.
 [29] R. Breton, D. Housset, C. Mazza, J. C. Fontecilla-Camps, *Structure* **1996**, *4*, 905–915.
 [30] J. L. Martin, J. Begun, A. Schindeler, W. A. Wickramasinghe, D. Alewood, P. F. Alewood, D. A. Bergman, R. I. Brinkworth, G. Abbenante, D. March, R. C. Reid, D. Fairlie, *Biochemistry* **1999**, *38*, 7978–7988.
 [31] D. Bossemeyer, R. A. Engh, V. Kinzel, H. Ponstingl, R. Huber, *EMBO J.* **1993**, *12*, 849–859.

Ungewöhnliche Bildung eines Azaphospholens aus 1,3,4,5-Tetramethylimidazol-2-yliden und Diisopropylaminophosphaethin**

F. Ekkehardt Hahn,* Duc Le Van, Michelle C. Moyes, Thorsten von Fehren, Roland Fröhlich und Ernst-Ulrich Würthwein

Bei ersten Studien zur Reaktivität von N-heterocyclischen Carbenen^[1] gegenüber Phosphaalkinen fanden wir vor kurzem, dass die anellierte Verbindung *N,N'*-Bis(2,2-dimethylpropyl)benzimidazolin-2-yliden **1**^[2] mit $\text{P}=\text{CtBu}$ **3a** quantitativ zum Triphosphol **4a** und mit $\text{P}=\text{CNiPr}_2$ **3b** zum 1,2,3-Triphospheten **5** reagiert.^[3] Parallel hierzu erhielten Nixon et al.^[4] das Triphospholderivat **4b** durch die Umsetzung des

Arduengo-Carbens 1,3,4,5-Tetramethylimidazol-2-yliden **2** mit 2,4,6-Tri-*tert*-butyl-1,3,5-triphasphabenzol. Diese Ergebnisse deuten auf eine vorzugweise Bindungsverknüpfung des



Carbens mit dem ungesättigten Kohlenstoffatom der P-C-Mehrfachbindungssysteme hin. Auch die Reaktion des Bertrandischen Phosphanylsilylcarbens mit dem Phosphaalkin **3a** verläuft in diesem Sinn: Das primär gebildete 2-Phosphanyl-2*H*-phosphiren stabilisiert sich durch Ringöffnung unter Erhaltung der C-C_{Carben}-Bindung und Bruch der P-C_{Carben}-Bindung zum 1,2,3-Diphospheten.^[5] Wir berichten hier über die Synthese eines isolierbaren 1:1-Adduktes aus dem N-heterocyclischen Carben **2** und dem Phosphaalkin $\text{P}=\text{CNiPr}_2$ **3b**, die unter Bildung einer P-C_{Carben}-Bindung verläuft, und schlagen auf der Basis quantenchemischer Berechnungen einen Mechanismus für diese neuartige Reaktion vor.

Während die Reaktion des Carbens **2** mit **3a** in Toluol bei Raumtemperatur das bekannte Triphosphol **4b**^[4] als Hauptprodukt (ca. 70 %) liefert, entsteht bei der analogen Umsetzung von **2** mit dem Aminophosphaalkin **3b** der Bicyclus **7** in nahezu quantitativer Ausbeute (Schema 1).

Verbindung **7** wurde in Form hellgelber, luft- und feuchtigkeitsempfindlicher Mikrokristalle isoliert. Im ¹³C-NMR-Spektrum tritt das Signal des Carbenzentrums bei $\delta = 178.9$ (¹J(P,C) = 81.5 Hz) auf und ist damit im Vergleich zu dem von **2** stark zu höherem Feld verschoben ($\Delta\delta = 33.8$). Das ³¹P{¹H}-NMR-Signal erscheint als Singulett bei $\delta = -64.8$ und liegt damit im charakteristischen Bereich für C-Diaminophosphaalkene.^[6] Die Bindungssituation im neuartigen Azaphospholen **7** lässt sich am besten mit den beiden Grenzformeln **7A** und **7B** beschreiben, wobei **7B** ein invers polarisiertes Phosphaalken mit π -Elektronendichteverteilung $\text{P}^{\delta-}=\text{C}^{\delta+}$ ist.^[7] Eine wichtige Stütze für diese Interpretation liefert die Derivatisierung von **7** mit der Lewis-Säure $\text{BH}_3 \cdot \text{THF}$. Dabei entsteht ausschließlich der Komplex **8** mit zwei am Phosphoratom koordinierten BH_3 -Gruppen (Schema 2). Ein analoger Di(boran)-Komplex mit einem acyclischen Diaminophosphaalken wurde von Cowley et al.^[8] beschrieben.

[*] Prof. Dr. F. E. Hahn, Dr. D. Le Van, M. C. Moyes, T. von Fehren
 Anorganisch-chemisches Institut der Universität
 Wilhelm-Klemm-Straße 8, 48149 Münster (Deutschland)
 Fax: (+49)251-833-3108
 E-mail: fehahn@uni-muenster.de

Dr. R. Fröhlich, Prof. Dr. E.-U. Würthwein
 Organisch-chemisches Institut der Universität
 Corrensstraße 40, 48149 Münster (Deutschland)
 Fax: (+49)251-833-3293

[**] Diese Arbeit wurde von der Deutschen Forschungsgemeinschaft und vom Fonds der Chemischen Industrie unterstützt.

Hintergrundinformationen zu diesem Beitrag sind im WWW unter <http://www.angewandte.de> zu finden oder können beim Autor angefordert werden.