

Ligandengestützte Homologiemodellierung von G-Protein-gekoppelten Rezeptor-Bindungstaschen: Modelle für erfolgreiches virtuelles Screening**

Andreas Evers und Gerhard Klebe*

Durch die Fortschritte in den Genomsequenzierungsprojekten werden in nächster Zeit Tausende von Proteinsequenzen entschlüsselt werden, die als Zielproteine für die Arzneimittelforschung dienen können.^[1] Dies verlangt nach Computermethoden, die es ermöglichen, die Struktur eines Proteins so genau vorherzusagen, dass sie als Grundlage für strukturbasiertes Wirkstoffdesign verwendet werden kann. Wir haben eine Methode zur Erstellung von Proteinbindungstaschen-Modellen entwickelt. Durch die Berücksichtigung von Liganden während der Homologiemodellierung gelingt es, Modelle höherer Genauigkeit und Relevanz zu erzeugen.^[2] Zunächst generierte, grobe Homologiemodelle des Zielproteins werden in einem iterativen Verfeinerungsverfahren optimiert, indem Informationen über bioaktive Liganden als zusätzliche Randbedingung in den Modellierungsprozess einbezogen werden (Abbildung 1).

Die Parametrisierung und Validierung des MOBILE-Verfahrens (*Modeling Binding Sites Including Ligand Information Explicitly*) erfolgte durch die Vorhersage von Proteinbindungstaschen mit bekannter Raumstruktur. Diese Methode wendeten wir anschließend zum Entdecken neuer Antagonisten des Neurokinin-1(NK1)-Rezeptors an. Für dieses Protein, das zur Familie der G-Protein-gekoppelten Rezeptoren (GPCRs) gehört, wurde ein neuer Ligand mit submikromolarer Bindungsaffinität entdeckt.

Die Kenntnis der Raumstruktur des Zielproteins ist eine Grundvoraussetzung für das strukturbasierte Wirkstoffdesign. Die Strukturaufklärung erfolgt durch Röntgenstrukturanalyse und/oder NMR-Spektroskopie. Allerdings ist die Geschwindigkeit, mit der die Primärsequenzen neuer Proteine aufgeklärt werden, um ein Vielfaches höher als die der experimentellen Bestimmung ihrer Raumstrukturen. Somit wird für einen Großteil neuer potenzieller Zielproteine die Raumstruktur für das Wirkstoffdesign nicht unmittelbar zur Verfügung stehen. In einer solchen Situation ist die Homologiemodellierung die zuverlässigste computergestützte Methode zur Vorhersage der Raumstruktur.^[7] Sie baut allerdings nur auf den Informationen auf, die durch verwandte Proteinstrukturen zur Verfügung stehen. Dementsprechend ist dieser Ansatz bis zu einem gewissen Grad approximativ, vor allem wenn in der Bindungstasche des

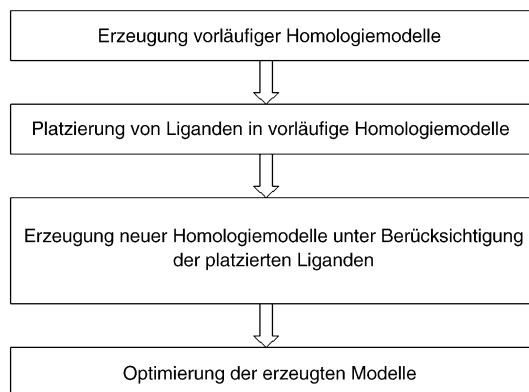


Abbildung 1. Schematischer Überblick über das Konzept des MOBILE-Ansatzes. Beginnend mit der Struktur eines oder mehrerer Protein-Vorlagen im Kristall werden einige vorläufige Homologiemodelle des Zielproteins mit MODELLER^[3] erstellt. Danach werden ein oder mehrere Liganden, von denen bekannt ist, dass sie an das Zielprotein binden, in die gemittelten Repräsentationen des generierten Ensembles überlagert Modelle platziert.^[4] Anschließend werden verbesserte Proteinmodelle mit MODELLER generiert, wobei die platzierten Liganden explizit als zusätzliche Randbedingung einbezogen werden. Die Liganden werden in Form von wissensbasierten Potentialen berücksichtigt, die aus der Protein-Ligand-Bewertungsfunktion DrugScore^[5] extrahiert wurden. Nach einer Bewertung der generierten Komplexe mit DrugScore wird das Modell ausgewählt, das die Ligandenbindung am besten reproduziert. Dieser modellierte Komplex kann anschließend durch Selektion und Kombination von Fragmenten verschiedener Modelle optimiert werden. Zum Schluss wird das zusammengesetzte Modell durch Optimierung der Seitenketten-Ligand-Wechselwirkungen mit einem Kraftfeld verfeinert.^[6]

Zielproteins mehrere Aminosäuren gegenüber dem Vergleichsprotein ausgetauscht sind. Normalerweise stehen bei einem Wirkstoffdesign-Projekt vor Aufklärung der Raumstruktur des Zielproteins Bindungsdaten von Liganden zur Verfügung, die etwa aus dem experimentellen Hochdurchsatz-Screening stammen. Im Laufe eines solchen Projekts werden ligandenbasierte Struktur-Aktivitäts-Beziehungen (mit 3D-QSAR-Methoden)^[8] abgeleitet, die helfen, Eigenschaften zu extrahieren, die Trends der Bindungsaffinität erklären. Diese Methoden erzeugen einen generischen Abdruck der Bindungstasche des Zielproteins, der keine explizite Informationen über ihre Aminosäurezusammensetzung enthält. Sie unterstützen die Ligandenoptimierung, sind aber im Hinblick auf die Entdeckung alternativer Molekülgerüste stark eingeschränkt. Unsere Methode versucht, die Vorteile der Homologiemodellierung und der ligandenbasierten 3D-QSAR-Analysen zu kombinieren.

Als ein praxisrelevantes Anwendungsbeispiel wurde ein Homologiemodell des NK1-Rezeptors mit MOBILE erstellt. Rezeptoren aus der Familie der GPCRs stimulieren Signalkaskaden, die z. B. durch optische, olfaktorische, hormonelle oder Neurotransmitter-vermittelte Signale ausgelöst werden. Die GPCR-Familie ist derzeit eine der wichtigsten Proteinfamilien in der Wirkstoffentwicklung. Etwa die Hälfte aller neuen Arzneistoffe richtet sich gegen GPCRs.^[9] Die Expression, Reinigung, Kristallisation und Strukturaufklärung dieser membranständigen Proteine gestalten sich schwierig. Bisher konnte nur die Struktur eines einzigen

[*] A. Evers, Prof. Dr. G. Klebe
Institut für Pharmazeutische Chemie der Universität Marburg
Marbacher Weg 6, 35032 Marburg (Deutschland)
Fax: (+49) 06421-28-28994
E-mail: klebe@mail.uni-marburg.de

[**] Die Autoren danken Dr. Arne Svensson (AstraZeneca, Mölndal) für das Messen der Bindungsaffinitäten.

GPCR, des Rinderrhodopsins, mit hoher Auflösung bestimmt werden.^[10] Diese Struktur dient heute als Vorlage bei der Homologiemodellierung aller Mitglieder der GPCR-Familie. Wegen der Ungenauigkeit der so erhaltenen Modelle ist bisher kein erfolgreiches prädiktives virtuelles Screening an einem auf Rhodopsin basierenden GPCR-Modell beschrieben worden. Hingegen konnten neue GPCR-Liganden erfolgreich durch Anwendung ligandenbasierter virtueller Screeningtechniken entdeckt werden.^[11] Ein Schritt in Richtung Wirkstoffsuche auf der Basis eines Modells wurde von Bissantz et al. unternommen.^[12] Die Gruppe zeigte vor kurzem, dass ihre Homologiemodelle des Dopamin-D3-, des muscarinischen M1- und des Vasopressin-V1a-Rezeptors verlässlich genug waren, um bekannte Antagonisten durch virtuelles Screening in einer Datenbank wiederzufinden, die mit wirkstoffähnlichen Molekülen komplettiert wurde. Eine andere GPCR-Modellierungstechnik, die nicht auf die Struktur von Rinderrhodopsin angewiesen ist, ist PRE-DICT.^[13] Mit ihrem so erzeugten Modell erreichten die Autoren beim Screening von Datenbanken Trefferquoten von 10 bis 24 %.

Der NK1-Rezeptor gehört zu den Tachykinin bindenden Rezeptoren (NK1, NK2 und NK3), die selektiv an die peptidischen Neurotransmitter Substanz P, Neurokinin A und Neurokinin B binden. Substanz P spielt eine wichtige Rolle bei Schmerzübertragung, Immunantworten und Entzündungsprozessen. Der wahrscheinlich am besten untersuchte NK1-Antagonist ist CP-96345. Erkenntnisse über dessen bioaktive Konformation wurden durch eine Reihe von theoretischen und experimentellen Untersuchungen gewonnen.^[14] Durch Mutationsstudien^[15] und vergleichende Affinitätsbestimmungen, basierend auf der Bindung von CP-96345, konnten die an der Ligandenbindung beteiligten Aminosäuren identifiziert und in ein topographisches Wechselwirkungsmodell übersetzt werden (siehe Abbildung 3b). Unserem Homologiemodell entsprechend überlappt die Bindungstasche, die CP-96345 aufnimmt, mit der von Retinal im Rinderrhodopsin.

Die Sequenzidentität zwischen Rinderrhodopsin und dem NK1-Rezeptor beträgt 21 % (Abbildung 2). Betrachtet man lediglich die Transmembranregion, steigt die Identität auf 27 %. Zur Retinal-Bindungsstelle besteht allerdings keinerlei Sequenzidentität. Dies stellt eine besondere Herausforderung für die Homologiemodellierung dar. Basierend auf der Struktur von Rinderrhodopsin wurden zunächst 100 vorläufige Homologiemodelle erstellt. Im nächsten Schritt wurde CP-96345 mit AutoDock in die modellierten Bindungsta-

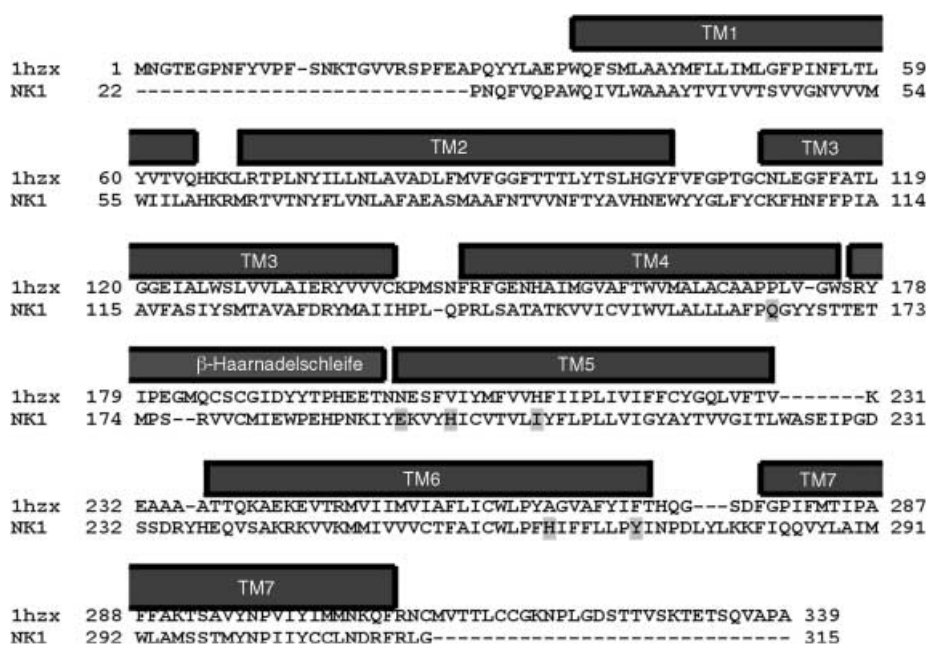


Abbildung 2. Sequenzalignement zwischen Rinderrhodopsin und dem NK1-Rezeptor. Die Reste, die nach Mutationsstudien in die Ligandenbindung involviert sind, sind grau schattiert. Diese Reste sind Gln 165 (nach der Ballesteros-Weinstein-Nomenklatur^[16] 4.60), Glu 193 (5.35), His 197 (5.39), Ile 204 (5.46), His 265 (6.52) und Tyr 272 (6.59).

schen platziert. Dabei wurde der Ligand rigide in eine Konformation aus der kristallographisch ermittelten Struktur (CSD-Referenzcode: LEWCUL) eingesetzt. Aus dem Satz erzeugter Ligandenorientierungen wählten wir die Lösung aus, die die Schlüsselwechselwirkungen des topographischen Modells (Abbildung 3b) am besten reproduzierte. Anschließend wurden unter Berücksichtigung des platzierten Liganden CP-96345 neue Homologiemodelle erstellt. Die am besten bewerteten Aminosäuren der individuellen Proteinmodelle wurden kombinatorisch zusammengefügt, sodass ein Modell ohne ungünstige Van-der-Waals-Wechselwirkungen entstand.

Abbildung 3c zeigt die Ligandenbindungsstelle des resultierenden Komplexes. Die im topographischen Modell (Abbildung 3b) vorgeschlagenen Wechselwirkungen werden gut reproduziert.^[14] Die Relevanz des so erzeugten NK1-Modells wurde durch die Vorhersage der Bindungsmodi weiterer bekannter NK1-Antagonisten aus strukturell unterschiedlichen Verbindungsklassen bestätigt. Mutationsdaten, die auf Schlüsselwechselwirkungen der Liganden mit dem NK1-Rezeptor verweisen, sowie strukturelle Gemeinsamkeiten bekannter NK1-Antagonisten wurden anschließend in eine Pharmakophor-Hypothese übersetzt (Abbildung 3d). Basierend auf dieser Pharmakophor-Hypothese wurden Liganden aus sieben Datenbanken mit insgesamt 800 000 Einträgen herausgesucht und auf ihre potenzielle Rezeptorbindung überprüft.

Wir verwendeten eine Suchstrategie, die in ähnlicher Form bereits für die Suche nach Enzyminhibitoren beschrieben wurde.^[17] Zunächst wurde ein topologischer Filter eingesetzt, um Verbindungen mit mindestens zwei Phenylringen und einem Wasserstoffbrückenacceptor herauszusuchen.

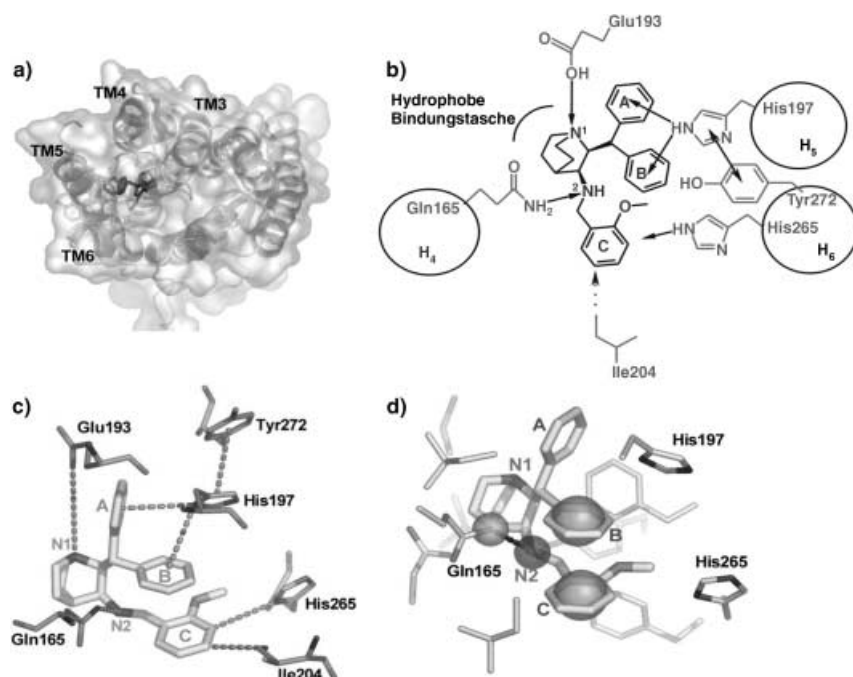
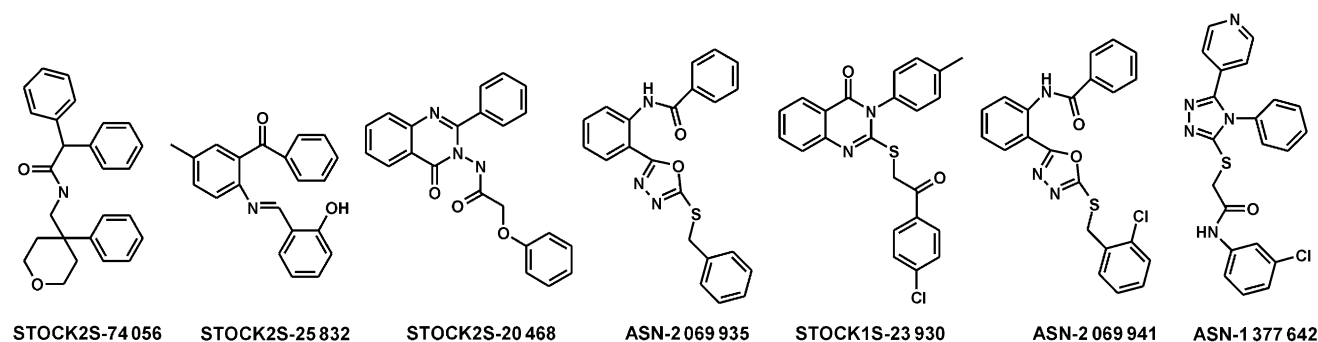


Abbildung 3. Der NK1-Rezeptor und sein Antagonist CP-96345. a) Blick von extrazellulär auf das Homologiemodell des NK1-Rezeptors, komplexiert mit CP-96345. b) Schematische Darstellung der postulierten Wechselwirkungen zwischen dem NK1-Rezeptor und CP-96345, erstellt nach Analyse der publizierten Mutationsdaten und Affinitätsmessungen von NK1-Antagonisten. Die Pfeile kennzeichnen die postulierten Schlüsselwechselwirkungen zwischen Rezeptor und Ligand. Die wichtigsten, für die Ligandenbindung essenziellen Wechselwirkungen sind 1) die Wasserstoffbrücke zwischen Gln 165(4.60)-NH und einem entsprechenden Acceptor im Antagonisten und 2) eine amino-aromatische Wechselwirkung zwischen His 197(5.39) und der aromatischen Gruppe B. Eine Analyse von NK1-Antagonisten aus strukturell unterschiedlichen Klassen ergab, dass die aromatische Gruppe C ebenfalls für die Ligandenbindung benötigt wird. c) Modellierter Komplex des NK1-Rezeptors mit CP-96345. Die gestrichelten Linien zeigen die im topographischen Wechselwirkungsmodell (b) vorgeschlagenen Schlüsselwechselwirkungen an. d) Strukturbasierte Pharmakophorhypothese. Der Pharmakophor wurden nach Analyse bekannter NK1-Antagonisten erstellt. Die Hypothese ist durch folgende drei Hauptmerkmale gekennzeichnet: 1) eine Wasserstoffbrücke zwischen Gln 165(4.60)-NH und einem entsprechenden Wasserstoffbrückenacceptor (hier N2) des Liganden; 2) die aromatische Gruppe B, die mit His 197(5.39) eine amino-aromatische Wechselwirkung eingeht, und 3) eine weitere aromatische Gruppe (C).

Dabei betrachteten wir nur Verbindungen mit maximal sieben drehbaren Bindungen und einem Molekulargewicht von weniger als 450 Da als geeignete Leitstrukturen. Anschließend wurde mit dem Programm UNITY die relative

räumliche Orientierung der aromatischen Ringe zum Wasserstoffbrückenacceptor eingeschränkt (Abbildung 3b und 3d). Der vom Protein besetzte Volumenbereich wurde für die Platzierung der Testliganden ausgeschlossen. Die nach diesem Schritt verbleibenden 11 109 Treffer wurden flexibel mit dem Programm FlexX-Pharm^[18] in die modellierte NK1-Bindungstasche platziert. Der in Abbildung 3d dargestellte Pharmakophor wurde dabei als zusätzliche Bedingung berücksichtigt. Die zu erwartenden Bindungsaffinitäten wurden mit DrugScore abgeschätzt. Nach einer lokalen Kraftfeldminimierung der resultierenden Komplexe mit MOLOC^[6] wurden die besten Treffer einer visuellen Inspektion unterzogen. Dabei wurden die Bindungsmodi hinsichtlich ihrer Übereinstimmung mit dem topographischen Wechselwirkungsmodell (Abbildung 3b) und hinsichtlich der Oberflächenkomplementarität zwischen Protein und Ligand beurteilt. Insbesondere sehr flexible Liganden und solche mit unangesättigten Wasserstoffbrückenkontakten in der Bindungstasche oder unbesetzten Hohlräumen entlang der Protein-Ligand-Grenzfläche wurden von der weiteren Betrachtung ausgeschlossen. Letztlich wurden sieben Verbindungen (Schema 1) für einen In-vitro-Test ausgewählt. Der verwendete Bindungsassay basiert auf der Verdrängung von radioaktiver Substanz P als Ligand und arbeitet mit CHO-Zellen (CHO = Chinese Hamster Ovary). Mit diesem Test können nur Liganden mit einer Mindestaffinität von 1 μM detektiert werden. Von den sieben Testverbindungen

zeigte eine (ASN-1377642 von AEPC, Abbildung 4) eine Bindungsaffinität im submikromolaren Bereich mit einem K_i -Wert von 251 nM. ASN-1377642 erfüllt die Einschlusskriterien unserer Suchanfrage.



Schema 1. Liste der Verbindungen, die auf Inhibition des NK1-Rezeptors getestet wurden. ASN-1377642 hatte einen K_i -Wert von 251 nM.

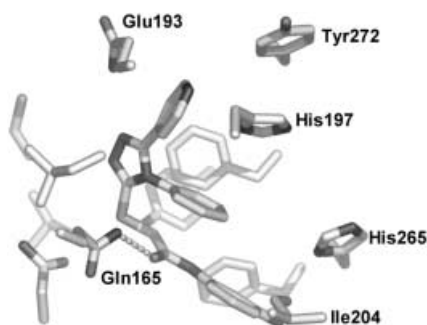


Abbildung 4. Modellierter Bindungsmodus von ASN-1377642, das durch virtuelles Screening an unserem Homologiemodell als potenter NK1-Antagonist entdeckt wurde. Die Pharmakophorhypothese wird gut erfüllt. Die essenzielle Wasserstoffbrücke zu Gln165(4.60)-NH geht vom Carboxylsauerstoffatom einer Amidgruppe aus.

Unsere ligandenunterstützte Homologiemodellierung generiert Proteinbindungstaschen-Modelle, die erfolgreich als Grundlage für strukturbasiertes Wirkstoffdesign verwendet werden können. Ein Hauptvorteil des beschriebenen Verfahrens besteht darin, dass Informationen über bekannte Liganden nicht nur für die Homologiemodellierung, sondern auch in den darauf folgenden Suchschritten verwendet werden konnten: Strukturelle Eigenschaften bekannter NK1-Antagonisten wurden zur Erstellung eines Pharmakophormodells eingesetzt, das als Referenz für unsere 2D- und 3D-Suchen diente. Dabei definierte das Proteinmodell die Randbedingungen für die 3D-Suche und steuerte die Generierung der Ligandenorientierung in der Bindungstasche. Am Ende wurden die erhaltenen Docking-Lösungen durch Evaluierung der Protein-Ligand-Wechselwirkungen bewertet. CP-96345 und ASN-1377642 erfüllen ähnliche Pharmakophoreigenschaften, sind aber hinsichtlich ihrer molekularen Gerüste deutlich verschieden. Daher glauben wir, dass ASN-1377642 mit einer reinen ligandenbasierten Suchstrategie nicht als bestbewerteter Treffer entdeckt worden wäre.

Man darf jedoch nicht vergessen, dass der Erfolg unserer Methode vom Grad der Übereinstimmung von Vorlagen- und Zielproteinsequenz abhängt. Ist diese gering, so sind zusätzliche Informationen notwendig, wie Mutationsdaten oder eine nativ-ähnliche Konformation eines hochaffinen Liganden.

Eingegangen am 3. September 2003 [Z52776]

Online veröffentlicht am 2. Dezember 2003

Stichwörter: Computerchemie · G-Proteine · Proteine · Proteinmodelle · Rezeptoren

- [6] P. R. Gerber, *J. Comput.-Aided Mol. Des.* **1998**, *12*, 37–51.
- [7] D. Baker, A. Sali, *Science* **2001**, *294*, 93–96.
- [8] a) *3D QSAR in Drug Design. Ligand-Protein Complexes and Molecular Similarity* (Hrsg.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer/ESCOM, Leiden, **1998**; b) *3D QSAR in Drug Design. Recent Advances* (Hrsg.: H. Kubinyi, G. Folkers, Y. C. Martin), Kluwer/ESCOM, Leiden, **1997**.
- [9] T. Klabunde, G. Hessler, *ChemBioChem* **2002**, *3*, 928–944.
- [10] a) D. C. Teller, T. Okada, C. A. Behnke, K. Palczewski, R. E. Stenkamp, *Biochemistry* **2001**, *40*, 7761–7772; b) K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, *Science* **2000**, *289*, 739–745.
- [11] a) S. Flohr, M. Kurz, E. Kostenis, A. Brkovich, A. Fournier, T. Klabunde, *J. Med. Chem.* **2002**, *45*, 1799–1805; b) D. P. Marriott, I. G. Dougall, P. Meghani, Y. J. Liu, D. R. Flower, *J. Med. Chem.* **1999**, *42*, 3210–3216.
- [12] C. Bissantz, P. Bernard, M. Hibert, D. Rognan, *Proteins Struct. Funct. Genet.* **2003**, *50*, 5–25.
- [13] O. M. Becker, S. Shacham, Y. Marantz, S. Noiman, *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 353–361.
- [14] a) Y. Takeuchi, E. F. B. Shands, D. D. Beusen, G. R. Marshall, *J. Med. Chem.* **1998**, *41*, 3609–3623; b) C. J. Swain, E. M. Seward, M. A. Cascieri, T. M. Fong, R. Herbert, D. E. MacIntyre, K. J. Merchant, S. N. Owen, A. P. Owens, V. Sabin, M. Teall, M. B. VanNiel, B. J. Williams, S. Sadowski, C. Strader, R. G. Ball, R. Baker, *J. Med. Chem.* **1995**, *38*, 4793–4805; c) G. J. Boks, J. P. Tollenaere, J. Kroon, *Bioorg. Med. Chem.* **1997**, *5*, 535–547.
- [15] a) T. M. Fong, M. A. Cascieri, H. Yu, A. Bansal, C. Swain, *Nature* **1993**, *362*, 350–353; b) T. M. Fong, H. Yu, M. A. Cascieri, D. Underwood, C. Swain, C. D. Strader, *J. Biol. Chem.* **1994**, *269*, 14957–14961; c) T. M. Fong, H. Yu, M. A. Cascieri, D. Underwood, C. Swain, C. D. Strader, *J. Biol. Chem.* **1994**, *269*, 2728–2732.
- [16] J. A. Ballesteros, H. Weinstein, *Methods Neurosci.* **1995**, *25*, 366–428.
- [17] a) S. Gruneberg, B. Wendt, G. Klebe, *Angew. Chem.* **2001**, *113*, 404–408; *Angew. Chem. Int. Ed.* **2001**, *40*, 389–393; b) S. Gruneberg, M. T. Stubbs, G. Klebe, *J. Med. Chem.* **2002**, *45*, 3588–3602; c) R. Brenk, L. Naerum, U. Gradler, H. D. Gerber, G. A. Garcia, K. Reuter, M. T. Stubbs, G. Klebe, *J. Med. Chem.* **2003**, *46*, 1133–1143.
- [18] S. A. Hindle, M. Rarey, C. Buning, T. Lengauer, *J. Comput.-Aided Mol. Des.* **2002**, *16*, 129–149.

- [1] a) A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730; b) J. Drews, *Science* **2000**, *287*, 1960–1964.
- [2] A. Evers, H. Gohlke, G. Klebe, *J. Mol. Biol.* **2003**, *334*, 327–345.
- [3] A. Sali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779–815.
- [4] a) A. Schafferhans, G. Klebe, *J. Mol. Biol.* **2001**, *307*, 407–427; b) F. Osterberg, G. M. Morris, M. F. Sanner, A. J. Olson, D. S. Goodsell, *Proteins Struct. Funct. Genet.* **2002**, *46*, 34–40.
- [5] H. Gohlke, M. Hendlich, G. Klebe, *J. Mol. Biol.* **2000**, *295*, 337–356.