b) D. B. Kitchen, J. E. Jackson, L. C. Allen, *J. Am. Chem. Soc.* **1990**, *112*, 3408–3414; c) A. Skancke, D. Van Vechten, J. F. Liebman, P. N. Skancke, *J. Mol. Struct. (THEOCHEM)* **1996**, *376*, 461–468.

[19] A. L. Spek, PLATON. A multipurpose crystallographic tool, Utrecht University, The Netherlands, **2000**.

[20] P. T. Beurskens, G. Admiraal, G. Beurskens, W. P. Bosman, S. Garcia-Granda, R. O. Gould, J. M. M. Smits, C. Smykalla, The DIRDIF97 program system, Technical Report of the Crystallography Laboratory, University of Nijmegen, The Netherlands, **1997**.

[21] G. M. Sheldrick, SHELXL97, Program for crystal structure refinement, University of Göttingen, Germany, **1997**.

Virtual Screening for Bioactive Molecules by Evolutionary De Novo Design[**]

Gisbert Schneider,* Odile Clément-Chomienne,
Laurence Hilfiger, Petra Schneider, Stefan Kirsch,
Hans-Joachim Böhm, and Werner Neidhart

One of the goals of computational chemistry is to be able to design novel molecular structures with biological activity comparable to a defined known drug (the "template") and with a significantly different architecture. There is a need to develop techniques that can be applied if a high-resolution receptor model is unavailable, complementing more conventional structure-based molecular modeling and design. Such algorithms would be particularly helpful for combinatorial library shaping and fast-follower strategies in drug development. Two major problems must be solved to successfully address this task. First, the tool must be able to systematically navigate and efficiently search in a vast chemical space containing billions of virtual compounds, as it is impossible to exhaustively screen all possible molecules for a desired biological activity. Second, the conceptual view of a molecule must be defined in such a way that allows rapid comparison of the template with novel compounds, yielding a measure of functional rather than structural similarity.

For the sake of "backbone (scaffold)-switching" the various pharmacophore concepts provide a straightforward approach.[1] Adaptive stochastic searching, notably evolutionary algorithms, represents a practicable solution to optimization in a high-dimensional search space. This is especially true when no or only little information is available about the ruggedness of the associated "fitness landscape".[2] Here we present an efficient computational molecular design strategy that implements pharmacophore-guided evolutionary searching in chemical space. Experimental proof of the concepts is demonstrated by the successful de novo design of a new structural class of potent K^+-channel inhibitors. The algorithm is referred to as TOPAS (*TOP*ology-*A*ssigning *S*ystem) in this report.[3] TOPAS provides a solution to template-based de novo design, in which novel molecules are assembled taking a given bioactive compound as the reference point (template structure).

A challenging task in de novo design is the generation of molecular structures with "druglike" properties. To reduce the risk of generating molecular architectures with inadequate features, or which are difficult to synthesize—a problem encountered by many de novo design procedures[4]—TOPAS was equipped with a stock of building blocks obtained from retrosynthetic fragmentation of compounds listed in the Derwent World Drug Index (WDI).[5] The idea is that assembly of such building blocks by a limited set of chemical reactions might lead to chemically feasible novel structures, from both the medicinal chemistry and the synthesis planning perspective. All structures contained in the WDI, which had an entry related to "mechanism" or "activity", were subjected to retrosynthetic fragmentation. This approach is similar to the RECAP procedure developed by Hann and co-workers.[6] Eleven reaction schemes were applied to perform fragmentation (for details, see the original publication).[6] We used the Daylight tool kit for implementation of reaction schemes.[7] A total of 24563 unique building blocks was obtained from our analysis.[3] The same eleven reactions were used by TOPAS to assemble novel molecules.

TOPAS is based on a special evolutionary algorithm, a $(1,\lambda)$ evolution strategy.[2, 8] This type of evolutionary algorithm has been successfully applied to peptide design experiments in previous applications.[9] The common feature of evolutionary algorithms is a cyclic variation-selection process. "Parents" breed "offspring", and the fittest of each "generation" becomes the parent of the subsequent optimization cycle. Starting from an arbitrary point in search space (a randomly selected chemical structure), TOPAS generates a set of λ variant structures which satisfy a bell-shaped distribution centered in the chemical space coordinates of the parent structure. This means that most of the variants will be very similar to their parent, and the number of offspring at a certain distance point will be less as the distance-to-parent increases.[3] Distance in chemical space is defined by a topological pharmacophore measure (see below).

Evolution strategies are often applied to real-valued function optimization problems.[2] Algorithms generally operate directly on the parameters to be optimized, in contrast to genetic algorithms, which usually operate on a separately coded transformation of the objective variables (the so-called "chromosome").[10] In addition, evolution strategies include a second-level optimization of "strategy parameters", that is, tunable variables that in part determine how each parent will generate offspring. This enables an adaptive stochastic search to be performed and guarantees a finite probability of breeding very dissimilar structures ("snoopers" in chemical space).[9] The width of the variant distribution is determined by the variance or standard deviation, σ, of the bell-shaped curve reflecting the distance-to-parent probability. The σ

---

[*] Priv.-Doz. Dr. G. Schneider, O. Clément-Chomienne, L. Hilfiger,
P. Schneider, S. Kirsch, H.-J. Böhm, W. Neidhart
Hoffmann-La Roche Ltd.
Pharmaceuticals Division
4070 Basel (Switzerland)
Fax: (+41) 61-68-8904
E-mail: gisbert.schneider@Roche.com

[**] Special thanks to Neil R. Taylor for his help in preparation of the manuscript.

Supporting information for this article is available on the WWW under http://www.wiley-vch.de/home/angewandte/ or from the author.

1433-7851/00/3922-4130 $ 17.50+.50/0

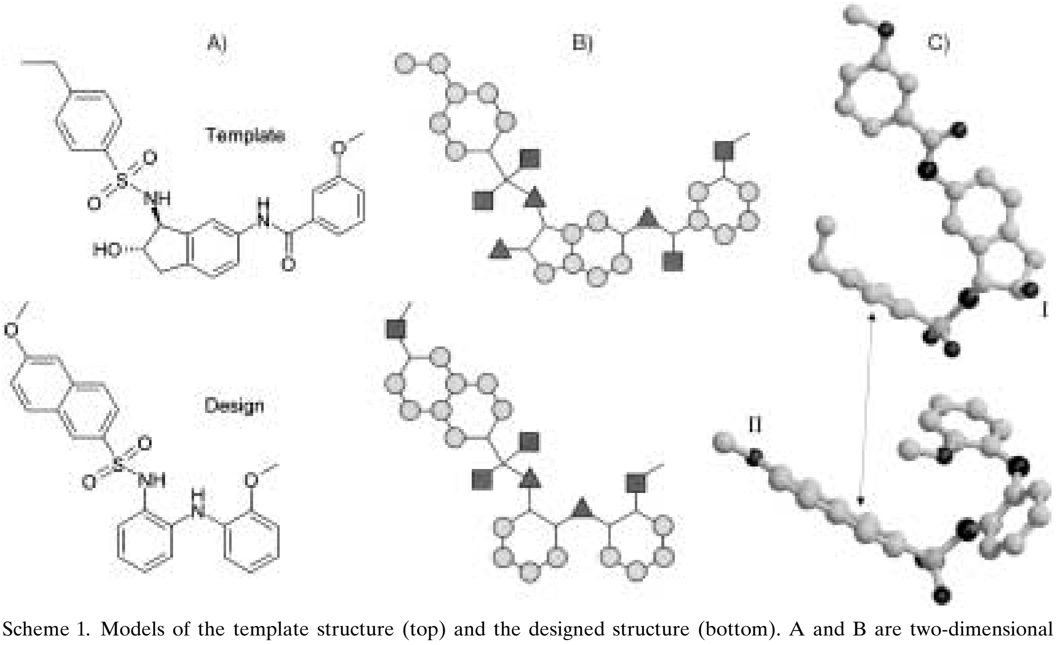
value may be regarded as a measure of library diversity: Small values lead to narrow distributions of variants, while large values often result in novel structures that are very dissimilar to the parent molecule.

At the start of a virtual screening experiment, σ should be set to a large value to explore chemical space for potential optima ("universal" library). At later stages σ will have to have small values to facilitate local hill-climbing (focused or biased libraries). In contrast to related techniques, there exists no predefined "cooling schedule" or a fixed decaying function for σ. Its value may freely change during optimization to be able to adapt to the actual requirements of the search.[2] Large σ values will automatically evolve if large steps are favorable, and small values will result when only a small move in chemical space leads to success on the fitness landscape. Further details on this mechanism, alternative implementation schemes, and a comparison of evolutionary algorithms can be found elsewhere.[2, 10]

In TOPAS, fitness is defined as the pairwise similarity between the template and the variant structures. This is done using a measure of topological pharmacophore similarity.[3] Molecules are encoded by two-point pharmacophores using generalized atom types (hydrogen-bond donors/acceptors, positive/negative charge, lipophilic). This is similar to correlation vector representation of molecular structures and properties.[11] According to this abstraction, a molecule is a combination of pairs of generalized atom types topologically connected by bonds. Similarity between two molecules was defined as the Euclidian distance between correlation vectors. The main advantage of this approach is that molecules are similar if they exhibit a similar distribution of generalized atom types in their two-dimensional structure, irrespective of the chemical way their atoms are connected.[12] A downside is the fact that important three-dimensional and full topological structural information is neglected. Whenever such information is crucial for bioactivity, the design algorithm may fail to produce useful results. Recently, novel potent Ca^{2+}-channel blocking agents were retrieved from a large database (CATS program) using the same definition of pharmacophore similarity.[13] This finding motivated us to apply the correlation vector scheme for de novo design by TOPAS. In this work, additional penalty terms were added to the fitness function to avoid undesired structures. The fitness value was decreased if a) the total number of atoms exceeded 50, b) the sum of oxygen and nitrogen atoms was greater than 12, or c) if there were more than seven potential hydrogen-bond donors present in a given molecule. These crude filters are thought to facilitate the design of bioavailable molecules.[14]

For a challenging test of TOPAS we selected a known potent K^+-channel blocking agent as the template molecule: *trans*-(2*S*,3*S*)-*N*-[3-(4-ethylbenzenesulfonylamino)-2-hydroxy-indan-5-yl]-3-methoxy-benzamide (Scheme 1 a, top).[15] In our electrophysiological studies this compound had an IC_{50} of 0.11 ± 0.02 μM for the inhibition of the human Kv1.5 potassium channel (hKv1.5). Inhibitors of voltage-dependent potassium channels induce a decrease in potassium ion movement across cell membrane. In cardiac cells this decrease leads to the prolongation of the action potential. Increasing myocardial refractoriness by prolonging the action potential can be useful for the treatment of cardiac arrhythmia.[16] Blocking potassium channels and depolarizing the resting membrane potential has been shown to regulate a variety of biological processes, like T-cell activation under immune-reactive conditions.[17]

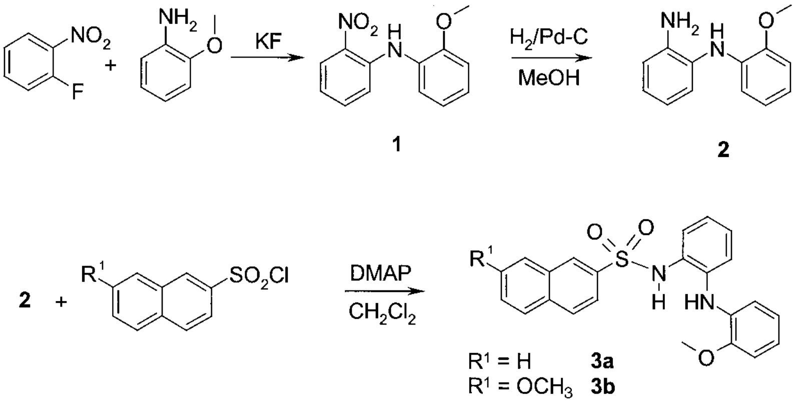
We performed several design runs with TOPAS, yielding slightly different "best" molecular structures after 100 generations (not shown). Starting from a randomly assembled molecular structure, 100 variants were virtually synthesized



Scheme 1. Models of the template structure (top) and the designed structure (bottom). A and B are two-dimensional representations. In B, generalized atom types are shown (squares, hydrogen-bond acceptors; triangles, hydrogen-bond donors; circles, lipophilic atoms). Three-dimensional models are shown in C.[20] These structures were superimposed by a pharmacophore matching routine. For clarity, they are given in a side-by-side view. Two functional groups are labeled by **I** and **II** (heteroatoms are shown in dark gray).

per generation and the best of each generation was the parent of the next cycle. The fittest molecule obtained was a naphthylsulfonamide (Scheme 1a, bottom). From a purely structural perspective, there is no obvious similarity between the template and the designed structure besides a common sulfonamide group. However, assignment of generalized atom types indicates a resemblance of the two molecules, and a mutual pharmacophore pattern is uncovered (Scheme 1b). This example demonstrates that simplistic abstractions from chemical structure can sometimes help reveal "hidden" molecular similarities. To get a more precise picture of relevant features, three-dimensional structural models were superimposed by using a pharmacophore matching routine. Both molecules, the template and the design, seem to favor a U-shaped structure, resulting in a good fit of the planar aromatic ring systems and the adjacent sulfonamide moieties (Scheme 1c). From this we concluded that the hydroxy group **I** does not significantly contribute to the relevant pharmacophore motif. In addition, omitting methoxy group **II** of the designed structure might be beneficial for bioactivity, a very lipophilic system seems to be required in this position. Indeed, after completion of our work experimental evidence was published to support this view.[18]

Two molecules were synthesized according to Scheme 2. Compound **3a** was synthesized to investigate the effect of the methoxy group **II** on the biological activity; compound **3b** is the original design recommended by TOPAS. Accordingly, 1-fluoro-2-nitrobenzene was allowed to react with *o*-anisidine in an aromatic nucleophilic substitution to give (2-methoxyphenyl)-(2-nitrophenyl)amine (**1**) in near-quantitative yields. Subsequent reduction of the nitro derivative **1** was performed by catalytic hydrogenation with palladium as catalyst to yield the diamine **2** in 87.9% yield. Condensation of **2** with 7-methoxynaphthalene-2-sulfonyl chloride or naphthalene-2-sulfonyl chloride in methylene chloride with 4-(dimethylamino)pyridine (DMAP) as a base gave rise to the sulfonamides **3a** and **3b** as colorless solids in 42 and 35% yields, respectively, after chromatography.



Scheme 2. Synthetic scheme for the molecules **3a** and **3b** designed by TOPAS. DMAP = 4-(dimethylamino)pyridine.

Electrophysiological measurement undoubtedly proved K^+-channel blocking activity by **3a** and **3b** (IC_{50} = 0.47 and 7.34 μM, respectively; Figure 1). This finding demonstrates that the TOPAS design technique can be used for generation of novel molecular structures exhibiting biological activity that is comparable to a structurally different template molecule. It must, however, be stressed that the original design **3b** was approximately two orders of magnitude less active (IC_{50} = 7.34 ± 0.22 μM) than the template (IC_{50} = 0.11 ± 0.02 μM). Nevertheless, it represents a valuable novel chemical class, which is currently being further explored. Straightforward structural modifications—as suggested by our molecular modeling studies—already led to the improved compound **3a** exhibiting K^+-channel blocking activity comparable to the original template (IC_{50} = 0.47 ± 0.09 μM).

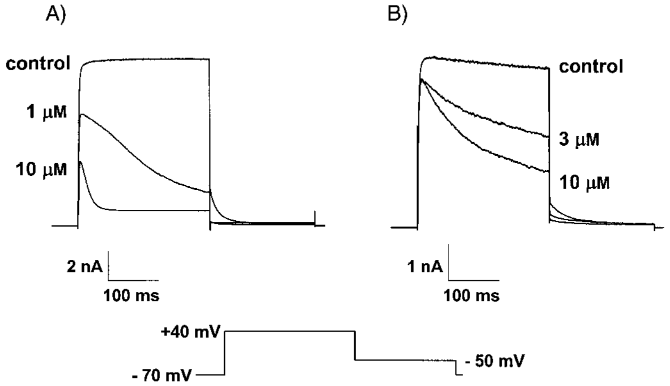


Figure 1. Inhibition of voltage-gated K^+ flux via the hKv1.5 channel in transfected mammalian CHO cells. Representative families of whole-cell currents are shown, evoked by 250 ms test pulses to +40 mV followed by a 200 ms step to −50 mV. A) Effect of structure **3a** at 1 μM and 10 μM. B) Effect of structure **3b** at 3 μM and 10 μM.

Our TOPAS strategy offers a number of advantages: 1) an adaptive stochastic search is performed in virtual chemical space; 2) the molecules generated are not restricted to a predefined combinatorial class (like e.g., peptides or Ugi reaction products); 3) novel structures are assembled from drug-derived building blocks using a set of "simple" chemical reactions; 4) a large diversity of molecular fragments can be explored; and 5) the design algorithm may be applied when a high-resolution receptor structure is unavailable. A particularly appealing future application is ligand design to influence G-protein-coupled receptor activity.

A significant advantage of using building blocks derived from known drugs is that reassembly of fragments potentially produces "druglike" structures. However, this technique may limit exploration of hitherto unknown "druglike" regions in chemical space. In similar fashion, exploration of chemical space may be further limited in the use of a small set of reaction schemes. An additional concern is that it is rarely known a priori which parts of a template molecule are relevant or irrelevant for a particular biological function. This can easily result in inactive molecular designs. One must not forget that it is comparably easy to find novel structures exhibiting some desired bioactivity using virtual screening tools, whereas the de novo design of real drugs is only slowly emerging.[19] We are convinced that fragment-based evolutionary design approaches like TOPAS that are able to optimize for several fitness functions in parallel—including predictions of metabolic and pharmacokinetic parameters—will lead to further significant progress in this area of computational chemistry and virtual screening.

Received: April 13, 2000
Revised: July 5, 2000 [Z14984]

---

[1] J. S. Mason, M. A. Hermsmeier, *Curr. Opin. Chem. Biol.* **1999**, *3*, 342–349.

[2] I. Rechenberg, *Evolutionsstrategie*, Frommann-Holzboog, Stuttgart, **1973**.

[3] G. Schneider, M.-L. Lee, M. Stahl, P. Schneider, *J. Comput. Aided Mol. Des.* **2000**, *14*, 487–494.

[4] H.-J. Böhm, D. W. Banner, L. Weber, *J. Comput. Aided Mol. Des.* **1999**, *13*, 51–56.

[5] Derwent World Drug Index, as of November 1998, Derwent Information, London, UK.

[6] X. O. Lewell, D. B. Judd, S. P. Watson, M. M. Hann, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 511–522.

[7] The Daylight Toolkit; Daylight Chemical Information Systems Inc., Irvine, CA, USA.

[8] a) J. H. Holland, *Adaptation in Natural and Artificial Systems*, M.I.T. Press, Cambridge, **1975**; b) D. E. Goldberg, *Genetic Algorithms in Search, Optimization & Machine Learning*, Addison-Wesley, Reading, **1989**; c) B. Levitan, S. Kauffman, *Mol. Diversity* **1995**, *1*, 53–68.

[9] a) G. Schneider, J. Schuchhardt, P. Wrede, *Biophys. J.* **1995**, *68*, 434–447; b) G. Schneider, W. Schrödl, G. Wallukat, J. Müller, E. Nissen, W. Rönspeck, P. Wrede, R. Kunze, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12179–12184; c) P. Wrede, O. Landt, S. Klages, A. Fatemi, U. Hahn, G. Schneider, *Biochemistry* **1998**, *37*, 3588–3593.

[10] a) P. Willett, *Trends Biotechnol.* **1995**, *13*, 516–521; b) R. C. Glen, A. W. Payne, *J. Comput. Aided Mol. Des.* **1995**, *9*, 181–202; c) M. G. Bures, Y. C. Martin, *Curr. Opin. Chem. Biol.* **1998**, *2*, 376–380; d) N. Saravanan, D. B. Fogel, K. M. Nelson, *Biosystems* **1995**, *36*, 157–166.

[11] a) G. Moreau, P. Broto, *Nouv. J. Chim.* **1980**, *4*, 757; b) M. Wagener, J. Sadowski, J. Gasteiger, *J. Am. Chem. Soc.* **1995**, *117*, 7769–7775; c) S. Anzali, W. W. R. K. Mederski, M. Osswald, D. Dorsch, *Bioorg. Med. Chem. Lett.* **1997**, *8*, 11–16; d) G. Schneider, P. Wrede, *Prog. Biophys. Mol. Biol.* **1998**, *70*, 175–222.

[12] P. Willett, J. M. Barnard, G. M. Downs, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983–996.

[13] G. Schneider, W. Neidhart, T. Giller, G. Schmid, *Angew. Chem.* **1999**, *111*, 3068–3070; *Angew. Chem. Int. Ed.* **1999**, *38*, 2894–2896.

[14] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.

[15] N. A. Castle, S. P. Hollinshead, P. F. Hughes, J. S. Mendoza, J. W. Wilson, G. Amato, S. Beaudoin, M. Gros, G. McNaughton-Smith (Icagen Inc., Eli Lilly and Company), WO-A 98/04521, **1998**.

[16] A. Naie, A. O. Grant, *Cardiovasc. Drugs Ther.* **1997**, *11*, 149.

[17] C. S. Lin, R. C. Boltz, J. T. Blake, M. Nguyen, A. Talento, P. A. Fischer, M. S. Springer, N. H. Sigal, R. S. Slaughter, M. L. Garcia, *J. Exp. Med.* **1993**, *177*, 637–645.

[18] M. F. Gross, N. A. Castle (Icagen Inc.), WO-A 99/37607, **1999**.

[19] V. J. Gillet, A. P. Johnson in *Designing Bioactive Molecules—Three-dimensional Techniques and Applications* (Eds.: Y. C. Martin, P. Willett), ACS, Washington, **1998**, pp. 149–174.

[20] Molscript v2.1.2; P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946.

Chirally Modified *n*-Butyllithium: Tuning the Composition, Structure, and Enantioselectivity with Modular Fencholates**

Bernd Goldfuss,* Melanie Steigelmann, and
Frank Rominger

*This work is dedicated to Prof. Dr. Günter Helmchen
on the occasion of his 60th birthday.*

Chirally modified organolithium reagents are eminent as reagents in enantioselective syntheses[1] and as initiators in asymmetric–anionic polymerization.[2] The control of reactivity and selectivity by suitable moderators is crucial for successful application.[3] Studies in solution[4] and in the solid state[1a,g,h, 5] yield precious information on the nature of chiral organolithium systems[6] and open possibilities for a rational design of new reagents.[7] Although *n*-butyllithium (*n*BuLi) is among the most widely employed organolithium reagents and several X-ray diffraction studies have been performed on achiral *n*BuLi complexes,[8] there is little structural information on chirally modified *n*BuLi species.[5b,c, 9]

Recently, we reported the synthesis and the X-ray crystal structure of the enantiopure *n*BuLi lithium fencholate complex **1**.[9] Complex **1** forms spontaneously as a colorless precipitate on mixing anisylfenchole (**2**) and a solution of *n*BuLi in hexane and contains one equivalent of *n*BuLi as well as three lithium fencholate moieties **2**-Li (Figure 1).[9] The

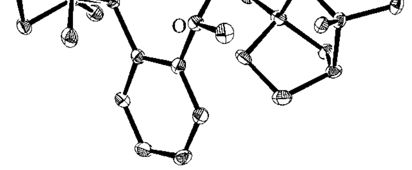


Figure 1. Molecular structure of **1**. The hydrogen atoms are omitted.

---

[*] Dr. B. Goldfuss, Dipl.-Chem. M. Steigelmann, Dr. F. Rominger[+]
Organisch-Chemisches Institut der Universität Heidelberg
Im Neuenheimer Feld 270, 69120 Heidelberg (Germany)
Fax: (+49) 6221-54-4205
www.uni-heidelberg.de/institute/fak12/OC/goldfus
E-mail: Bernd.Goldfuss@urz.uni-heidelberg.de

[+] X-ray crystal structure analysis

[**] This work was supported by the Fonds der Chemische Industrie (Sachbeihilfe and Liebig-Stipend for B.G.), the Deutsche Forschungsgemeinschaft (GO 930/1-1 and Habilitationsstipendium), the Research Pool Foundation (University Heidelberg), the Degussa-Hüls AG, and the BASF AG. B.G. is grateful to Prof. P. Hofmann for support at Heidelberg and thanks Prof. G. Boche (Marburg), Prof. D. Hoppe (Münster), and Prof. D. Seebach (Zürich) for fruitful suggestions. Parts of this work were presented at the 37th IUPAC Congress/27th GDCh General Meeting in Berlin (1999) and at the SFB-424 Symposium in Münster (2000).