doi:10.2533/chimia.2014.472 Chimia 68 (2014) 472–482

Three Decades of Structure- and Property-based Molecular Design

Klaus Müller*

KGF-SCS Distinguished Industrial Investigator Award 2013

Abstract: Roche has pioneered structure- and property-based molecular design to drug discovery. While this is an ongoing development, the past three decades feature key events that have revolutionized the way drug discovery is conducted in Big Pharma industry. It has been a great privilege to have been involved in this transformation process, to have been able to collaborate with, direct, guide, or simply encourage outstanding experts in various disciplines to build and further develop what has become a major pillar of modern small-molecule drug discovery. This article is an account of major events that took place since the early decision of Roche to implement computer-assisted molecular modeling 32 years ago and is devoted to the key players involved. It highlights the internal build-up of structural biology, with protein X-ray structure determination at its core, and the early setup of bioinformatics. It describes the strategic shift to large compound libraries and high-throughput screening with the development of novel compound storage and ultra-high-throughput screening facilities, as well as the strategic return to focused screening of small motif-based compound libraries. These developments were accompanied by the rise of miniaturized parallel compound property analytics which resulted in a major paradigm shift in medicinal chemistry from linear to multi-dimensional lead optimization. The rapid growth of huge collections of property data stimulated the development of various novel data mining concepts with 'matched molecular pair' analysis and novel variants thereof playing crucial roles. As compound properties got more prominent in molecular design, exploration of specific structural motifs for property modulation became a research activity complementary to target-oriented medicinal chemistry. The exploration of oxetane is given as an example. For the sake of brevity, this account cannot detail all further developments that have taken place in each individual area of structure- and property-based drug discovery and it can only hint at important developments in other disciplines that have equally contributed to major paradigm shifts in Roche's small-molecule drug discovery efforts.

Keywords: Drug discovery · Molecular design · Paradigm shifts · Property-based · Structure-based

Klaus Müller has been Extraordinary Professor at the University of Basel since 1991. Prior to his retirement in 2009 he occupied leading positions at F. Hoffmann-La Roche AG, among them as Head of '*Pharma New Technologies*' and of '*Science and Technology Relations*' for the development of innovative technologies and scientific collaborations with external groups. He was Secretary-General and Board Member of the *Roche Research Foundation* until its conclusion in 2008 and since then manages the *Roche Postdoc Fellowship Program*. Educated as an organic chemist at ETH Zurich, he obtained his PhD in 1970, undertook a 1-year postdoctoral stay in the US in physical-organic chemistry, and was then lecturer at Harvard University, before returning to ETHZ in late 1974. He habilitated in 1977 and remained there as a junior staff member before joining Roche in 1982 to develop computer-assisted molecular modeling and implementing structural biology and bioinformatics. He fostered the development of novel compound storage and ultra-HT screening systems, helped build a comprehensive miniaturized and parallelized platform for fast compound property analytics, and guided the development of a first chemistry-biology information management system. Structure- and property-based molecular design, both conceptual and experimental, has been a major area of interest in which he is still active today.

Correspondence: Prof. Dr. K. Müller
Pharmaceutical Research & Early Development
Roche Innovation Center Basel
F. Hoffmann-La Roche AG
CH-4070 Basel, Switzerland
Tel.: +41 61 688 4075
E-mail: klaus.mueller@roche.com

Over the last three decades structure- and property-based modeling has matured to an essential pillar of modern drug discovery in pharmaceutical research. What is well established nowadays has been developed in a gradual manner starting in Roche at a time when almost all necessary elements were lacking. When, in 1982, the first X-ray crystal structures of a therapeutically relevant protein, the bacterial dihydrofolate reductase (DHFR) became available,[1] Roche shared the vision that computer-assisted molecular modeling (CAMM) would be an indispensable tool to build on three-dimensional structure information, as an essential element in rational approaches to drug discovery. At that time, programs for nice computer graphics display of large molecules, proteins and DNA already existed, that toured the world, but could not be regarded as a solid foundation for practical modeling work in a pharmaceutical industrial environment. Likewise, no CAMM tools existed that could cope with the structurally complex and diverse world of medicinal chemistry. Molecular force field methods were still in their early stages, focusing mainly on peptides or small prototypic organic molecules,[2] or being further refined on hydro-

carbons.[3] Collections of small-molecule X-ray crystal structures, the Cambridge Structural Database (CSD),[4] or large biomolecules, the Protein Data Bank (PDB),[5] were far from being structurally comprehensive, holding some 50,000 entries for small-molecule and fewer than 200 protein X-ray crystal structures, respectively. However, it was clear from the outset that meaningful CAMM would require a highly functional, efficient, and trustworthy modeling platform with lean access to experimental structure information. To go for this at that early time was a highly courageous but farsighted decision by Roche. It was observed with critical interest and soon followed by many of its competitors; however, for quite some time it was considered with much skepticism in academia, particularly regarding the internal implementation of protein X-ray crystal structure determination.

For those who attempt to predict the future some 10 to 20 years ahead, this whole exercise would have looked like a daunting expedition. Thus, it is sometimes helpful not to look too far into the future and never promise too much, but start in a sensible way, developing modules stepwise to build and expand a new system gradually in a logical and structured way; always keeping close to productive applications where new tools can be immediately validated, their utility demonstrated, and customers' needs adequately addressed so that successive applications would visibly document continuous improvements of functionality and performance.

Roche offered excellent conditions for the development of a strong structure-based modeling environment, not only by providing and sustaining a generous dedicated technological infrastructure, but also by granting much operational freedom (something that one has to fight for permanently, but has to be recognized by senior management as an important success factor in creative pharmaceutical research) and empathic interest by both senior management and many colleagues in drug discovery research with application needs that could be addressed early on even with relatively modest modeling tools at their early stages of implementation. However, the successful development of what has become one of the most powerful structure- and property-based molecular design platforms in pharmaceutical industry would not have been possible without the close and sustained collaboration of outstanding scientific and technological experts in various disciplines and their excellent contributions of key elements to this platform. This account is therefore dedicated to them and their achievements.

The initial developments were done using a Digital VAX 11/780, at that time the most powerful machine for interactive computing (Fig. 1). To put it into perspective, this computer required an air-conditioned room of at least 3×5 m^2. Its price was approximately 1'000 times that of a standard personal computer today, but its computing power some five orders-of-magnitude less! Nevertheless, having such a computer dedicated for both development and application was paradise! In order to be able to concentrate fully on method developments and applications, it was essential that this computer was competently and efficiently managed. This was achieved in an outstanding manner by my first collaborator, Dr. Hans-Jakob Ammann (Fig. 1), a chemist by training who quickly evolved into a VAX/VMS specialist highly respected not only within Roche, but also externally, in particular by the experts at Digital itself. It has been due to Dr. Ammann's continuous superb management of the whole CAMM computer environment that an initial reasonably functional modeling system, the Roche Interactive Molecular Graphics (RIMG) system, could be developed relatively swiftly. The graphic display capabilities used the Evans & Sutherland MPI

Fig. 1. Dr. Hans-Jakob Ammann turning on the VAX-11/780, the most powerful machine for interactive computing at the beginning of the 1980s. By comparison, a contemporary PC, at a cost of about 1'000 times less, would offer 100'000 times more computing power and some 3 orders-of-magnitude more disk storage capacity and CPU memory. The VAX computer was intimately linked with the Multi-Picture System (MPS) of Evans & Sutherland (bottom left), the then state-of-the-art vector graphics platform for interactive molecular structure display. A typical vector-graphics rendering of the experimental structure of the complex of *E. coli* dihydrofolate reductase with bound methotrexate, as determined by protein X-ray crystallography[1] is given in the center below. The insert on the bottom right illustrates one of the first examples of successful structure-based inhibitor design, displayed in high-resolution raster graphics that replaced vector graphics technologies in the mid-1990s.

System, the gold standard for interactive molecular structure display based on vector graphics technologies that dominated the 1980s. For good picture perception, one had to work in a completely dark room, which had both positive and negative aspects: The obvious disadvantage was that one could not perform molecular modeling in a standard office with ambient illumination; the advantage, however, was that the dark room eliminated the sense of day and night so that development work would continue unimpeded late into night and early morning hours.

The early modeling system was based on a novel generalized united-atom force field, in which even polar N–H and O–H units were condensed into united heavy atoms. In combination with novel geometrically anisotropic hydrogen bonding potentials, this force field had the great advantage to (i) realistically model hydrogen bonding interactions, (ii) be independent of torsional potentials involving polar hydrogen atoms, (iii) allow easy protonation or deprotonation, as this is handled simply by means of proton counts rather than explicit hydrogen atom positions, and (iv) provide hydrogen bond network analyses, including cooperative effects and surrounding water molecules, on the fly during dynamic molecular docking into the pockets of target proteins.

While this initial modeling system was already quite powerful,[6] it still lacked important functionality. The real breakthrough to a fully functional and highly versatile modeling system came when Dr. Paul Gerber (Fig. 2), a highly creative theoretical physicist with great understanding of molecular structures and energetics, joined the still young modeling group and started to rework and expand the original united-atom force field on sound physicochemical grounds, resulting in what is now known as the MAB force field.[7] One key aspect of this unique force field is the fact that it is based solely on atom types as given by the periodic system, deriving local variations of atom or bond types from the molecular topology of the local covalent environment using a minimum amount of parameters calibrated against experimental structures and conformational energy differences. Apart from many graphics display techniques and molecular structure handling modules, all novel at their time, a series of very important functionalities were developed, such as ring shape analysis,[8] multiple simultaneous structure superposition methods,[9] and highly efficient molecular docking and energy-minimization methods with interactive modalities for fixing or relaxing parts of the molecular system under optimization as well as active user interference during the process. As this new modeling system, MOLOC,[10] attracted not only the interest of medicinal chemists but also of X-ray crystallographers and NMR specialists, the program suite expanded quickly into the domains of structural biology by X-ray diffraction and NMR spectroscopy with unanticipated benefits for structure modeling by the CAMM group. Many of the original ideas developed and incorporated in MOLOC have eventually become available in commercial modeling systems; however, there are a number of very valuable functionalities which are still unique to MOLOC.[10]

During the 1990s, high-speed raster graphics technologies of sufficient resolution became available and eventually replaced vector graphics systems. Paul Gerber successfully moved the MOLOC system to the new graphics technology platforms, thus making it available on various high-end as well as low-cost graphics systems, including also a plethora of molecular structure display options that are only available, but in part also necessary for raster graphics systems.

Models need reality checks! A beautiful picture of a structural model may be deceptive and modeling errors are easily overlooked, particularly when a molecular structure fits very nicely into a pharmacophore model or a target protein pocket. For this reason, lean access to and efficient mining in collections of experimental structures of both small and large molecules and their complexes are essential.[6a,11] The CSD has grown over three decades to an overwhelming size of currently more than 650'000 crystal structures of small organic compounds, covering a wide range of structural diversity (Fig. 3). While there are still uncharted areas of interesting structural motifs, much highly relevant information can be extracted from this resource.

For quite some time in the past, however, substructural searches were based on non-interactive and time-consuming batch modes. Thanks to a mutually rewarding cooperation, Roche received permission to use the CSD raw data to build its own internal database. At that time, Hans-Jakob Ammann was just completing an external course in relational database concepts and successfully set up a novel database, ROCSD, in a remarkably short time. With some efforts to convince him to deviate from purist relational database concepts, his developments met with great success providing us with the much needed interactive access to substructural information[6a,12]

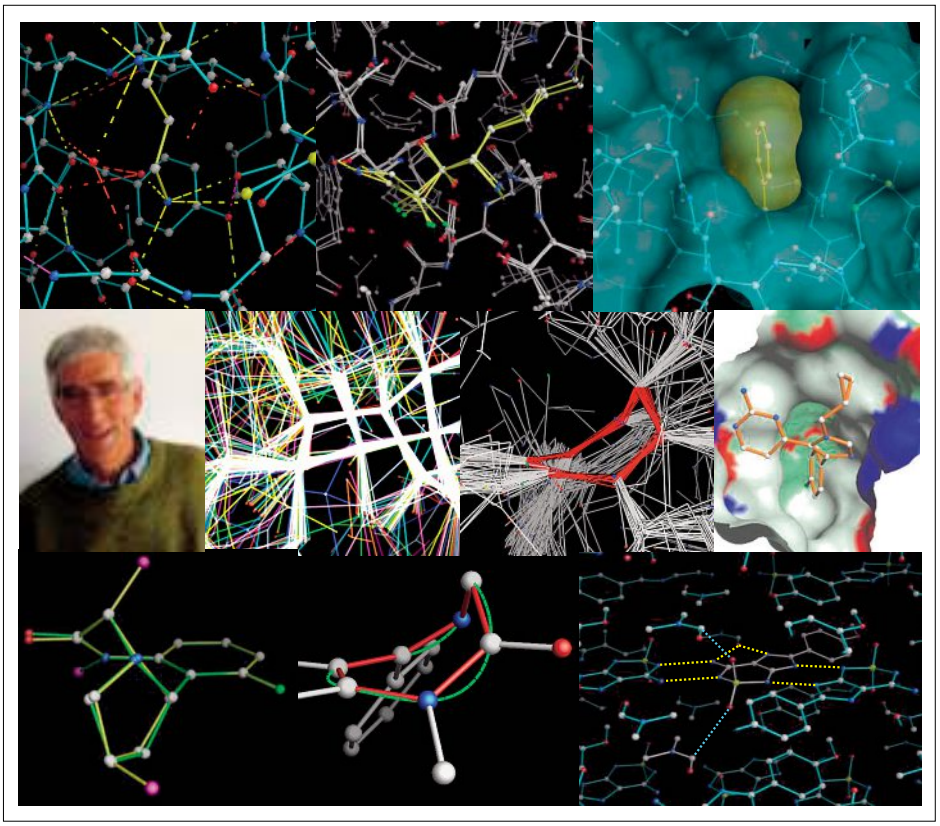

Fig. 2. Dr. Paul Gerber is the father of the molecular modeling platform MOLOC,[10] which emerged from the concepts of a novel united-atom force field originally developed for the Roche Interactive Molecular Graphics System (RIMG), but was then significantly developed further into a novel generic atom- and bond-based (MAB) force field;[7] it has been dramatically expanded to include most efficient structural modeling, superposition, and docking facilities with flexible and powerful energy-minimization procedures, ring shape, non-bonded interaction and crystal packing analyses, making use of various state-of-the-art graphics rendering options for optimal display of complex structural information.

for many years. In the meantime, the CCDC (Cambridge Crystallographic Data Center) Group developed its own interactive version that became so convincingly functional and interactive that, towards the turn of the century, the time for change had come, abandoning the internal system in favor of the much more powerful and functional ConQuest System for substructural searches with its companion tool, Vista, for subsequent structure-statistical analyses. While important developments of the CSD program suite are still ongoing today, the mining and analyses tools offered at that time already established a gold standard in structure-based molecular design. With its unabridged growth in content and structural diversity, the CSD system has become not only an indispensable resource for structure-based design, but also a rich source for chemical inspiration and education. The latter would call for a full integration at all levels of undergraduate and graduate chemical education.

The number of published protein crystal structures remained very low during the first decade of CAMM development (Fig. 4), and experimental structures of therapeutically relevant target proteins were rare due to the difficulties in isolation, purification, and crystallization of proteins from natural sources. The situation changed dramatically with the advent of heterologous gene expression technologies during the 1990s. Combined with improved purification and crystallization methodologies, in particular cryo-crystal structure determination techniques, and the availability of synchrotron light beams of high focus and brilliance, the number of available protein X-ray structures has risen dramatically since the mid-1990s, reaching a level of approximately 10'000 entries around the turn of the century and having passed the 100'000 mark during this year, with over 90'000 X-ray structures of proteins from all areas of the Life Sciences and a rich population of proteins of high therapeutic relevance.[13] The enormously beneficial enrichment of the structural contents in both PDB and CSD has been greatly fostered by the policy of peer-reviewed scientific journals that publication of molecular structures be complemented by deposition of the experimental structures in these worldwide sole data repositories.[14]

Despite the enormous growth of publicly accessible protein structures, the continuous need for a strong internal protein X-ray structure group, with the capacity to express, isolate, purify, and crystallize target proteins and selected mutants, with or without small molecules bound to them, has been recognized as a key success factor in structure-based drug discovery. Thanks to Dr. Fritz Winkler (Fig. 4), a distinguished

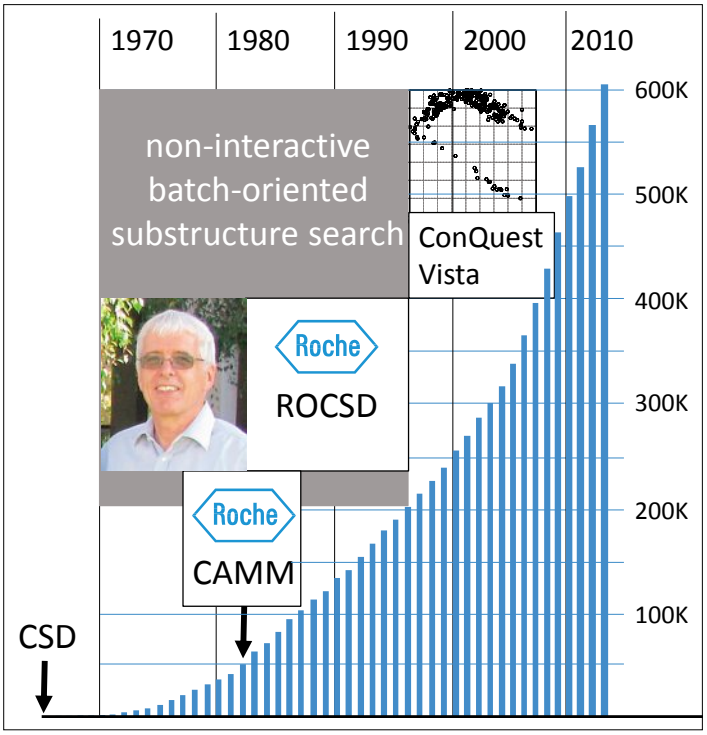

Fig. 3. The Cambridge Structural Database (CSD),[4] founded in 1965, contained some 50'000 small-molecule crystal structures at the time of CAMM development at Roche, but has been growing steadily over the decades, containing over 650'000 structural entries today. During the first decade of our CAMM development and applications, use of the CSD was hampered by the non-interactive mode of substructural searches. By special permission from the CSD, Roche received the CSD data and updates at regular intervals for internal use, which allowed Dr. Hans-Jakob Ammann to develop a Roche-internal relational database, ROCSD, with special enhancements for highly efficient substructure search and retrieval. ROCSD provided a competitive basis for efficient structure-based modeling until towards the end of the second decade of Roche's CAMM developments when the program suite of ConQuest and Vista by CSD had matured to the extent that ROCSD could be replaced by the highly versatile new tools by CSD. The latter nowadays represent a most powerful and indispensable platform not only for structure-based modeling, but also for teaching at all levels of chemical education.

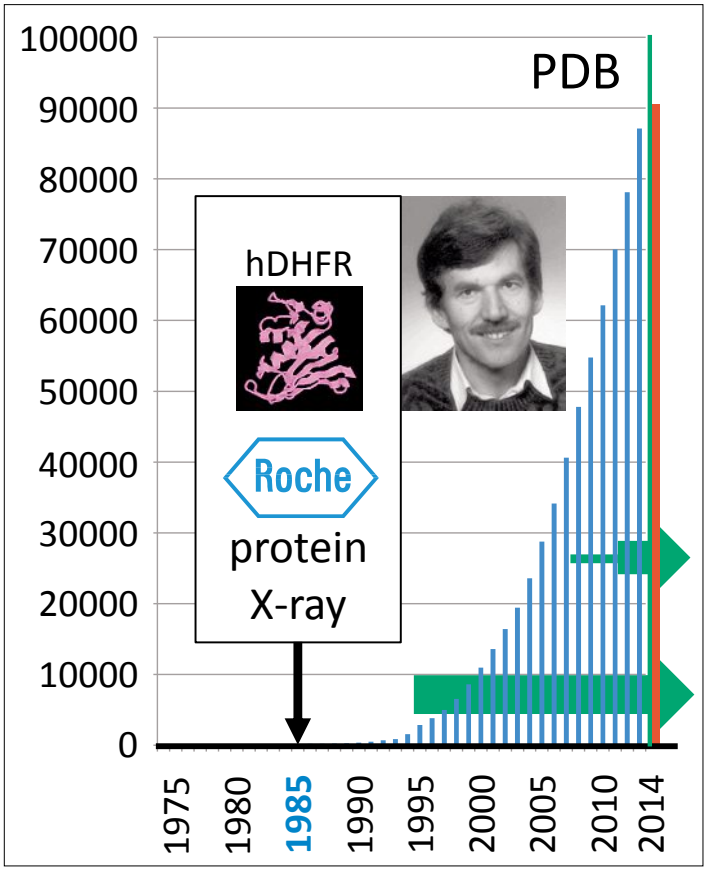

Fig. 4. The growing size of the Protein Data Bank (PDB) over four decades documented by the numbers of macromolecular structure depositions per year. When Dr. Fritz Winkler (center) started in 1985 to set up a Roche-internal group for protein X-ray crystal structure analysis with his inaugural structure determination of the human dihydrofolate reductase (hDHFR), the PDB contained less than 200 entries with only very few pharmaceutically relevant structures. The situation changed about 10 years later with the advent of heterologous protein expression, improved protein purification and crystallization techniques (bottom green solid arrow). The membrane-bound proteins remained a big hurdle until approximately 10 years ago (green arrow above). While each new membrane protein structure determination still presents significant challenges today, it appears that membrane protein crystallography may become more routine in the foreseeable future as novel purification, stabilization, and crystallization techniques are being invented.

scientist and expert in both small-molecule and protein X-ray crystal structure determination, a very successful and productive internal structure biology group could be built soon after the start of CAMM developments. This group documented its competence by solving the crystal structure of human DHFR less than a year after implementation of the internal protein X-ray diffraction setup, a problem that had not been solved externally for many years, but was essential in attempts to understand bacterial-to-human selectivity of DHFR inhibitors. Since then the group has solved, in a highly project-oriented focus, thousands of highly relevant target protein structures, both soluble and membrane-associated, that have contributed enormously to understanding inhibitor binding, creative molecular design, and opening of new discovery research opportunities. Many of these structures have been made publicly available; and the Group has engaged in much collaboration with many academic groups all around the world in solving challenging protein crystallographic problems jointly, building an extended international network and augmenting the Group's expertise and visibility in the field of structural biology.

The dramatic rise of protein X-ray structures was preceded by an upsurge in protein sequence data that stimulated much early work in sequence analyses and comparisons, as well as secondary structure and fold predictions. These activities triggered our curiosity early on, although direct practicable applications would not be immediately obvious. However, farsighted support by senior management, especially Dr. Conradin von Planta†, allowed us to hire a highly talented biochemist with excellent informatics training, Dr. Clemens Broger, to engage in various novel protein sequence analyses[15] while embarking at the same time on the systematic setup of what would later become 'bioinformatics' (Fig. 5). The internal build-up of well-structured protein and gene sequence databases with efficient search tools was seen by most colleagues internally as a luxury in the late 1980s. However, when Roche acquired the patents for the Polymerase Chain Reaction (PCR) methodology[16] and started to develop it into robust and reliable technology platforms for diagnostic applications, the urgent needs for such database systems came as an unforeseen early justification.

The initial phase of these 'bioinformatics' activities were fully dedicated to sequence analyses, multi-sequence alignments, and 1D-to-3D structure correlations. The latter was fueled by a then novel interactive tool, developed by Clemens Broger, which allowed us to visualize, interrogate, and interactively manipulate sequence alignments with two displays side-by-side always showing the 1D-alignment as well as the 3D-consequences for the superimposed sequences on the fly. Likewise, the 3D-matching and topology modification of a modeled sequence to a reference structure was directly mapped in the 1D-alignment display. This important tool uniquely integrated the one- and three-dimensional worlds of protein structures and enabled proper fold predictions of whole proteins or characteristic domains. The stage was thus set for efficient classification of proteins and domains, as well as the reverse, the comprehensive search for proteins belonging to families of similar biological functions or to classes of similar folds based on whole-sequence or 3D-structure-assisted non-contiguous sequence pattern alignments.

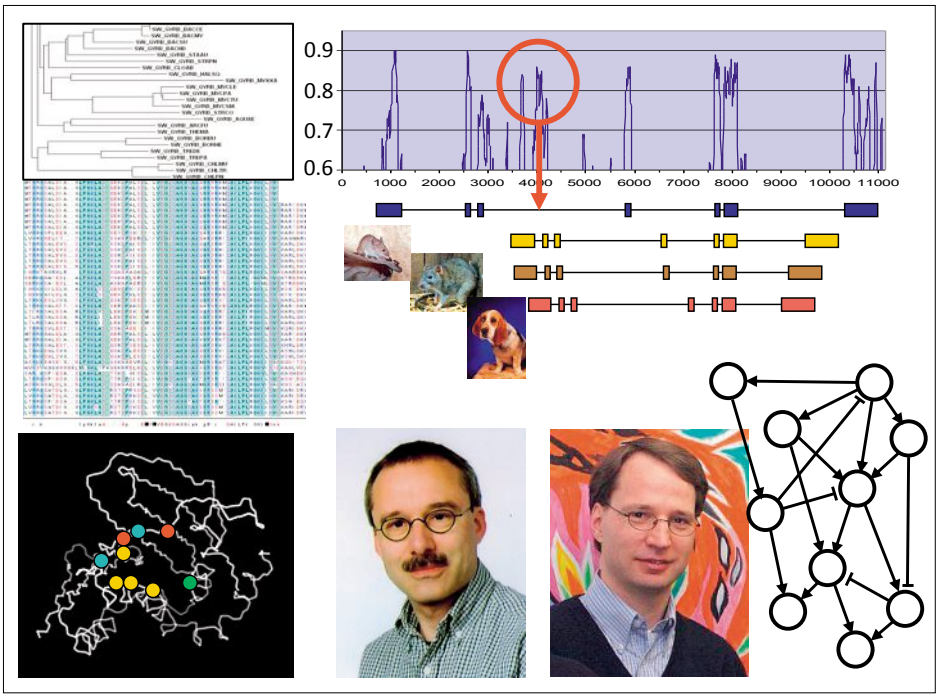

Fig. 5. Dr. Clemens Broger (bottom center) started in 1989 to set up a Bioinformatics Group at Roche; the focus then was on sequence analysis, multiple-sequence alignment, and sequence-to-3D-structure correlation for protein classification and fold prediction, as well as the set-up of consistent comprehensive Roche-internal protein and DNA sequence databases with efficient mining systems; Dr. Martin Ebeling (bottom right), current head of the Roche Bioinformatics Group, developed a highly interactive Comparative Genome Analysis tool which resulted in the successful identification of regulatory elements in non-coding gene sequences, as illustrated on the top. One current focus of the Bioinformatics Group at Roche is the identification and analysis of gene regulatory networks by computational modeling (bottom right).

In parallel to these developments, the sequencing of whole genomes of microorganisms and small animals took place, which, around the turn of the century, culminated in the first draft of the human genome.[17] Likewise, the genomes of some important animals, like mouse, rat, dog, became available. This stimulated a new collaborator in the still young 'Bioinformatics Group', Dr. Martin Ebeling, a highly creative biochemist with excellent bioinformatics expertise (Fig. 5), to develop a then novel graphic-interactive tool for comparative genome analyses. This tool allowed us to interrogate and comparatively analyze whole genome sequences of different species regarding both coding and non-coding regions. It had its first outstanding success in the correct prediction of gene-regulatory elements in non-coding gene sequence domains already some 10 years ago. This tool has been expanded ever since its first implementation and has proved to be an indispensable tool in coping with an ever increasing amount and diversity of genomic information that have become available over the last decade.

A major and well accomplished task of the 'Bioinformatics Group' has been to provide a highly functional platform with integrated databases and diverse data resources, with powerful analysis tools and efficient navigation systems so that challenging biological discussions can be carried out in highly productive interactive sessions. As new areas in biology and bioinformatics open, such as the fascinating, but complex worlds of RNA biology or epigenetics, this highly integrated and well-structured 'bioinformatics platform' is continuously expanded to incorporate new information contents, functionality, and navigational outreach. Our internal 'bioinformatics' efforts have expanded into the management and analysis of extended gene networks, and the modeling and simulation of complex gene transcrip-

tion regulation. We are witnessing the emergence of 'computational biology', which will put us on a qualitatively and quantitatively new level of comprehending intra- and intercellular communication and regulation processes. This will undoubtedly play an important role as a powerful integrator of experimental and theoretical biology as well as biophysics and become another key pillar in early pharmaceutical research to drug discovery.

When the hype of 'One Design – One Drug' did not materialize towards the end of the first decade, general interest turned to screening approaches, dismissed before as 'non-imaginative' by comparison to 'rational design'. Large, eventually ultra-large compound libraries and (ultra) high-throughput screening became the name of the game. This called for innovations in both areas of compound repository and screening systems. Regarding the former, novel technologies were required that would enable high-throughput random access to individual samples and pick and place operations all at low temperatures, without exposure to air and moisture or intervening thaw-freeze cycles. At that time sample handling and bio-screening were dominated by the concepts of multi-well plate handling techniques so that even the most advanced technology drivers in the field would reject our collective requirements as 'mission impossible'. This was one of those typical situations where Dr. Christof Fattinger, a highly talented physicist of extraordinary experimental and theoretical skills, became creative. He found the basic solution to the problem by substituting conventional well-plates by doubly-open racks filled with mini-tubes (Fig. 6).[18] The difficult fit, or typical misfit, of cylindrical objects into square holes has been a small, but ingenious detail, in which mini-tubes, by natural friction, can be firmly held in the rack, filled with aliquots of sample solutions under dry inert gas, sealed in arrays using plastic-coated aluminum foil, applying a very short welding step, followed by stamping out the individual seals in a one-step cutting process and pushing the whole mini-tube array into its final rack position. Racks cooled to –20 °C with frozen samples in the sealed mini-tubes are then properly placed by a robot into the appropriate positions in the large compound repository, providing room for some 30 million samples. Individual sample pick and place operations would involve the robot-assisted retrieval of the appropriate rack, its proper positioning above a receiver (customer) rack and pushing the requested mini-tube by a properly designed piston vertically from the storage into the receiver rack without violating the seal. All operations are computer-controlled and take place at low temperature (–20 °C) with an average performance rate of 6'000 random sample pick and place operations per day. Only at the time of the assay, filled frozen customer racks would be thawed, the seals punched open in parallel and all required liquid-handling operations performed in standard parallel ways at room or assay temperature.

The successful development of this 'smart-RCD' (smart Roche Compound Depository) system had to overcome several other critical hurdles, such as high-precision low-temperature robotics, continuous sample replacement and position optimization, system maintenance without warming the storage facility, as well as fully integrated computer control, logging, and error handling procedures. All these critical matters could be successfully solved in close collaboration with the then Swiss company REMP AG specialized in all aspects of high-precision robotics. Another important issue that had to be addressed from the beginning was the smooth integration of the new system into the working and chemo-informatics environment of the medicinal chemists. The successful deployment of the new 'smart-RCD' has become a most instructive demonstration of the importance and value of a timely planned and executed system programming and informatics integration into the targeted application environment, ensuring productive use of the new system essentially at the time of its physical setup. With this successfully established, the new system has enabled medicinal chemists to interrogate a huge collection of existing compounds by computer, search for structural motifs in connection with given protein targets, easily select and electronically submit lists of interesting compounds for retrieval, assembling, and bioassays. Thus, biological results for selected sets of compounds and often first exploratory structure–activity relationships have been available prior to starting any synthetic work in the lab, thus establishing a new paradigm in medicinal chemistry.

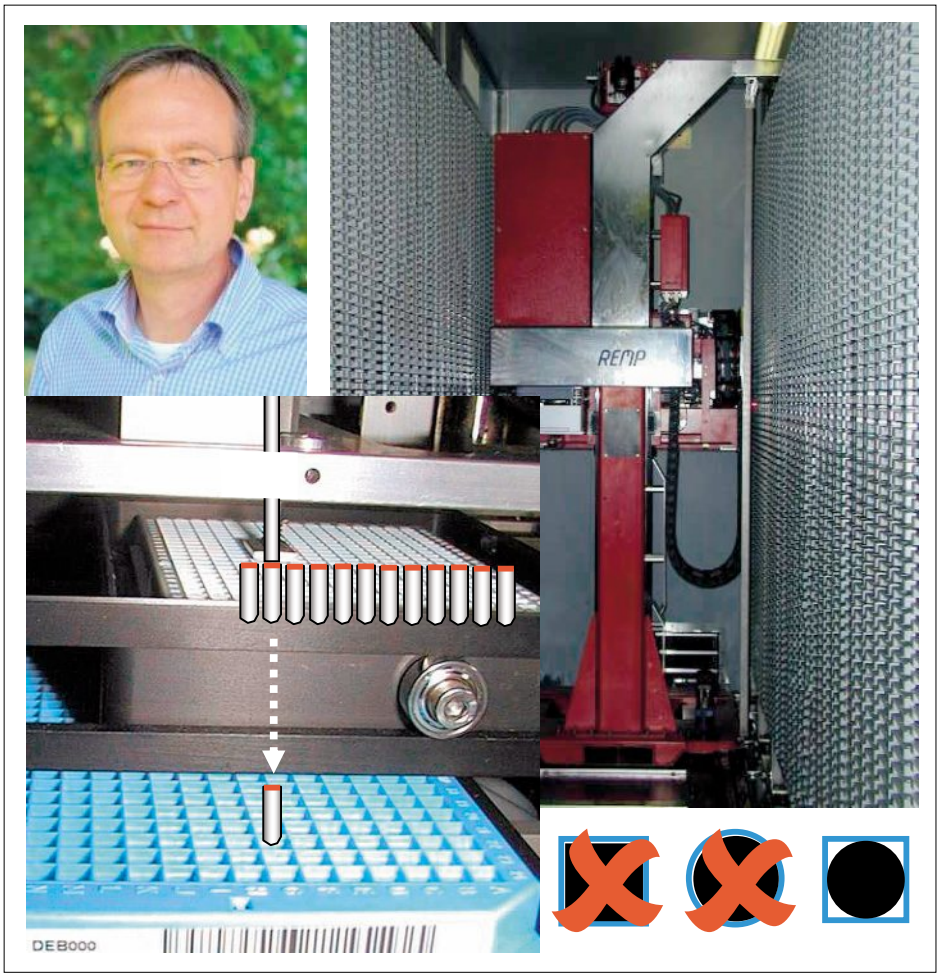

Fig. 6. Dr. Christof Fattinger developed the 'smart Roche Compound Depository' ('smart RCD'), which is a low-temperature (–20 °C) robot-operated depository holding some 30 million sealed mini-tubes of frozen compound samples in open racks. The computer-controlled picking and placing of a specific mini-tube (bottom left corner) is by a precise vertical push of the sealed mini-tube by a piston from the storage to a properly positioned receiver (customer) rack (blue) without violating the seal; up to 6'000 pick-and-place operations for randomly selected samples are possible. The mini-tubes are firmly held in both the storage and the receiver rack due to friction between a square container hole and a cylindrical tube (bottom right).

While biological screening of compounds could typically handle some 100 samples per day at the beginning of the 1990s, enhanced read-out methodologies, in particular fluorescence-based assay methodologies, improved parallel pipetting stations, and whole-plate imaging readout systems, increased the screening capacity at least an order of magnitude to high-throughput (HT) screening.[19] Again, the enormous pressure for increased capacity and performance of biological screening, spurred the creativity of Christof Fattinger in collaboration with the Carl Zeiss Group in Germany.

I vividly remember a brain-storming discussion with a leading delegation of the Carl Zeiss Group in the middle of the 1990s about optical instruments in the bio-analytics area when a tentatively placed idea of splitting an intense broad light beam into 95 parallel and equally focused light beams was spontaneously rejected as 'mission impossible', but then taken up by the head of the Zeiss delegation as something worth trying, just to come up with the multi-lens plate solution only a few months later (Fig. 7).[20] This parallel-beam concept offered many advantages. It allowed instant read-out of a 95-well plate, similar to plate imaging technologies; but unlike the latter, the parallel-beam methodology allowed different modes of light focusing into the samples according to specific needs, and for higher-density plates, such as 384-well plates, the parallel-beam methodology avoided the problem of well-to-well cross-talk, by taking four readouts in the 95-beam format repositioning the plate four times in rapid succession.

Another important innovation was the invention of turn-tables with four orthogonal entry/exit ports for efficient plate transport allowing plates to cross. This innovation, jointly by Christof Fattinger and Hansjörg Tschirky (Fig. 7), a creative expert and at that time Head of the Mechanical Workshop at Roche Basel, enabled a radical change in the configuration of screening systems. The conventional linear arrangement of robot stations with their inherent limitation of processing capacity could be abandoned in favor of two-dimensional arrangements of multi-tasking workstations along a central plate highway from which plates could be taken into and simultaneously exported from the workstations, with concomitant increase in processing capacity and parallel complex multi-tasking. Combined with the high-speed parallel readout system, the nominal capacity of this novel screening system has been estimated to be in the order of 100'000 samples per day, so that huge compound collections could now be screened within a couple of days.

It is not without irony, though, that at the time when this novel ultra-HT screening system became operational, the initial hype of super-large compound collections from combinatorial chemistry had largely faded away yielding to convincing arguments for chemical quality of compound libraries in terms of both purity and drug-like properties. This implied generation and screening of smaller and well-designed compound collections that would generally exhibit much higher hit rates with more useful results than purely random screening of big compound libraries. Focused screening of smaller compound collections, sometimes containing no more than 40–100 members, had become the favorite mode of operation. Such libraries are typically designed around a target-specific structural motif with well-designed diversity for preliminary structure–property analyses in order to assess early on the potential for lead optimization and, in favorable cases, even to gain structural insight into possible target-binding modes by X-ray structure analyses of selected protein-molecule complexes. In spite of the conceptual return from super-large to small focused compound libraries, the need for large random screening campaigns has remained; and having a platform for ultra-HT bio-analytics in place has remained an invaluable asset.

The cumulative experience with plate design and manipulation, miniaturization and parallelization[19] provided ideal foundations for much innovation in the area of physicochemical and biophysical compound property analytics. A successful start was given just before the turn of the century by the development of the PAMPA method (Parallel Artificial Membrane Permeation Assay[21]). While this HT-analytical device for preliminary assessment of membrane permeation of compounds met with critical reservation by the expert pharmacologists,[22] it has been gradually adopted by the medicinal chemistry community as a simple, cheap, and fast methodology for a first qualitative assessment of passive membrane permeability.[23] Various important PAMPA formats have since been developed to mimic the properties of different types of biological membranes,[24] and PAMPA-type assays now represent a standard across the pharmaceutical industry. Several other miniaturized and parallelized medium- to high-throughput assays have been developed in fast succession, such as for measuring compound lipophilicity, acidity and basicity, solubility, chemical and metabolic stability, and many other important key properties for the early identification of potential problems when moving compounds from early discovery stages into *in vivo* pharmacological testing. The key players in these developments were Drs. Manfred Kansy and Holger Fischer (Fig. 8), two pharmaceutical scientists with ex-

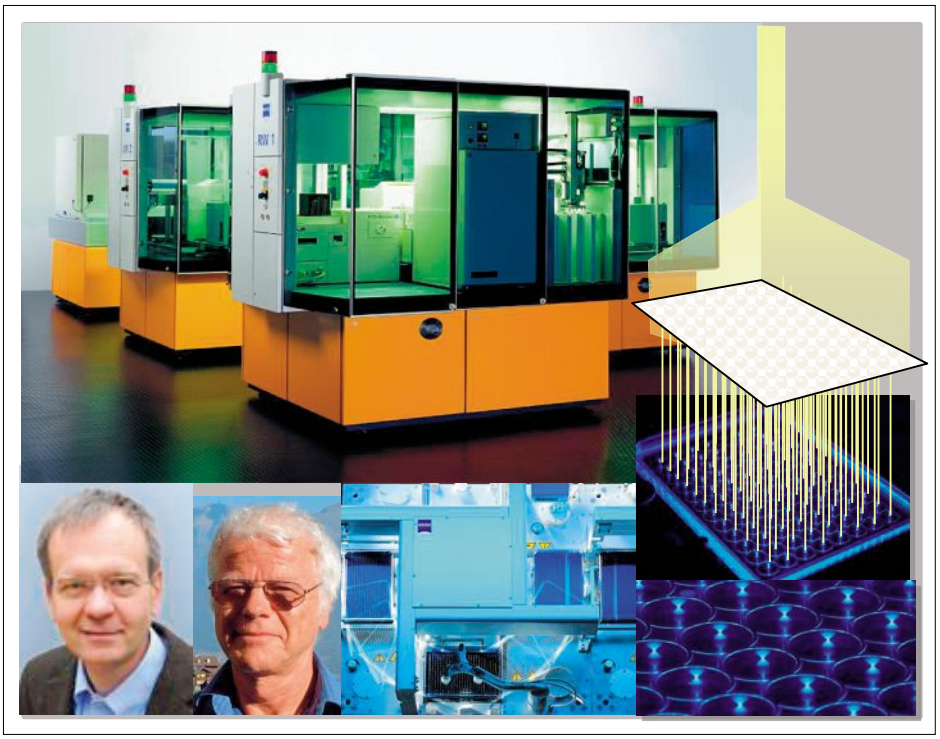

Fig. 7. The ultra-High-Throughput Screening (uHTS) system (top) developed at Roche under the guidance of Dr. Christof Fattinger (bottom left) is based on two main innovations: (1) the splitting of a broad parallel input light beam into 96 parallel and equally focused light beams with special lens plates (right) for instant parallel read-out, successfully developed in collaboration with the Carl Zeiss Group in Germany, and (2) the plate turn-table for plate access and export, allowing incoming and outgoing plates to cross (bottom), successfully developed in collaboration with the then Head of the Roche Mechanics Workshop, Hansjörg Tschirky (bottom, second from left).

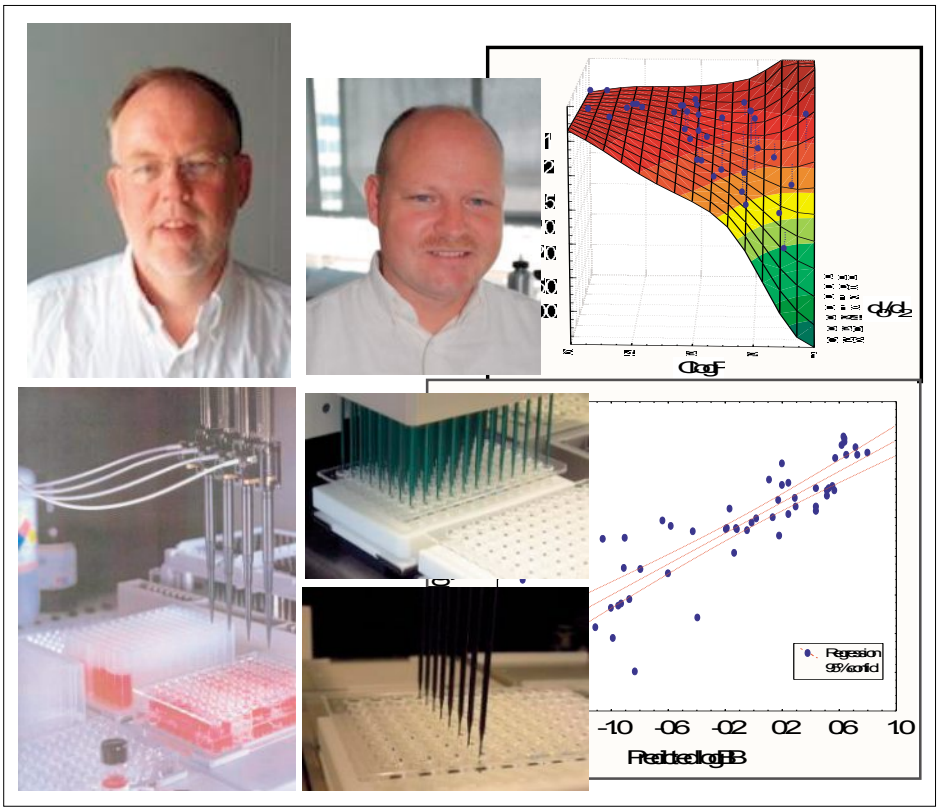

Fig. 8. Drs. Manfred Kansy (top left) and Holger Fischer (top middle) have been the leading scientists involved in the development of novel miniaturized parallel analysis techniques for the measurement of important physicochemical properties for chemical lead optimization; they played also key roles in data interpretation, structure–property correlation, and cautious attempts to develop property prediction tools.

cellent chemical, theoretical and experimental backgrounds as well as interest and commitment in structure–property relationship analysis.

The advent of technological platforms for the rapid measurement of important bio-physicochemical and pharmacologically relevant properties, using only minute compound samples caused a paradigm shift in medicinal chemistry (Fig. 9). Before the turn of the century, early drug discovery was largely a linear sequential process where medicinal chemistry would primarily focus on compound potency and selectivity, then moving forward with a set of selected compounds into pharmacological testing. Critical findings and concomitant failures were the rule of the game, and compounds surviving the 'cruel' testing by the pharmacologists represented rare and celebrated singularities. After the turn of the century, with the new technology platform for rapid measurements of critical compound properties in place, 'multi-dimensional optimization' (MDO) strategies allowed medicinal chemists to design and synthesize new compounds with simultaneous consideration of both potency and drug-like properties.[25] Since then the progress of compounds from early discovery stages into pharmacological *in vivo* testing has become more and more an event of confidence and rare surprises.

The mastering of this early hurdle in drug discovery and development has its prospective parallel in the current attempts to overcome the difficult second and third hurdles from *in vitro* and *in vivo* animal pharmacology to early- and later-stage clinical studies. Much resembles the developments around the turn of the century. Again novel concepts and methodologies, spurred by great technological advances, promise to provide major breakthroughs. In fact, genomics and proteomics, imaging, HT-sequencing, modeling and simulation, as well as, and again, miniaturized parallel devices for highly specific and efficient large-scale diagnostics are contributing much to biomarkers research and translational medicine, holding great promise to create new paradigm shifts in drug discovery and reduce many of the hurdles and risk factors in the early and later phase transitions of drug development.[26]

The availability of compound property data has promoted structure–property correlation activities over many decades. Many attempts of property predictions have been made, but with moderate success,[27] largely due to the unavailability of sufficiently large and diverse data sets in the public domain and insufficient consideration of differentiating structural aspects. The continuous accumulation of an immense and ever-growing collection of experimental physicochemical and biophysical property data on vast and diverse compound collections may help in the future to provide significantly im-

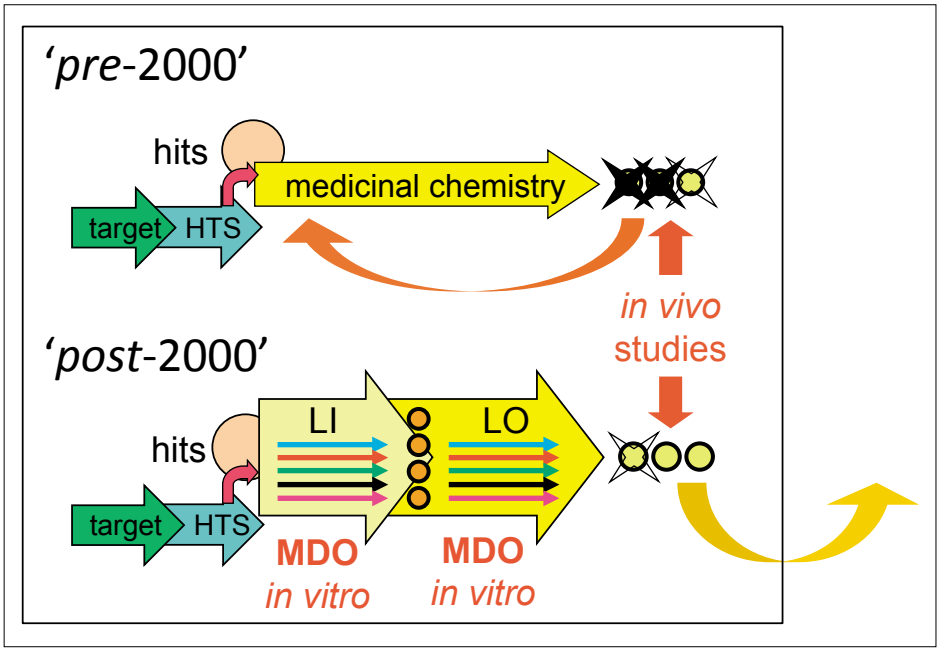

Fig. 9. Around the turn of the century miniaturized parallel analytical methodologies for efficient measurement of physicochemical and biophysical compound properties using only minute amounts of samples became available, resulting in a paradigm shift for medicinal chemistry. The essentially sequential workflow from screening to hit identification to lead optimization regularly ran into adverse findings in subsequent *in vivo* pharmacological studies. With efficient and high-capacity analytical tools in place, transitions from hits to lead identification (LI) and optimization (LO) are conducted in a parallel manner by multi-dimensional optimization (MDO), resulting in significantly improved quality of candidate compounds for *in vivo* testing, and hence reduced frequency of compound rejections.

proved tools. However, in the meantime, another concept, the Molecular Matched Pair (MMP) analysis[28] has come into play. While the MMP concept is basically not new, but had to be applied by hand in rather time-consuming endeavors (see *e.g.* ref. [29]), the development of automated and easy-to-use MMP tools by different groups during the last decade, mostly in industry where large compound databases are available, the situation has dramatically changed. MMP analysis is a powerful concept by which data are retrieved specifically for pairs of compounds which differ structurally exclusively in one prespecified structural replacement. If many diverse pairs of compounds with measured data are available, the resulting information provides insight into both the property change due to the structural replacement and its dependence on the structural context around the site of substitution. Key people in the Group of Dr. Manfred Kansy were Drs. Stefanie Bendels and Gregori Gerebtzoff, both excellent informatics specialists with solid statistical and chemical backgrounds (Fig. 10). Originally called ComPair, with a further twist in its logo (Fig. 10), the concept of MMP has now been significantly expanded to a platform called LUCID and developed by Gregori Gerebtzoff. It represents a next major innovation, by enabling data comparison not only for single matched pairs, but essentially unlimited matched series of compound pairs, with reference to only a single specified seed structure or substructural unit. This novel concept has turned out to be extremely powerful and highly stimulating in all its applications, immediately producing much relevant lateral insight in the context of the results of prime interest. A particularly innovative feature of LUCID is its capability to respond to user requests regarding desired changes in one or more properties for a compound at hand, quickly retrieving and displaying in a highly structured and easily interpretable way all experimental cases across the whole database, *i.e.* across all past or present projects and therapeutic departments of the company, thus suggesting possible structural solutions by analogy to the case at hand.

An important aspect of MMP or LUCID analyses is that they ought to be and are based on consistent experimental, not predicted or computed property data. Thus, such tools are currently particularly powerful in Big Pharma industry where typically millions of compounds with well curated property data are available. However, to the extent that academic institutions would also commit themselves to measuring compound properties, rather than resorting to computational schemes or simply ignoring them, one could imagine

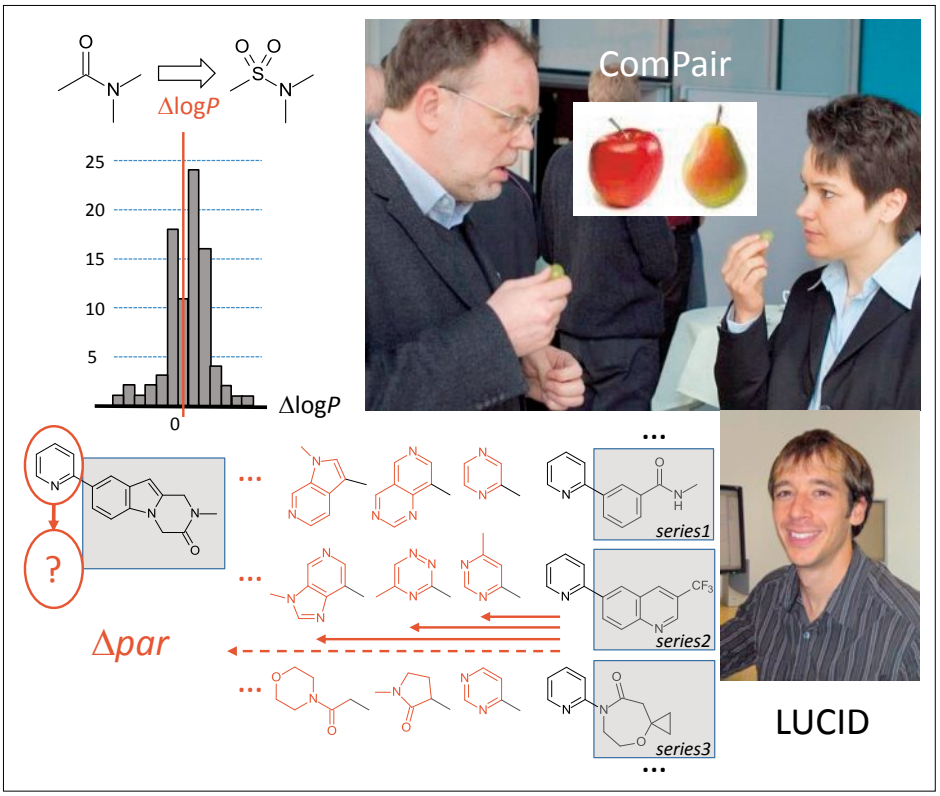

Fig. 10. The availability of vast amounts of experimental properties of an ever-growing and structurally diverse set of compounds has triggered the development of various data mining methods. Among them, the Matched Molecular Pair (MMP) analysis is a very powerful tool to explore the effect on selected compound properties upon exchange of a distinct structural element (example at the top). This methodology has been developed early at Roche, with Drs. Stefanie Bendels (top right) and Gregori Gerebtzoff (bottom right) as the key scientists involved in the development of ComPair, the first implementation of an interactive MMP analysis tool, and the further development of LUCID by Gregori Gerebtzoff, expanding the 'Matched Pair' to a 'Matched Series' concept. The latter is a very powerful and easily manageable data mining tool that provides structure- and property-based guidance to medicinal chemists by mining through the accumulated data pool within a company; the bottom part illustrates one of many possible applications requesting suggestions for structural modifications to effect a desired property modulation by analogies from other projects.

that collective efforts by academia worldwide could, within less than a decade, produce a compound and property database that could even dwarf the collections in individual industries.

With large databases of 3D-structural information and compound properties with lean access and mining tools in place, the stage is now set for combined structure–property-guided molecular design, which brings us an important step closer to 'rational drug design', a term that has been carelessly used since the dawn of CAMM. Apart from the fact that there is hardly any 'non-rational design', we are still a long way from the ultimate goal of 'drug design', not only regarding critical issues in safety and toxicology assessments at the preclinical stage, but also, and in particular in view of much needed research and technological developments in translational medicine and personalized healthcare.

Data mining in structure and property databases is sometimes frustrated by missing data or matching structures. While the former can be corrected by resubmission of still existing compounds to corresponding analytical measurements, the latter requires explicit exploration of new compound series with novel substructural elements, an endeavor that is increasingly difficult at times of limited chemical resources and increased project focus. As an example, we recall the successful introduction of oxetane as a promising property-modulating unit into medicinal chemistry.[30] While the concept of oxetane as a polar substitute for a *gem*-dimethyl group was rejected internally by pharmacologists as metabolically too unstable, and skeptically regarded by medicinal chemists as chemically too reactive and synthetically cumbersome, it needed again the farsighted support by senior management, specifically Dr. René Imhof (Fig. 11), then Head of the Roche Pharma Research Center, Basel, to allow exploration of this concept in a collaboration with Prof. Erick M. Carreira, ETH Zürich. With an excellent PhD student, Georg Wuitschick and

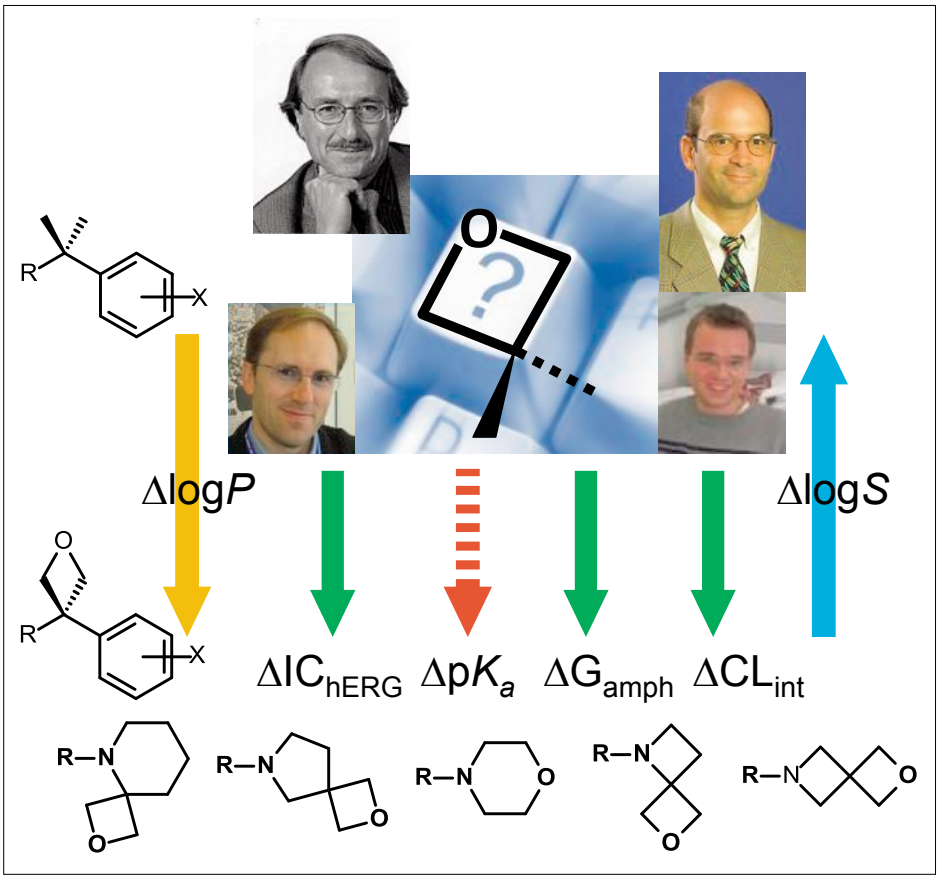

Fig. 11. The oxetane as an interesting unit in medicinal chemistry could be explored in a fruitful collaboration between Roche (Dr. Mark Rogers-Evans as Mentor, bottom left) and Prof. Erick M. Carreira at ETH Zurich (top right), supported by Dr. René Imhof (top left), then Head of the Roche Research Center Basel, and involving Georg Wuitschik as PhD student, who, after successful conclusion of his PhD thesis and a postdoctoral stay with Prof. Steven Ley at University of Cambridge, joined the Synthesis and Process Research Group at Roche Basel. By contrast to wide-spread skepticism, the oxetane unit turned out to be a very promising property-modulating group worth trying in medicinal chemistry. This could be shown through detailed explorations of prototypic acyclic and cyclic compound series which also established novel short and robust synthetic approaches for the incorporation of the oxetane unit into structurally diverse molecular substrates.

a Roche Mentor, Dr. Mark Rogers-Evans, a highly innovative and skilled medicinal chemist of insatiable curiosity, two major compound series with acyclic and spirocyclic amine scaffolds, respectively, offering opportunities to introduce the oxetane unit at different locations relative to the amine group, were synthesized and the relevant physicochemical and biophysical properties measured (Fig. 11).

To our great satisfaction, the oxetane derivatives turned out to be considerably more stable, both chemically and metabolically, than ever anticipated and exhibited many beneficial effects, such as reduced lipophilicity, increased solubility, reduced amphiphilicity and hERG liability, as well as interesting basicity modulation as a function of topological distance between the oxetane and amine functions. Quasi as a by-product of these study efforts, many novel and high-yielding synthetic access routes for the incorporation of the oxetane unit into complex molecular substrates had been identified,[31] which opened the gate for many applications in both academia and industry; the latter manifesting itself in an exponential upsurge of the numbers of patent applications including the oxetane either as a modulating unit or the critical core structure. Likewise, the whole repertoire of key building blocks is now offered commercially by various chemical providers[32] enabling rapid synthesis of compounds containing the oxetane unit.

Some of the smaller members of the cyclic amines with spiro-connected oxetane units exhibited not only interesting physicochemical properties, but also particularly promising structural features that merited follow-up studies of such constructs, also replacing the oxetane by other saturated 4-membered heterocyclic units, as well as a variety of specifically substituted analogues as 'compact scaffolds' for medicinal chemistry.[33] These studies have been pursued by Mark Rogers-Evans both internally and in continued collaboration with the group of Prof. E. M. Carreira; and many further interesting applications of the oxetane as a remarkable structure- and property-modulating unit have been and are currently being explored.[34]

This account provides only a very sketchy and necessarily incomplete description of major events of a fascinating journey of developing and implementing key elements of structure- and property-based molecular design to drug discovery at Roche over three decades. To have been associated and collaborate with outstanding colleagues and experts in my own or neighboring disciplines throughout all these years has been one of my greatest privileges at Roche. However, I am enjoying yet another great privilege that comes with the first like a twin: to remain associated with outstanding colleagues who are always ready to pick up ideas and carry them further on their own. Many of them have been explicitly mentioned in this short account. However, there have been many more involved in these endeavors which have not been included here for the sake of brevity. To all of them I wish to express my whole-hearted gratitude for their support, collaboration, joint and individual contributions that have dramatically changed the way pharmaceutical research at Roche is being conducted.

Received: June 30, 2014

[1] J. T. Bolin, D. J. Filman, D. A. Matthews, R. C. Hamlin, J. Kraut, *J. Biol. Chem.* **1982**, *257*, 13650.

[2] a) A. T. Hagler, E. Huler, S. Lifson, J. *Am. Chem. Soc.* **1974**, *96*, 5319; b) T. Philip, R. L. Cook, T. B. Malloy, Jr., N. L. Allinger, S. Chang, Y. Yuh, *J. Am. Chem. Soc.* **1981**, *103*, 2151; c) H. Dodziuk, H. Von Voithenberg, N. L. Allinger, *Tetrahedron* **1982**, *38*, 2811; d) J. H. Lii, S. Gallion, C. Bender, H. Wikstroem, N. L. Allinger, K. M. Flurchik, M. M. Teeter, *J. Comput. Chem.* **1989**, *10*, 503.

[3] a) O. Ermer, S. Lifson, *J. Am. Chem. Soc.* **1973**, *95*, 4121; b) N. L. Allinger, *J. Am. Chem. Soc.* **1977**, *99*, 8127; c) S. Yamamoto, M. Nakata, T. Fukuyama, K. Kuchitsu, D. Hasselmann, O. Ermer, *J. Phys. Chem.* **1982**, *86*, 529.

[4] *http://www.ccdc.cam.ac.uk/pages/home.aspx*

[5] *http://www.wwpdb.org/*

[6] a) K. Müller, H. J. Ammann, D. M. Doran, P. R. Gerber, K. Gubernator, G. Schrepfer, in 'QSAR: Quantitative Structure-Activity Relationships in Drug Design', J. L. Fauchere, Ed. A. R. Liss, **1989**, p. 219; b) K. Müller, in 'Entwicklungstendenzen in der Arzneimittelforschung', Ed. H. J. Dengler, Gustav Fischer, New York **1989**, p. 21.

[7] P. R. Gerber, K. Müller, *J. Computer-Aided Mol. Design* **1995**, *9*, 251.

[8] P. R. Gerber, K. Gubernator, K. Müller, *Helv. Chim. Acta* **1988**, *71*, 1429.

[9] P. R. Gerber, K. Müller, *Acta Cryst., Section A* **1987**, *A43*, 426.

[10] *http://www.moloc.ch/*

[11] C. Bissantz, B. Kuhn, M. Stahl, *J. Med. Chem.* **2010**, *53*, 5061.

[12] a) K. A. Brameld, B. Kuhn, D. C. Reuter, M. Stahl, *J. Chem. Info. Mod.* **2008**, *48*, 1; b) C. Schärfer, T. Schulz-Gasch, H.-C. Ehrlich, W. Guba, M. Rarey, M. Stahl, *J. Med. Chem.* **2013**, *56*, 2016.

[13] *http://www.rcsb.org/pdb/statistics/contentGrowthChart.do?content=molType-protein&seqid=100*.
[14] a) *http://www.nature.com/authors/policies/availability.html;* b) *http://www.wwpdb.org/policy.html;* c) *http://www.iucr.org/home/leading-article/2011/2011-06-02.*
[15] C. Broger, K. Müller, *Struct. Correl.* **1994**, *2*, 685.
[16] a) R. K. Saiki, S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, N. Arnheim, *Science* **1985**, *230*, 1350; b) H. A. Erlich, G. Horn, R. K. Saiki, S. Stoffel, K. B. Mullis, F. C. Lawyer, D. H. Gelfand, Eur. Pat. Appl. **1988**, EP 258017 A2 19880302; c) K. B. Mullis, H. A. Erlich, D. H. Gelfand, G. Horn, R. K. Randall, US Patent 4965188 A 19901023; d) *http://molecular.roche.com/pcr/pages/history.aspx*.
[17] E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.-F. Cheng, A. Olsen, S. Lucas, C.Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J.-A. Dubois, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A.Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H.-C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. R. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. A. Smit, E. Stupka, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S.-P. Yang, R.-F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, A. Patrinos, M. J. Morgan, *Nature* **2001**, *411*, 720.
[18] C. Fattinger, H. Tschirky, US-Pat. 6'827'907 B2, **2004**.
[19] C. Fattinger, G. Dernick, in 'Exploiting Chemical Diversity for Drug Discovery', Eds. P. A. Bartlett, M. Entzeroth, Royal Society of Chemistry, Cambridge, **2006**, p. 203.
[20] F. Müller, C. Fattinger, in 'Single-Photon Imaging', Eds. P. Seitz, A. J. P. Theuwissen, Springer, Heidelberg, Berlin, **2010**, p. 327, 331.
[21] a) M. Kansy, F. Senner, K. Gubernator, *J. Med. Chem.* **1998**, *41*, 1007; b) M. Kansy, A. Avdeef, H. Fischer, *Drug Discovery Today* **2004**, *1*, 349.
[22] D. Galinis-Luciani, L. Nguyen, M. Yazdanian, *J. Pharm. Sci.* **2007**, *96*, **2886**.
[23] A. Avdeef, S. Bendels, L. Di, B. Faller, M. Kansy, K. Sugano, Y. Yamauchi, *J. Pharm. Sci.* **2007**, *96*, 2893.
[24] a) L. Di, E. H. Kerns, K. Fan, O. J. McConnell, G. T. Carter, *Eur. J. Med. Chem.* **2003**, *38*, 223; b) G. Ottaviani, S. Martel, P. A. Carrupt, *J. Med. Chem.* **2006**, *49*, 3948; c) B. Sinko, T. M. Garrigues, G. T. Balogh, Z. K. Nagy, O. Tsinman, A. Avdeef, K. Takacs-Novak, *Eur. J. Pharma. Sci.* **2012**, *45*, 698.
[25] a) K. H. Bleicher, H.-J. Böhm, K. Müller, A. I. Alanine, *Nature Reviews Drug Discovery* **2003**, *2*, 369; b) P. Gribbon, A. Sewing, *Drug Discovery Today* **2005**, *10*, 17.
[26] a) M. Day, J. L. Rutkowski, G. Z. Feuerstein, *Adv. Exp. Med. Biol.* **2009**, *655*, 1; b) C. S. Fishburn, *Drug Discovery Today* **2013**, *18*, 487.
[27] a) R. Mannhold, G. I. Poda, C. Osetrmann, I. V. Tetko, *J. Pharma. Sci.* **2009**, *98*, 861; b) I. V. Tetko, G. I. Poda, C. Ostermann, R. Mannhold, *QSAR Comb. Sci.* **2009**, *28*, 845; c) L. Settimo, K. Bellman, R. M. A. Knegtel, *Pharm. Res.* **2014**, *31*, 1082; d) M. J. Waring, *Expert Opin. Drug Discov.* **2010**, *5*, 235.
[28] a) J. Hussain, C. Rea, *J. Chem. Inf. Model.* **2010**, *50*, 339; b) G. Papadatos, M. Alkouri, V. J. Gillet, P. Willett, *J. Chem. Inf. Model.* **2010**, *50*, 1872; c) E. Griffen, A. G. Leach, G .R. Robb, D. J. Warner, *J. Med. Chem.* **2011**, *54*, 7739.
[29] H.-J. Boehm, D. Banner, S. Bendels, M. Kansy, B. Kuhn, K. Müller, U. Obst-Sander, M. Stahl, *ChemBioChem* **2004**, *5*, 637.
[30] a) G. Wuitschik, M. Rogers-Evans, K. Müller, H. Fischer, B. Wagner, F. Schuler, L. Polonchuk, E. M. Carreira, *Angew. Chem. Int. Ed.* **2006**, *45*, 7736; b) G. Wuitschik, M. Rogers-Evans, A. Buckl, M. Bernasconi, M. Märki, T. Godel, H. Fischer, B. Wagner, I. Parrilla, F. Schuler, J. Schneider, A. Alker, W. B. Schweizer, K. Müller, E. M. Carreira, *Angew. Chem. Int. Ed.* **2008**, *47*, 4512; c) G. Wuitschik, E. M. Carreira, B. Wagner, H. Fischer, I. Parrilla, F. Schuler, M. Rogers-Evans, K. Müller, *J. Med. Chem.* **2010**, *53*, 3227.
[31] J. A. Burkhard, G. Wuitschik, M. Rogers-Evans, K. Müller, E. M. Carreira, *Angew. Chem. Int. Ed.* **2010**, *49*, 9052.
[32] a) *http://www.activate-scientific.com*; b) *http://synthonix.com*; c) *http://www.spirochem.com*.
[33] a) J. A. Burkhard, B. Wagner, H. Fischer, F. Schuler, K. Mueller, E. M. Carreira, *Angew. Chem. Int. Ed.* **2010**, *49*, 3524; b) J. A. Burkhard, C. Guerot, H. Knust, M. Rogers-Evans, E. M. Carreira, *Org. Lett.* **2012**, *14*, 66.
[34] a) J. A. Burkhard, G. Wuitschik, J.-M. Plancher, M. Rogers-Evans, E. M. Carreira, *Org. Lett.* **2013**, *15*, 4312; b) D. B. Li, M. Rogers-Evans, E. M. Carreira, *Org. Lett.* **2013**, *15*, 4766.; c) E. M. Carreira, ETH Zurich personal communication.