Multi-component reactions and evolutionary chemistry

Lutz Weber

Multi-component reactions (MCRs) provide a new approach towards the efficient synthesis of diverse compounds and compound libraries. MCRs are evolving from being purely a chemistry curiosity to being recognized as having increasing relevance for drug discovery in terms of lead discovery and optimization. The combination of MCR-type chemistry-planning with evolutionary *in silico* and *in vitro* prediction of drug potential or biological properties is a new and powerful tool in drug discovery.

Lutz Weber
Morphochem
Gmunderstr. 37
D-81379 München
Germany
tel: +49 89 7800 50
fax: +49 89 7800 5555
e-mail: lutz.weber@
morphochem.de

▼ Multi-component reactions (MCRs) can be distinguished from classical, sequential two-component chemistry synthesis processes in that they use three or more chemical starting materials as the input for product formation. Up to seven starting components have been used, and MCRs have often been shown to produce higher product yields than classical chemistry [1]. Ideally, all reaction equilibria in the complex MCR mixture are reversible, and the last, product-forming reaction step is irreversible, thus providing the driving force to shift all intermediates and starting materials towards a single final product.

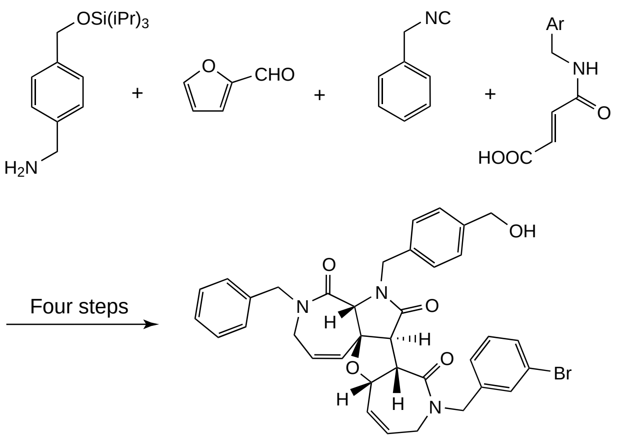
Examples of such an 'experimental' design can be found in Nature, in which many building blocks of life are believed to be the result of MCRs. For example, adenine may be formed by the addition of five molecules of isocyanic acid [2]. Although the history of MCRs dates back to the 19th century with reactions such as the Doebner or Strecker three-component reactions, the more complex nature of MCRs have prevented their wider application during the past century. The outcome of an MCR is often crucially dependent on the nature of the solvents, catalysts, concentrations and excess of reagents used, making the optimization of ideal reaction conditions more demanding when compared with sequential reaction schemes. However, the unparalleled atom efficiency as well as the often very mild reaction conditions also mean that MCRs are of interest for environmentally friendly, large-scale production of chemical entities. The recent demands in drug discovery for more diverse small molecules and their efficient synthesis has now shifted the attention of the chemist to this as-yet largely unexplored area of chemistry. In the past few years, a variety of novel MCRs has been discovered and applied to the synthesis of biologically active molecules [3,4].

Large compound libraries of diverse molecules

The low diversity of traditional combinatorial libraries in both biological and chemical terms has been recognized, and has led to the concept of 'diversity oriented synthesis', developed by Stuart Schreiber [5]. His method of choice for generating increased chemical diversity in fewer synthetic steps are MCRs. Using this technique, more complex, natural-product-like structures, leading to libraries of up to two million compounds, have been generated in fewer steps than was previously possible (Fig. 1) [6]. Similarly, many biotech and large pharma companies are now working to exploit MCR chemistry to generate chemical diversity. Organic chemists are increasing their efforts to obtain improved control of the outcome of MCRs by introducing novel catalysts and reaction conditions [7].

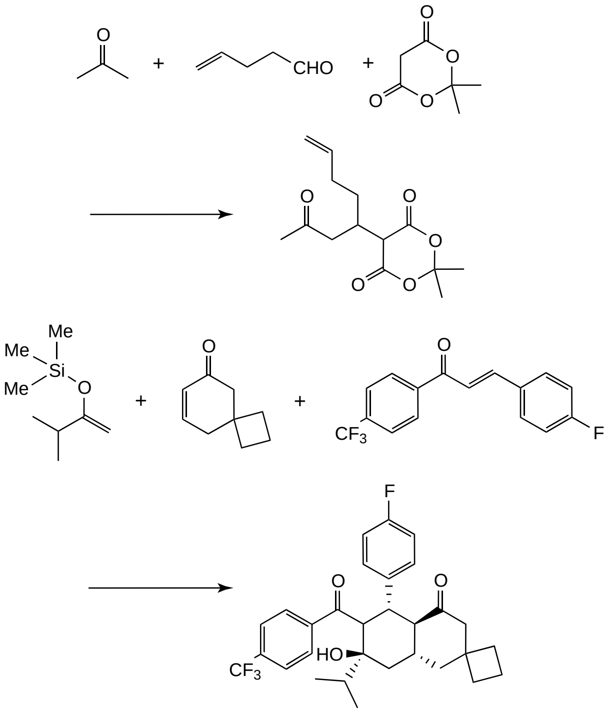
Natural products have proved to be a rich source of drug candidates in the past, as highlighted by comparing the ratio of known natural products with synthetic molecules. However, the ability to produce these compounds synthetically was problematic for hits found by the screening of either pure natural products or natural-product mixtures. Often, traditional classical synthesis schemes for

1359-6446/02/$ – see front matter ©2002 Elsevier Science Ltd. All rights reserved. PII: S1359-6446(01)02090-6



Drug Discovery Today

Figure 1. Diversity-generating synthesis scheme using a multi-component reaction and building 'natural product-like' molecules.



Drug Discovery Today

Figure 2. Novel multi-component reaction synthesis of carba-acetalization products and an intermediate for a cholesterin ester transfer protein inhibitor yielding products that do not contain nitrogen atoms.

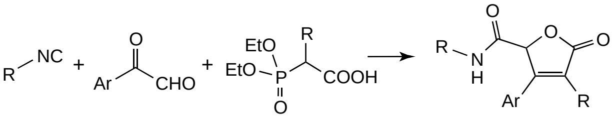
these molecules are too long or too expensive. MCRs can provide an alternative to this problem, either by using MCRs to synthesize the natural product, or by generating molecules that are 'artificial' but that bear a similar high density of the structural features that make natural products unique in the chemical universe of small molecules. The latter concept yields natural-product-like molecules.

A recent comparison of the structural characteristics of commerical drugs and natural products showed that the latter tend to have fewer H-bond donors and fewer nitrogen atoms, but more oxygen atoms per molecule [8]. The discovery of novel MCRs clearly follows this idea, exploring methods to build such non-nitrogen-containing molecules [9,10] (Fig. 2) to arrive at new chemical structure spaces.

At the same time, more classical drug-like, heterocyclic core structures can be created through MCR synthesis. For example, butenolides are found in several natural products as chemical core structures. Using a novel three-component MCR (Fig. 3), it is now possible to synthesize these molecules that are either structurally close or more distant to their natural progenies [11]. Recent research encompasses many examples of the application of known and novel MCRs in the synthesis of natural products [3].

In practical terms, MCRs also facilitate the generation of very large compound libraries [12]. For example, by using a particular four-component reaction and only 100 starting materials for each component, a library of 10^8 products can be generated, exceeding the size of all existing compound libraries. Several MCRs share the same starting materials, offering the opportunity to create different chemical scaffolds by reusing the same building blocks in different synthesis routes. This concept of varying the reactions rather than the starting materials enables simultaneous variation of both the chemical scaffold and its substituents in the same chemical library; something that has been impossible with previously used combinatorial chemistry methods. For example, it is possible to synthesize a variety of different chemical core structures using different combinations of just nine starting materials (Fig. 4).

This diversity-oriented approach is similar to building a multi-dimensional crossword in which a letter (i.e. chemical starting material) takes a different role in a different context (i.e. chemical reaction). For example, using just nine



Drug Discovery Today

Figure 3. The synthesis of natural-product-like butenolides was achieved using a three-component Ugi-type multi-component reaction and a subsequent cyclization step.

starting materials it should be possible to build 27 different chemical scaffolds using only three-component reactions. Each of the 3 × 3 × 3 starting materials in this 'chemical Rubik's Cube' would participate in at least three different MCRs. This chemical 3D crossword was proposed at the the 1st International Conference on Multi Component Reactions, Combinatorial and Related Chemistry (Munich, Germany, 4–6 October 2000; http://www.mcr2000.org).

In summary, MCRs are of particular interest for three reasons: efficiency, diversity and their large, unexplored chemical space. The number of steps needed for generating a particular chemical structure or chemical diversity is a measure of the efficiency of chemical synthetic. MCRs have been shown to enable the most efficient approach to a variety of chemical synthetic problems, especially for the synthesis of natural-product-like molecules. The chemical space that is available through MCRs is virtually unlimited. However, drug discovery scientists realized that it is neither practical nor possible to synthesize all interesting molecules and put them into a large compound repository. The high complexity of the MCR product space clearly requires novel methods to select and synthesize the most promising candidates for biological screening in a way that was neither possible nor required before. Ultimately, the combination of the new opportunities of MCR chemistries with the recent concept and developments of virtual screening of compound libraries requires such an integrated chemoinformatics approach.

In silico selection methods

Historically, productivity levels resulting from the generation of large compound libraries of small molecules for lead discovery and optimization by combinatorial chemistry have been low, and this has stimulated the development of *in silico* prediction methods for both biological activity and drug potential. These methods enable large libraries to be condensed into information-rich subsets that can then be submitted for synthesis and testing. Given the potential of MCRs to generate both larger and more diverse compound libraries, such methods are required to reduce the number of theoretical possibilities to a practical level for synthesis and testing.

Several predictive *in silico* filtering methods has been developed in recent years. Reference [13] gives an excellent overview of how these tools can be used to assess structural, physicochemical and pharmacological parameters, including high-throughput docking methods (if the 3D structure of the target is available), or quantitative structure–activity relationships (if the 3D structure is not available) [13]. These filter functions can then be applied for either: (1) existing and available compound libraries that are often in the range of one million compounds; or (2) virtual libraries of compounds that could be made in principle.

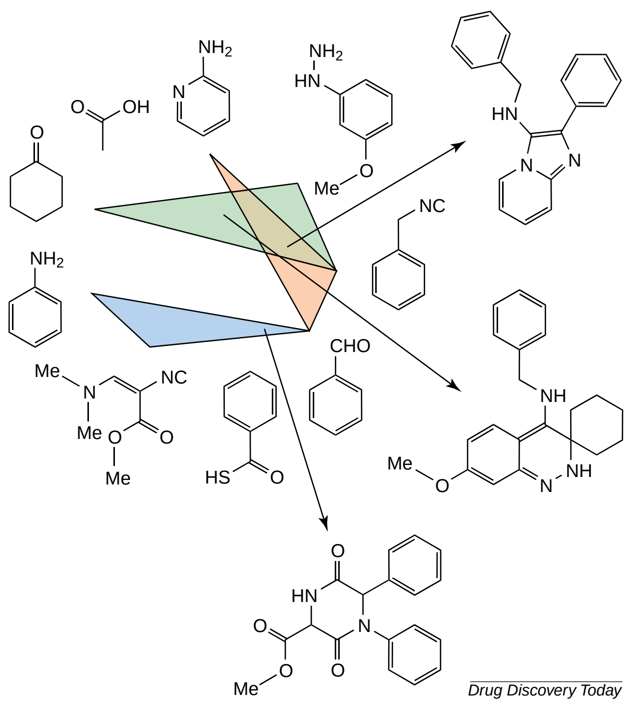


Figure 4. The principle of diversity oriented combinatorial reaction chemistry. By using different combinations of just nine starting materials a variety of chemical scaffolds can be produced, all through different known reaction types (only three shown).

For evaluation of the hitherto unexplored space of MCR-type products, the latter method is clearly of greater interest for drug discovery. Considering the vast numbers of small molecules that could possibly be made by MCR chemistry, we are dealing with so-called very-large libraries: libraries that are so large ($>10^{10}$) that it is not possible, in practical terms, to calculate their properties [14].

Evolutionary search methods

Finding a solution in a very-large search space is addressed by heuristic, evolutionary methods [15,16]. The application and development of such methods for drug design is an emerging discipline, as described recently [17].

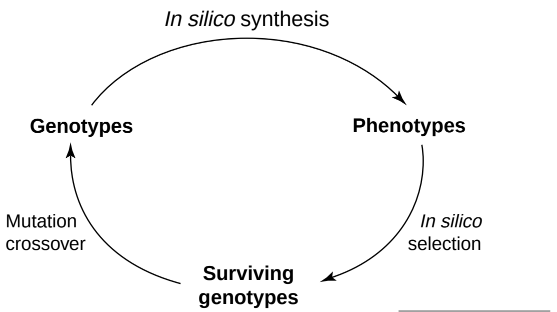
Evolutionary methods for screening virtual libraries contain the following basic steps:

(1) Generate virtual molecules by a 'synthesis algorithm'.
(2) Evaluate molecule fitness by a filter function.
(3) Select best molecules.
(4) Return to step (1), transfer of genomes (3).

A large variety of methods are possible by different implementations of this technique. In its simplest form, a large library of compounds is generated *in silico* and the respective products are evaluated for desired properties.

Only the best molecules are then subjected to synthesis and biological screening. Most drug discovery companies have set up such a discontinuous process for molecules made in their chemistry departments.

More sophisticated methods are those in which the virtual synthesis algorithm is influenced by the selection process. These are ‘learning evolutionary algorithms’ and depend heavily on the quality of the filter functions. In addition, the practical feasibility of the virtual synthesis algorithm is crucial. If the best virtual molecules cannot be synthesized by an established, facile chemistry route, then the likelihood of success becomes negligible.



Drug Discovery Today

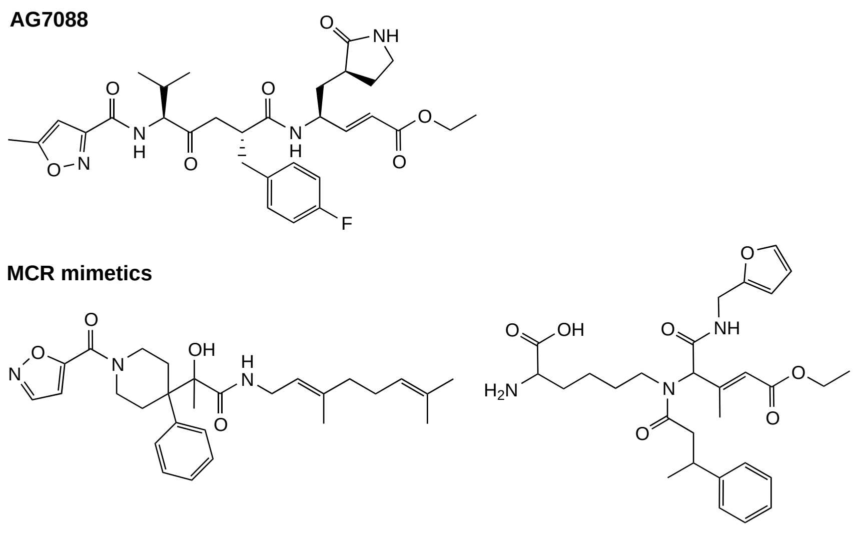
Figure 5. A generalized representation of an evolutionary system. New genotypes are generated according to the fitness of the phenotypes of the first generation of genotypes by mutation and crossover.

Chemistry and evolutionary selection

Several methods have been developed to integrate aspects of practical synthesis and *in silico* selection, and also aim to lower the barriers between computational and synthetic chemistry. These methods include NewLead [18], MCDLNG [19], MCSS [20], LUDI [21] and GROW [22].

The considerations outlined previously render the integration of any easy-to-implement chemistry (such as MCRs in virtual synthesis algorithms) with evolutionary *in silico* selection methods ideal. In one example, a facile reductive amination has been used to generate new products within an enzyme active site to select only 10 molecules for synthesis. All of the selected and synthesized molecules showed activity down to the nanomolar range [23]. The ability to generate diverse molecules with MCR-type chemistry, and to evaluate product properties for drug discovery, has recently been introduced by the new program MolMind [24]. MolMind combines chemical synthesis planning with filter functions for desired drug-like properties.

The principles of an artificial feedback-driven evolution of small molecules can be generalized in a simplified scheme that displays the basic elements for its implementation (Fig. 5). In general, new generations of genotypes with better fitness are evolved, based on the fitness of the phenotypes of previous generations in a feedback loop. After several rounds of evolution, the best surviving molecules can be submitted to synthesis and biological testing. In this context, a genotype represents the respective starting material, and the reactions and reaction conditions are used to generate a phenotype (the final product) with its 3D coordinates. Optimal parameters (mutation and crossover rate) for the evolutionary algorithm can be determined [25]. Using existing starting materials and a database of established MCRs and other basic reaction chemistry, a first generation of products is synthesized *in silico* and evaluated by a series of filter functions. The starting materials and reactions of the best products are then used as the parents of a new, mutated series of products. Ultimately, products that have been found to satisfy a series of criteria after several rounds of feedback cycles, including, for example, predicted bioavailability,



AG7088

MCR mimetics

Drug Discovery Today

Figure 6. MolMind (Morphochem; Munich, Germany) proposals for mimetics of AG7088 using a simple 2D similarity Tanimoto function and a polar surface area limit of 120 Å. The mimetics have a 2D similarity of 40–50% to the goal compound, and were found after evaluating just 18 reactions and 1570 starting materials. Further expansion of the reaction database and the use of all commercially available starting materials enables the generation of even more similar proposals.

biological activity or lack of toxicity, can be made quickly and then subjected to biological testing.

Heuristic evolutionary methods are stochastic and each run might produce a different result, providing a set of good solutions instead of the best solution. Early examples of evolutionary compound selection have selected new polymeric molecules that mimic either a given target polymer [26], or general molecules of all structural classes with a genetic algorithm [27]. MolMind can be used in drug discovery to mimic known biologically active molecules, such as screening hits or natural products. Figure 6 details MolMind-derived proposals for making analogues of Agourons AG7088 human rhinovirus protease inhibitor by using MCRs and standard reactions in just two steps, which were evolved after 200 generations. A simple 2D similarity Tanimoto function and a polar surface area of 120 Å were used as the selection criteria.

Concluding remarks

Combinatorial chemistry is able to generate a systematic chemical space and has found a complementary partner in novel computational methods such as evolutionary *in silico* selection algorithms. These novel methods promise to yield small molecules with drug-like properties in a similar way to biologicals such as phage-display libraries, or the generation of antibodies using the mechanisms of Nature for molecular evolution. The value of such molecules can be viewed almost instantly when combined with automated synthesis and screening. Whether this technique will yield drug candidates will become apparent in the next few years, and will be dependent on a closer integration of computation, synthesis, and screening in the drug discovery process.

References

1 Weber, L. *et al.* (1999) Discovery of new multi component reactions with combinatorial methods. *Synlett.* 3, 366–374

2 Ugi, I. *et al.* (1994) Multicomponent reactions in organic chemistry. *Endeavour* 18, 430–435

3 Tietze, L.F. (1996) Domino reactions in organic synthesis. *Chem. Rev.* 96, 115–136

4 Weber, L. (2000) High-diversity combinatorial libraries. *Curr. Opin. Chem. Biol.* 4, 295–302

5 Schreiber, S.L. (2000) Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 287, 1964–1969

6 Lee, D. *et al.* (2000) Pairwise use of complexity-generating reactions in diversity-oriented organic synthesis. *Org. Lett.* 2, 709–712

7 Grigg, R. (1999) Palladium catalysed cascade cyclisation-anion capture, relay switches and molecular queues. *J. Organomet. Chem.* 576, 65–87

8 Lee, M-L. and Schneider, G. (2001) Scaffold architecture and pharmacophoric properties of natural products and trade drugs: application in the design of natural product-based combinatorial libraries. *J. Comb. Chem.* 3, 284–289

9 Paulsen, H. *et al.* (1999) Cyclization as the key step in the synthesis of pentasubstituted arenes: an efficient access to highly active inhibitors of cholesteryl ester transfer protein (CETP). *Angew. Chem., Int. Ed. Engl.* 38, 3373–3375

10 List, B. *et al.* (2001) Proline-catalyzed asymmetric aldol reactions between ketones and unsubstituted aldehydes. *Org. Lett.* 3, 573–575

11 Beck, B. *et al.* A novel three component butenolide synthesis. *Org. Lett.* (in press)

12 Tan, D.S. *et al.* (1998) Stereoselective synthesis of over two million compounds having structural features both reminiscent of natural products and compatible with miniaturized cell-based assays. *J. Am. Chem. Soc.* 120, 8565–8566

13 Böhm, H-J. and Schneider, G. eds (2000) *Virtual Screening for Bioactive Molecules*, Wiley-VCH

14 Weber, L. and Almstetter, M. (1999) Diversity in very large libraries. In *Molecular Diversity in Drug Design* (Dean, P.M. and Lewis, R.A., eds), pp. 93–114, Kluwer

15 Rechenberg, I. (1973) *Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution.* Frommann-Holzboog

16 Goldberg, D.E. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley

17 Weber, L. (2000) Molecular diversity anaylysis and combinatorial library design. In *Evolutionary Algorithms in Molecular Design* (Clark, D., ed.) pp. 137–158, Wiley-VCH

18 Tschinke, V. and Cohen, N.C. (1993) The NEWLEAD program: a new method for the design of candidate structures from pharmacophoric hypotheses. *J. Med. Chem.* 36, 3863–3870

19 Gehlhaar, D.K. *et al.* (1995) *De novo* design of enzyme inhibitors by Monte Carlo ligand generation. *J. Med. Chem.* 38, 466–472

20 Caflish, A. *et al.* (1993) Multiple copy simultaneous search and constructions of ligands in binding sites: application to inhibitors of HIV-1 aspartic protease. *J. Med. Chem.* 36, 2142–2167

21 Böhm, H-J. (1992) The computer program LUDI: a new method for the *de novo* design of enzyme inhibitors. *J. Comput.-Aided Mol. Design* 6, 61–78

22 Moon, J.B. and Howe, W.J. (1991) Computer design of bioactive molecules: a method for receptor-based *de novo* ligand design. *Proteins Struct. Funct. Genet.* 11, 314–328

23 Böhm, H-J. *et al.* (1999) Combinatorial docking and combinatorial chemistry: design of potent non-peptide thrombin inhibitors. *J. Comput.-Aided Mol. Design* 13, 51–56

24 Almstetter, M. *et al.* (2001) *Molmind – An Evolutionary Drug Discovery Game.* Daylight User Meeting, MUG 2001, 6–9 March 2001, Santa Fe, NM, USA (www.daylight.com)

25 Illgen, K. *et al.* (2000) Simulated molecular evolution in a full combinatorial library. *Chem. Biol.* 7, 433–441

26 Glen, R.C. and Payne, A.W.R. (1995) A genetic algorithm for the automated generation of molecules within constraints. *J. Comput.-Aided. Mol. Design* 9, 181–202

27 Venkatasubramanian, V. *et al.* (1995) Evolutionary design of molecules with desired properties using the genetic algorithm. *J. Chem. Inf. Comput. Sci.* 35, 188–195