Discovery of New Multi Component Reactions with Combinatorial Methods

Lutz Weber*, Katrin Illgen, Michael Almstetter
*Morphochem AG, Am Klopferspitz 19, D-82152 Martinsried, Germany
Received 8 February 1999

Abstract: The discovery of novel chemical reactions or reaction sequences that are able to generate useful chemical products may be regarded as the heart of organic chemistry. We present here concepts and methods on how to find and explore new multi component reactions, especially with automated combinatorial methods. This "combinatorial reaction finding" provides also a powerful tool to the understanding of the rules of organic chemistry, especially structure-reactivity relationships.

Keywords: multi component reaction, combinatorial chemistry, synthetic strategy, reaction design

1 Introduction

Organic chemistry is the science of the rules of how chemical entities react with each other to form new molecules. The importance of a specific reaction may be judged on its ability to deliver interesting products with high yield, chemo-, regio-, stereo- or enantioselectivity. Further requirements may include generality over a wide range of starting materials as well as productivity aiming at more diverse and complex products using a rather low number and complexity of starting materials. An "ideal" organic reaction would fulfill at least one of these criteria.\(^7\)

Most organic reactions in the textbooks of organic chemistry are reactions of either single or two starting materials. Reactions that use more than two different starting materials are called multi component reactions (MCRs). In the light of chemical productivity and generation of molecular diversity an "ideal" MCR should not only comprise more than two starting materials but also these starting materials would be different and all or most of the atoms of those starting materials would be incorporated into the final product.

Thus, with respect to their productivity, yield, convergence and facile execution such MCRs occupy an outstanding position among all other reactions - making them especially interesting for the concept of combinatorial chemistry.\(^3\) Therefore, the discovery and development of novel MCRs is receiving a growing interest from industrial chemistry research groups. However, we believe that the goal to discover and design novel and useful MCRs unravels also a new challenge for organic chemists and the basic understanding of organic chemistry itself.

In this review we would like to present a conceptual overview on methods that are aiming at the discovery and design of novel MCRs. This overview is neither comprehensive nor exhaustive - it is rather intended to ignite the interest of organic chemists for the concept of multi component reactions and to introduce the emerging field of "rational" or systematic reaction finding.

2 The Logic of Multi Component Reactions

2.1 Defining MCRs

For our goal to discover new MCRs it is necessary to understand the specific characteristics and logic of these reactions. Ugi, the most productive protagonist and inventor of MCRs, distinguishes 3 idealised types of MCRs considering the reversibility of reactions leading to intermediary products \(P^1\), \(P^2\) and the final product \(P^N\).

<table>
<thead>
<tr>
<th>Type</th>
<th>Reaction Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>(A + B = P^1 + C = P^2 + D, \ldots = P^N)</td>
</tr>
<tr>
<td>II</td>
<td>(A + B = P^1 + C = P^2 + D \Rightarrow O \rightarrow P^N)</td>
</tr>
<tr>
<td>III</td>
<td>(A + B \rightarrow P^1 + C \rightarrow P^2 + D \rightarrow \ldots \Rightarrow O \rightarrow P^N)</td>
</tr>
</tbody>
</table>

While reaction yields of type I MCRs depend on the position of the reaction equilibria, types II and III may give yields of up to 100% and are thus especially interesting. However, as we will show later, type I MCRs can be converted virtually into type II if the product can be withdrawn from the equilibrium.

True multi component reactions should be distinguished from the so-called tandem, cascade, domino or zipper reactions\(^1\)\(^6\) where one starting material bears with several functionalities that react in several consecutive steps, e.g.: \(A + B \Rightarrow P^1 \rightarrow P^2 \rightarrow \ldots \rightarrow P^N\). These one-pot, multi step synthesis procedures rely on a high conversion of \(A + B \Rightarrow P^1\) that will allow to add a third reactant \(C\), eventually under different reaction conditions whereas "pure" MCRs are run under the same conditions by adding the starting materials at the same time. However, MCRs may well include reaction mechanisms that could be described with the terms domino, tandem or zipper if e.g. some of the starting materials contain multiple functional groups that are involved in the reaction sequence.

Ugi has also developed the concept of unifying multi component reactions in order to obtain new MCRs.\(^7\) Thus, the Ugi four component reaction may be viewed as the combination of a type I Mannich three component reaction (3CR) using amine \(A\), aldehyde or ketone \(B\) and an acid \(C\) with the Passerini type II 3CR that uses aldehyde or ketone \(B\), carboxylic acid \(C\) and isonitrile \(D\).
Since the Passerini reaction is of type II, one has to expect Passerini type products $P_4$ as side products (e.g. apparent when using dichloromethane as solvent) of the Ugi 4CR while the Mannich product $P_2$ will be withdrawn through the equilibrium. Unions of MCRs may be viewed as logical nets with – ideally – one or more outputs. Scheme 1 gives a simplified view on such a Ugi 4CR net based on the Mannich and Passerini reactions and neglecting other possible reactions.

### 2.2 The Rational Design of New MCRs

Using the above logic and otherwise known reactions one can start to design novel MCRs. Thus, ideally in a type II reaction sequence it would be favourable if starting material $C$ would react only with $P_1$ but not with $A$ or $B$ or, alternatively, a reaction $C$ with $A$ or $B$ should be reversible. A nice example for this rational design strategy is the 3CR of aromatic amines with aldehydes and the subsequent aza-Diels-Alder cycloaddition of the resulting azomethine with electron rich dienophiles (Scheme 2) giving tetrahydroquinolines under mild conditions.

Both parts of this 3CR, the formation of azomethines from aldehydes and amines and the hetero cycloaddition of azomethines and dienophiles, were known as individual reactions beforehand. The idea to carry out these two reactions in one step as a 3CR however required the recognition that the dienophile does not react either with the amine nor the aldehyde but only with their azomethine product under the given reaction conditions.

### Biographical Sketches

Katrin Illgen was born in Zwickau, Germany in 1970. After studying chemistry at the University of Leipzig she obtained her Dr. rer. nat. under the supervision of Prof. B. Schulze in 1996. She moved then to Hoffmann-La Roche AG, Basel as a post-doctoral fellow until 1998 and joined then Morphochem AG in Munich in 1998.

Michael Almstetter was born in 1967 in Ingolstadt, Germany. He studied chemistry at the Technical University of Munich and is now completing his doctoral thesis work under the supervision of Prof. I. Ugi.

Lutz Weber was born in Halle/Saale, Germany in 1958. He studied chemistry at the Technical University of Budapest, Hungary. After completing his Dr. rer. nat. under supervision of Prof. Muhlstadt at the University of Leipzig in 1986 he worked on artificial enzyme modelsystems that led to Habilitation in 1992. From 1991 until 1998 he joined the Preclinical Research Division of Hoffmann-La Roche AG, Basel Switzerland, lately as Vice Director, responsible for pharmaceutical projects and combinatorial chemistry.
Following these ideas, the currently known MCR with the most starting materials, the seven component reaction published by Dömling and Ugi, was constructed as a combination of a Asinger 4CR and a Ugi 5CR (Scheme 3).

While the Asinger reaction yields cyclic azomethines, the Ugi reaction is known to proceed especially well when using pre-condensed azomethines with isonitriles and carboxylic acids. Considering the synthetic breadth of this 7CR only 3 components are highly variable. However, the field of isonitrile based MCRs has been and is still the source of many elegant novel synthetic routes to all kinds of molecules (see reviews given by Ugi and Torroba) and is based on the unique properties of the exothermic C\textsuperscript{II} to C\textsuperscript{IV} conversion of isonitriles.

In an intriguing example of a “designer” MCR reaction Kelly elaborated a synthesis of highly substituted pyridines. Merging the concepts of the Knoevenagel route to pyridines and enolate technologies, starting materials for a highly complex natural product, rubrolone, were prepared (Scheme 4) in a one-pot reaction.

Using the existing knowledge about reactions with lower dimensions one can construct MCRs of higher dimensions. In general terms, the construction principle for “ideal” MCRs can be described therefore as follows:

Find a reaction that yields a product that may react with an other starting material that does not react irreversibly with the starting materials of the previous reaction.

By repeating this construction sequence one may obtain in principle reactions with almost any number of starting materials. It is certainly possible and represents a unique challenge to the creativity of organic chemists to design novel MCRs by using their trained knowledge about known reactions. We would like to describe here methods that may help chemists to meet this challenge of playing “chess” with organic reactions.

3 Computational Methods

The above outlined MCR construction guidelines can be in principle automated by using suitable computational reaction database searching techniques. The first step in such an implementation is to derive from the pool of available reaction data “reaction prototypes” that describe both the reacting centres (atoms) as well as those neighbour atoms that have a marked influence on the reactivity of the reacting centres. E.g. an azomethine formation using a primary amine and an aldehyde can be described as follows in the Daylight SMIRKS-notation:

\[
[C.c,#1:60][C:1](=\text{[N:3][C.c,#1:62]})(\text{[C.c,#1:61]}).[O:2](\text{[\#1:11]})\text{[\#1:12]}]\rightarrow[C.c[0:2][](\text{[\#1:11]})[\#1:12]]\text{[C.c,#1:60][C:1][N:3][C.c,#1:62]}(\text{[C.c ,#1:61]}).[O:2](\text{[\#1:11]})[\#1:12]
\]

In the second step, the right part (the reaction product \([C.c[0:2][](\text{[\#1:11]})[\#1:12]]\text{[C.c,#1:60][C:1][N:3][C.c,#1:62]}(\text{[C.c ,#1:61]}).[O:2](\text{[\#1:11]})[\#1:12]\)) of this first equation may be used as a starting material in a reaction substructure search over a reaction database. An eventually found reaction should comprise an additional starting material that is then checked whether it undergoes also an irreversible reaction with one of the first starting materials. If not, we may have found a candidate for a new MCR. This method may be viewed as a combinatorial search to find MCRs through our knowledge of chemical reactions – a task that is best accomplished by a computer. For the described strategy and using a program based on the Daylight reaction toolkit program suite such a newly proposed reaction is given in Scheme 5. This recently reported hetero Diels-Alder reaction is catalysed by Lewis acids and may work also when performed as a 3CR.

Similarly, the reaction of alkenyl boronic acids with azomethines can be found. The corresponding 3CR was indeed recently used by Petasis for an enantioselective synthesis of \(\alpha\)-amino acids starting from amines, \(\alpha\)-keto acids and alkenyl boronic acids.

In a different implementation, one may enter a set of starting materials into the program that has a list of known...
"prototypic" reactions as a knowledge base. The program than constructs all possible products by applying the prototypes in iterative cycles. Sometimes the results are surprising or unexpected as shown in Scheme 6 for a novel synthesis of a piperazinone. The validity of such proposed reactions however has to be verified by the experiment.

The outcome and use of this approach depends on how many and how exact the underlying reaction prototypes are. The conversion of the abundant knowledge on reaction data into such more abstract prototype reactions requires immense work and can only poorly be automated.

4 Serendipity

Most of the known older multi component reactions have been found by serendipity rather than by rational planning. It is interesting to note that many of them are constructed from only a few "prototypic" reactions that include amines and carbonyl functions of different reactivities like aldehydes, ketones, carboxylic acids, esters and amides – the main repertoire of chemists of those days.

With the emerging automation of synthesis, purification and analysis in the area of combinatorial chemistry we are entering into a new phase of our synthetic capabilities. New techniques like coupling liquid chromatography, mass spectroscopy and nuclear magnetic resonance (LC-MS-NMR) allow to perform the fast and exhaustive investigation of rather crude reaction products that may lead to new insights into "old" reaction mechanisms.

We have recently investigated the reaction of aromatic amines, aromatic aldehydes and isonitriles with LC-MS-NMR. The reaction of 4-nitro-benzaldehyde, 4-amino-benzonitrile and benzylisonitrile in methanol gives a rather complex mixture of products as revealed by LC (Figure 1) with an acetonitrile/D_2O gradient with 0.04% trifluoroacetic acid as the eluent.

Through the parallel use of NMR and MS, 6 out of those products could be identified in the reaction mixture (Scheme 7) in a rather facile way. A typical stopped-flow NMR is shown for the unstable intermediate 1 in Figure 2.

The expected main product in this reaction is 4, molecules 1 and 2 are known intermediates whereas 5 represents a side product that was already observed by McFarland when using secondary amines in the Ugi 4CR.

Molecules 3 and 6 however have not been observed so far, but can be explained from the postulated imino-ether intermediate of a Ugi 4CR (Scheme 8). Most probably, the imino-ether is cleaved by nucleophilic attack of traces of water to yield 3 (Scheme 8) resulting in a novel "Strecker"-type reaction sequence where the isonitrile provides only one carbon atom.

Obviously, by using other nucleophiles than water or amines one can try to extend the scope of this novel 3CR reaction. Also, the reaction may provide an alternative to the use of the cleavable cyclohexenylisonitrile that has been introduced by Armstrong. Attempts are currently under way in our laboratory to optimise both the yield and applicability of this reaction.

The parallel synthesis of large libraries of pure single compounds has been established in many pharmaceutical and agrochemical companies over the last years. Many otherwise "known" reactions have been investigated and used with a so far unmet wide range of starting materials, providing both insights into the breadth of these reactions as well as tales of the unexpected. Thus, by using aminopyridines as the amine component in the Ugi 4CR the facile formation of imidazo-pyridines was observed by chance (Scheme 9):
Acetic acid, the fourth component of the Ugi 4CR, is required as a catalyst. This novel reaction was discovered and published almost at the same time by three industrial research groups!

The in-depth evaluation of this new finding resulted in the extension of the 4+1 insertion reaction of isonitriles for a broad variety of aldehydes and five- and six-membered amines that contain an imino-amine substructure.\(^{21-25}\)

### 5 Automated Combinatorial Reaction Finding

The availability of high performance and fast analytical tools together with novel computational techniques that analyse and abstract the large volume of raw data has an impact on the way we deal with reaction data. Thus, an alternative way to find new MCR reactions was
introduced by us through the concept of combinatorial reaction finding. Ten different starting materials were selected for this experiment as shown in Scheme 10. All possible multi component reactions were carried out by the combinatorial variation of two up to ten starting materials (2CR to 10CR) in parallel in methanol at room temperature and using a robotic dispensing system.

Overall there are \(2^n - n = 1013\) possibilities for two-, three-, four- up to the one ten-component mixture. With the aid of automated liquid chromatography and data evaluation one can now search for products that are unique to a specific mixture by comparing the retention times with the starting materials and over all other mixtures that contain the sub-combinations.

Using a minimal peak height requirement of 30% (compared to the sum of all peaks in the respective LC chromatogram) for this novel and unknown reaction product, a new MCR was found. In the reaction of cyclohexanone, benzylisocyanide, 4-methoxy-phenylhydrazine and catalysed by acetic acid a dihydro-cinnoline was formed (Scheme 11).

This reaction appears to be similar to the above imidazopyridine formation (Scheme 9) probably via a 5+1 insertion reaction of the isonitrile into the corresponding hydrazone (Scheme 12). This reaction mechanism is likely since only electron rich aromatic hydrazines yield this product. The Ugi 4CR reaction with phenylhydrazine has been reported to give the expected Ugi-type 4CR product.

The same procedure and principles may also be used with conventional combinatorial libraries made by parallel synthesis and their analysis by mass spectroscopy.

Contrary to the above example, in combinatorial libraries the same reaction is repeatedly applied in an exhaustive manner with starting materials of the same chemical type but different substitution patterns. Constructing all expected products and their molecular weights in the computer and the subsequent analysis of all combinatorial products by mass spectroscopy gives us a complete picture on the substituent dependency of the respective reaction.

We have used the Daylight reaction toolkit and a list of prototypic reactions (see Section 3) to construct "structure-reactivity" relationships.

In a first example, the well-known Doebner-type synthesis of quinoline from anilines, various aldehydes and pyruvic acid derivatives were used, a three component reaction carried out in ethanol at room temperature. Surprisingly, the reaction yielded not only the expected substituted quinoline-4-carboxylic acids 7 but also 3-arylamino-dihydro-pyrrol-2-ones 8 (Scheme 13) - 4 component product.

After inspection of the literature one can find a few examples of the reaction of special anilines, aldehydes and 2-keto carboxylic acids that yield 3-arylamino-dihydro-pyrrol-2-ones, However, this picture is still far from being complete. A full combinatorial library comprising anilines with electron withdrawing and donating groups, aromatic and aliphatic aldehydes as well as various pyruvic acid derivatives gave further products with different structures! Most of these products precipitated from the reac-
tion mixtures with high purity – an example of how MCR reactions may have different “outputs” due to the withdrawal of products from the reaction equilibrium network by a physical effect like insolvibility.

A small matrix of 4^4=4 starting materials (anilines, aldehydes and pyruvic acid derivatives) illustrates this behavior – examining the four reaction products on diagonal throughout this 3-dimensional combinatorial reaction cube gave rise to four different reaction products (Scheme 14).

Scheme 14

The formation of these products depends on defined property combinations of the respective starting materials. Electron donating substituents R¹ favor the formation of 7, while electron withdrawing substituents R¹ and R² the formation of 9, electron withdrawing R¹ and electron donating R² 8 and so forth (Scheme 13 and 14).

In another example, we synthesized a library of 16'896 products in 384 well plates using the Ugi 3CR shown in Scheme 7. To our current knowledge, about 12 different reaction types may occur when using various structurally different amines, aldehydes and isonitriles under the used reaction conditions (methanol as solvent at room temperature). High throughput automated mass spectroscopy of the crude reaction mixtures together with our Daylight based evaluation of the corresponding mass data was used to establish large structure-reactivity relationships for all 16'896 reaction products. For example, Figure 3 shows such a MS analysis of a representative 384-well plate. The association of three product types out of the known 12 is given by panels A, B and C whereby 4-amino-benzamidine dihydrochloride as the amine component throughout the whole plate, 16 different isonitriles and 24 aldehydes are distributed in the horizontal and vertical direction, respectively. Green points in Panel A correspond to wells with a found and desired product of type 4 as revealed by its high intensity mass peak in the ion spray spectrum. Panel B corresponds to the azomethine type 2, whereas panel C shows the occurrence the aminal formed from the respective aldehyde and 2 equivalents of the 4-amino-benzamidine.

As the side products 2 and the aminal occur preferentially in the columns of the 384 well plate, their formation depends mostly on the nature of the respective isonitriles. Evaluating the mass spectra of all reaction products against all known reactions and starting materials also opens the opportunity to search for remaining and systematically occurring molecular masses that can not be associated with an expected product. Such products may have been formed by a novel, yet unknown MCR reaction.

6 Optimisation of Multi Component Reactions

Multi component reactions are networks of various reactions with individual mechanisms that for the most part also require different reaction conditions. However, it is not very likely that the actual experimental MCR conditions will suit these reaction mechanisms. Thus, finding the right reaction conditions for a novel MCR as e.g. solvent, concentration, reaction time is likely to be more difficult than for conventional reactions and probably represents a compromise. Recently we have introduced the application of genetic algorithms to solve complex multi dimensional problems in MCR chemistry.31

Based on automated parallel synthesis that assures exactness and consciousness of the preparative execution and high-throughput LC we have now performed a genetic algorithm (GA) driven optimization of MCR reaction conditions (Figure 4).32

The automated optimization cycle starts with the examination of the crude reaction mixture via LC-MS (or LC-MS-NMR) whereby the target peak of the desired MCR-product is specified. The first generation of experiments with different reaction conditions is created randomly based on a given set of the possible parameter space. By this means a population of a number of N experiments is generated containing the same starting components but different reaction conditions. A control and set-up unit adjusts the amount of reagents assuring the comparability of the LC-spectra and coordinates the automated execution of the experiments. After a fixed reaction time the resulting mixtures are analyzed. The obtained results - i.e. the yield of the desired reaction product - are evaluated and the GA sorts the experimental parameter sets, the parent genomes, for their fitness. After mating and mutating the best parent sets a new set of experiments, the children population, is generated. This iterative feedback cycle is stopped when a set of reaction parameters is found under which the desired MCR-product is formed with sufficient selectivity and yield.

The method was tested with the well-known Ugi 4CR. The starting materials were made available in eight different solvents, depending on their solubility. Only considering solvent mixtures, 8^2 = 4096 combinations are possible.
The time for the addition of a third and a fourth component was varied from 0 to eight hours, which allows the investigation of whether the respective MCR is a "real" MCR or a domino-type reaction. Combining, thus, all possible ways of performing the synthesis, we got over $1,572,864$ combinations. The experiment was then performed with 16 parallel experiments ($N = 16$) in each cycle as shown in Figure 5. Already after 13 feedback cycles of synthesis and screening the yields of the desired 4CR product was increased from 13% to nearly 78%.

The best best parameter set of reaction conditions found by the GA for the Ugi 4CR were started with the amine and aldehyde diluted in CH$_2$Cl$_2$ and adding the isonitrile and the acid both in methanol after two hours. Thus, only $13 \times 16 = 208$ experiments out of the pool of 1,572,864 were needed to select a good yielding and selective method to perform the respective MCR.

Figure 3 MS analysis of a Ugi 3CR combinatorial library in a 384 well plate. Panel A corresponds to found products of type 4, B to 2 and C to an aminal.

Figure 4 Optimisation of reaction parameters with a genetic algorithm, based on the desired product yield.

Figure 5 The evolution of product yield during the course of the genetic algorithm based optimisation.
References

(1) This article is based on a lecture held at the XVth EFMC International Symposium, “Discovery of New Multi Component Reactions by Combinatorial Approaches”, Edinburgh (U.K), 6-10 September 1998.
(3) Bienaymé, H.; Hulme, C.; Oddon, G. Chemistry, in press.