Summary

Systems chemistry has flourished over the past decade and yielded promising results. Systems chemistry is predominantly about exploring the connections between the individual components in complex systems and the emergent properties that result from the interaction of these components. The emergence of systems chemistry has also opened new research directions into of the origin of life and the *de novo* synthesis of life. A better understanding of the principles of assembly and function in complex systems may aid in revealing the origin of biological complexity. The spontaneous emergence of complexity such as self-replication and the formation of ordered structures, undoubtedly play an important role in the origin of life. On the early Earth, self-replication not only acted as a mechanism for the self-amplification of certain molecules in the chemical mixture, but also a means of transmitting information to the descendants of these molecules. In addition, the emergence of self-synthesizing complex folded molecules may shed new light on the mystery of the origin of life suggesting a potential role for primitive proteins. This thesis describes new methods through which replicators and foldamers can be accessed.

In Chapter 1, we briefly introduced the basic concepts of systems chemistry and the origin of life. In order to mimic some of life’s important features and to achieve life-like behavior, it is particularly important to construct complex systems that combine the processes of self-replication and folding. While self-replication can provide a means of information transfer, complex folded molecules can often bring functions. We highlighted the current state of the art of synthetic folded systems and self-replicating systems. Most currently developed synthetic folding systems rely on multi-step synthesis and feature a limited number of folded backbones. The majority of synthetic self-replicating systems cannot achieve exponential growth. We then introduced dynamic combinatorial chemistry and its application to the synthesis of self-replicating molecules and folded molecules. The dynamic combinatorial approach reduced the amount of synthetic work, side-stepping challenging organic synthesis. Our group discovered the emergence of self-replicating molecules from DCLs. We have found that peptide-based self-replicating molecules can emerge from a DCL of short peptides functionalized by an aromatic dithiol. These replicators exhibit exponential growth. Keeping the aromatic dithiol core of the building block unchanged, we have synthesized a series of new building blocks (Table S1). These building blocks not only allow us to access self-synthesizing complex folded structures, but also provide the possibility to link self-replicating and folded molecules in a single system.
Table S1 Chemical structures of building blocks 1, 2, 3 and 4

In Chapter 2, we described how to use DCC as a strategy to identify folded molecules with unprecedented complex structures. We have found that building block 3, consisting of aspartic acid and nucleobase residues, can selectively assemble into a remarkably complex foldamer. The spontaneously emerging folded molecule consists of 15 identical units of building blocks 3 and forms in almost quantitative yield in aqueous solution. Single crystal X-ray data and NMR spectroscopy studies have shown that the formation of the folded molecule is mainly driven by non-covalent π–π stacking, hydrophobic interactions and hydrogen bonding. Non-covalent interactions involve residues that are widely separated in the extended structure, testing to the presence of secondary and tertiary structure motifs. These results establish DCC as an effective methodology for synthesizing complex foldamers outside the realm of biology.

In order to further expand and promote this methodology, we have modified the structure of the building blocks. In Chapter 3, we synthesized a series of novel dipeptide building blocks 2, composed of substituted phenylalanine and lysine residues. By introducing different groups on the para position of the phenyl ring of the phenylalanine amino-acid residue, DCLs selectively yielded foldamers consisted of 9, 12, 13, 16, and 23 identical units. Single crystal structures show that these complex folded molecules adopt tertiary structures and that their formation is driven by intramolecular non-covalent interactions, including hydrophobic interactions, π–π stacking and hydrogen bonding. Since the numbers of building blocks constituting these folded molecules are different, they are folded in different ways. In addition, simple modification of the building blocks can have a huge impact on the
entire folded skeleton, which makes it difficult to predict the emergence of folded molecules from
DCLs. We believe that in the near future, more complex folded structures will be discovered, and that
eventually predictions of the structure may also become possible. Although the folded molecules we
currently synthesized still lack functions, or we have not yet discovered their functions, the
emergence of self-synthesized functional foldamers would be an important piece of the puzzle of the
emergence of early life.

Chapter 4 described the emergence of self-replicating molecules that contain both amino acids and
nucleobases from DCLs. In order to achieve the combination of peptides and nucleobases, two of the
most important components of life, we have designed two strategies: the first is to construct a mixed
DCL consisting of peptide building block 1a and nucleobase building block 3; the other is to build a
DCL consisting of PNA functionalized building block 4. The results show that the emergence of
chimeric nucleobase-peptide self-replicating molecules in the mixed DCLs is determined by the ratio
of the two building blocks. Only at a specific ratio, self-replicating molecule consisting of two units of
building blocks 1a and one unit of building block 3 emerges. The emergence of self-replicating
molecules from DCLs made from PNA-functionalized building block is strongly affected by the
structure of the amino acid. In summary, the emergence of nucleobase-peptide self-replicating
molecules from DCLs depends on the formation of ordered supramolecular structures, rather than
base pairing as observed for previously nucleic-acid based self-replicating systems.

In Chapter 5 we described for the first time the self-sorting between self-replicating and folded
molecules in a DCL composed of building blocks 1b and 3a. Furthermore, we developed a transient
self-replicating system using the same building blocks. In the DCL consisting of equimolar amounts of
building blocks 1b and 3a, the emergence of a replicator composed of both of these building blocks
was observed during the early stages of the experiment. After all of building block 1b is consumed,
the remaining building block 3a assembles into a foldamer. Thus, to some extent, the emergence of a
self-replicator in these DCLs drives the formation of folded assemblies. Self-sorting between self-
replicating molecules and complex folded molecules is driven by a balance of non-covalent
interactions such as intra- and inter-molecular π-π stacking and hydrogen bonding. When reducing
the concentration of building block 3a in these DCLs, we unexpectedly observed transient self-
replication. A metastable self-replicator emerged and was amplified during an initial period. Over
time, it gradually degraded and transformed into a thermodynamically more stable self-replicator.
The discovery of such transient self-replication systems paved the way to study self-replication far
from equilibrium.