Self-Sorting in Dynamic Combinatorial Libraries Leads to the Co-Existence of Foldamers and Self-Replicators

Bin Liu, Meagan A. Beatty, Charalampos G. Pappas, Kai Liu, Jim Ottelé, and Sijbren Otto*

Abstract: Nature segregates fundamental tasks such as information storage/transfer and catalysis between two different compound classes (e.g., polynucleotides for replication and folded polymides for catalysis). This division of labor is likely a product of evolution, raising the question of how simpler systems in which replicators and folded macromolecules co-exist may emerge in the transition from chemistry to biology. In synthetic systems, achieving co-existence of replicators and foldamers in a single molecular network remains an unsolved problem. Previous work on dynamic molecular networks has given rise to either self-replicating fibers or well-defined foldamer structures (or completely un-sorted complex systems). We report a system in which two cross-reactive dithiol foldamer structures (or completely un-sorted complex systems) interact, forming a polynucleotide that stores and transmits genetic information. Nucleic acids are orthogonal to proteins, bonds to form a polynucleotide that stores and transmits genetic information. Nucleic acids are orthogonal to proteins, which have a folded polyamide backbone, bind biomolecules and catalyze reactions. The building blocks of these two classes of biomacromolecules (nucleotides and amino acids) mostly do not cross-react (except in t-RNA-amino acid conjugates). This raises the question of how this division of labor over two compound classes can come about during the emergence and evolution of life and whether it is necessary to have two different connecting chemistries, or whether simpler systems can form based on a single connecting chemistry in which replicators and folded macromolecules co-exist. The “RNA world hypothesis” attempts to address these questions by assuming that RNA can both store/transmit information and catalyze reactions. These functions are accomplished by adopting significantly different conformations. For example, RNA-based enzymes or ribozymes exhibit three-dimensional folded structures, just like modern proteins require folding for function, yet in order to be replicated RNA needs to adopt an unfolded structure. These different conformational requirements are difficult to reconcile and limits the ability of RNA to self-replicate and have ribozymal activity at the same time.

We reasoned that it should be possible to achieve the parallel emergence and co-existence of self-replicators and unfolded structures through a self-sorting process between compound classes in a single molecular network using a single connecting chemistry. Self-sorting of mixtures of interacting monomers is a relatively well-established method for achieving the co-existence of complex assemblies in both organic solvents and aqueous medium. Yet such sorting tends to be limited to different assemblies within a single class of structures, such as coordination cages, foldamers, nanofibers, or knots. Self-sorting of complex structures into two dissimilar structure classes is still a largely unmet challenge in systems chemistry.

We also reasoned that dynamic combinatorial chemistry (DCC) should allow access to the desired self-sorting behavior. Dynamic combinatorial libraries (DCLs) are generated upon mixing building blocks that react with one another through reversible covalent bonds and library members continuously interconvert by exchange of these building blocks. Although the possible number of library members can be vast, molecular recognition events involving added molecules (templates) or between library members can direct the product distribution towards a specific thermodynamically stable system, including self-replicating molecules and foldamers.

Specifically, we have shown that oxidation of aryl dithiol building block 1a, functionalized with a pentapeptide sequence, gives rise to a library of macromolecules, from which a specific ring size was amplified as a result of assembling into fibers stabilized by β-sheet interactions. This assembly process is autocatalytic so the assembling macromolecules can be
considered self-replicators. In other work, the same aryl dithiol motif, functionalized with different moieties (nucleobases as in building block 2 or short peptides), were found to form specific, unexpectedly large macrocycles (ranging from a 9mer to a 23mer).\[11b,13\] These structures form by virtue of folding into well-defined three-dimensional conformations, akin to the folding exhibited by proteins.

Our first attempts at achieving the parallel emergence of foldamers and self-replicators, by mixing the nucleobase building blocks that produced foldamers with the peptide building blocks that produced replicators, failed as a result of the dominance of self-replicators that incorporated both building block types at a specific building block ratio and the absence of self-sorting when varying building block ratios. A representative example of the observed diversity in products obtained by mixing equimolar amounts of building blocks 1a and 2 is shown in Figure 1a.\[14\]

We now report a system in which two cross-reactive building blocks self-sort into a replicator fiber and a foldamer that both emerge spontaneously and stably co-exist subsequently. The two structures rely on reversible disulfide chemistry and in the self-sorting process replicator fiber formation is followed by the emergence of the folded structure. These results demonstrate that a mixture of nucleobase and peptide building blocks can autonomously self-sort.

Results and Discussion

In the previously reported system\[14\] based on peptide-based building blocks 1a and nucleobase-containing building block 2 self-sorting did not occur, despite the fact that 1a and 2, individually, can give rise to replicators\[12\] and foldamers,\[11b\] respectively. In order to favor self-sorting we choose to replace 1a with the somewhat more hydrophobic building block 1b (Scheme 1), featuring a β-alanine instead of a glycine residue, which should enhance the stability of the stacks of replicators and facilitate their emergence across a wider range of building block stoichiometries. When 1b would still form a replicator also in the presence of excess 2 then foldamer formation from the latter would be possible at the same time.

Indeed, two distinct species emerge from a DCL made from an equimolar solution of 1b and 2 ([1b] = [2] = 1.0 mM) in aqueous borate buffer (50 mM Na₂B₄O₇, pH 8.0, 1.0 M NaCl) upon stirring at room temperature in the presence of air. The reaction progress was monitored by UPLC-MS at a wavelength where building blocks 1b and 2 have comparable molar absorptivities. Control experiments showed that these absorptivities are independent of the macrocycles in which the building blocks reside (see Figure S6), so that peak areas directly reflect concentrations (in terms of units of building block). The first 24 hours revealed a mixture of differently sized mixed building block macrocycles (Figure 1b), however, after 7 days, the DCL was dominated by two major species: cyclic trimer (1b)₃₂₂₁ and foldamer 2₁₁₅.
accompanied by a small amount of another cyclic trimer (1b)2 (Figure 1c). This contrasts with the results of a DCL prepared from 1b alone which forms only the tetramer (1b)4 (Figure 1d). Interestingly, 215 was the only foldamer that emerged from a plethora of foldamer possibilities and the ring size was preserved as 2 alone also generates a foldamer consisting of 2 (Figure 1c).\[^{[10b]}\]

Self-sorting was found to be robust to changes in pH in the analyzed range of pH 6–8, while high ionic strength (0.5–1.0 M NaCl) promoted foldamer formation (see Figure S22). The self-sorting behavior was monitored over time at different building block ratios and revealed two phases of emergence. A series of libraries was set up with a total concentration of 2.0 mM with different molar fractions of 1b (100%–0%) in borate buffer and monitored by UPLC. In cases where 1b was present at a molar fraction of less than 60%, the kinetic traces revealed two phases of emergence: the immediate formation of (1b)2, until all of building block 1b was consumed, which was followed by the second phase of 2 growth until finally both species co-exist. The time of completion of (1b)2, formation (and the subsequent initiation of foldamer formation) is dependent on the concentration of building block 1b available in the library (Figure S1, S2). For example, at 50% in 1b, the (1b)2 species forms in 96 h (Figure 2a, shaded area) corresponding to the onset of 1b growth (shaded area). The (1b)2 species forms in 48 h, thereby promoting the early formation of 2. This suggests that (1b)2 indirectly promotes the formation of only a single foldamer in the two-component system by removing 1b from the exchange pool, reminiscent of the behavior observed previously for coordination cages.\[^{[5a]}\]

The self-sorting of the system into (1b)2 and 2 is preserved across a wide range of building block ratios. Only when the molar fraction of building block exceeds 70% in 1b foldamer formation is no longer observed, as all 2 is now used for replicator formation (Figure 2c, S1 and S2). We were thus interested in characterizing (1b)2, given its important role in the self-sorting process.

Seeding experiments reveal (1b)2 and 1b, are both self-replicators that form fibers. The addition of 6 mol% of preformed (1b)2 (seed) into a fresh DCL made from 1b (1.0 mM) and 2 (0.50 mM) accelerated the formation of (1b)2, in comparison to the non-seeded library, which is a hallmark of autocatalytic self-replication (Figure 3a).\[^{[11a]}\]

Acceleration of (1b)4 formation was also observed upon the addition of (1b)2 seed to a fresh solution of 1b (Figure 3b). Previous reports by our group show that peptide-based replicators owe their autocatalytic behavior to the formation and breakage of fibrous structures formed by β-sheet stacking.\[^{[11b]}\]

Note that, as expected, the formation of foldamer 2 did not show any signs of autocatalysis (Figure S24).

Figure 2. a) Kinetics of emergence of replicator (1b)2, and foldamer 2, from an equimolar solution of 1b (1.0 mM) and 2 (1.0 mM) show that 2, emerges only after formation of (1b)2, is near complete (shaded area, 96 h). b) Earlier emergence of 2 is observed in an analogous experiment starting from a different building block ratio (30% in 1b, as (1b)2 growth is complete early at 48 h (shaded area). c) Experiments starting from different ratios of 1b and 2 show that replicators and foldamers dominate the DCL.

Figure 3. Seeding experiments reveal accelerated formation of (1b)2, or (1b), indicative of self-replication when a) (1b)2, (6 mol%) seed is added to a solution of 1b and 2 (1b 1.0 mM, 2 = 0.5 mM, red trace) and when b) (1b)4, (10 mol%) seed is added to 1b alone (1b = 1.0 mM, green trace) in comparison to non-seeded samples (black traces). TEM images reveal bundled fibers from samples of c) (1b)2, and d) (1b)4.

Negative staining transmission electron microscopy (TEM) micrographs of the libraries dominated by replicators (1b)2, and 1b, showed bundled fibrous nanostructures (Figures 3c and d). β-sheet stacking of the pentapeptide sequence contributes to the formation of these fibers as evident from both Circular Dichroism (CD) spectroscopy and Thioflavin T fluorescence (Figure S3a). The CD spectra of replicator (1b)4 showed strong positive helicity at 206 nm and negative helicity around 220 nm, indicative of the presence of β-sheet structure\[^{[12]}\] (Figure S3b). Similar yet weaker helicities were observed for replicator (1b)2.

The emergence of self-replicating fibers, such as (1b)2, operates via a fiber elongation/breakage mechanism and is heavily affected by mechanical agitation.\[^{[13]}\] In the absence of mechanical agitation, a fresh library was set up by dissolving...
equimolar building blocks 1b and 2. UPLC analysis revealed that self-sorting was less efficient in the absence of agitation (Figure S4). A range of different small rings composed of building block 1b and 2 formed. 15mer rings formed as well but incorporated one or two units of building block 1b (e.g. (1b)2(1a)). We conclude that, in the absence of efficient self-replication by (1b)2, the degree of self-sorting is diminished. Furthermore, the autocatalytic (1b)2 fiber is more stable than a mixture of 2a and (1b). The addition of 1b (1.0 mM) to a solution containing 2a (1.0 mM) promoted the partial degradation of the foldamer and, following the transient formation of (1b)4, the dominant species was (1b)2, consuming both 2a and (1b)4 (Figure S5).

The self-sorting observed in the 1b/2 system was absent in the 1a/2 system. The difference between the two systems is relatively small: 1b contains an additional methylene group. Yet this difference makes 1b more hydrophobic and more efficient at assembling as compared to 1a, as evident from the fact that cyclic oligomers of 1b are capable of self-assembling at a smaller ring size than those of 1a (tetramer vs. hexamer; we know from previous work that the minimal ring size required for assembly is inversely correlated to building block hydrophobicity). [12] We speculate that this enhanced tendency for 1b-containing rings to self-assemble translates into (1b)2 assemblies being more stable than the corresponding (1a)2 assemblies. This notion is supported by the fact that (1b)2 assemblies are more robust to changes in building block stoichiometry than the corresponding (1a)2 assemblies, as evident from Figure S23.

**Conclusion**

We have shown how a DCL made from a nucleobase-containing and a peptide-based building block self-sorts into a foldamer incorporating 15 units of the nucleobase building block and a self-replicator that contains both building block types. Intriguingly, the foldamer is formed from the nucleobase building block while the peptide building block is fully incorporated into the self-replicator, which is opposite to the distribution of these compound classes in Nature, where peptides make folded proteins and nucleic acid oligomers are replicated. The emergence of the self-replicator promotes self-sorting and indirectly facilitates foldamer formation by homogenizing library composition: by taking the peptide building block out of the exchange pool the latter becomes relatively enriched in the remaining nucleobase building block, which is then more readily incorporated into the foldamer. These results show for the first time that an initially diverse mixture of compounds can transform spontaneously and selectively into discrete and stably co-existing foldamers and self-replicators. Since self-sorting occurs also when the system is closed (i.e. not agitated; cf. Figure S4), albeit less efficiently, the sorting process is thermodynamically favorable, indicating that the strength of the non-covalent interactions in the folded and self-assembled structures are sufficient to overcome the entropic cost that is inevitably associated with sorting. These findings show that assembly processes are a powerful force to counteract the inherently divergent nature of chemistry and push chemical systems to the formation of a few specific structures where many alternatives are also potentially accessible. Such changeover from divergence to specificity is key to the transition from chemistry to biology in the contexts of the origin and de-novo synthesis of life.

**Acknowledgements**

This research was supported by the EU (ERC AdG 741774, MCIF 745805-DSR; MCIF 786350-PSR), NWO (VICI grant) and the Dutch Ministry of Education, Culture and Science (Gravitation program 024.001.035).

**Conflict of interest**

The authors declare no conflict of interest.

**Keywords:** dynamic combinatorial chemistry · foldamers · self-assembly · self-replicators · self-sorting

---

[1] a) F. H. C. Crick, J. Mol. Biol. 1968, 38, 367 – 379; b) L. E. Orgel, J. Mol. Biol. 1968, 38, 381 – 393; c) W. Gilbert, Nature 1966, 215, 618 – 618; d) A. Szilágyi, B. Königü, T. Czárán, Sci. Rep. 2020, 10, 51.

[2] Á. Kun, A. Szilágyi, B. Königü, G. Boza, I. Zachar, E. Szathmáry, Ann. N. Y. Acad. Sci. 2015, 1341, 75 – 95.

[3] a) M. P. Robertson, G. F. Joyce, Chem. Biol. 2014, 21, 238 – 245; b) T. A. Lincoln, G. F. Joyce, Science 2009, 323, 1229 – 1232.

[4] a) M. M. Safont-Sempere, G. Fernández, F. Wurtzner, Chem. Rev. 2011, 111, 5784 – 5814; b) M. Mukhopadhyay, A. Wu, L. Isaacs, J. Org. Chem. 2004, 69, 6157 – 6164; c) S. Ghosh, L. Isaacs, Complex self-sorting systems, Wiley-VCH, Weinheim, 2010; d) K. Aratsu, R. Takeya, B. R. Pauw, M. J. Hollamby, Y. Kitamoto, N. Shimizu, H. Takagi, R. Haruki, S.-i. Adachi, S. Yagai, Nat. Commun. 2020, 11, 1623; e) J.-F. Ayme, S. Dhers, J.-M. Lehn, Angew. Chem. Int. Ed. 2020, 59, 12484 – 12492; Angew. Chem. 2020, 132, 12584 – 12592; f) A. Ghosh, M. Schmittel, Beilstein J. Org. Chem. 2020, 16, 2831 – 2853.

[5] a) F. J. Rizzuto, J. R. Nitschke, J. Am. Chem. Soc. 2020, 142, 7749 – 7753; b) Y. Lei, Q. Chen, P. Liu, L. Wang, H. Wang, B. Li, X. Lu, Z. Chen, Y. Pan, F. Huang, H. Li, Angew. Chem. Int. Ed. 2021, 60, 4705 – 4711; Angew. Chem. 2021, 133, 4755 – 4761.

[6] a) C. Tsimantas, X. de Hatten, C. Douat, B. Kauffmann, V. Maurizot, H. Ibara, M. Takafuji, N. Metzler-Nolte, I. Huc, Angew. Chem. Int. Ed. 2016, 55, 6846 – 6852; Angew. Chem. 2016, 128, 6962 – 6966.

[7] R. Kubota, K. Nagao, W. Tanaka, R. Matsumura, T. Aoyama, K. Urayama, I. Hamachi, Nat. Commun. 2020, 11, 4100.

[8] J. Zhong, L. Zhang, D. P. August, G. F. S. Whitehead, D. A. Leigh, J. Am. Chem. Soc. 2019, 141, 14249 – 14256.

[9] a) C. W. Hsu, O. Š. Miljanić in Dynamic Covalent Chemistry (Ed.: W. Zhang, Y. Jin), John Wiley & Sons Ltd, Chichester, UK, 2017, pp. 253 – 286; b) Ž. Š. Miljanić, Chem. Rev. 2017, 2, 502 – 524; c) A. Herrmann, Chem. Soc. Rev. 2014, 43, 1899 – 1933; d) M. Kołodziejski, A. R. Stefankiewicz, J.-M. Lehn, Chem. Sci. 2019, 10, 1836 – 1843; e) S. Klotzbach, F. Beuerle, Angew. Chem. Int. Ed. 2015, 54, 10356 – 10360; Angew. Chem. 2015, 127, 10497 – 10502.

[10] a) P. Frei, R. Hevey, B. Ernst, Chem. Eur. J. 2019, 25, 60 – 73; b) E. Mattia, S. Otto, Nat. Nanolett. 2015, 10, 111 – 119.
[11] a) J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J.-P. Peyralans, S. Otto, Science 2010, 327, 1502–1506; b) B. Liu, C. G. Pappas, E. Zangrando, N. Demitri, P. J. Chmielewski, S. Otto, J. Am. Chem. Soc. 2019, 141, 1685–1689; c) T. Kosikova, D. Philp, J. Am. Chem. Soc. 2019, 141, 3059–3072; d) T. Kosikova, D. Philp, Chem. Soc. Rev. 2017, 46, 7274–7305.

[12] M. Malakoutikhah, J. J. Peyralans, M. Colomb-Delsuc, H. Fanlo-Virgos, M. C. Stuart, S. Otto, J. Am. Chem. Soc. 2013, 135, 18406–18417.

[13] C. G. Pappas, P. K. Mandal, B. Liu, B. Kauffmann, X. Miao, D. Komáromy, W. Hoffmann, C. Manz, R. Chang, K. Liu, K. Pagel, I. Hue, S. Otto, Nat. Chem. 2020, 12, 1180–1186.

[14] B. Liu, C. G. Pappas, J. Ottelé, G. Schaeffer, C. Jurissek, P. F. Pieters, M. Altay, I. Marić, M. C. A. Stuart, S. Otto, J. Am. Chem. Soc. 2020, 142, 4184–4192.

[15] P. W. J. M. Frederix, J. Idé, Y. Altay, G. Schaeffer, M. Surin, D. Beljonne, A. S. Bondarenko, T. L. C. Jansen, S. Otto, S. J. Marrink, ACS Nano 2017, 11, 7858–7868.

[16] M. Colomb-Delsuc, E. Mattia, J. W. Sadownik, S. Otto, Nat. Commun. 2015, 6, 7427.

Manuscript received: January 22, 2021
Version of record online: May 5, 2021