Showcasing research from Professor Alain Jonas’ laboratory, Université catholique de Louvain, Belgium, and Professor Mathieu Surin’s laboratory, Université de Mons, Belgium.

Dynamic self-assembly of supramolecular catalysts from precision macromolecules

The self-assembly of two complementary precision oligomers leads to a collection of linear and cyclic oligomers in which di(oligomeric) macrocycles are favored at low concentration. In the macrocycles, the five catalytic groups needed for the Copper(I)/TEMPO-catalyzed aerobic oxidation of alcohols are brought closely together as shown by molecular dynamics simulations and network analysis, resulting in high catalytic turnover frequencies even at high dilution (cover image). The self-assembly of the complementary oligomers and their catalytic activity are mathematically described, which provides guidelines to further improve self-assembling catalytic systems.
Dynamic self-assembly of supramolecular catalysts from precision macromolecules†

Qian Qin,‡a Jie Li,‡a David Dellemme,b Mathieu Fossépré,b Gabriella Barozzino-Consiglio,a Imane Nekkaa,a Adrian Boborodea,c Antony E. Fernandes,a,ac Karine Giinel,b Mathieu Surin,c,b and Alain M. Jonasd,†a

We show the emergence of strong catalytic activity at low concentrations in dynamic libraries of complementary sequence-defined oligomeric chains comprising pendant functional catalytic groups and terminal recognition units. In solution, the dynamic constitutional library created from pairs of such complementary oligomers comprises free oligomers, self-assembled d(oligomeric) macrocycles, and a virtually infinite collection of linear poly(oligomeric) chains. We demonstrate, on an exemplary catalytic system requiring the cooperation of no less than five chemical groups, that supramolecular d(oligomeric) macrocycles exhibit a catalytic turnover frequency ca. 20 times larger than the whole collection of linear poly(oligomers) and free chains. Molecular dynamics simulations and network analysis indicate that self-assembled supramolecular d(oligomeric) macrocycles are stabilized by different interactions, among which chain end pairing. We mathematically model the catalytic properties of such complex dynamic libraries with a small set of physically relevant parameters, which provides guidelines for the synthesis of oligomers capable to self-assemble into functionally-active supramolecular macrocycles over a larger range of concentrations.

Introduction

In biomolecular catalysts, the various chemical groups involved in the catalytic site self-assemble in spatially close positions, leading to optimal availability and often adaptive activity. This optimized self-assembly of active groups is intrinsically dynamic, allowing accessibility to and from the reactive site, and the realization of the reaction transition state.1,2 Reaching such an efficiency in the dynamic self-assembly of active groups and resulting catalytic activity remains challenging in synthetic molecular catalysts. Here, we show by molecular dynamics simulations and experiments that, by designing sequence-defined oligomeric chains equipped with functional catalytic groups and complementary recognition units, it is possible to induce the dynamic self-assembly of supramolecular cyclic structures displaying unusually high catalytic efficiency at low concentrations. Importantly, we also show that the catalytic properties of such oligomers can be accurately predicted by a compact mathematical model, despite the multiplicity of their possible assembly configurations.

Attempts to form catalytic sites by self-assembly of synthetic macromolecules have been reported using peptides,3–7 DNA scaffolds,8–14 single-chain polymeric nanoparticles (SCNPs),15–20 peptides,21 metal organic frameworks,22,23 supramolecular polymers,24 foldamers,25 and helical (supramolecular) polymers.26–29 Additionally, supramolecular and covalent macrocycles have been synthesized to create catalytic sites involving the concerted action of different functional groups,30–34 because cyclic configurations decrease the accessible conformational space and thereby increase the probability to bring catalytically-active groups in spatially-close positions.

In our approach, we combine these ideas by designing two precision oligomers (Oa and Oa1 in Fig. 1D) that can self-assemble in a variety of configurations, including d(oligomeric) macrocycles (Fig. 1E). All these supramolecular species possess the five cooperative catalytic groups involved in the copper-catalyzed aerobic oxidation of alcohols, which are...
distributed between the two oligomers. We demonstrate the emergence of catalytic activity from the formed dynamic constitutional library at low concentrations, and that this activity mainly results from supramolecular di(oligomeric) macrocycles that enable the ideal positioning and connection of the cooperative catalytic residues within a confined space. Additionally, a mathematical model is developed to quantitatively predict the self-assembly of catalytic species and the catalytic properties of the constitutional library, thereby providing designing rules for self-assembled macromolecular catalytic systems.

Results and discussion

To demonstrate the possibility to self-assemble two properly-designed precision oligomers into a catalytically-active superstructure, we selected a system developed for the sustainable aerobic oxidation of alcohols based on 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) and Cu(I). Although the catalytic mechanism is still discussed, it was proposed that the catalytic cycle requires the formation of a binuclear copper complex involving two bidentate nitrogen ligands (such as bipyridine) and two auxiliary ligands (such as N-methylimidazole, NMI) (Fig. 1A). In support of this hypothesis, recent findings have shown the effectiveness of dinuclear copper complexes in the TEMPO-catalyzed aerobic oxidation of alcohols. As shown in ESI† (Section 12), our results also indicate that the catalytic cycle involves two Cu(I) atoms, two pyridyltriazole bidentate copper ligands (P), and two imidazole auxiliary copper ligands (I), in addition to the TEMPO radical (M).

These five needed functional groups were precisely distributed between two strands Oa and Ob, each strand having
intrinsically a very low catalytic activity, if any. Additionally, the two strands were equipped with complementary hydrogen-binding recognition units at their ends to favor their assembly into a variety of configurations, including supramolecular di(oligomeric) cycles, while avoiding the formation of homo-oligomeric supramolecular structures (Fig. 1B–E).

The functional groups are attached as side chains to an oligo(urethane triazole) backbone (Fig. 1B); the precise sequence of these complementary oligomers is displayed in Fig. 1D. The general synthetic procedure (Scheme 1) was demonstrated before to provide access to sequence-defined stereocontrolled oligomers. Here, we considerably extended the range of available side chains to include complementary nucleobase analogs (cytosine C, guanine G, thymine T and 2,6-diamidopyridine D) acting as recognition units through G–C and D–T pairing (Fig. 1C). The catalytic groups comprise two pyridytriazole (P)–copper complexes, two imidazole co-ligands (I and I′ having different spacer lengths for optimal accessibility), and one TEMPO radical (M). Additionally, hexyl side chains (C₆) were added at both ends of the chains to favor solubility in the acetonitrile : dimethylsulfoxide 95 : 5 v/v solvent system used for the experiments and possibly contribute to stabilization of the self-assembled structures. The synthetic procedures and characterization of the intermediary and final products by NMR, mass spectrometry and SEC chromatography are described in ESI (Sections 1–5).

The first strand, Oₐ, consists of a C₆GMPTC₆ sequence (catalytic groups underlined), complementary to the second strand, Oₐ₁, with a C₆ClIPDC₆ sequence (Fig. 1D). When self-assembled in a di(oligomeric) cycle in the presence of Cu(I), the two strands should provide together the five groups needed for catalysis (Fig. 1A). To demonstrate the completeness of this self-assembled active site, two other Oₐ strands were also synthesized, both of which can only form an incomplete catalytic site in combination with Oₐ:aC₆CIIIPDC₆ (Ob₂) sequence that lacks a P group, and a C₆ClIPDC₆ (Ob₃) sequence that lacks an I group (Fig. 1D). Due to the removal of one unit in one strand of each of these two systems, the Ob₂ and Ob₃ chains are one unit shorter than the Ob₁ chain of the complete system. The libraries of self-assembled configurations that form from these three systems are shown in Fig. 1E.

Molecular dynamics (MD) simulations of one Oₐ and one Oₐ₁ chains placed in an effective solvent indicate the rapid assembly of the oligomers in a stable heteromolecular complex of ca. 1 nm radius of gyration (Fig. 2A; details in ESI Section 10†). Inside this compact globule, the two strands remain flexible and highly dynamic, allowing the different units to rearrange in space due to a vast network of interactions involving functional

---

**Scheme 1** Synthesis of the oligomers. The compounds (R)-1, (R)-2 and (R)-3 were prepared according to the literature and our previous work.© 2023 The Author(s). Published by the Royal Society of Chemistry Chem. Sci. 2023, 14, 9283–9292 | 9285

© 2023 The Author(s). Published by the Royal Society of Chemistry
moeities (catalytic groups and nucleobase analogs) as well as the carbamate and triazole moieties of the backbone. Interestingly, the catalytic units are among the less buried residues and are accessible at the periphery of the globule (Fig. S164†).

Network theory applied to the collection of conformations generated over a 10 μs timescale indicates that interchain connectivity arises dominantly through interactions between complementary bases, whereas intrachain connectivity replicates the primary structure of the chains (Fig. 2B). Other parameters of the network such as betweenness and average shortest path length provide further support to the formation of a globule stabilized by interactions between nucleobases (Fig. S166†). A modular representation of the network shows that the catalytic modules (M, I′, I, P in Fig. 2C) are connected via the connectivity between the bases. Analysis of MD simulations indicates that the formation of this supramolecular complex is driven both by hydrogen bonding and π-stacking interactions (Fig. 2D).

Heatmaps enumerating the frequency of hydrogen bonds (Fig. 2E) show that C⋯G and D⋯T pairings provide an important but not exclusive contribution to the persistence of the globular assembly, with D⋯G pairing providing a significant but less frequent mechanism of stabilization. Stacking interactions are far less specific, involving numerous residues including triazole groups from the backbone and aromatic catalytic units (I′, I, P) (Fig. S163†).

Heatmaps of the enthalpy of binding per residue provide a detailed view on the stabilization per residue pair, globally showing the stabilization of the supramolecular complex by interchain interactions between the different bases. This indicates that when forming the complex, intermolecular interactions outperform the intramolecular interactions of each strand when alone. The same analysis was performed on the Oa/Ob2 system and provided similar results, with the same type of interactions stabilizing the formation of a globular complex (Fig. S165†).

The catalytic systems were tested in the aerobic oxidation of benzyl alcohol (BnOH) into benzaldehyde (BzH), for different temperatures between 30 °C and 60 °C and at different molar concentrations of catalyst. Cu(I) was introduced in stoichiometric amount with respect to P groups, and catalyst concentration is expressed as the content in M units relative to the
molar concentration of introduced alcohol (0.2 mol L$^{-1}$). Gas chromatography was used to monitor the appearance of BzH; no other oxidation product could be found. Previous studies$^{35}$ indicated that the competition of urethane and triazole motifs for Cu complexation is limited compared to P; additionally, competition of base analogs for copper binding is also limited according to NMR (ESI Section 7†). Results from a typical catalytic experiment are displayed in Fig. 3A; similar graphs for all other conditions are in ESI Fig. S158.†

The turnover frequencies (TOFs) measured between 30 and 60 °C for the different oligomeric systems at different concentrations are plotted in Fig. 3B–D. The TOF is defined as the slope versus time of the BzH molar concentration at moderate conversion rates, divided by the molar concentration in catalyst; it represents the catalytic efficiency per mole of catalyst at low to moderate conversion. The catalytic activity of single oligomer chains is very low (diamonds in Fig. 3D), since they lack part of the needed catalytic groups; when combined with their complementary oligomer, they display very different catalytic activities depending on whether the two chains taken together comprise or not a complete catalytic system. Indeed, the collection of poly(oligomers) of the incomplete systems (such as in Oa/Ob$_{b2}$ solutions) exhibits a very low TOF despite comprising all the groups needed for catalysis (Fig. 3C); in contrast, the library of poly(oligomers) of the complete Oa/Ob$_1$ system exhibits a much higher TOF at low concentration (Fig. 3B). Additionally, a catalytic system made of all monomer units at the same concentration as in the complete Oa/Ob$_1$ system was measured for reference (crosses in Fig. 3D); its performance rapidly decreases upon dilution, owing to the decrease of the probability of chance encounter of the catalytic groups. This is in marked contrast with the complete Oa/Ob$_1$ catalytic system that resists dilution and intriguingly exhibits an optimum close to 1 mol% (Fig. 3B). Likewise, the system consisting of the Oa and Ob$_1$ oligomers devoid of H-binding G, T, D and C units and of hydrophobic C$_6$ units (the MP dimer and the II'P trimer, pentagons in Fig. 3D), although very effective at high concentrations due to its limited steric crowding, rapidly decreases with concentration and becomes less efficient than the hydrogen-bonded system below 1 mol% of catalyst.

The complete Oa/Ob$_1$ system is thus peculiar in its specific ability to maintain catalytic activity at low dilutions. Whereas higher catalytic activities can certainly be obtained for the other systems simply by introducing larger amounts of catalytic components (such as for the system made of separated monomer units or the MP/II'P system, Fig. 3D), the complete system is unique in its preservation of catalytic efficiency when diluted. We ascribe this resistance to dilution to the formation of self-assembled catalytic di(oligomeric) cycles.

The oligomer chains dynamically bind in a variety of configurations (Fig. 1E), resulting in the coexistence of different species including free oligomeric strands, linear poly(oligomer) s of varying length and chain ends, and di(oligomeric) cycles. In the incomplete Oa/Ob$_{b2}$ and Oa/Ob$_{b3}$ systems, self-assembly configurations comprising at least three oligomers generally contain all groups needed for catalysis, whereas di(oligomeric) configurations do not; as shown in Fig. 3C, these self-assembly configurations only provide mediocre catalytic efficiency. The fully different behavior of the complete Oa/Ob$_1$ system thus necessarily arises from di(oligomer) assemblies, which now comprise all needed catalytic groups. Since these di(oligomeric) assemblies tend to easily form self-assembled macrocycles as shown by molecular dynamics simulations, it follows that the high catalytic efficiency at low concentrations of the Oa/Ob$_1$ system arises from the formation of di(oligomeric) cycles.

Despite the virtually-infinite number of species in each constitutional library, their composition and catalytic activity can be modeled in a compact form by a mathematical model comprising a small set of physically-meaningful parameters, common to all libraries, temperatures and concentrations (ESI Section 9†). These parameters include two binding constants,
$K_1$ and $K_2$ associated to chain end pairing and a standard entropy penalty $\Delta S^e_2$ for cycle formation (Fig. 1E). The concentration in each species can be computed numerically from this model, as well as their contribution to the turnover frequencies at each temperature and concentration (ESI Section 9†). This compact mathematical model was simultaneously fitted to all experimental catalytic turnover frequencies measured at different temperatures and concentrations, for the two Oa/Ob1 and Oa/Ob2 systems taken together, leading to a quantitative reproduction of the experimental data (lines in Fig. 3B and C); parameters are reported in Table S2 (ESI Section 9†).

The possibility to represent at once with a compact mathematical model the catalytic properties of such complex self-assembling systems is an important outcome of this study. However, the model needs to be validated by an independent measurement of the concentration of each species in the solution (Fig. 1E), especially d(oligomeric) cycles which are of interest to us. These species are in constant dynamic equilibrium and cannot be isolated; furthermore, the oligomers they contain are involved in similar interactions in all species. Their concentrations are thus very difficult to measure by any available technique. However, each species possesses its own molecular mass and hydrodynamic volume. The diffusion coefficient of a proton attached to an oligomer included in a species of the constitutional libraries depends directly on the hydrodynamic volume of this species, which can be computed from its composition in each oligomer and from the hydrodynamic volume of the oligomers (ESI Section 9†). Additionally, the diffusion coefficient averaged over all co-existing species only depends on the hydrodynamic radii and concentrations in solution of these species. Hence, measuring the average diffusion coefficient of the protons provides an elegant way to check in situ the compositions of the constitutional libraries predicted by our model.

The diffusion coefficient of the protons of the complete Oa/Ob1 system was thus measured by DOSY NMR. A single diffusion band was seen at all concentrations by DOSY proton NMR performed in deuterated acetonitrile/DMSO 95 : 5 v/v at 300 K in the absence of copper (ESI Section 6†), indicating that fast exchange of the oligomers between the different co-existing species occurs within the NMR timeframe. In these circumstances, the average diffusion coefficient can be computed from the concentration of each co-existing species by the Stokes–Einstein equation (ESI Section 9†), with as input parameters $K_1$, $K_2$, $\Delta S^e_1$ and the hydrodynamic radii of each isolated oligomer, which were measured by DOSY to be 0.50 ± 0.02 and 0.57 ± 0.02 nm for Oa and Ob1, respectively (Table S1, ESI Section 6†).

The measured and predicted diffusion coefficients of the complete Oa/Ob1 system were then compared (Fig. 4), using as parameters for the prediction the fit parameters obtained from the analysis of the catalytic data (Fig. 3). The experimental data agree very well with the predicted values assuming good solvent conditions (continuous line in Fig. 4).

This model was thus used to compute the distribution of Oa oligomers between the different species for the Oa/Ob1 systems, depending on temperature and concentration (Fig. 5 top row). Oligomers are essentially free at very low concentrations, whereas they dominantly belong to linear poly(oligomeric) chains above 1 mol% of catalyst. The decrease of the binding constants $K_1$ and $K_2$ with temperature explains why the average degree of poly(oligomerization) of the linear chains slightly decreases with temperature (inset in top row of Fig. 5); however,
the total amount of linear chains is very insensitive to temperature (Fig. 5, top right). Di(oligomeric) cycles slightly dominate the constitutional libraries at intermediate concentrations and lower temperatures, but this dominance is progressively lost as temperature increases. The decreased amount of cycles as temperature increases essentially results from the increased free energy cost associated to the entropy penalty for cycle formation ($\Delta S^\circ$).

The contributions of each species to the total catalytic turnover frequency of the complete Oa/Ob1 system, which is the product between the amount of Oa in each species and the turnover frequency per Oa (ESI Section 9†), are collected in the bottom row of Fig. 5. The contribution of supramolecular di(oligomeric) cycles to the total turnover frequency is the dominant factor that explains the high catalytic activity of the complete system, even when the supramolecular cycles do not dominate the composition (e.g., at 60 °C). At this higher temperature, the relative proportion of cycles decreases but their catalytic activity increases due to thermal activation, resulting in an overall dominance of the cycles in the catalytic performance.

Cycles are thus much better than linear chains to reach a high catalytic activity, despite both configurations having all required active groups. At the concentration maximizing di(oligomeric) cycles, the ratio between the contributions to the catalytic activity of cycles versus all linear chains is ca. 30 and 10, at 30 and 60 °C, respectively. This is due to the much higher probability for the catalytic groups to be in an active site when assembled in a supramolecular cycle, compared to when in linear chains. This demonstrates that, by assembling the complementary strands in a supramolecular cycle, and thereby maximizing the probability that the five required groups meet in space, dramatic improvements of the catalytic efficiency by more than one order of magnitude can be obtained compared to linear configurations. In contrast, for the incomplete Oa/Ob2 system, supramolecular cycles cannot form complete catalytic sites and therefore do not contribute to the catalytic activity; consequently, this system only exhibits low activity associated to the low probability of formation of complete catalytic sites in linear poly(oligomeric) chains.

It is tempting to compare the efficiency of our supramolecular catalyst to other systems based on the same composition, in which the different catalytic groups are positioned differently in space.64 Depending on the sequence of units along chains and on their grafting configurations when heterogeneized, catalysts based on M, I and P units have been shown to exhibit very different catalytic efficiencies, as judged from their initial TOF. Compared to systems grafted on porous silica beads, the current supramolecular cycles are generally slightly inferior, probably because of a smaller accessibility of the substrates in the crowded globules of the cycles. In contrast, at 60 °C and below 1 mol%, the current supramolecular catalysts are better than any other homogeneous system we tested so far; this advantage is rapidly lost as concentration increases, because supramolecular cycles are disfavoured at higher concentrations while the probability of encounter is increased for other systems. However, our model can be used to predict directions for a larger range of conditions favouring the formation of the catalytically-active self-assembled di(oligomeric) cycles. For instance, increasing the chain end pairing constants by a factor of ca. 10 compared to their current geometric average value at 30 °C would result in a solution consisting almost exclusively of supramolecular cycles below 1 mol% catalyst. To favour cycles at higher concentrations, which would result in higher absolute catalytic rates, a supplementary stabilization mechanism must be given to cycles compared to linear chains. One way to achieve this is to reduce the entropy cost of cycle formation ($\Delta S^\circ$), which might be attained by increasing the rigidity of the oligomer backbones; however, this might be counter-productive with respect to the probability to have the catalytic groups meet in space. Another way would be to provide a supplementary enthalpy of stabilization specific to macrocycles, for example by using pairs of recognition units at each end of each oligomer, which would create a directional interaction in the pairing of ends and could thereby favour cycles over linear chains.

Conclusions

Our study investigated the self-assembly of complementary oligomers into a dynamic library of poly(oligomeric) species, most of which contribute little to catalytic activity. In this dynamic library, di(oligomeric) cycles were demonstrated to be highly catalytically-active, provided the two chains together comprise all the needed elements for catalysis, as in the Oa/Ob1 system. When only one element is missing as in Oa/Ob2 and Oa/Ob3, the activity drops very significantly, despite all needed elements are present in the solution. The activity can be recovered by adding the missing unit as a free component in solution (Fig. S16†). These results support the assumption that five functional groups, including a dinuclear copper complex, are involved in the catalytic cycle.

Molecular dynamics simulations demonstrate that these self-assembled cycles form globules in which different interactions contribute to stabilization, even though the base analogs at the ends of the strands provide a continuous stabilization mechanism and connectivity between the catalytic groups, as shown by network analysis. Measurement of hydrodynamic radius also provides experimental support for the formation of di(oligomeric) cycles.

Hence, the primary structure of the oligomers is important for the emergence of catalytic activity, because the presence or absence of the needed catalytic groups on the strands directly affects the efficiency of the di(oligomeric) cycles. However, it is likely that the sequence order is not a critical factor at this stage, considering the limited length of the investigated oligomers. Nevertheless, quadraternary structures are crucial for the emergence of catalytic activity, which here arises dominantly from the self-assembly of the strands in di(oligomeric) cycles. Interestingly, the self-assembly of di(oligomeric) cycles happens in a relatively dilute range of concentrations, in which cycles totally outperform simple mixtures of catalytic groups or any combination of the catalytic groups in linear chains. This demonstrates that an efficient catalytic activity can be obtained at rather high dilutions in the presence of properly self-assembled structures.
Our study also demonstrates that a small number of experimental parameters is sufficient to describe the complexity of the dynamic constitutional library of species resulting from the assembly of complementary oligomers. Indeed, the reproduction of catalytic activities and diffusion coefficients by a compact mathematical model is an important outcome of the present work, showing that the complexity of this system is still amenable to a deterministic description. Additionally, the interest of using network theory to interpret molecular dynamics simulations, and to support assumptions made in the mathematical model of the system, is another important outcome of this study.

Data availability

GC spectra and a table of values extracted therefrom (yields, turnover frequencies) are available at https://doi.org/10.14428/DVN/RYTG0V. Tables of DOSY diffusion coefficients are available at https://doi.org/10.14428/DVN/1XCNOK.

Author contributions

Q. Q. and L. J. performed the synthesis and characterization of the molecules, the catalytic experiments, and the analysis of the spectroscopic results. A. E. F. contributed to define the project and provided guidance for synthesis and analysis. K. G. contributed to the definition of the project and to project guidance. I. N. contributed to the synthesis of molecules. A. B. performed the SEC analysis of the oligomers. G. B. C. performed the DOSY experiments and extracted the diffusion coefficients. M. S. led the research on the molecular dynamics simulations. D. D. and M. F. performed the molecular modeling simulations and interpretation. A. M. J. led the research and performed the modelling of the catalytic and diffusion data.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Fonds de la Recherche Scientifique - FNRS and the Fonds Wetenschappelijk Onderzoek, under EOS Project No. 30650939. Computational resources are provided by the Consortium des Equipements de Calcul Intensif (CECI), funded by FNRS (Grant No. 2.5020.11) and Wallonia Region. The authors acknowledge Raoul Rozenberg for APCl-MS, Hervé Degand for ESI-MS/MS analysis, and Laurent Collard for GC analysis. K. G. is Senior Research Associate and M. S. is Research Director of the F. R. S.-FNRS.

Notes and references


