A versatile strategy for the synthesis of sequence-defined peptoids with side-chain and backbone diversity via amino acid building blocks†

Shixue Wang,ab Yue Tao,ac Jianqun Wang,ac Youhua Taoac and Xianhong Wangac

Designing artificial macromolecules with absolute sequence order is still a long-term challenge in polymer chemistry as opposed to natural biopolymers with perfectly defined sequences like proteins and DNA. Herein, we combined amino acid building blocks and iterative Ugi reactions for the de novo design and synthesis of sequence-defined peptoids. The highly efficient strategy provided excellent yields and enables multigram-scale synthesis of perfectly defined peptoids. This new strategy furnishes the broad structural diversity of side chains, as well as backbones. Importantly, the overall hydrophobicity and lower critical solution temperature (LCST) behaviours of these precisely defined peptoids can be logically altered by variation of the sequence. By following the same Ugi chemistry, these peptoids are also conjugated to DNA in a simple way, facilitating the development of novel therapeutics.

Introduction

The sequence of biopolymers such as DNA, RNA and proteins is elaborately controlled, and functional substituents at well-defined positions cooperatively play an essential role in their special and unique functions.1,2 Recently, sequence control of synthetic polymers has been drawing great interest towards creation of more advanced functions like those of natural polymers,3,4 but the control remains a great challenge in polymer science.

Some concepts or methodologies for controlling the monomer sequence in synthetic polymers have been reported, including controlled radical polymerization,7–11 ring-opening metathesis polymerization,12,13 DNA-templated synthesis,14,15 and iterative synthesis.16–35 One-by-one iterative methods, which are generally carried out on a solid support, enable complete regulated monomer sequences for artificial peptides and DNA. Other notable examples include the method for the synthesis of triazene-based polymers,36 poly(alkoxylamine amide)s,37 oligothioetheramides,38–40 and thiolactone aminalysis,41,42 and the multicomponent reaction.43–45 Although solid-phase synthesis enables simple purification, solution-phase synthesis is desired due to its homogenous reaction kinetics and non-necessity for removal of the products from the solid support.46,47

Peptoids (or poly-N-substituted glycines) are a class of biomimetic sequence-defined polymers that have found a wide range of biotechnological applications.48–57 Peptoids are routinely synthesized using the solid-phase submonomer approach developed by Zuckermann et al.58 Up to now, several hundred commercially available amines have been applied to synthesize peptoids with a wide array of side chains.53,54 Although significant progress has been made in peptoids, developing a novel strategy towards the de novo synthesis of sequence-defined peptoids is still imperative. To ensure sufficient molecular diversity and complexity, it is highly desirable to synthesis backbone-variants of traditional α-peptoids, such as β-peptoids, which can expand the structural and functional repertoire of peptoids.52,53,59,60 However, submonomer synthesis similar to that used for α-peptoids was determined to be less efficient for β-peptoids.61 This challenge and the need for increased molecular diversity prompt us to develop a new approach to generate sequence-defined peptoids with broad structural diversity, both of the side chains and the backbones.

Herein, we describe a general route for the assembly of sequence-specific polymers through iterative Ugi reactions of amino acid building blocks (Scheme 1; see the ESI† for the detailed mechanism of the Ugi reaction). We hypothesized that amino acids, which supply the carboxylic acid and amino functions, could sequentially react with an aldehyde and an isocyanide to form a peptoid structure. In order to achieve scalability, we employed the efficient solution-phase synthesis, which theoretically allowed scalable reactions.23–25 The Ugi reaction43–47 was chosen as a coupling strategy due to its ability to introduce broad structural diversity via simply employing
To verify the feasibility of our approach, we initiated the synthesis of a short peptoid with glycine tert-butyl ester, 2a. The amino functionality of 2a was reacted in an Ugi reaction with acetic acid 1 as the starting substrate and benzaldehyde, and tert-butyl isocyanide. The use of a 0.1-fold excess of aldehyde and isocyanide is necessary to push the reaction to quantitative transformation (Fig. 1A and B). In addition, purification of the resultant product was realized by flash chromatography from silica using ethyl acetate/dichloromethane/petroleum ether (1:1:1) as the eluent. The resultant material of the Ugi reaction bears a tert-butyl ester moiety, which can be used for another Ugi reaction after quantitative hydrolysis with trifluoroacetic acid (TFA) and simple workup by evaporation. The structure of the product after deprotection was confirmed by $^1$H NMR via the disappearance of the tert-butyl groups at 1.43 ppm (Fig. S2†) and by MALDI-TOF-MS (Fig. S3†). The isolated yield over the two steps was 90%. After this first round of the Ugi reaction and deprotection, we continued peptoid synthesis with another round of reactions. Again, both reactions proceeded smoothly as determined by $^1$H NMR and MALDI-TOF-MS (Fig. 1C, and S4†). Additionally, it should be noted that the obtained peptoid based on building block 2a lacks chirality, due to the racemic characteristic of the Ugi product.66,69

**Results and discussion**

**Amino acid building blocks**

Our approach for achieving synthetic sequence control involves the synthesis of an amino acid-based building block. The building block framework includes a reactive amino group that can undergo Ugi reactions, and a tert-butyl ester protected carboxylic acid function that can undergo efficient deprotection. The amino acid building block can be synthesized from readily available amino acid starting materials (Fig. S1†). A unique feature of the amino acid platform is that a rich variety of amino acids enable the introduction of broad structural diversity into peptoid backbones. For this proof-of-concept study, we synthesized seven different amino acid building blocks 2a–g (Scheme 1, Fig. S1 and S16–S54†).

**A model short peptoid by iterative Ugi reactions**

To verify the feasibility of our approach, we initiated the synthesis of a short peptoid with glycine tert-butyl ester, 2a. The sequence-defined peptoids with side-chain diversity via iterative Ugi reactions

Having demonstrated the efficient synthesis of the model short peptoid, we proceeded with the synthesis of 9-mer and 10-mer sequence polymers, peptoid 1 and peptoid 2 (Fig. 2A and B). Both polymers were synthesized with the same amino acid building block 2a but have different side chain sequences. The synthesis of peptoid 1 and peptoid 2 was followed and confirmed at intermediate steps with MALDI-TOF-MS and SEC (Fig. 2, S5–S6 and S85–S86†). Further structural confirmation of both peptoids was obtained via MALDI-TOF-MS. The observed parent ions of the two peptoids were identical at 2252.4 Da and 2534.1 Da (Fig. 2A and B), and matched with the theoretical value of 2252.0 Da and 2533.9 Da, respectively. Furthermore, by the application of different aldehydes for the Ugi reaction, we could incorporate ten different pendant groups into the sequence-specific peptoid 3, such as aliphatic, aromatic, and olefinic motifs (Fig. 2C, 3, S7 and S71†). We were pleased to note that the strategy for these sequence-defined peptoids is scalable, and multiple grams of linear peptoids were obtained. For instance, peptoid 3 (4.01 g) was produced in a 40% isolated yield over 19 steps as an off-white powder material (Fig. S8†), indicating the high efficiency of this chemistry. Therefore, using the above chemistry can afford sequence-defined polymers for valuable applications, facilitating the exploration of structure–property relationships accurately. Starting from bis-COOH–functionalized PEG, the sequence-defined block polymer (peptoid 4) was synthesized using the above chemistry. The SEC curve (Fig. 2D) of the obtained peptoid 4 shifted to higher molecular weights ($M_n \sim 4.1$ kDa), which was consistent with the successful formation of the desired symmetrical sequence-specific products.

To illustrate the versatility of this strategy, we synthesized a 5-mer α-peptoid 5 (Fig. S9†) consisting of l-alanine (2b) and...
Fig. 1 (A) Synthesis of a model short peptoid by iterative Ugi reactions. (B) MALDI-TOF-MS spectrum of a 1-mer peptoid. (C) MALDI-TOF-MS spectrum of a 2-mer peptoid.

Fig. 2 (A) Chemical structure and MALDI-TOF-MS spectrum of a 9-mer sequence polymer (peptoid 1). (B) Chemical structure and MALDI-TOF-MS spectrum of a 10-mer sequence polymer (peptoid 2). (C) Chemical structure and MALDI-TOF-MS spectrum of a 10-mer sequence polymer with ten different pendant groups (peptoid 3). (D) Chemical structure and SEC traces of a sequence-defined block polymer (peptoid 4).
benzyl-protected L-serine (2c) building blocks. SEC and MALDI-TOF-MS spectra revealed its structural integrity. The corresponding peptoid 6 (Fig. S9 and S75†) with pendant hydroxyl groups was recovered by hydrolysis of tert-butyl ester followed by removal of the benzyl protecting group. The presence of hydroxyl groups in its side chains would enhance the hydrophilicity of the resultant peptoids, facilitating its biomedical applications.

Sequence-defined peptoids with backbone diversity via iterative Ugi reactions

Faced with the ever-present challenge of β-peptoid preparation using the submonomer method,54 we further synthesized a 10-mer β-peptoid 7 with β-alanine tert-butyl ester 2d as the building block, according to the aforementioned procedure (Fig. 4A, S76 and S87†). Importantly, the chemistry for β-peptoids is high efficient, and the isolated yield in each cycle over two successive steps (that is the Ugi reaction and deprotection) was ~85%.

The versatility and robustness of the above chemistry for building sequence-defined polymers were further demonstrated by constructing backbone-variants of traditional α-peptoids. We synthesized a 5-mer peptoid (Fig. 4B, S77 and S88†) consisting of all five amino acid building blocks. The resulting peptoid 8 was characterized by the MALDI-TOF-MS spectrum, and the doubly charged [M – O’Bu]⁺, and [M – O’Bu – Ac]⁻ ions were clearly observed. Thus, the iterative Ugi reaction combined with amino acid building blocks is indeed a very powerful new approach towards sequence-defined peptoids with backbone diversity.

Effect of the sequence structure of peptoids on materials properties

The above chemistry can afford an array of sequence-defined polymers for accurately exploring the structure–property relationship between sequence and material properties. The SEC traces obtained after Ugi reactions in the synthesis of the sequence-defined peptoid 3 (in THF, 35 °C, polystyrene as a standard).

Fig. 3  SEC traces obtained after Ugi reactions in the synthesis of the sequence-defined peptoid 3 (in THF, 35 °C, polystyrene as a standard).

Fig. 4  (A) Synthetic route and MALDI-TOF-MS spectrum of 10-mer β-peptoid 7 obtained by iterative Ugi reactions using β-alanine tert-butyl ester as the building block. (B) Chemical structure and MALDI-TOF-MS spectrum of a 5-mer peptoid consisting of all five amino acid building blocks (peptoid 8).
In particular, the study of the properties of polymers with similar chemical structures but different sequence structures is a fundamental topic in the polymer science area. For this purpose, we employed the tert-butyl isocyanide and oligo-ethylene-glycol isocyanide to make 12-mer peptoids containing hydrophobic and hydrophilic side chains (Fig. 5A and B, S10–S12, and S78–S81†). SEC measurements demonstrate the successful synthesis of the desired peptoids by the shifting towards higher molecular weights after each step (Fig. S89–S92†). Each of these peptoids differs in the placement fashion of the hydrophobic and hydrophilic units, which were grouped into a block size of one (peptoid 12), two (peptoid 11), three (peptoid 10), and six (peptoid 9). The resultant peptoids were soluble in cold water. The lower critical solution temperature (LCST) behaviours of these peptoids were investigated. As shown in Fig. 5C, a gradual increase in LCST was detected as the block size decreased from six in peptoid 9 to one in peptoid 12, demonstrating a gradual increase in the overall hydrophilicity. This may be due to the fact that the more water-soluble peptoid 12 may adopt a structure in which the hydrophilic OEG moieties provide better shielding of the hydrophobic portion from water.

Furthermore, to elucidate the backbone monomer sequence on the polymer properties, we applied the glycine tert-butyl ester and oligo-ethylene-glycol based amino acid tert-butyl ester as building blocks to make a series of main-chain sequence-regulated peptoids (Fig. 6).
Peptoids with ABB and AABBB sequence-regulated backbone structures, and peptoids with a block sequence in the backbone were verified by MALDI-TOF-MS and SEC measurements (Fig. 6B, S13–S14, and S93–S95†). The solubility of these peptoids was then assessed in the mixed solvents of EtOH and H2O at different temperatures via determining the light transmittance of the peptoid solutions. As shown in Fig. 6C, the peptoids with the same composition but different monomer sequences in the backbones, i.e., peptoid 13, peptoid 14 and peptoid 15, showed different solubilities; the peptoid with an ABB alternating sequence in the backbone (peptoid 15) was more soluble than the others: it became turbid at $T_c = 45 \degree C$ (cloud point), whereas its block counterpart (peptoid 13) did so at a higher temperature, $T_c = 53 \degree C$. These experiments again suggested that not only the side-chain sequences but also the backbone sequences were important for determining the solubility of the peptoids and that the solubility could be gradually adjusted by both backbone and side-chain monomer sequences. Actually, the interaction preference between the polymer and solvent might be changed depending on the placement fashion of monomer units.

**Bioconjugation of peptoids**

One possible application of the resultant sequence-specific peptoids may be in nucleic acid delivery. In addition, research on the combination of synthetic materials and biologics through bioconjugation has been the subject of increased study because this integration enables the development of novel therapeutics. Therefore, we evaluated whether our peptoids could be coupled to DNA for biomedical applications. By applying the same Ugi reaction and 3-azidopropan-1-amine as a component, we attached an azide functionality to a peptoid (Fig. 7A and S15†). The azide-modified peptoid 16 was coupled to DNA with a dibenzocyclooctyne (DBCO) group under ambient conditions without catalysts with acetonitrile and PBS as solvents. DBCO-DNA specifically conjugated to peptoid 16 based on gel electrophoresis shifts, and a 24 h reaction with DNA: peptoid molar ratio = 1 was found to be efficient based on the disappearance of the unconjugated DNA gel band (Fig. 7B). Further verification of conjugation was obtained via MALDI-TOF-MS (Fig. 7C).

**Conclusions**

In summary, we have described herein the development of amino acid building blocks coupled with iterative Ugi reactions for the efficient and multigram-scale assembly of sequence-defined peptoids. This efficient chemistry provides much feasibility for structural diversity, synthetically varying and sequencing both the side chains and the backbones. Activating agents were unnecessary, rendering this strategy advantageous over other stepwise methods, such as the traditional peptide or peptoid synthesis. We further demonstrated that the alteration in the overall hydrophobicity and LCST behaviours of these precisely defined peptoids could be accordingly changed by variation of the sequence. Regulation of sequence specific hydrophobic aggregation within a polymer is a significant result. We also showed that these materials could be conjugated to biomacromolecules, allowing the development of novel therapeutics. We anticipate that this versatile strategy may afford new peptoids for application in therapeutics and as supramolecular foldamers or simple protein mimics for the
investigation of advanced self-assembly driven by hydrophobic or other supramolecular interactions.

Conflicts of interest
There are no conflicts to declare.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant No. 21474101, 21704097 and 51873211), Jilin Science and Technology Bureau (Grant No. 20160414001GH and 20180201070GX), the Opening Foundations of Key Laboratory of Synthetic and Biological Colloids, Ministry of Education, Jiangnan University (Grant No. JDSJ2017-03) and State Key Laboratory of Supramolecular Structure and Materials, Jilin University (Grant No. sklssm201817).

Notes and references

64 B. Yang, Y. Zhao, Y. Wei, C. Fu and L. Tao, *Polym. Chem.*, 2015, 6, 8233.
65 B. Yang, Y. Zhao, S. Wang, Y. Zhang, C. Fu, Y. Wei and L. Tao, *Macromolecules*, 2014, 47, 5607.