Anion mediated, tunable isoguanosine self-assembly: decoding the conformation influence and solvent effects†‡

Mengjia Liu,a Ying He,a Chuan Shan,a Lukasz Wojtas,a Ion Ghiviriga,c Omar Fathalla,a Yu Yan,a Xiaopeng Liab and Xiaodong Shi*a

Systematic investigations were performed with various substituted groups at C8 purine and ribose. A series of isoG analogs, C8-phenyl substituted isoG were synthesized and applied for Cs+ coordination. The structural proximity between purine and ribose limited pentaplex formation for C8-phenyl substituted isoG derivatives. Based on this observation, deoxy isoG derivative with modification on ribose (tert-butyldimethylsilyl ether) was applied to assemble with the Cs+ cation. Critical solvent (CDCl3 and CD3CN) and anion (BPh4−/C0, BARF−/C0, and PF6−/C0) effects were revealed, leading to the controllable formation of various stable isoG pentaplexes, including singly charged decamer, doubly charged decamer, and 15-mer, etc. Finally, the X-ray crystal structure of [isoG20Cs3]3+(BARF−/C0)3 was successfully obtained, which is the first example of multiple-layer deoxy isoG binding with the Cs+ cation, providing solid evidence of this new isoG ionophore beyond two-layer sandwich self-assembly.

Introduction

The core value of supramolecular chemistry is the efficient construction of highly ordered structures at the molecular level through controllable building block interactions.1 The fundamental scientific tasks in supramolecular chemistry are (A) to develop building blocks that could conduct efficient non-covalent interactions; (B) to understand the driving forces between molecules to reach controllable molecular architecture; and (C) to apply these systems as potential solutions for important scientific problems that are difficult and/or challenging when using alternative covalent approaches.2 To reach these goals, development of basic supramolecular building blocks and understanding the driving forces toward the formation of specific non-covalent interactions are of great importance for tackling challenging scientific questions in alternative covalent approaches.3

Isoguanosine (isoG, also known as 2-hydroxy-adenosine) is a structural isomer of guanosine (G).4 IsoG pentaplex is an interesting supramolecular scaffold with intrinsic H-bond donors and acceptors. The 108° H-bond angle allows isoG to form a larger self-assembled core beyond tetrameric structures.

Scheme 1 IsoG self-assembly with unique C3 symmetry. (A) Formation of isoG-star from isoG self-assembly; (B) tunable deoxy isoG self-assembly.
formed by guanosine (G).

In 2000, the Davis group reported the first two X-ray crystal structures that provided the evidence for the formation of the H-bonded isoG₅ (isoG-star) self-assembly. As shown in Scheme 1, the isoG-star holds an overall planar structure, while the sp³ configuration at the C¹ position gives a twisted conformation between ribose and purine, producing a well-defined bowl-shape assembly, which further led to the formation of a stable sandwiched structure of isoG₁₀.

Unlike the G₄-quartet (square planar geometry), the isoG₅ gives a star-shaped building block with C₃ symmetry and larger pore. This unique feature of isoG has enabled it to selectively bind with large size cation Cs⁺ (radii = 174 pm) in high affinity. The selective binding of Cs⁺ over other alkali and alkali earth metal cations makes isoG₅ a potential host for effective caesium-137 extraction in radioactive waste treatment.

However, little progress has been made in isoG self-assembly studies since the initial breakthrough. This is probably due to the close distance between N9-ribose and N3-purine. Unlike G-quartet, where N3 is not involved in the H-bonding network, the N3 position is crucial in the isoG self-assembly as a key H-bond receptor. In addition, the crystal structure of isoG-star revealed interesting “extra H-bonds” between the ribose O²⁺ and purine N6-H₆₈ (Scheme 1A). The purine–ribose interaction in the isoG pentamer makes it much more challenging for precise conformation control. Thus, new systems that could achieve isoG self-assembly with good stability and structural diversity are highly desirable. Herein, we report a systematic study of the influence of the isoG C8 substitution and ribose modification on the formation of isoG-star. Conditions to achieve controllable isoG₅ pentaplex have been investigated, which revealed the critical solvent and anion effects (Scheme 1B). The successful preparation of singly charged decamer, doubly charged decamer and pentaplex isoG₁₀ through the anion and ribose conformation control was undeniably characterized by ¹H NMR, diffusion NMR and ESI-MS. The X-ray crystal structure of [isoG₂₀Cs₃]⁺ (BARF⁺) was successfully obtained, confirming the formation of pentaplex isoG₁₀ beyond the previously reported sandwich structure. These results not only highlighted the critical solvent and anion effects in this supramolecular assembly, but also provided new isoG-based structural motifs for molecular recognition and related potential applications.

**Results and discussion**

As reported by Davis, isoG₁ has been used to react with CsBPh₄ for the formation of isoG-star (Fig. 1). The relatively rigid ribose modification successfully controlled the formation of a well-defined bowl-shaped structure, which led to the sandwiched isoG₁₀. Notably, as an H-bond receptor, the N3 position is crucial for the isoG-star formation. However, the structural proximity between N3 and ribose limited the choice of modification on ribose or purine. In addition, the crystal structure of isoG-star from isoG₁ revealed interesting “extra H-bonds” between the ribose O²⁺ and purine N6-H₆₈, which makes one wonder if this purine–ribose interaction is necessary for the stable isoG-star formation. With all these structural analyses, the purine C8 position is the ideal synthetic handle to introduce new functionality without interrupting the needed H-bond in the isoG-star formation. The 8-phenyl modified compound isoG₂ was synthesized (see detailed synthesis in the ESI†). To evaluate the importance of purine–ribose O²⁺–N6-H₆₈ binding, 2-deoxy isoG₃ and 4 were also prepared. All these new isoG derivatives were applied in the self-assembly studies. It is known that, with the large cation radii, Cs⁺ provides the best match toward the isoG₅ pore. As reported previously, treating isoG₁ in CDCl₃ solution with CsBPh₄ (aqueous solution) gave effective extraction of Cs⁺ into the organic layer, forming stable [(1)₁₀Cs⁺(BPh₄⁻)]₂ complexes. However, conducting similar experiments with isoG₂ or isoG₃ gave no formation of Cs⁺ complexes.

Fortunately, the X-ray single crystal structures of both compounds were successfully obtained (Fig. 1). As shown in the crystal structures of isoG₂ (from 5'-TBS-deprotection of 2), the formation of the free-rotatable C-N bond is influenced by substitution of the purine C8 position. The anti-conformation is energetically favored by avoiding steric repulsion. As a result, O⁵⁺ was placed close to N3, blocking critical H-bonding in the isoG₅ formation. With the less rigid deoxy-ribose with no O²⁺, isoG₃ gave a H-bonding dimer since N3 is not accessible for H-bonding, which is critical for the isoG-star formation. These results highlighted the challenge associated with developing new systems for the preparation of isoG-star from small molecular assembly. As revealed by the crystal structures, C8 functionalization prevents N3 from being accessible for H-bonding to form isoG-star.

To examine whether O²⁺–N₆-H₆₈ H-bonding is necessary for the stable isoG-star formation, deoxy-isoG₄ with no C8 phenyl group and O²⁺ was applied for self-assembly with Cs⁺. As shown in Fig. 2E, treating isoG₄ with CsBPh₄ (excess amount) led to the formation of multiple sets of new signals in the ¹H NMR spectrum which indicated the formation of H-bonded complexes. Interestingly, monitoring the reaction over time suggested the formation of one major set of signals when treating the ligand and the cation for 30 minutes (Fig. 2D). The ¹H NMR spectrum integration confirmed that the ratio between isoG₄ and Cs⁺ is 10 : 1. This result is exciting since it suggested the formation of isoG-star could be achieved with deoxy-isoG.
interactions in this deoxy-isoG complexes. The biggest
Cs\(^{3+}\)(BARF\(^{-}\)) observed in the isoG
\(\text{˚}\) between 3.20 and 3.37
formed with the deoxy-ribose ligand. The Cs\(^{+}\) cations bind to
isoG\(^{\text{4}}\) monomer; (D) treating isoG\(^{\text{4}}\) with CsBPh\(_{4}\) for 30 min; (E) treating
isoG\(^{\text{4}}\) with CsBPh\(_{4}\) for 12 h.

without the extra ribose–purine H-bond. Various conditions
were used to obtain X-ray crystal structures. Fortunately, by
switching the anion to BARF\(^{-}\), a stable complex was obtained in a
CHCl\(_{3}\)–CH\(_{3}\)CN solvent mixture with structure confirmed by X-
ray crystallography as \([\left[4\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BARF}^{-}]_3}\) (Fig. 3). Notably, this
is not only the first crystal structure of an isoG-assembly from
deoxy-ribose but also the first multiple-layer isoG pentaplex
beyond the sandwich structure reported so far.

As shown in the crystal structure, the isoG\(^{20}\) assembly was
formed with the deoxy-ribose ligand. The Cs\(^{+}\) cations bind to
isoG\(_{\text{5}}\) between layers through ion–dipole interactions, similar to
the previously reported isoG\(_{\text{10}}\) from isoG\(^{\text{4}}\) monomer; (D) treating isoG\(^{\text{4}}\) with CsBPh\(_{4}\) for 30 min; (E) treating
isoG\(^{\text{4}}\) with CsBPh\(_{4}\) for 12 h.

The distances between isoG\(_{\text{5}}\) layers in \([\left[4\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BARF}^{-}]_3}\) are
between 3.20 and 3.37 \(\text{˚}\), almost identical to the distance
observed in the isoG\(^{\text{1}}\) complex, suggesting similar ion–dipole
interactions in this deoxy-isoG complexes. The biggest
difference between these two types of isoG-assemblies is the
ribose conformation. In \([\left[1\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BPh}_4^{-}]_3}\), the two isoG\(_{\text{5}}\) pentamers stack in a tail-to-tail configuration. With more flexible
deoxy ribose, isoG\(_{\text{5}}\) from isoG\(^{\text{4}}\) is less sterically hindered,
allowing ribose side chain interdigitation between layers. As
shown in complex \([\left[1\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BARF}^{-}]_3}\), the opposite H-bond orientation is
crucial for the formation of strong ion–dipole interactions (Cs\(^{+}\)
to O2). Similar stacking pattern was observed in \([\left[4\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BARF}^{-}]_3}\),
where the reversed H-bond orientation was observed between two adjacent layers, giving either a head-to-head or tail-
to-tail configuration throughout the structures. This result is
important since it revealed the possibility to further extend
the structure along the c-axis with this new deoxy-ribose isoG
for potential applications, such as the Cs\(^{+}\) ion-channel.

With the solid-state structure confirmed by X-ray crystallog-
raphy, we then focused on the self-assembly in the solution. As
discussed above, since this deoxy isoG-star could further extend
the assembly along the vertical direction, achieving a controllable
pentaplex is critical for potential future applications. To
study how deoxy-isoG 4 binds with Cs\(^{+}\) in solution, we first
dissolved \([\left[4\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BARF}^{-}]_3}\) in CDCl\(_{3}\) and monitored by \(^{1}\)H
NMR. As shown in Fig. 4A, dissolving \([\left[4\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BARF}^{-}]_3}\) in
CDCl\(_{3}\) gave a mixture of various H-bonded complexes (see full
spectra in Fig. S2†). To identify the structures of these
complexes, we conducted a series of titration experiments based
on \(^{1}\)H NMR.

First, unlike the BPh\(_{4}^{-}\) anion (Fig. 2E), treating isoG 4 with excess of CsBARF gives the formation of a complex with 5 : 1 ratio between isoG 4 and Cs\(^{+}\) (Fig. 4B). Notably, treating isoG 4 with limited amount of Cs\(^{+}\) could form \([\left[4\right]_{\text{isoG}}\text{Cs}\text{][\text{BARF}^{-}]_3}\) (Fig. 4D).
Interestingly, titration experiments (addition of isoG 4 into the
5 : 1 complex) revealed the formation a new complex with three
sets of NMR signals in 1 : 1 : 1 ratio (complex A, Fig. 4C). To

Fig. 2 \(^{1}\)H NMR spectra showing formation of isoG monomers and complexes in CDCl\(_{3}\) (A) isoG 1 monomer; (B) \([\left[1\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BPh}_4^{-}]_3}\); (C) isoG 4 monomer; (D) treating isoG 4 with CsBPh\(_{4}\) for 30 min; (E) treating
isoG 4 with CsBPh\(_{4}\) for 12 h.

Fig. 3 X-ray single crystal structure of the deoxy isoG\(^{20}\). (A) modes of organization in the formation of the complex; (B) side view of \([\left[4\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BARF}^{-}]_3}\); (C) ion channel in \([\left[4\right]_{\text{isoG}}\text{Cs}_{3}\text{][\text{BARF}^{-}]_3]\); (D) outer and inner layer.
elucidate the composition of the stable 5 : 1 complex as ([4]2Cs)‡ or ([4]0Cs2)‡, diffusion ordered spectroscopy (DOSY) experiments were performed (see the ESI†), which was firstly reported by Cohen.12 The diffusion coefficients of ribose H1′ were determined as 2.87 ± 0.02 × 10⁻¹⁰ m² s⁻¹ (5 : 1 complex) and 3.37 ± 0.01 × 10⁻¹⁰ m² s⁻¹ (10 : 1 complex) (Fig. S7 and S8†). These results strongly suggested the formation of ([4]0Cs2)‡ instead of ([4]0Cs)‡. To further confirm this composition analysis, direct comparisons between 5 : 1 and 10 : 1 complexes formed from deoxy isoG 4 and isopropylidene isoG 1 were performed. The diffusion coefficients of these four complexes are summarized in Table S1.† The diffusion coefficient values of H1′ were observed for the 5 : 1 complex from isoG1 as 3.11 ± 0.06 × 10⁻¹⁰ m² s⁻¹ and the 10 : 1 complex as 3.33 ± 0.09 × 10⁻¹⁰ m² s⁻¹. These results strongly suggested the formation of ([1]10Cs2)‡ with the BARF⁻ anion in CDCl₃ solution, which further confirmed the formation of ([4]0Cs2)‡ as described above.

It has been previously reported that treating isoG 1 with excess CsBPh₄ gave [1]10Cs⁺ (Fig. 5B). A similar decamer assembly was formed with the PF₆⁻ anion (Fig. 5C). As described above, switching the anion to BARF⁻ gives the formation of a new set of signals (Fig. 5A), suggesting the special anion effect of BARF⁻. Comparing 1H NMR spectra of ([4]10-Cs2)‡[BARF⁻]₂ with free BARF⁻ solution (cryptand complex, Fig. 5G), different chemical shifts were observed, indicating the interaction of the BARF⁻ anion toward the pentaplex in solution (Fig. 5H), which was further confirmed by the diffusion coefficients of BARF⁻ (see Fig. S7 and S8†).

With the ability to control the formation of singly charged decamer and doubly charged decamer, we turned our attention to determine the structure of complex A. Considering that the solid-state stable ([4]10Cs2)‡[BARF⁻]₂ was observed from evaporation of CHCl₃ and CH₃CN mix solvents, we hypothesized that the polarity difference of solvents could be crucial for the formation of complex A. Notably, the Meijer group reported the G-quadruplex self-assembly controlled by the coulombic interaction. The authors suggested that the ion pair separation could be regulated through tuning solvent polarity.13 Inspired by this work, polar solvent (like CH₃CN) would cause the dissociation of the ion-pair between isoG-star and the BARF⁻ anion. Therefore, tuning the polarity of solvents might be critical to reach the optimal conditions for the formation of complex A. Solvent titration was performed using a mixture of CDCl₃ and CD₃CN. The self-assembly process was monitored by 1H NMR (Fig. 6).

In less polar CDCl₃, doubly charged decamer was formed (Fig. 6A). Addition of CD₃CN caused increase in polarity with gradual formation of complex A. Finally, in the 1 : 1 mixture of CDCl₃ and CD₃CN, complex A was achieved as the dominant structure. 1H NMR spectra integration confirmed the proposed ([4]10Cs2)‡[BARF⁻]₂, consistent with the observed three set of signals (Fig. 6G). In addition, this structure was further confirmed using ESI-MS, giving the doubly charged peak with
m/z at 3850.8621 (Fig. S6‡). These experiments confirmed that the more polar solvents which can separate ion pairs favor the formation of larger assemblies. Notably, conducting the similar reaction with isoG 1, gave discrete isoG10 decamer in CD3CN with no formation of multiple-layer pentaplex, which highlighted the versatile coordination ability of deoxy-isoG 4 in the formation of pentaplex and critical solvent effect in the isoG self-assembly process.

Conclusions

In conclusion, in influence of the purine C8 position and ribose functionalization in the isoG derivative self-assembly with the Cs+ cation was explored. Although C8 substituents prevent N3 from being accessible for H-bonding formation in the isoG-star, we successfully discovered that 2′-deoxy isoG with no O2′ could serve as a new building block to achieve controllable supramolecular assemblies. This finding suggests that O2′-···N6-H8 H-bonding is not necessary for the stable isoG-star formation. Based on the coordination with the BARF+ anion and solvent mixtures, complexes such as isoG10, isoG15, and isoG20 were achieved with structures characterized by 1H NMR, diffusion NMR, ESI-MS and X-ray crystallography. This study not only revealed the important anion and solvent effects in the formation of various isoG self-assemblies, but also paved the crucial foundation for the isoG-star as a new scaffold in molecular architecture for future applications.

Author contributions

X. S. conceived the original idea and research direction and supervised the research. X. S. and M. L. designed the study. M. L. carried out the material synthesis and characterization; M. L. and Y. H. analyzed the results and data; C. S. and W. L. performed the X-ray measurements and analyzed the data; Y. Y. and X. L. contributed to the ESI-MS; I. G. performed the DOSY experiments. All authors discussed and commented on the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We are grateful to the NSF (CHE-1665122) and NIH (1R01GM120240-01) for financial support. This work was supported in part by the University of South Florida Interdisciplinary NMR Facility and the Chemical Purification, Analysis, and Screening (CPAS) Core Facility.

Notes and references


