CrystEngComm

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *CrystEngComm*, 2024, **26**, 2775

Sugar–nucleobase hydrogen bonding in cytidine 5′-monophosphate nucleotide-cadmium coordination complexes†

Yaqoot Khan, [a] Ismail Ismail, [b] Hongwei Ma,[c] Zhongkui Li[a] and Hui Li *[a]

Received 19th February 2024,
Accepted 23rd April 2024

DOI: 10.1039/d4ce00150h

rsc.li/crystengcomm

Hydrogen bonds are the fundamental factors stabilizing DNA and RNA macromolecules. Based on their base-pair sequences, DNA and RNA perform various biological functions. A key feature of these sequences is their linkage *via* hydrogen bonds. The hydrogen bonding between sugars and nucleobases in RNA sequences is one of the major reasons behind several mutagenic disorders that can cause numerous genomic instabilities, various genetic diseases, and RNA rearrangement problems. Chemists explore hydrogen bonding stability in nucleic acids, which is crucial for understanding molecular-level differentiations with potential applications in the early diagnosis of hereditary diseases. In this work, five types of coordination polymers of CMP and dCMP, {[Cd(CMP)(bpa)(H_2O)$_3$]·2H_2O}*n* (**1**), {[Cd(dCMP)$_2$(bpa)(H_2O)$_2$]·4H_2O}*n* (**2**), {[Cd(azpy)(H_2O)$_4$](CMP)·3H_2O}*n* (**3**), {[Cd(dCMP)$_2$(azpy)(H_2O)$_2$]·4H_2O}*n* (**4**), and {[Cd(CMP)(bpe)(H_2O)$_3$]·2H_2O}*n* (**5**) (azpy = 4,4′-azopyridine, bpa = 1,2-bis(4-pyridyl)ethane, and bpe = 1,2-bis(4-pyridyl)ethylene), were designed and studied. All complexes were fully characterized by employing the single-crystal X-ray diffraction method. Complex **1** is a 2D coordination polymer, whereas complexes **2–5** are 1D coordination polymers. Significantly, a novel sugar–nucleobase hydrogen bonding interaction was discovered in complexes **1**, **3**, and **5** for the first time, thus introducing a new supramolecular interaction that can be used in self-assembly and molecular recognition. The chirality in the supramolecular assemblies of the five complexes was comprehensively analyzed using single-crystal and solid-state CD spectra.

Introduction

Nucleotides are the fundamental units of nucleic acid, which are categorised into ribonucleotides and deoxynucleotides based on their pentose group. They are the essential components of DNA and RNA and are composed of pyrimidine or purine bases, ribose or deoxyribose, and phosphoric acid. They play a crucial role in storing and transferring genetic information and synthesizing proteins. Moreover, nucleotides are extremely desirable ligands owing to their various chirality and metal binding sites. The most crucial characteristic of nucleic acids is their ability to self-assemble into well-ordered structures built on hydrogen bonds and stacking interactions between nucleobases, despite the charge repulsions between the strands.[1] Nucleotides classically form a double helix structure *via* Watson–Crick (guanine–cytosine and adenine–thymine) base-pairing, as predicted by Watson and Crick.[2] However, during metabolic processes, including RNA transcription and DNA replication, the DNA double helix is partially broken down into a single-stranded sequence. Nucleobases may flip into new helix positions as a result of unstable mismatches in nucleic acid sequences, which is a significant phenomenon observed when small-molecule ligands bind to RNA or DNA duplexes.[3] These recombination structures induce genomic instability, causing DNA strand breakage; rearrangements; and implications for human diseases, including genetic alterations.[4] Understanding unusual DNA duplex structures and mismatched pairing interactions in structural details is crucial for unraveling genetic mutation and developing targeted treatments for specific hereditary diseases.[5] In contrast to B-DNA, a typical right-handed double helix with Watson–Crick base-pairing, repetitive DNA sequences can adopt non-B structures, including Hoogsteen base-paring,[6] G-quadruplex,[7] left-handed Z-DNA,[8] triplex (H-DNA),[9] tetraplex,[10] cruciform,[11] hairpin,[12] i-motif (cytosine–cytosine base-pair

[a] *Key Laboratory of Cluster Science of Ministry of Education, School of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing 100081, P. R. China. E-mail: lihui@bit.edu.cn; Tel: +86 10 68912667 602*

[b] *Department of Chemistry and Center for Diagnostics and Therapeutics, Georgia State University, Atlanta, Georgia 30303, USA*

[c] *Analysis & Testing Center of Beijing Institute of Technology, Beijing 102488, P. R. China*

† Electronic supplementary information (ESI) available. CCDC 2266320–2266323 and 2296547. For ESI and crystallographic data in CIF or other electronic format see DOI: **https://doi.org/10.1039/d4ce00150h**

mismatch),[13] and A-motif (adenine–adenine base-pair mismatch).[14]

Metal–nucleotide coordination complexes involve significant (M–L) supramolecular interactions. Such complexes serve as valuable biomimetic techniques, aiding the comprehension of metal ion-nucleic acid interactions.[15] Exploring new patterns of supramolecular interactions is a long-term goal for chemists, benefiting molecular recognition, self-assembly, peptide folding, and host-guest chemistry.[16,17] We studied nucleotide supramolecular assemblies using X-ray diffraction, thereby gaining expertise in the tunable synthesis of desired structures.[18] This innovative approach provides precise structural insights into sugar–nucleobase interactions and creates a suitable microenvironment for hosting hydrogen bonds, laying the groundwork for designing functional supramolecular assemblies and coordination complexes. Additionally, they are versatile building blocks for diverse complexes, MOFs, and coordination polymers with functions such as biosensing,[19] luminescence,[20] chirality,[21] and gas and solvent storage. The X-ray analysis of a 3D copper(II)-cytidine coordination polymer shows the use of cytidine nucleoside as a building block for polymer assembly.[22–26] In nucleotide chemistry, the sugar–nucleobase hydrogen bonding is a special supramolecular interaction, and achieving a well-defined structure for this type of hydrogen bonding is challenging but crucial for understanding genomic DNA. To the best of our knowledge, there is no literature report on single crystals relative to sugar–nucleobase hydrogen bonding in nucleotide coordination polymers.

In this work, we systematically studied the construction and supramolecular assemblies of unnatural sugar–nucleobase hydrogen bonding interaction and unnatural base-pair i-motif, and five Cd(II) coordination polymers with cytidine 5′-monophosphate (CMP) and 2′-deoxycytidine-5′-monophosphate (dCMP) were obtained, which are $\{[Cd(CMP)(bpa)(H_2O)_3]\cdot 2H_2O\}n$ (**1**), $\{[Cd(dCMP)_2(bpa)(H_2O)_2]\cdot 4H_2O\}n$ (**2**), $\{[Cd(azpy)(H_2O)_4](CMP)\cdot 3H_2O\}n$ (**3**), $\{[Cd(dCMP)_2(azpy)(H_2O)_2]\cdot 4H_2O\}n$ (**4**), and $\{[Cd(CMP)(bpe)(H_2O)_3]\cdot 2H_2O\}n$ (**5**), (azpy = 4,4′-azopyridine, bpa = 1,2-bis(4-pyridyl)ethane, and bpe = 1,2-bis(4-pyridyl)ethylene) (Scheme 1), respectively. In the supramolecular assembly, the function of the pentose of the hydroxyl group was studied. Particularly, solid-state and solution-state CD spectra were used to investigate the chirality of the crystal formation. Meanwhile, bipyridyl bridging ligands played a crucial role in designing specific molecular and supramolecular structures.

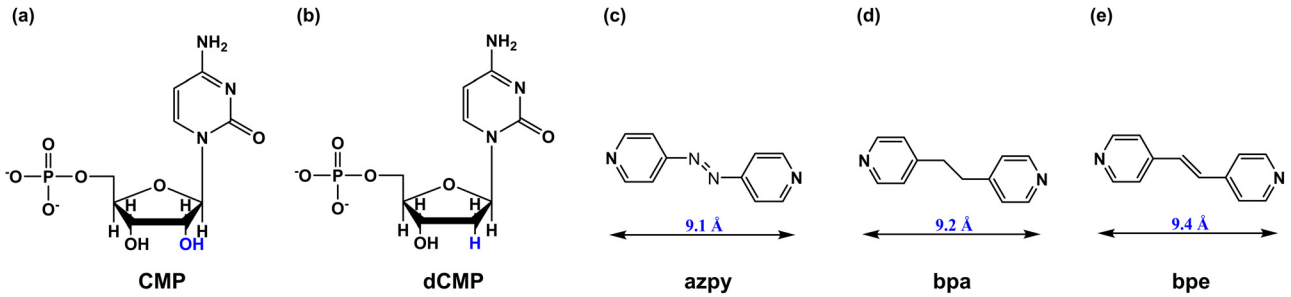
They were employed to fine-tune non-covalent interactions, adjust cytosine base orientation, and inhibit the non-enzymatic hydrolysis of phosphate groups mediated by metal ions. Oxynucleotide and deoxynucleotide could effectively compete with each other to coordinate with the metal ion. Thus, it might be possible to improve the nucleotide complex crystallization.

Experimental section

Materials and measurements

All chemicals and organic solvents were obtained from commercial sources and used without further purification. 1,2-Bis(4-pyridyl)ethene, 1,2-bis(4-pyridyl)ethylene, and 4,4′-azopyridine were purchased from Adamas, $Cd(NO_3)_2\cdot 4H_2O$ from Aladdin and cytidine 5′-monophosphate disodium and deoxycytidine 5′-monophosphate disodium salts from Alfa Aesar. A Thermo IS5 FT-IR spectrometer was used to record the FT-IR spectra (KBr pellets) in the range from 400–4000 cm^{-1}. A Bruker D8 Advance X-ray diffractometer at a scanning rate of 5° min^{-1} (2θ angle measurement range: 5–50°) was used for acquiring powder X-ray diffraction (PXRD) patterns. An X-ray diffractometer was used to examine a single crystal with graphite-monochromatized Mo Kα radiation (λ = 0.71073 Å). An EA3000 elemental analyzer was used to obtain the elemental composition (C, H, and N). A PHB-1PH meter was used to determine the pH level of the solution. Absorption spectra in the UV-visible range were recorded using a Persee TU-1950 spectrophotometer. Circular dichroism spectra of solids were recorded at room temperature using a JASCO J-1500 spectropolarimeter with KBr pellets (1 : 200) under a stable nitrogen stream purge. The spectra of circular dichroism in solutions were recorded at a concentration of 0.025 mmol l^{-1}. Thermogravimetric analyses (TGA) were performed using a DTG-60H thermal analyzer for complexes 1–4 and a NETZSCH STA 449 F5 Jupiter thermal analyzer for complex 5 in a nitrogen atmosphere at 50 °C and 700 °C at a heating rate of 10 °C min^{-1}.

$\{[Cd(CMP)(bpa)(H_2O)_3]\cdot 2H_2O\}n$ (1). First, 4 ml aqueous solution of Na_2CMP (18 mg, 0.05 mmol) was mixed with a 4 ml aqueous solution of $Cd(NO_3)_2\cdot 4H_2O$ (15 mg, 0.05 mmol)



Scheme 1 Structural formulas of (a) CMP, (b) dCMP, (c) azpy (d) bpa and (e) bpe ligands.

and stirred for 15 minutes. Subsequently, 9 mg (0.05 mmol) of the auxiliary ligand bpa in 4 ml ethanol was added to the mixture. After 30 minutes, a white precipitate was obtained. Then 1 M HNO_3 was added until the solution became transparent, reaching a pH of around 5.7. The mixture was stirred at room temperature for an hour before filtration. After a week of room-temperature evaporation, a white needle-like single crystal suitable for single-crystal X-ray diffraction was obtained. The yield was 53%. Anal. calc. (%) for $C_{21}H_{30}CdN_5O_{12}P$ (Mr. 687.88): C, 36.67; H, 4.40; N, 10.18. Found (%): C, 36.43; H, 4.01; N, 10.04. Selected IR (KBr pellet, cm^{-1}): 3414vs, 2899w, 1647vs, 1611vs, 1529m, 1499s, 1433m, 1382m, 1329w, 1292m, 1256w, 1226m, 1111s, 1078vs, 1059s, 1016m, 971s, 906w, 875w, 814m, 787m, 753m, 632m, 597m, 548m.

{[Cd(dCMP)$_2$(bpa)(H$_2$O)$_2$]·5H$_2$O}*n* (2). Complex **2** was synthesized similarly to **1**, using 2′-deoxycytidine-5′-monophosphate (dCMP, 15 mg, 0.05 mmol) instead of CMP, and auxiliary ligand bpa (9 mg, 0.05 mmol) in an aqueous solution (pH = 5.6). After one week, a white bulk single crystal appropriate for X-ray diffraction was obtained by evaporation at room temperature. The yield was 58%. Anal. calc. (%) for $C_{30}H_{38}CdN_8O_{21}P_2$ (Mr 1028.12): C, 35.08; H, 4.32; N, 10.91. Found (%): C, 34.82; H, 3.94; N, 10.46. Selected IR (KBr pellet, cm^{-1}): 3409vs, 1722s, 1654vs, 1611s, 1559m, 1493m, 1424m, 1287m, 1197w, 1087vs, 1016m, 972s, 876w, 832m, 782m, 615m, 549m.

{[Cd(azpy)(H$_2$O)$_4$](CMP)·3H$_2$O}*n* (3). Complex **3** was synthesized similarly to **1**, using auxiliary ligand azpy (9 mg, 0.05 mmol) instead of bpa; 30 min later, a yellow color solution was obtained (pH = 5.3). Orange block crystals suitable for single-crystal X-ray diffraction analysis were obtained by evaporation at room temperature after one week. The yield was 51%. Anal. calc. (%) for $C_{19}H_{34}CdN_7O_{15}P$ (Mr 743.90): C, 30.68; H, 4.61; N, 13.18. Found (%): C, 30.23; H, 4.05; N, 12.92. Selected IR (KBr pellet, cm^{-1}): 3454vs, 1645s, 1598m, 1533w, 1492s, 1418m, 1288m, 1108s, 1063vs, 1013w, 972s, 845w, 803s, 786m, 567w, 541s.

{[Cd(dCMP)$_2$(azpy)(H$_2$O)$_2$]·4H$_2$O}*n* (4). Complex **4** was synthesized similarly to **3**, using 2′-deoxycytidine-5′ monophosphate (dCMP, 15 mg, 0.05 mmol) instead of CMP (pH = 4.6). After one week, orange block single crystals suitable for single-crystal X-ray diffraction analysis were obtained by evaporation at room temperature. The yield was 55%. Anal. calc. (%) for $C_{28}H_{46}CdN_{10}O_{20}P_2$ (Mr 1017.09): C, 33.07; H, 4.56; N, 13.77. Found (%): C, 33.05; H, 4.10; N, 13.20. Selected IR (KBr pellet, cm^{-1}): 3374s, 3088m, 2930m, 1722vs, 1657s, 1599m, 1528w, 1491s, 1418m, 1285m, 1200m, 1092s, 977m, 947w, 925m, 836w, 808s, 782w, 613w, 568m, 540w.

{[Cd(CMP)(bpe)(H$_2$O)$_3$]·2H$_2$O}*n* (5). Complex **5** was synthesized similarly to **1**, using auxiliary ligand bpe (9 mg, 0.05 mmol) instead of bpa; 30 min later, a yellow color solution was obtained (pH = 5.7). White block single crystals suitable for single-crystal X-ray diffraction analysis were obtained by evaporation at room temperature after one week. The yield was 51%. Anal. calc. (%) for $C_{21}H_{32}CdN_5O_{13}P$ (Mr 705.88): C, 30.68; H, 4.61; N, 13.18. Found (%): C, 30.31; H, 4.52; N, 13.13. Selected IR (KBr pellet, cm^{-1}): 3372vs, 1647m, 1604s, 1556w, 1494m, 1417m, 1383s, 1283m, 1249w, 1069s, 974s, 828s, 801m, 758w, 627w, 548s.

Crystallographic data collection and refinement

Coordination complexes with suitable dimensions were chosen for the analysis of single-crystal X-ray diffraction. Data were acquired at room temperature (296.15 K) using a Bruker DUO APEX II-CCD area diffractometer with graphite-monochromatized Mo Kα (λ = 0.71073 Å) radiation. The XSCAN software was used to identify peaks and determine the unit cell at 45 KV and 30 mA. In order to acquire diffraction data, the ω–2θ scanning mode was used, and an empirical absorption correction was implemented to all the scanned data. By employing Olex2,[27] the crystal structure was successfully solved using the direct methods of the SHELXT program.[28] Subsequently, full-matrix least-squares techniques were employed using the SHELXL program[29] for refinement of F^2. The non-hydrogen atoms in the complexes were obtained from various Fourier maps and refined anisotropically. The positions of all hydrogen atoms were determined using a combination of Fourier maps and geometric analysis and then refined isotropically. The refinement process for the five heavy-atom structures was carried out without factoring in the H-atoms, as there were no clear electron-density peaks observed in the various maps that would indicate an acceptable location for the water oxygen sites. Tables 1 and S1† provide the crystallographic data and CCDC numbers for complexes.

Results and discussion

Structure description

{[Cd(CMP)(bpa)(H$_2$O)$_3$]·2H$_2$O}*n* (1). The CMP ligands in complex **1** coordinated with Cd^{2+} ions *via* the phosphate terminal in the bridge mode, while bpa served as a bridging ligand making complex **1** a two-dimensional coordination polymer with the chiral space group $P2_1$. Each Cd^{2+} ion is six-coordinated in the asymmetric unit of **1** with two oxygen atoms of (O3, O5) from the phosphate group of CMP, two nitrogen atoms (N1, N2) from bpa bridging ligands, and another two oxygen atoms from coordinated water molecules (O1, O2) (Fig. 1a). The average Cd–O and Cd–N bond lengths are 2.32 Å and 2.30 Å, respectively. The distances between Cd^{2+} ions linked through bridging phosphate and bpa ligands are 5.918 (11) and 13.895 (31) Å, respectively (Fig. 1b and S5†). There is a π–π stacking interaction (3.757 Å) between the nucleobase of CMP and bpe ligands in the adjacent layer (Fig. 1c and S2†). Moreover, there are hydrogen bonding interactions between the layers. Here, sugar–nucleobase hydrogen bonding has formed between the imine nitrogen atom of the base and the sugar-ring hydroxyl group to a carbonyl oxygen atom (O9–H9A···O8, 1.81 Å, 2.62 Å, 171°; O10–H10A···N4, 2.04 Å, 2.78 Å, 148°) (Fig. 1c and d, S1 and S4†).

Table 1 Crystallographic data for complexes **1**–**5**

Complex	**1**	**2**	**3**	**4**	**5**
CCDC no.	2266320	2266321	2266322	2266323	2296547
Chemical formula	$C_{21}H_{30}CdN_5O_{12}P$	$C_{30}H_{43}CdN_8O_{21}P_2$	$C_{19}H_{34}CdN_7O_{15}P$	$C_{28}H_{46}CdN_{10}O_{20}P_2$	$C_{21}H_{32}CdN_5O_{13}P$
Formula weight (g mol^{-1})	687.87	1026.06	743.90	1017.09	705.88
Colour	Colourless	Colourless	Orange	Orange	Colourless
T (K)	296(2)	296.15	296.15	296.15	296.15
Crystal system	Monoclinic	Orthorhombic	Monoclinic	Orthorhombic	Monoclinic
Space group	$P2_1$	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P2_1$
a (Å)	8.6015(17)	13.813(5)	9.3779(13)	13.5898(15)	8.310(5)
b (Å)	11.121(2)	14.472(6)	12.2117(18)	14.6787(17)	11.901(7)
c (Å)	13.895(3)	20.721(7)	13.6250(19)	21.008(2)	14.109(8)
α (°)	90.00	90.00	90	90	90
β (°)	90.292(7)	90.00	106.071(4)	90	92.58(2)
γ (°)	90.00	90.00	90	90	90
V (Å^3)	1329.1(4)	4142(3)	1499.4(4)	4190.8(8)	1393.9(14)
Z	2	4	2	4	2
ρ_{cald} (g cm^{-3})	1.719	1.645	1.648	1.612	1.682
μ (Mo Kα) (mm^{-1})	0.954	0.696	0.861	0.687	0.914
F(000)	700	2108	760	2088	720
θ range/°	4.692–62.49	3.432–51.386	3.11–63.394	3.57–63.39	2.89–62.822
Range of h, k, l	$-12 < h < 12$	$-16 < h < 16$	$-13 < h < 13$	$-19 < h < 18$	$-11 < h < 12$
	$-16 < k < 15$	$-7 < k < 17$	$-17 < k < 18$	$-20 < k < 21$	$-16 < k < 17$
	$-19 < l < 20$	$-21 < l < 20$	$-19 < l < 19$	$-30 < l < 30$	$-19 < l < 20$
Reflections collected	18 337	13 987	20 743	58 229	18 764
Independent reflections	7880	7089	8886	13 047	8061
Unique data (R_{int})	0.0584	0.0535	0.0221	0.0562	0.0194
$R_1{}^a$/w$R_2{}^b$ [$I > 2(I\sigma)$]	0.0493/0.0785	0.0509/0.1203	0.0281/0.0642	0.0437/0.0970	0.0183/0.0442
$R_1{}^a$/w$R_2{}^b$ [all data]	0.0775/0.0872	0.0603/0.1263	0.0330/0.0665	0.0593/0.1029	0.0194/0.0448
GOF	1.012	1.058	1.016	1.022	1.038
H atom treatment	Constr	Mixed	Constr	Constr	Mixed
Residuals (e Å^{-3})	0.61, −0.76	1.88/−0.64	1.07/−0.27	1.63/−0.47	0.40/−0.35
Flack parameter	0.003(18)	0.003(17)	0.002(18)	−0.04(2)	0.000(12)

^{a} $R_1 = \sum||F_o| - |F_c||/\sum|F_o|$. ^{b} $wR_2 = [\sum w(F_o{}^2 - F_c{}^2)^2/\sum w(F_o{}^2)^2]^{1/2}$.

Another is a hydrogen bonding interaction between phosphate and pyrimidine base (N5–H5C···O4, 1.99 Å, 2.83 Å, 166°) (Fig. 1c and Scheme 2 (blue dotted line) and Fig. S1 and S4†). Therefore, the layers of 2D coordination polymer of complex **1** are linked by multi-hydrogen bonding and π–π stacking interaction along the c axis. The X-ray structure showed that within a coordination polymer, each ribose formed hydrogen bonds with its neighbor nucleotide base. Moreover, the existence of the 2′-OH hydroxyl group makes RNAs capable of particular interactions engaging the ribose part of the polynucleotide structure. Non-Watson–Crick (non-

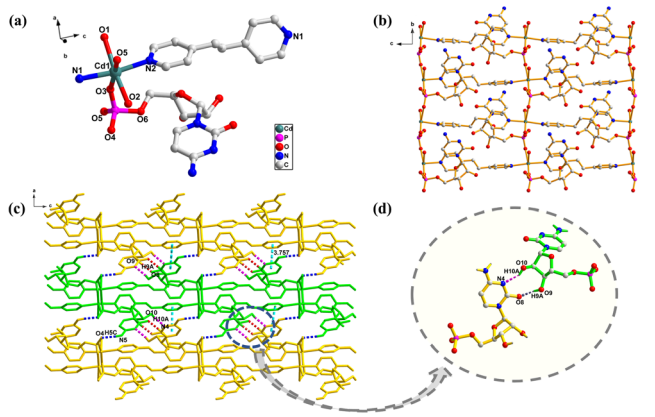
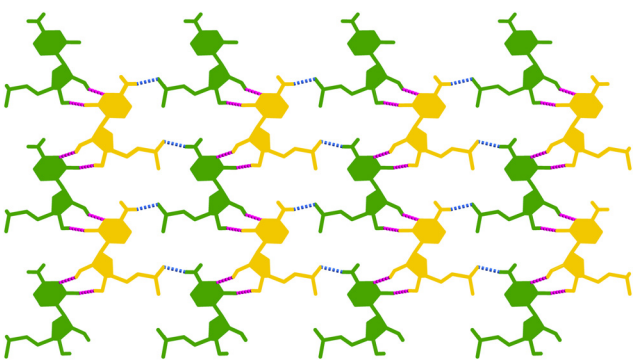


Fig. 1 (a) Coordination environment of Cd^{2+} in complex **1**; hydrogen atoms are omitted for clarity. (b) Two-dimensional coordination network of the complex viewed down from the *a* axis; uncoordinated water molecules and part of hydrogen atoms are omitted for clarity. (c) Three-dimensional structure based on the π–π stacking interaction viewed down from the *b* axis. (d) Details of the formation of sugar–nucleobase hydrogen bonding interaction (pink and red dotted-line, O9–H9A···O8, 1.81 Å, 2.62 Å, 171°; O10–H10A···N4, 2.04 Å, 2.78 Å, 148°) in complex **1**.



Scheme 2 Schematic of the formation of a sugar–nucleobase H-bonding framework in complexes **1**, **3** and **5**.

WC) base pairs play a significant role in influencing the stability and folding of RNA architectures.[30] The formation of hydrogen bonds between the sugar moiety and nucleobase arises from mutations and structural rearrangements within DNA/RNA configurations. Contrary to the hydrogen bonds between nucleobase-nucleobase in the G-quartets and DNA, sugar–nucleobase hydrogen bonds can only occur in a homochiral complementary pair,[30–32] pioneering the conveyance of stereochemical information from one nucleotide to its adjacent counterpart, which is a distinctive contribution highlighted in this work.

{[Cd(dCMP)$_2$(bpa)(H$_2$O)$_2$]·4H$_2$O}*n* (2). In comparison with complex **1**, the dCMP ligands used in complex **2** replaced CMP ligand, which is a one-dimensional coordination polymer with the chiral space group $P2_12_12_1$. Each Cd^{2+} ion is also six-coordinated with two dCMP (O6, O8), two bpa bridging ligands (N1, N8) and two coordinated water molecules (O11, O14) (Fig. 2a). The ranges of Cd–O and Cd–N bond lengths are 2.239–2.331 Å and 2.272–2.282 Å, respectively. The Cd···Cd distance bridged by bpa ligands in 1D chains is 13.8130(51), while the Cd···Cd distance between neighboring chains is 7.3052(31) Å (Fig. 2b and S9†). Clearly, no sugar–nucleobase hydrogen bonds exist in complex **2** since one OH– group is absent in dCMP. However, complex **2** exhibits an i-motif structural feature, which has been observed and reported by our group previously.[13] Complex **2** has confirmed that the molecular configuration of the auxiliary ligand bpa is suitable for the construction of the i-motif, which can alter the orientation of the cytosine base for the development of the i-motif *via* intra- and inter-chain π–π-stacking interactions between the cytosine base and pyridine ring. The i-motif structural feature is very strong H-bonding (N5–H5A···O3, 1.96 Å, 2.82 Å, 173°, N2–H2A···N3, 2.10 Å, 2.87 Å, 168°, N7–H7A···O7, 2.11 Å, 2.97 Å, 171°) (Fig. 2c and d) and coplanar with the dihedral angle of the two cytosine bases of 5.0(1)° (Fig. S7†). Nucleotides are arranged along the one-dimensional chain (Fig. 2b and S9†) based on the π–π stacking with distances of 3.695 Å and 3.698 Å within the chain and of 3.781 Å and 3.969 Å between the chains (Fig. 2c and S7†). The i-motif exists between the adjacent chains. In addition, the donors of the pyrimidine base, as well as the phosphate oxygen atoms, create hydrogen bonds to form the 1D supramolecular assembly, (N5–H5B···O4, 2.12 Å, 2.97 Å, 168°, N7–H7B···O12, 2.15 Å, 3.01 Å, 173°), (Fig. 2c and d and S7†). A homochiral supramolecular framework would be produced as a result of the coordination and transmit the chirality of the dCMP ligands to the 3D supramolecular assembly (Fig. S8a†).

{[Cd(azpy)(H$_2$O)$_4$](CMP)·2H$_2$O}*n* (3). Relative to the complex **1**, complex **3** has been designed by 4,4′-azopyridine (azpy) replacing bpa (1,2-bis(4-pyridyl)ethane). Based on our previous research, the auxiliary ligands play an important role in tuning the structure according to their Lewis basicity and the flexibility of structural conformation. Indeed, complex **3** is different from the complex **1**, in which CMP did not coordinate with Cd(II) metal ions. We classified this type of complex as a supramolecular coordination polymer indicating the nucleotides as uncoordinated species. Each metal Cd^{2+} ion is hexa-coordinated with four coordinating water molecules and two azpy ligands in the axial position. CMP linked to a one-dimensional chain through a strong hydrogen bond formed by the coordinated water and the phosphate group of CMP (O1–H1A···O5, 0.85 Å, 1.81 Å, 171°; O2–H2B···O6, 0.85 Å, 1.88 Å, 159°) (Fig. 3b). Moreover, the sugar–nucleobase hydrogen bonds were found in complex **3**, which are stronger and more symmetrical than those in complex **1** (O11–H11···N6, 0.82 Å, 2.05 Å, 168°; O10–H10A···O12, 0.82 Å, 1.78 Å, 164°) (Fig. 3c and d). The freedom of the CMP in complex **3** is larger than that in complex **1**, which is beneficial for forming the symmetrical sugar–nucleobase hydrogen bonds. The novel hydrogen bonds play an important role to link the one-dimensional

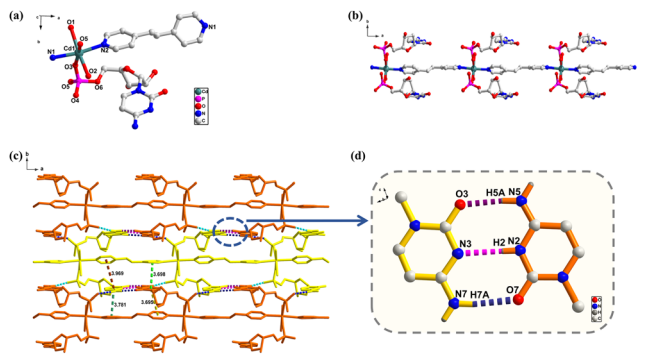


Fig. 2 (a) Coordination environment of Cd^{2+} in complex **2**; hydrogen atoms are omitted for clarity. (b) One-dimensional structure of uncoordinated water molecules; part of hydrogen atoms are omitted for clarity. (c) Two-dimensional supramolecular assembly based on π–π-stacking interactions and complementary H-bonding between nucleobases. (d) The i-motif in complex **2** (N5–H5A···O3, 1.96 Å, 2.82 Å, 173°, N2–H2A···N3, 2.10 Å, 2.87 Å, 168°, N7–H7A···O7, 2.11 Å, 2.97 Å, 171°).

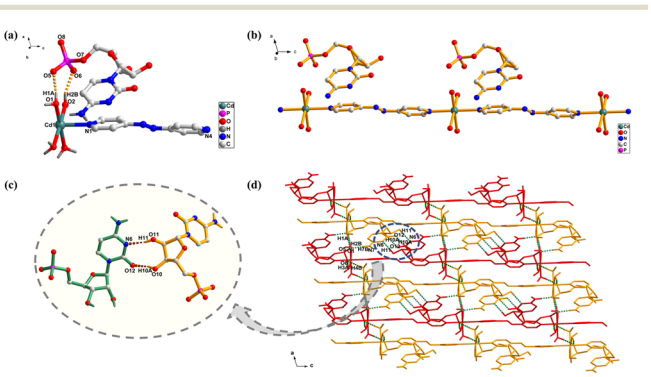


Fig. 3 (a) Coordination environment of Cd^{2+} in complex **3**; hydrogen atoms are omitted for clarity. (b) One-dimensional coordination network of the complex viewed down from the *b* axis; uncoordinated water molecules and part of hydrogen atoms are omitted for clarity. (c) Sugar–nucleobase hydrogen bonds (O11–H11···N6, 0.82 Å, 2.05 Å, 168°; O10–H10A···O12, 0.82 Å, 1.78 Å, 164°). (d) Two-dimensional structure based on the interaction viewed down from the b axis.

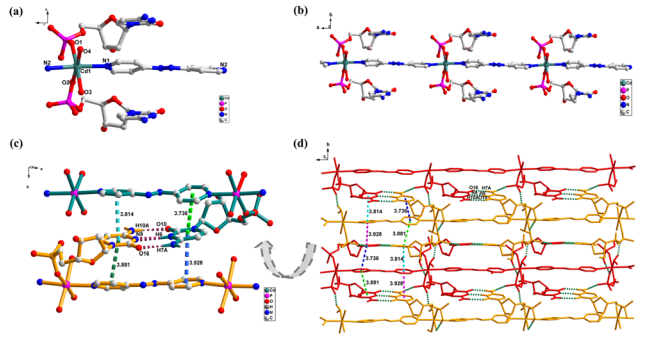


Fig. 4 (a) Coordination environment of Cd^{2+} in complex **4**. (b) One-dimensional structure of uncoordinated water molecules; part of hydrogen atoms are omitted for clarity. (c) The i-motif in complex **4** with strong and symmetrical hydrogen bonding interactions (N10–H10A···O10, 0.86 Å, 2.07 Å, 173°, N6–H6···N9, 0.86 Å, 1.99 Å, 173°, N7–H7A···O16, 0.86 Å, 1.92 Å, 175°) (d) two-dimensional supramolecular assembly based on π–π-stacking interactions and hydrogen bonds.

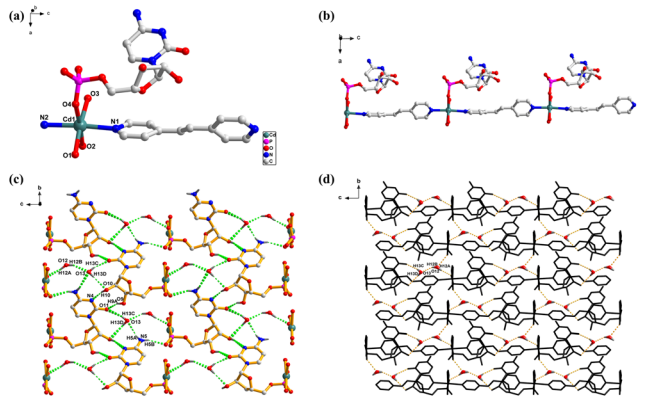


Fig. 5 (a) Coordination environment of Cd^{2+} in complex **5**. (b) One-dimensional structure of uncoordinated water molecules; part of hydrogen atoms are omitted for clarity. (c) Two-dimensional coordination network of complex **5** formed by hydrogen bonding between water molecules and sugar–nucleobases viewed down from the *a* axis; the formation of sugar–nucleobase hydrogen bonding with interactions of uncoordinated water molecules (green dotted-line). (d) Details of the channels formed in the 3D polymeric structure of **5** that are filled with water molecules.

coordination chains to the two-dimensional structure (Fig. 3d and S13†).

{[Cd(dCMP)$_2$(azpy)(H$_2$O)$_2$]·4H$_2$O}*n* (4). For a systematic study, complex **4** was constructed. It is isostructural with complex **2** (Fig. 4a and b). The Cd···Cd spacing of the ancillary azpy bridge ligand is 13.5898(15) Å (Fig. 4c and d) and a little bit shorter than that in complex **2** with bpa, while the distance between the neighboring chains in this layer measured in Cd···Cd is 7.4150(9) Å and a little bit longer than that in complex **2**. As expected, i-motif existed in complex **4** with strong hydrogen bonding interactions (N10–H10A···O10, 0.86 Å, 2.07 Å, 173°, N6–H6···N9, 0.86 Å, 1.99 Å, 173°, N7–H7A···O16, 0.86 Å, 1.92 Å, 175°) (Fig. 4c and d and S16†). Compared with complex **2**, the i-motif in complex **4** is stronger and more symmetrical. The π–π stacking effect (3.736 Å, 3.881 Å) and hydrogen bonds of the phosphate oxygen to the imine nitrogen atom of the base (N7–H7B···O12, 0.86 Å, 2.07 Å, 170°) and coordinated water molecules (O4–H4A···O11, 0.96 Å, 1.77 Å, 150°) are responsible for the formation of two-dimensional and three-dimensional structures (Fig. 4d).

{[Cd(CMP)(bpe)(H$_2$O)$_3$]·2H$_2$O}*n* (5). In comparison with complex **1**, complex **5** was designed by 1,2-bis(4-pyridyl) ethylene (bpe) replacing bpa (1,2-bis(4-pyridyl)ethane). Complex **5** is a 1D coordination polymer with the CMP terminal coordination to the central metal ion. It is different to complex **1** and **3**. However, we are excited to find the sugar–nucleobase hydrogen bonds in complex **5** again. In the asymmetric unit of **5**, each Cd^{2+} ion is six-coordinated with one oxygen atom (O4) from the phosphate group of CMP, two nitrogen atoms (N1 and N2) from bpe bridging ligands, and other three oxygen atoms (O1, O2, and O3) from coordinated water molecules. The average Cd–O and Cd–N bond lengths are 2.32 Å and 2.30 Å, respectively. The distances between Cd^{2+} ions linked through bridging phosphate and bpe ligands are 6.539 (45) and 14.109 (8) Å, respectively. The details of sugar–nucleobase hydrogen bonding are O10–H10A···N4 (2.08 Å, 2.88 Å, 166°) and O9–H9A···O11 (1.88 Å, 2.62 Å, 150°), which are a little bit weaker than that in complex **3**. In addition, there are uncoordinated water molecules forming hydrogen bonding with sugar ring-OH and the nucleobase (O13–H13C···O11, 2.16 Å, 2.92 Å, 155°; O13–H13D···O10, 2.31 Å, 3.06 Å, 147°; N5–H5A···O13, 2.29 Å, 3.12 Å, 160°; O12–H12B···O13, 2.00 Å, 2.80 Å, 150°) in complex **5** (Fig. 5c and S19†).

In summary, the availability of two OH hydroxyl groups in the sugar ring motif of CMP is favourable for the sugar–nucleobase hydrogen bonding formation. For dCMP, no hydrogen bond was found between the nucleobase and the sugar ring motif of dCMP. Furthermore, the complementary hydrogen bonds between nucleobases, that is an i-motif in this work, are competitive to sugar–nucleobase hydrogen bonding (Fig. 6). These findings contribute to our understanding of the nuanced interactions within nucleotide structures, holding implications for molecular design and recognition in biological systems.

Chirality and CD spectra

The chirality of coordination complexes has been studied and explained according to their crystal structures and crystallized solid CD spectra. Three different types of chirality exist in these complexes: i) inherent chirality of the nucleotide ligands, ii) supramolecular helical chirality, and iii) extended axial chirality (EAC) of the bridging ligand (bpa), (bpe) and (azpy).[18,33] In the solution-state CD spectra, the two negative peaks of CMP and dCMP were observed near 222 nm and 218 nm identified as the n–π* transition. The positive peak near 270 nm shows that CMP/dCMP is a D-ribonucleotide, and the UV/vis spectrum was found to be reliable (Fig. S27†).[34] It appears that the nucleotides are the

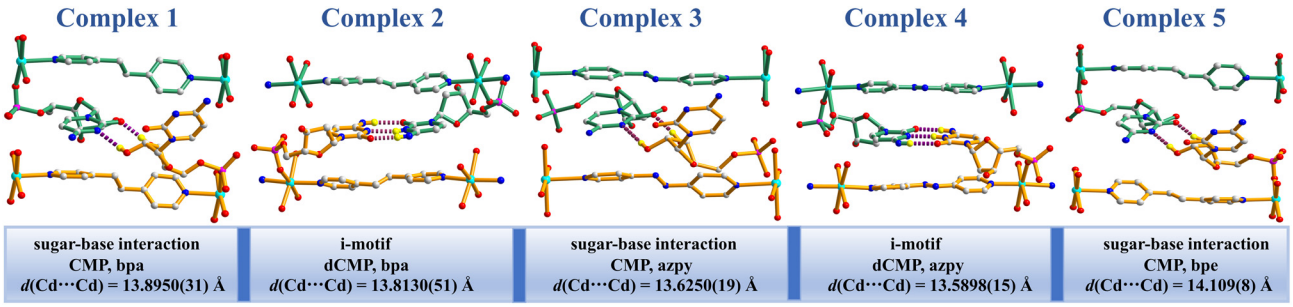


Fig. 6 Summary of the crystal structures presenting sugar–nucleobase hydrogen bonds and i-motif in complexes **1**–**5**.

main contributors to chirality in solutions based on the similarity of the CD spectra between CMP/dCMP and their complexes.

A signal emerged at around 206 nm in the CD spectrum of CMP due to the presence of an additional hydroxyl group during the n–σ* transition of the CMP molecule. The patterns of the CD spectra of the complexes **1**–**4** are similar to their nucleotide ligands. The main peaks of complexes **2** and **4** are red-shifted relative to the signals of nucleotide ligands, respectively. However, the peaks around 222 nm of complex **1** blue-shifted, which are almost the same as those for CMP of complex **3**, which confirmed the formation of sugar–nucleobase hydrogen bonds. It should be noted that the spectrum of complex **1** is different from that of complex **3**. It can be explained that the pre-organization of complex **1** in solutions is different from that of complex **3** according to the single-crystal structural analysis. In complex **2**, a positive peak observed near 244 nm was attributed to the n–π* transition and red-shifted (from 234 to 244 nm) compared to the dCMP ligand, which may be due to the existence of i-motif. A strong positive peak that appeared around 284 nm was linked to the π–π* transition and primarily due to the extended π–π stacking interactions within the crystal lattice. It should be noted that there is a small negative cotton signal around 310 nm for complex **4**, which indicates that there is a hydrogen bonding helix assembled in the solution.

In the crystallized solid-state CD spectra, the patterns of CMP, dCMP and the complexes **1**–**5** are obviously different (Fig. 7). There are two distinct broad peaks centered at 220 and 283 nm for CMP and three peaks centered at 235, 270, and 304 nm for dCMP, respectively.[35,36] The existence of a strong negative band at 304 nm indicates that the supramolecular helix is formed with the involvement of the dCMP ligand (Fig. 7 and S8†). Based on the above-mentioned crystal structural analysis, the CMP exhibits T-conformations, whereas dCMP exhibits E-conformations.[37,38] Hence, the positive and negative Cotton effect peaks between 220 and 235 nm can be attributed to the varying conformations of ribose or deoxyribose in CMP and dCMP.

For complex **1**, the supramolecular helix resulted in a weaker negative band at 312–320 nm compared to CMP (Fig. 7 and S3†). For complex **2**, a wide peak with positive Cotton effects in the 275–310 nm region was identified as the π–π* transition, and this peak was red-shifted compared to the dCMP ligand due to the strong π–π-stacking interaction present in complex **2**. The wider positive band near 311 nm can be attributed to the supramolecular helix in complex **2** (Fig. 7). In **3**, the negative CD band ranging from 218 to 208 nm appears to be blue-shifted, which can be attributed to the protonated CMP ligands in **3**, which obstructs the n–π* transition and causes the absorption to require a higher energy.[33,39] The opposite signals near 291 nm slightly red-shifted due to the strong π–π* transition (Fig. 7). Different from the complex **2**, the complex **4** shows a big alteration to others and its peak is opposite in the region of 200–375 nm with the dCMP ligand, and a new negative peak emerges at 227 nm. In **5**, the negative CD band ranging from 220 to 210 nm appears to be blue-shifted, which can be attributed to the n–π* transition. The positive signals near 265 nm blue-shifted owing to the strong π–π* transition. The negative band at 298–310 nm compared to CMP indicates that supramolecular helix formed by **5**. The CD signal at 227 nm originates from the excitation coupling of π–π* transitions in aromatic chromophores within the helix. This includes the intra- and intermolecular coupling of cytosine chromophores.

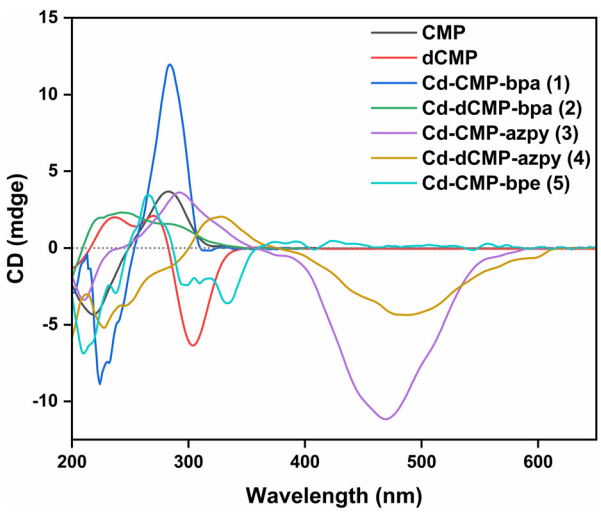


Fig. 7 Crystallized solid-state CD spectra of CMP, dCMP, and complexes **1**–**5** (KBr : [sample] = 200 : 1).

The negative signal near 340 nm, the positive signal near 330 nm, the negative cotton effect near 375 nm, the positive signal near 327 nm and the negative signal near 333 nm are related to the EAC of complexes **1**, **2**, **3**, **4** and **5** respectively. The EAC of azpy in complex **3** is *P*-chirality because all deoxynucleotide ligands surrounding the axis of azpy are arranged clockwise (Fig. S15 and S18†). The *P*-EAC of azpy should be the positive Cotton effect in the CD spectra. The symmetry of the ligands' electronic transition is altered by the inclusion of a supramolecular helix made of dCMP ligands (an extra chiral source).[40,41] Additionally, the chiral configuration of the EAC is certainly influenced by the axial chirality of the azpy ligand and the metal center's chirality, resulting in a definite *P* configuration and is probably reliant on the D-ribonucleotides.[40]

Importantly, a broad peak near 469 and 486 nm observed in the CD spectrum of **3** and **4** is assigned to the n–π* transition of the azpy ligand. The long-wavelength excitation that remains in accordance with the X-ray structure reveals a right-handed twist in the guest molecule (Fig. 7).[42] When complex formation occurs, absorption wavelengths in the CD spectra increase due to aligned nitrogen heterocycles allowing for electronic transitions to happen.

Complexes **1** and **5** presented *M*-EAC, while the other complexes **2**, **3**, and **4** are *P*-EAC, based on crystal structure analysis. The maximum torsional angle of bpa 62.3 and bpe 52.18 may indicate that the *M*-EAC is related to the formation of supramolecular sugar–nucleobase interaction in complexes **1** and **5**. Despite causing neighboring nucleotides to separate opposite to the axis of auxiliary ligands, the nucleotide ligand must be positioned counterclockwise to induce *M* supramolecular chirality (Fig. S5, S6, S20 and S21†).

The EAC in complex **3** has *P*-chirality, because the uncoordination of the CMP nucleotide makes the sugar–nucleobase interaction network similar to complexes **1** and **5**, as determined by the clockwise arrangement of nucleotide ligands surrounding the azpy axis with a 25.8 dihedral angle (Fig. S14 and S15†). Based on the further examination of crystal structures, it was revealed that the *P*-EAC resulted in the positioning of the nucleotide base towards the inner space between the chains, and this positioning is critical for the formation of the i-motif in complexes **2** and **4** (Fig. S9, S10, S17 and S18†). To observe a positive Cotton effect in the CD spectra, the *P*-EAC of the auxiliary ligand should be considered. The supramolecular helix of dCMP may affect the symmetry of electronic transitions and be responsible for changes in the Cotton effect.[40]

Luminescence emission spectra

In **1** and **2**, the red-shifts (20–22 nm) are observed with respect to bpa and nucleotide ligands. The results of luminescence emission spectra and electronic spectra upon addition of dCMP/CMP both indicate pre-organization of **1** and **2** in the solution. In **3** and **4**, the emission peaks of the complex around 375 nm are similar to that of CMP/dCMP,

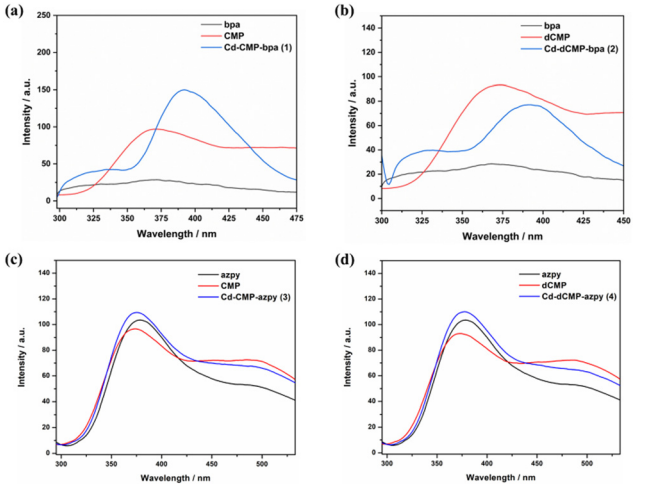


Fig. 8 Luminescence emission spectra (λ_{ex} = 270 nm) of the solution of (a) dCMP, bpa and complex **1**, (b) dCMP, bpa and complex **2**, (c) CMP, azpy and complex **3**, and (d) dCMP, azpy and complex **4**. The spectra were obtained by measuring 2.5 × 10^{-5} mol l^{-1} solution in a 1 cm cell, and the slit width was 2.5 nm.

which may indicate that the origin of this emission involves emissive state derived from the π–π* transition of CMP/dCMP. The emission around 450–550 nm is related to the azpy ligand (Fig. 8).

Thermogravimetric analysis

The thermal characteristics of the complexes were investigated by thermogravimetric (TG) analysis conducted on crystalline samples up to a temperature of 700 °C, and the resulting curves are depicted in Fig. S27.† The TG curves of **1**–**2** show a three-step weight loss process, while those of **3**, **4**, and **5** show a two-step weight loss process. In complexes **1**–**5**, the initial weight loss (**1**: obs. 4.30%, cal. 5.20%, **2**: obs. 3.73%, cal. 8.77%, **3**: obs. 10.0%, cal. 7.25%, **4**: obs. 7.99%, cal. 7.09%, **5**: obs. 3.50%, cal. 5.10%) in the temperature range of 40–220 °C is assigned to the loss of guest water solvates and coordinated water molecules per formula unit. In complexes **1** and **2**, the second step ranges from 200 to 400 °C corresponding to the loss of the 1,2-bis(4-pyridyl) ethane ligand (obs. 24.25%, cal. 26.78%, obs. 15.0%, cal. 17.95%). In complexes **3** and **4**, the second step ranges from 230 to 365 °C corresponding to the loss of the 4,4-azopyridine ligand (obs. 20.21%, cal. 24.76%, obs. 22.63%, cal. 18.11%). Complex **5** is stable up to 180 °C, and decomposition slows down around 500 °C, while the mass does not change on further heating to 700 °C. The residue of complexes **1**–**5** begins to decompose around 130–180 °C (130 °C for **1**, 145 °C for **2**, 140 °C for **3**, 125 °C for **4**, and 180 °C for **5**).

Conclusions

In conclusion, with the aim of better understanding the sugar–nucleobase hydrogen bonding interaction in supramolecular chemistry, we reported the effective synthesis of five crystalline

coordination polymers of CMP and dCMP nucleotides. It is the first time to report the solid-state structure of sugar–nucleobase hydrogen bonding, which is a new kind of supramolecular interaction. It has been found that the formation of sugar–nucleobase hydrogen bonding (complexes **1**, **3** and **5**) will impede the formation of i-motif (complexes **2** and **4**). Therefore, the research results of this work not only indicate the diversity of supramolecular interactions but also provide structural information and self-assembly conditions for different functional groups in nucleotide coordination polymers. Understanding the favorable chemical environment for expected supramolecular interactions and molecular recognition remains a challenge that necessitates future studies. Therefore, the design of increasingly more nucleotide coordination polymers may offer a platform to study and solve the problem, particularly *via* X-ray single-crystal diffraction analysis.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (no. 21071018, 21271026, 21471017).

Notes and references

1 S. Takahashi and N. Sugimoto, *Chem. Soc. Rev.*, 2020, **49**, 8439–8468.
2 J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 737–738.
3 L. R. Ganser, M. L. Kelly, D. Herschlag and H. M. Al-Hashimi, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**, 474–489.
4 A. Marnef and G. Legube, *Nat. Cell Biol.*, 2021, **23**, 305–313.
5 R. Satange, C.-K. Chang and M.-H. Hou, *Nucleic Acids Res.*, 2018, **46**, 6416–6434.
6 N. G. Abrescia, C. González, C. Gouyette and J. A. Subirana, *Biochemistry*, 2004, **43**, 4092–4100.
7 G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182–186.
8 L. Feng, A. Zhao, J. Ren and X. Qu, *Nucleic Acids Res.*, 2013, **41**, 7987–7996.
9 A. Jain, G. Wang and K. M. Vasquez, *Biochimie*, 2008, **90**, 1117–1130.
10 S. Daravath, A. Rambabu, N. Ganji, G. Ramesh and P. A. Lakshmi, *J. Mol. Struct.*, 2022, **1249**, 131601.
11 F.-T. Wang, Y.-H. Wang, J. Xu, K.-J. Huang, Z.-H. Liu, Y.-F. Lu, S.-Y. Wang and Z.-W. Han, *Nano Energy*, 2020, **68**, 104310.
12 M. Manghi and N. Destainville, *Phys. Rep.*, 2016, **631**, 1–41.
13 Q. M. Qiu, P. Zhou, L. Gu, L. Hao, M. Liu and H. Li, *Chem. – Eur. J.*, 2017, **23**, 7201–7206.
14 M. J. Iqbal, Z. Li, M. A. Khan, Y. Zhu, W. Hussain, H. Su, Q.-M. Qiu, R. Shoukat and H. Li, *CrystEngComm*, 2021, **23**, 4175–4180.
15 P. W. Laird, *Nat. Rev. Genet.*, 2010, **11**, 191–203.
16 Y. Xu, H. Zhang, F. Li, F. Shen, H. Wang, X. Li, Y. Yu and Y. Ma, *J. Mater. Chem.*, 2012, **22**, 1592–1597.
17 A. L. Boyle and D. N. Woolfson, *Supramol. Chem.*, 2012, 1639–1662.
18 T. Sawada and M. Fujita, *J. Am. Chem. Soc.*, 2010, **132**, 7194–7201.
19 B. Mohapatra and S. Verma, *Chem. Commun.*, 2017, **53**, 4748–4758.
20 G. Beobide, O. Castillo, J. Cepeda, A. Luque, S. Pérez-Yáñez, P. Román and J. Thomas-Gipson, *Coord. Chem. Rev.*, 2013, **257**, 2716–2736.
21 S. D. Dabhi, B. Roondhe and P. K. Jha, *Phys. Chem. Chem. Phys.*, 2018, **20**, 8943–8950.
22 A. Lopez and J. Liu, *ChemNanoMat*, 2017, **3**, 670–684.
23 P. Zhou, R. Shi, J.-F. Yao, C.-F. Sheng and H. Li, *Coord. Chem. Rev.*, 2015, **292**, 107–143.
24 D. Dutta, N. Nagapradeep, H. Zhu, M. Forsyth, S. Verma and A. J. Bhattacharyya, *Sci. Rep.*, 2016, **6**, 24499.
25 A. Terrón, L. Tomàs, A. Bauzá, A. García-Raso, J. J. Fiol, E. Molins and A. Frontera, *CrystEngComm*, 2017, **19**, 5830–5834.
26 N. Marino, D. Armentano, T. F. Mastropietro, M. Julve, F. Lloret and G. De Munno, *Cryst. Growth Des.*, 2010, **10**, 1757–1761.
27 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339–341.
28 G. M. Sheldrick, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 2008, **64**, 112–122.
29 G. M. Sheldrick, *Acta Crystallogr., Sect. C: Struct. Chem.*, 2015, **71**, 3–8.
30 N. Baidya and O. C. Uhlenbeck, *J. Biochem.*, 1995, **34**, 12363–12368.
31 J. E. Šponer, J. Leszczynski, V. Sychrovský and J. Šponer, *J. Phys. Chem. B*, 2005, **109**, 18680–18689.
32 A. Rich, *Acc. Chem. Res.*, 1977, **10**, 388–396.
33 P. Zhou, J.-F. Yao, C.-F. Sheng and H. Li, *CrystEngComm*, 2013, **15**, 8430–8436.
34 J. T. Davis and G. P. Spada, *Chem. Soc. Rev.*, 2007, **36**, 296–313.
35 P. Zhou, C. Wang, Q.-M. Qiu, J.-F. Yao, C.-F. Sheng and H. Li, *Dalton Trans.*, 2015, **44**, 17810–17818.
36 J. T. Davis, *Angew. Chem., Int. Ed.*, 2004, **43**, 668–698.
37 A. M. Ako, C. S. Hawes, B. Twamley and W. Schmitt, *CrystEngComm*, 2017, **19**, 994–1000.
38 K. Maeda, H. Mochizuki, M. Watanabe and E. Yashima, *J. Am. Chem. Soc.*, 2006, **128**, 7639–7650.
39 S. MacQuarrie, M. P. Thompson, A. Blanc, N. J. Mosey, R. P. Lemieux and C. M. Crudden, *J. Am. Chem. Soc.*, 2008, **130**, 14099–14101.
40 P. Zhou and H. Li, *Dalton Trans.*, 2011, **40**, 4834–4837.
41 S. B. Nielsen, T. Chakraborty and S. V. Hoffmann, *ChemPhysChem*, 2005, **6**, 2619–2624.
42 J. Alfuth, J. Chojnacki, T. Połoński and T. Olszewska, *New J. Chem.*, 2019, **43**, 5512–5517.