ChemComm
Chemical Communications

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: K. McQuaid, J. P. Hall, L. Baumgartner, D. J. Cardin and C. Cardin, *Chem. Commun.*, 2019, DOI: 10.1039/C9CC04316K.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemcomm

View Article Online
DOI: 10.1039/C9CC04316K

COMMUNICATION

Three thymine/adenine binding modes of the ruthenium complex Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ to the G-quadruplex forming sequence d(TAGGGTT) shown by X-ray crystallography

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Kane McQuaid,[a,b] James P. Hall,[b,c] Lena Baumgartner,[a] David J. Cardin[a] and Christine J. Cardin*[a]

Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ was crystallised with the G-quadruplex-forming heptamer d(TAGGGTT). Surprisingly, even though there are four unique binding sites, the complex is not in contact with any G-quartet surface. Two complexes stabilise cavities formed from terminal T·A and T·T mismatched pairs. A third shows kinking by a TAP ligand between T·T linkages, while the fourth shows sandwiching of a dppz ligand between a T·A/T·A quadruplex and a T·T mismatch, stabilised by an additional T·A base pair stacking interaction on a TAP surface. Overall, the structure shows an unexpected affinity for thymine, and suggests models for G-quadruplex loop binding.

Currently there are no structural models for the interactions of monomeric ruthenium polypyridyl complexes with the loop regions of nucleic acid assemblies such as the G-quadruplex and the i-motif. The biological importance of the G-quadruplex has become clear in the last few years,[1] and it has become an important drug target.[2] The DNA at the ends of human chromosomes, in the single stranded telomeric region, has upwards of 2,000 repeats of the sequence 5'-AGGGTT-3', and various versions of the sequence have been widely studied due to the G-quadruplex forming regions potential therapeutic significance.[2] These single stranded assemblies are often polymorphic in solution and probably for this reason have so far resisted nearly all attempts to crystallise them with metal complexes,[3,4] although NMR has been successfully used to provide elegant binding models for diruthenium complexes.[5] In that work, the binding mode of the diruthenium complex was clearly enantioselective, with only the Λ,Λ- enantiomer able to interact convincingly with the diagonal loop. This solution model is still the only one to define how this important class of metal complexes can interact with a unimolecular G-quadruplex, though strong 'light-switch' effects have been seen with related complex and known quadruplex-forming sequences.[6,7]

The binding mode of Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ to duplex DNA was recently described by us and showed that the inclusion of the 11-CN substituent in the dppz ligand resulted in the closing of the T·C/G·A terminal step, in contrast to a wealth of previous observations showing that the T·A base pair was readily flipped out when adjacent to dppz at a terminal step.[8,9] When racemic [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ was crystallised with d(TAGGGTTA), a G-quadruplex assembly was formed in which two Λ-enantiomers were bound at each end of the G-quartet stack.[10] In that work, two Δ-enantiomers (not the stoichiometric equivalent of four) were sandwiched between adjacent quadruplex assemblies and made little interaction with the DNA component, and two of the four 3' terminal adenine bases were not visible at all due to disorder, leading us to believe that this base was unimportant. The enantiomeric difference shown by this study reinforces the Λ preference previously reported by Thomas et al.[5] The structure also showed the stabilisation of the 5'-syn-guanine residues in the quadruplex assembly, leading to an overall antiparallel conformation. The disorder of the terminal adenine suggested an investigation of the truncated telomeric repeat sequence d(TAGGGTT) might give a more reproducible crystallisation. Unexpectedly, in this work we show for the first time that the parent Λ complex (without the 11-CN dppz substituent) can stabilise a T·T mismatch pair as part of a T·A/T·T cavity, and also for this first time that semi-intercalation (kinking) can be seen between thymine residues, whereas up till now we have only seen such kinking induced between two guanine residues, at a G·G/C·C step.[11] These are structural features which are most readily observed by X-ray crystallography, although very probably detectable in solution experiments and by single molecule approaches.[12,13] The lack of interaction with the parent dppz chromophore strongly supports our previous observation of the surprisingly powerful

[a] Department of Chemistry, University of Reading, Whiteknights, Reading, RG6 6AD, UK.
[b] Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, UK.
[c] Department of Pharmacy, University of Reading, Whiteknights, Reading, RG6 6AP, UK..

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

effect of 11-CN substitution in these Λ enantiomers on DNA binding.

different outcome.10 In our study of $[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]^{2+}$ crystallised with d(TAGGGTTA), each of the four

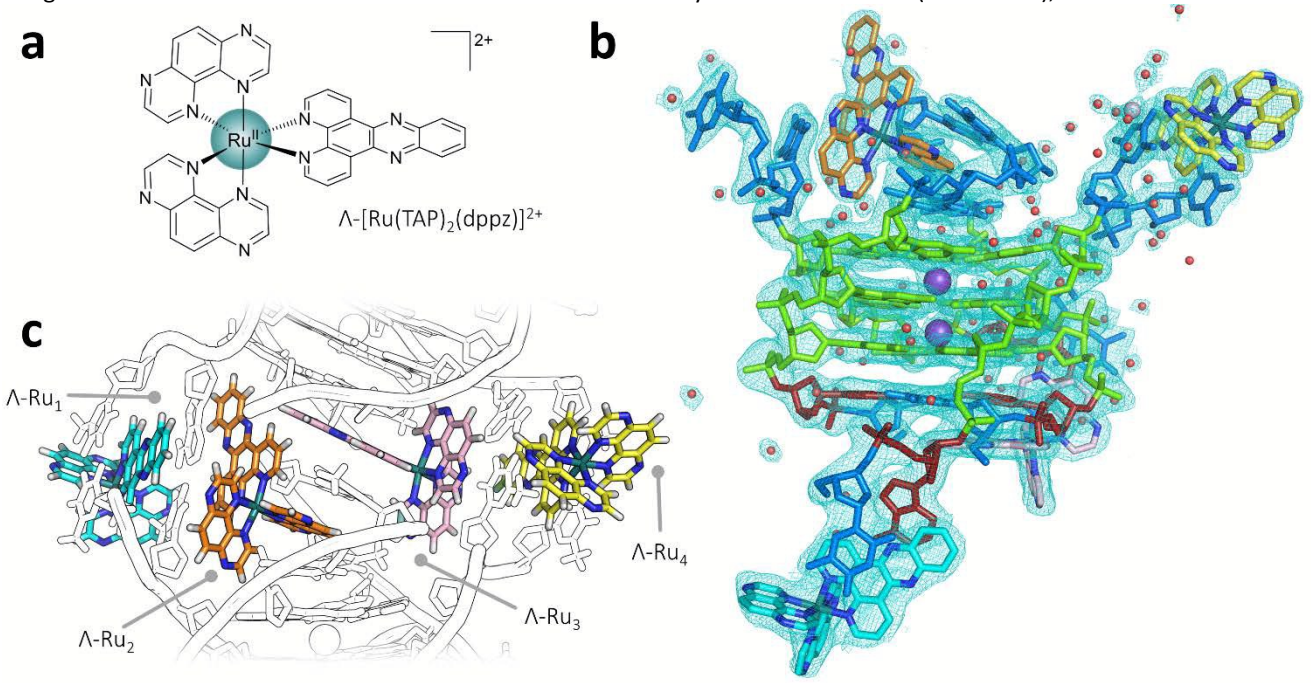

Figure 1 – (a) The Λ-$[Ru(TAP)_2(dppz)]^{2+}$ complex used in this study; (b) Overall view of the parallel stranded asymmetric unit of the structure reported here (PDB code 6RNL). Four strands of the sequence d(TAGGGTT) assembled with two K^+ ions and four crystallographically independent Λ-$[Ru(TAP)_2(dppz)]^{2+}$ cations. Colour code for residues throughout: guanine - green; adenine - red; and thymine - blue. Ruthenium complexes are coloured separately in cyan, orange, pink, and yellow for carbons; teal for ruthenium; and dark blue and white for nitrogen and hydrogen respectively. Alternative views are illustrated in figures S1 and S2. The map is contoured at 0.29 e $Å^{-3}$; (c) Generation of the four ruthenium complex environments at the interface between two nucleic acid assemblies. The numbering of the four ruthenium complexes corresponds to that used in the text. The kink in the DNA stack is generated by one of the TAP ligands of complex Ru_2, between a T·T mismatch and a water-bridged T·T mismatch.

The complex rac-$[Ru(TAP)_2(dppz)]^{2+}$ crystallised with the d(TAGGGTT) sequence and K^+ ions to give crystals containing only the lambda enantiomer (figure 1a). The structure was phased using SAD data measured above the Ru absorption edge at 22.26 KeV. Data collection and refinement parameters are given in table S1. The stoichiometric ratio in the resulting crystal is 1:1, giving four complexes per tetrameric assembly. This is the same ratio as in the previous study but giving an entirely crystallographically independent lambda complexes had an almost identical nucleic acid environment. In this crystal structure, each complex has a distinct environment within the crystal, and none makes contact with the central G-quadruplex unit (figure 1b and c). Here, we observe a parallel-stranded assembly, illustrated schematically in figure 2, held together by two K^+ ions. A Na^+ ion can also be identified (figure 3a/b). What was unexpected is the overall bend (figure 1c) introduced into an otherwise parallel stack by a semi-intercalative kinking motif similar to that we have we have previously observed in duplex structures. [9,11,14] The ~50° kink seen previously was always at GG steps of the sequence. Here, the kink is formed by a T·T mismatch pair and a second pair of thymine bases linked by water bridges (figure 4b). The motif generates an overall kink angle of about 28°, as can be seen by looking at the angle generated between the G-quartets of the assembly in Figure 1c, with a local kink of 34°, measured from the thymine base planes shown in Figure 3c. The packing diagrams of the structure viewed perpendicular to the long axis show the overall effect of this kinking on the assembly. The asymmetric units are packed together in spirals about the z direction in space group P65 (figure S4), giving head to tail stacking and generating the four ruthenium environments observed at the interface between the units. All four crystallographically independent complexes are bound in thymine-rich environments and hence suggest comparisons with the binding of metal complexes to loop regions in single stranded DNA, as thymine-thymine mismatched base pairs are situated adjacent to, and possibly

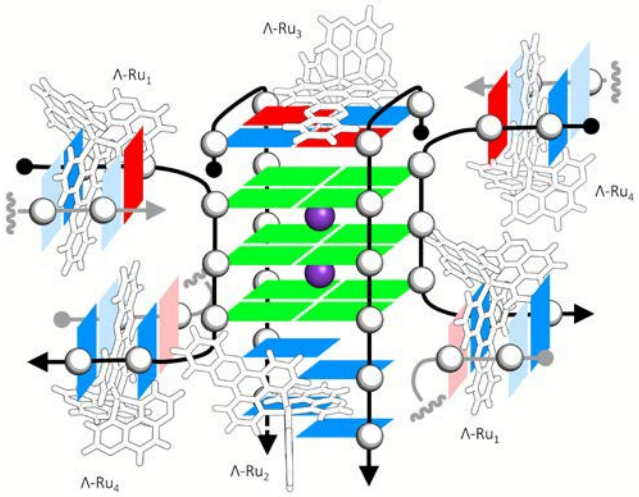

Figure 2 - Structural schematic. The four strands of the sequence d(TAGGGTT) are shown with arrows in the chain direction 5'-3'. The T·A and T·T base pairs formed with bases from symmetry related strands have thymine as pale blue and adenine as pale red. Potassium ions are shown as purple spheres.

stabilised by, all of the complexes, and stacking with both the TAP and the dppz ligands. For clarity, each will be described separately.

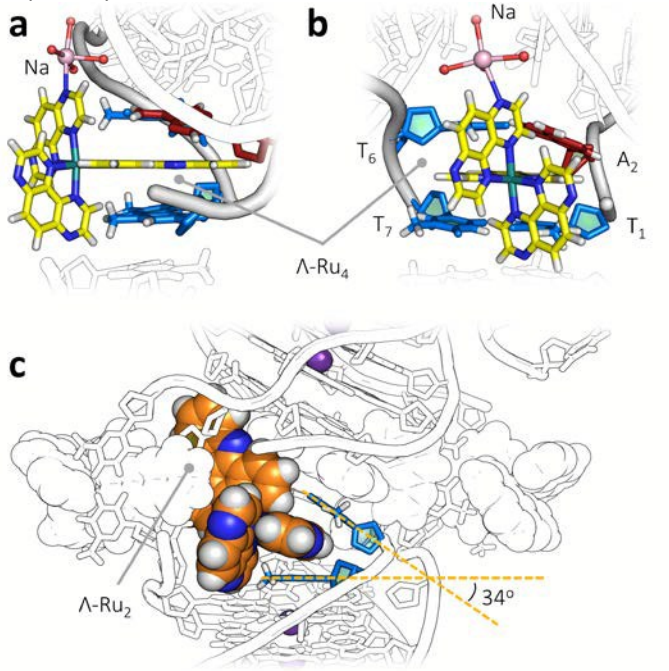

Figure 3 - Structural details of TAP ligands. (a) and (b) coordination of a Na^+ ion (pink) to a TAP ligand of Ru_4; and (c) the local 34° kink introduced by a TAP ligand of Ru_2. Potassium ions are shown as purple spheres, sodium in pink, and coordinated waters as red spheres.

Two of the four complexes (Ru_1 and Ru_4) have almost the same environments, at the ends of the overall assembly. T_1 and A_2 of one strand pair with T_6 and T_7 of another strand, as shown in the schematic of figure 2, to generate an intercalation cavity created by a standard A·T base pair and a T·T mismatch pair (figure 4a). These binding sites also provide a model for what would be major groove binding in duplex DNA. The sites are distinguished by coordination of a Na^+ ion to one TAP ligand of Ru_4 only. The ion is directly coordinated to one of the TAP ligands, and through a water bridge to a TAP ligand of Ru_3. The TAP ligands in these structures have previously been observed to be hydrated,[9] and this provides a convenient rationale for the sometimes relative ease of crystallisation when compared to the more hydrophobic but isosteric and isoelectronic phen analogue, the well-known 'light-switch' complex $[Ru(phen)_2(dppz)]^{2+}$.[15,16] In the overall assembly the Ru_1 and Ru_4 cavities are end-stacked on each other, generating a quasi-continuous stack running orthogonal to the main helix axis direction, which corresponds to the b axial direction in figure S4. The additional charge neutralisation by Na^+ is possibly an additional stabilising factor for this assembly. This monovalent ion coordination may also account for the asymmetry introduced by the differing environments of Ru_2 and Ru_3, since there is no corresponding ion linking Ru_1 and Ru_2.

The environments of the two central ruthenium complexes, Ru_2 and Ru_3, are distinctly different, thus generating the overall lack of quasi-twofold symmetry in this structure. Ru_3 appears almost completely surrounded by T·A and T·T base pairs, and the two faces of the complex are shown in Figures 4c and 4d. One dppz face contacts a T·A/T·A quartet surface formed by two T_1/A_2 ends, shown in the same figure. The other dppz surface contacts a T·T mismatch formed from two T_7 residues, shown in figure 4d. A further $A_2 \cdot T_6$ base pair contacts the TAP ligand, almost completely encapsulating it, shown in figure 4c.

The environment of Ru_2 is perhaps the most unexpected and as already stated generates the kink in the overall $P6_5$ packing shown in figures 3c and S4. The kink is generated at one of the TAP ligands, with a $T_7 \cdot T_7$ mismatch on one side of the TAP and two thymine residues, T_6 and T_6, with two water bridges on the other side of the TAP (figure 4b). Unexpectedly, the dppz is free, so that this complex is only held in place by this kinking interaction. This kinking site shows a remarkable overall resemblance to that seen in the original Λ-$[Ru(TAP)_2(dppz)]^{2+}$ structure, with the DNA duplex sequence d(TCGGCGCCGA) and in many structurally isomorphous examples since then.[11,9] In dilute solutions of B-DNA the thermodynamic binding constants show a relatively weak interaction compared with dppz intercalation.[9] In crystals and in other tightly packed environments, a combination of weak interactions can lead to environments which could not be predicted from any solution study, and what we are seeing here is perhaps a model for such cases.

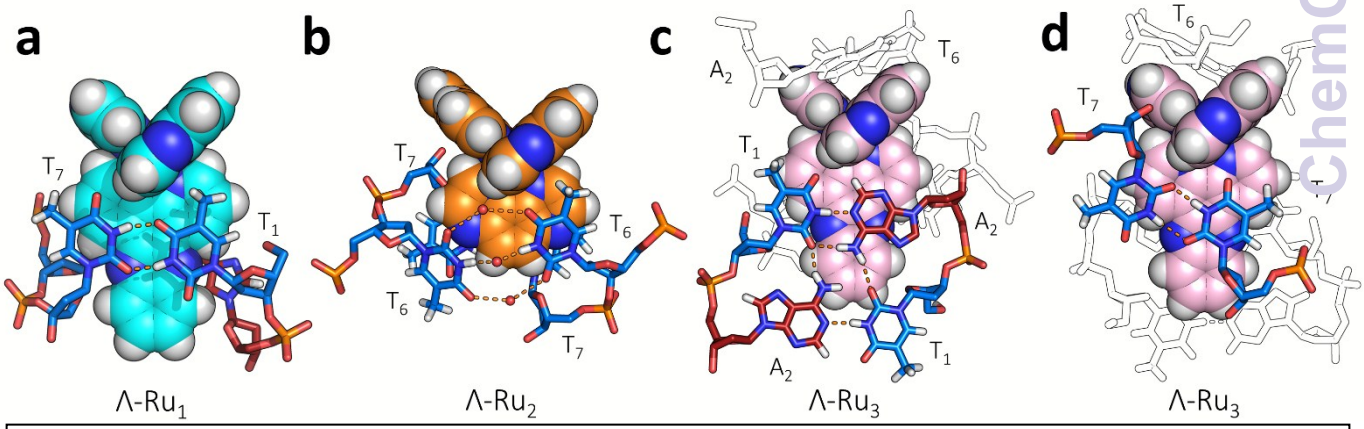

Figure 4 - The ruthenium complex environments: (a) T·T mismatch and A·T match form the cavity for Ru_1 and Ru_4 similarly; (b) kinking (semi-intercalation) at Ru_2, also generated by T·T mismatching. This complex is not bound by intercalation of the dppz chromophore, but just by the interaction of one of the TAP ligands as shown here; (b) and (c) illustrate the T·T mismatch cavity and additional T stacking and hydrogen bonding around Ru_3. Note that all ruthenium complex environments feature T·T mismatched base pairs. Coordinated waters are shown as red spheres. For clarity, interacting residues from neighbouring units are not coloured.

Both T·T mismatched base pairs and kinking by phen and TAP ligands may be important components of the binding of ruthenium polypyridyl complexes to higher-order DNA structures containing loop regions. Octahedral complexes have an inherently greater potential for specific interactions than square planar ones but not much is known about their binding modes.[17] So far the only structural evidence is provided by the work of the Thomas group as already stated.[5] In that work (in which the ancillary ligands were bpy) the diruthenium cation threaded through a diagonal loop, with the principal interactions being with the central thymine residue of the loop. We have previously shown that, of the well-known ancillary ligands in these systems, bpy is less likely to cause kinking and stacking than either phen or TAP.[18] More recently the TAP analogue of this diruthenium compound has been shown to have a range of useful properties in cell systems.[19] It was studied as an enantiomeric mixture and gave spectroscopic results clearly indicative of several binding modes. The specificity of these complexes does not just arise from end-stacking to the G-quadruplex chromophore but plausibly also includes the sort of thymine interactions revealed by the present work. There are several examples of ruthenium polypyridyl complexes which are luminescent when bound to what may well be thymine-adenine loop regions of G-quadruplexes, but there is no structural data for any of these. The binding modes seen in this work, which have no counterpart in duplex DNA, and would not be predictable from any modelling calculation, provide a useful springboard for understanding luminescence and other spectroscopic behaviour. Each of the binding modes shown in figure 4 would have different luminescence behaviour if it were the phen analogue, based on our previous work. Ru_3, being almost enclosed, would be most luminescent, with Ru_1 and Ru_4 expected to be similar, and Ru_2 the most exposed to quenching of the excited state via non-radiative pathways. A previous paper from our laboratory has considered the delta enantiomer/duplex DNA case in detail.[14]

In future we aim to provide a comparable interpretation of the binding of lambda complexes to G-quadruplex loop regions. We would also like to understand the crucial factor which determines whether the G-quadruplex is parallel, or antiparallel as in ref 10. It is not clear how much of the switch can be ascribed to the modification of the dppz ligand and how much the crystallisation is sequence dependent.

Acknowledgements

We thank Diamond Light Source and EPSRC for a studentship (to K. McQ.) and the BBSRC for support from grant nos. BB/019250/1 and BB/M004635/1 (To CJC, JAB, JPH and DJC). We gratefully acknowledge the provision of beamtime on beamline I03 at Diamond Light Source Ltd. (MX18745-1).

Conflicts of interest

There are no conflicts to declare.

References

View Article Online
DOI: 10.1039/C9CC04316K

1 G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182–186.

2 S. Asamitsu, S. Obata, Z. Yu, T. Bando and H. Sugiyama, *Mol.* , 2019, 24.

3 N. H. Campbell, N. H. A. Karim, G. N. Parkinson, M. Gunaratnam, V. Petrucci, A. K. Todd, R. Vilar and S. Neidle, *J. Med. Chem.*, 2012, **55**, 209–222.

4 C. Bazzicalupi, M. Ferraroni, F. Papi, L. Massai, B. Bertrand, L. Messori, P. Gratteri and A. Casini, *Angew. Chemie - Int. Ed.*, 2016, **55**, 4256–4259.

5 T. Wilson, P. J. Costa, V. Félix, M. P. Williamson and J. A. Thomas, *J. Med. Chem.*, 2013, **56**, 8674–8683.

6 L. Xu, X. Chen, J. Wu, J. Wang, L. Ji and H. Chao, *Chem. - A Eur. J.*, 2014, **21**, 4008–4020.

7 Q. Yu, Y. Liu, C. Wang, D. Sun, X. Yang, Y. Liu and J. Liu, *PLoS One*, , DOI:10.1371/journal.pone.0050902.

8 K. McQuaid, J. P. Hall, J. A. Brazier, D. J. Cardin and C. J. Cardin, *Chem. - A Eur. J.*, 2018, **24**, 15859–15867.

9 C. J. Cardin, J. M. Kelly and S. J. Quinn, *Chem. Sci.*, 2017, **8**, 4705–4723.

10 K. McQuaid, H. Abell, S. P. Gurung, D. Allan, G. Winter, T. Sorensen, D. J. Cardin, J. A. Brazier, C. J. Cardin and J. P. Hall, *Angew. Chemie Int. Ed.*, , DOI:10.1002/anie.201814502.

11 J. . Hall, K. O'Sullivan, A. Naseer, J. . Smith, J. . Kelly and C. J. Cardin, *Proc.Natl.Acad.Sci.USA*, 2011, **108**, 17610–17614.

12 P. Lincoln and B. Nordén, *J. Phys. Chem. B.*, 1998, **102**, 9583–9594.

13 W. Vanderlinden, P. J. Kolbeck, W. Frederickx, S. F. Konrad, T. Nicolaus, C. Lampe, A. S. Urban, C. Moucheron and J. Lipfert, *Chem. Commun.*, , DOI:10.1039/C9CC02838B.

14 J. P. J. P. Hall, P. M. P. M. Keane, H. Beer, K. Buchner, G. Winter, T. L. T. L. Sorensen, D. J. D. J. Cardin, J. A. J. A. Brazier and C. J. C. J. Cardin, *Nucleic Acids Res.*, 2016, **44**, 9472–9482.

15 J. P. Hall, D. Cook, S. R. Morte, P. McIntyre, K. Buchner, H. Beer, D. J. Cardin, J. A. Brazier, G. Winter, J. M. Kelly and others, *J. Am. Chem. Soc.*, 2013, **135**, 12652–12659.

16 H. Niyazi, J. P. Hall, K. O'Sullivan, G. Winter, T. Sorensen, J. M. Kelly and C. J. Cardin, *Nat Chem*, 2012, **4**, 621–628.

17 Q. Cao, Y. Li, E. Freisinger, P. Z. Qin, R. K. O. Sigel and Z.-W. Mao, *Inorg. Chem. Front.*, 2017, **4**, 10–32.

18 J. P. J. P. Hall, S. P. S. P. Gurung, J. Henle, P. Poidl, J. Andersson, P. Lincoln, G. Winter, T. Sorensen, D. J. D. J. Cardin, J. A. J. A. Brazier and C. J. C. J. Cardin, *Chem. - A Eur. J.*, , DOI:10.1002/chem.201605508.

19 S. A. Archer, A. Raza, F. Dröge, C. Robertson, A. J. Auty, D. Chekulaev, J. A. Weinstein, T. Keane, A. J. H. M. Meijer, J. W. Haycock, S. MacNeil and J. A. Thomas, *Chem. Sci.*, 2019, **2**, 3502–3513.

