









# Paramagnetic encoding of molecules

Jan Kretschmer<sup>1,2</sup>, Tomáš David <sup>1</sup>, Martin Dračinský <sup>1</sup>, Ondřej Socha <sup>1</sup>, Daniel Jirak <sup>3,4</sup>, Martin Vít <sup>3,5</sup>, Radek Jurok<sup>6,7</sup>, Martin Kuchař <sup>6,8</sup>, Ivana Císařová <sup>9</sup> & Miloslav Polasek <sup>1</sup>✉

Contactless digital tags are increasingly penetrating into many areas of human activities. Digitalization of our environment requires an ever growing number of objects to be identified and tracked with machine-readable labels. Molecules offer immense potential to serve for this purpose, but our ability to write, read, and communicate molecular code with current technology remains limited. Here we show that magnetic patterns can be synthetically encoded into stable molecular scaffolds with paramagnetic lanthanide ions to write digital code into molecules and their mixtures. Owing to the directional character of magnetic susceptibility tensors, each sequence of lanthanides built into one molecule produces a unique magnetic outcome. Multiplexing of the encoded molecules provides a high number of codes that grows double-exponentially with the number of available paramagnetic ions. The codes are readable by nuclear magnetic resonance in the radiofrequency (RF) spectrum, analogously to the macroscopic technology of RF identification. A prototype molecular system capable of 16-bit (65,535 codes) encoding is presented. Future optimized systems can conceivably provide 64-bit ( $\sim 10^{19}$  codes) or higher encoding to cover the labelling needs in drug discovery, anti-counterfeiting and other areas.

<sup>1</sup>Institute of Organic Chemistry and Biochemistry of the CAS, Flemingovo náměstí 542/2, 160 00 Prague 6, Czech Republic. <sup>2</sup>Department of Organic Chemistry, Faculty of Science, Charles University in Prague, Hlavova 2030/8, 128 43 Prague 2, Czech Republic. <sup>3</sup>MR Unit, Department of Radiodiagnostic and Interventional Radiology, Institute for Clinical and Experimental Medicine, Vídeňská 1958/9, 140 21 Prague 4, Czech Republic. <sup>4</sup>Institute of Biophysics and Informatics, First Faculty of Medicine, Charles University in Prague, Salmovská 1, 120 00 Prague 2, Czech Republic. <sup>5</sup>Faculty of Mechatronics Informatics and Interdisciplinary studies, Technical University of Liberec, Hálkova 917/6, 460 01 Liberec, Czech Republic. <sup>6</sup>Forensic Laboratory of Biologically Active Substances, Department of Chemistry of Natural Compounds, University of Chemistry and Technology Prague, Technická 5, 166 28 Prague 6, Czech Republic. <sup>7</sup>Department of Organic Chemistry, University of Chemistry and Technology Prague, Technická 5, 166 28 Prague 6, Czech Republic. <sup>8</sup>Department of Experimental Neurobiology, National Institute of Mental Health, Topolová 748, 250 67 Klecany, Czech Republic. <sup>9</sup>Department of Inorganic Chemistry, Faculty of Science, Charles University in Prague, Hlavova 2030/8, 128 43 Prague 2, Czech Republic. ✉email: [miloslav.polasek@uochb.cas.cz](mailto:miloslav.polasek@uochb.cas.cz)

Managing digital information is crucial not only for our civilization but for the existence of life itself. All known life forms depend on nucleic acids, which are in essence digital medium. The information capacity of molecules is enormous and very tempting for technological applications. Indeed, humans learned to hack the system of nucleic acids for the purpose of encoding data<sup>1–7</sup>. However, the principle of molecular recognition that works well within a biological microenvironment is difficult to connect with technology built primarily around the electromagnetism of inorganic materials and RF communication. Thus, alternative synthetic polymers have been proposed for information encoding<sup>8–12</sup>. Unfortunately, the current methods of reading often result in the destruction of the molecules. In contrast, modern civilization demands to communicate information wirelessly and repeatably, in an effort to identify and track objects<sup>13,14</sup>, monitor health<sup>15,16</sup>, and eventually merge the digital with the living<sup>17,18</sup>. To achieve this at the molecular level will require adapting the molecular systems to be more compatible with macroscopic information technologies.

Writing (encoding) information at the molecular level can be achieved with either (i) sequences of monomers concatenated within one molecular string<sup>7–12</sup> or with (ii) mixtures of individual, unique compounds<sup>19–23</sup>. For both, there is a trade-off between synthetic difficulty, material demands, and information capacity. Long sequences can be made from a few monomers, but synthetic writing is difficult, slow, and prone to errors<sup>7–9</sup>. On the other hand, mixtures are easy to make but require a high number of uniquely distinguishable components, which eventually become a limiting factor<sup>19–23</sup>. Theoretically, an alternative approach, better balancing the need to generate many codes with a small number of components and synthetic steps, is (iii) to first synthesize short sequences and then use them as unique components in mixtures. Although this approach is conceptually well-understood<sup>21</sup>, the difficulty of decoding mixed sequences is prohibitive.

Reading (decoding) nongenomic molecular codes currently heavily relies on mass spectrometry (MS), which can decode sequences as well as complex mixtures with high sensitivity<sup>8–10,19–21</sup>. However, mixtures of sequences containing permutations of the same monomers have the same molecular weight and are nearly impossible to decode. In addition, MS is a destructive method that does not allow for repeated readings that are desirable for long-term tracking and monitoring. On the other hand, spectroscopic methods which are capable of this can typically only detect individual components (monomers) but cannot distinguish their order in sequences<sup>12,24</sup>. Another challenge for spectroscopic methods are signal overlaps that limit the number of distinguishable codes. For example, fluorescent labels, as one of the most important tools for interrogation of biological systems, provide relatively broad signals within a narrow spectral window<sup>23–26</sup>. Up to ten different fluorescent tags can be distinguished from mixtures with advanced computational methods, but this still amounts only to  $2^{10} - 1 = 1023$  codes<sup>26</sup>. This limitation cannot be overcome without expanding the size along at least one dimension<sup>23,27</sup>. Indeed, various microscale particles have been proposed as barcodes, based on tuneable luminescent materials combined to generate the code<sup>28–31</sup>. However, the information capacity per mass is much lower for particles compared to molecules. A potential way out of this dilemma is to use non-optical methods of reading. For example, the spectral window of nuclear magnetic resonance (NMR) spectroscopy can accommodate much higher number of distinguishable signals usable for information encoding<sup>22</sup>.

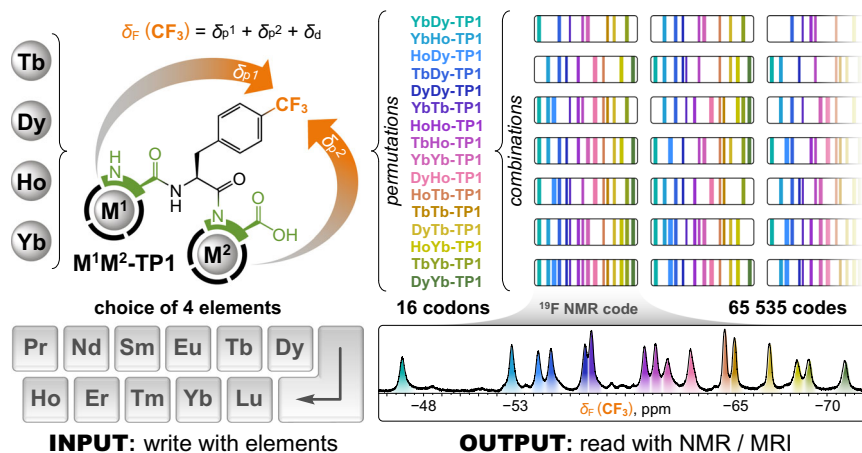
In the same way that magnetic media accelerated modern information technologies, it is recognized that molecular magnetism could be a major breakthrough for future technologies at

the molecular scale. Alas, single-molecule magnets so far function only at impractically low temperatures<sup>32</sup>. The next best magnetic alternative that works at the molecular level and over a broad temperature range simultaneously is paramagnetism. When placed within an external magnetic field, unpaired electrons of a paramagnetic metal ion generate their own local field described by a magnetic susceptibility tensor. This tensor has specific radial and angular dependencies dictated by the particular element and its coordination environment, and strongly influences the NMR frequencies of nearby nuclei<sup>33–35</sup>, an effect exploited in protein structure elucidation<sup>36</sup>. Lanthanides are ideally suited for this purpose, as they provide a range of chemically similar but magnetically different  $\text{Ln}^{3+}$  ions amenable to stable incorporation into organic molecules via coordination<sup>35,37</sup>. While distinguishable NMR signals can be obtained from chemically different diamagnetic molecules and formulations<sup>38–42</sup>, introducing paramagnetic lanthanides allows for decoupling of the NMR shifts from chemical properties. This principle has been previously used to create chemically nearly identical molecules that could be spectrally resolved with magnetic resonance imaging (MRI)<sup>43–45</sup>. However, with one lanthanide ion per molecule, the number of distinguishable signals remains limited to 12 usable lanthanides (excluding Pm, Gd; including 1 diamagnetic). So far, only a few works have investigated the possibility of combining paramagnetic effects of multiple lanthanides within one molecule<sup>46,47</sup>, presumably due to the difficulty in synthesizing well-defined multimetallic compounds<sup>24,48,49</sup>. We speculated that two or more lanthanide ions arranged in sequence within one molecule should combine their magnetic susceptibility tensors to produce a unique NMR-readable outcome for each permutation of the elements. This greatly increases the number of signals and molecular codes that can be generated with a limited number of elements (Fig. 1).

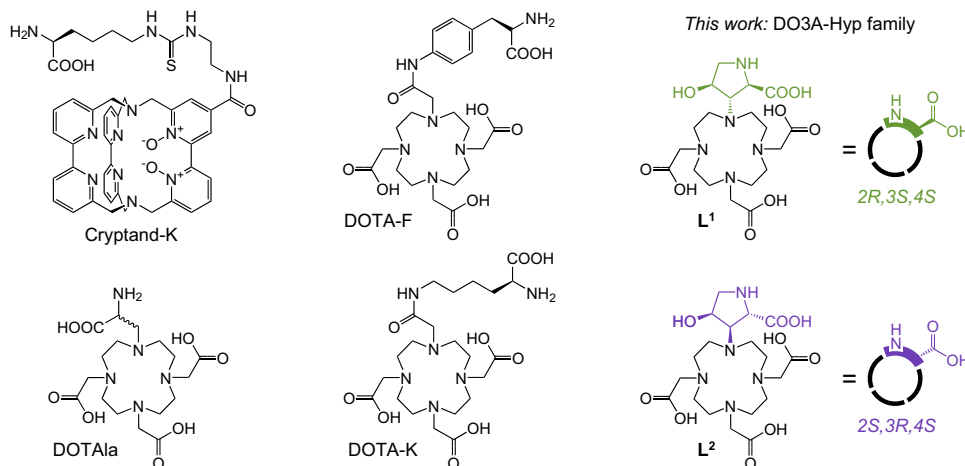
In this work, we present a prototype of paramagnetically encodable molecular architecture, demonstrate the encoding and decoding principles, provide practical examples of information encoding, and outline the possible future capabilities. The number of unique codes that can be generated in this way from a few chemical building blocks is compatible with the foreseeable needs in drug discovery<sup>1,2</sup> and anti-counterfeiting<sup>13</sup> applications. This work provides an alternative direction toward programmable digital molecular information systems.

## Results

**Design of building blocks.** Combining magnetic susceptibility tensors of lanthanide ions requires their placement in close proximity within one molecule, preferably in a modular way to enable encoding of specific sequences. Conformational freedom and isomerism must be limited, so that the metal ion positions are well defined and there is no averaging or multiplication of NMR signals. Peptides can offer the desired modular synthetic approach, provided that a suitable building block is available to act simultaneously as amino acid and a strong chelator for  $\text{Ln}^{3+}$  ions. Several amino acid/chelator building blocks have been described previously, but none were suitable for the task, falling short of the required control over conformations and/or isomerism (Fig. 2)<sup>24,49–51</sup>. Learning from these examples, we developed a family of building blocks collectively named DO3A-Hyp, where the chelator is based on macrocyclic DO3A (1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid) and the amino acid is derived from hydroxyproline (Hyp). Here, we present two members of the family, building blocks  $\text{L}^1$  and  $\text{L}^2$  (Fig. 2 and synthetic scheme in Fig. 3), that differ in the configuration of chiral centers in the Hyp moiety. The DO3A-Hyp family is structurally related to the chelator HP-DO3A that is clinically used in the MRI contrast agent *Gadoteridol*<sup>37</sup>. Although knowing



**Fig. 1** The principle of paramagnetic encoding presented in this work. A sequence of two lanthanide ions ( $M^1M^2$ ) is chemically written into a tripeptide molecule  $M^1M^2\text{-TP1}$ , which incorporates two stereochemically well-defined amino acid/chelator building blocks. The paramagnetic effects of the lanthanides combine to influence a reporter  $\text{CF}_3$  group. Each  $M^1M^2$  permutation of the elements (codon) results in a unique  $^{19}\text{F}$  NMR shift ( $\delta_F$ ) that is given by the sum of individual paramagnetic shift contributions from each  $M^1$  and  $M^2$  ( $\delta_{p1}$  and  $\delta_{p2}$ , respectively) and a constant diamagnetic shift ( $\delta_d$ ). The encoded molecules are further multiplexed to produce codes readable by NMR or MRI. A small number of elements allows the writing of a very high number of codes. The exact chemical structures of the amino acid/chelator building blocks used in this work are provided in Figs. 2 and 3.



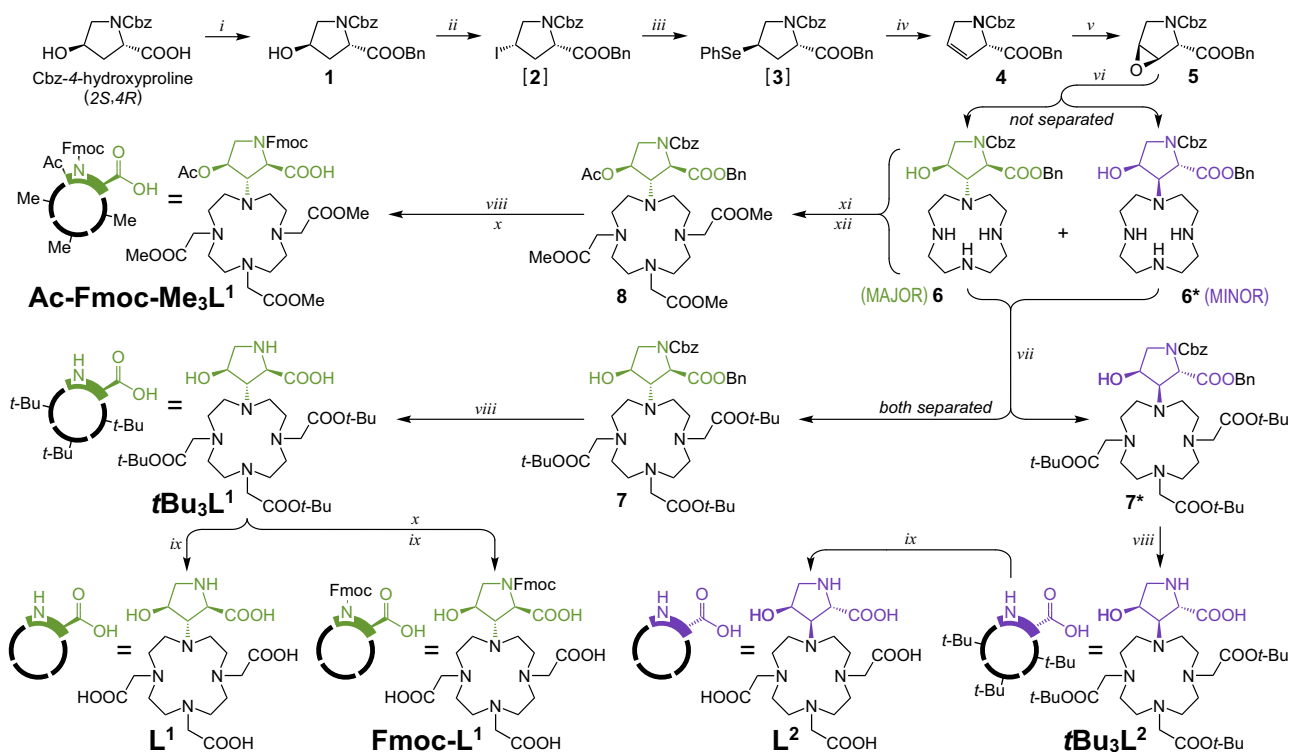
**Fig. 2** Chemical structures of amino acid/chelator building blocks for incorporation of  $\text{Ln}^{3+}$  ions into peptides discussed in this work. Structures were redrawn without protective groups and metal ions from references: Cryptand-K<sup>24</sup>, DOTA-K<sup>49,50</sup>, DOTA-F<sup>50</sup>, and DOTAAla<sup>51</sup>. The DO3A-Hyp family of building blocks presented in this work comprises two isomers that differ in the configuration of chiral centers in the Hyp (hydroxyproline) moiety: isomer 2R,3S,4S ( $L^1$ , green shading) and isomer 2S,3R,4S ( $L^2$ , violet shading).

the exact structure of the chelates was not necessary for the purpose of paramagnetic encoding, we found with X-ray diffraction on a  $[\text{Dy}(\text{L}^1)]$  crystal that inversion ( $2S \rightarrow 2R,3S,4S$ ) occurred in Hyp during synthesis of  $L^1$  (Fig. 4A and Supplementary Figs. 1–3). We explain this by deprotonation of the epoxide 5 during its reaction with basic cyclen (suggested mechanism in Supplementary Fig. 4). The same reaction step also opened a path to the isomer  $L^2$ , though with a much lower yield (9:1 ratio  $L^1$ :  $L^2$ ). The crystal structure of  $[\text{Dy}(\text{L}^2)]$  confirmed 2S,3R,4S configuration of Hyp in  $L^2$  (Fig. 4B and Supplementary Figs. 5–7). In both solid-state structures, the Hyp moiety coordinated with the 2R- or 2S-carboxyl, respectively, rather than with the 4S-hydroxyl group. Nevertheless, the 2R-carboxyl of  $[\text{Ln}(\text{L}^1)]$  chelates was reactive in peptide coupling reactions (in contrast to unreactive coordinated acetate arms), indicating coordination competition with 4S-hydroxyl in solution (Supplementary Fig. 8). Coordination of 4S-hydroxyl is therefore considered in figures to respect this reactivity. Our assumption was that the chiral Hyp moiety would impose a preference for one of several possible

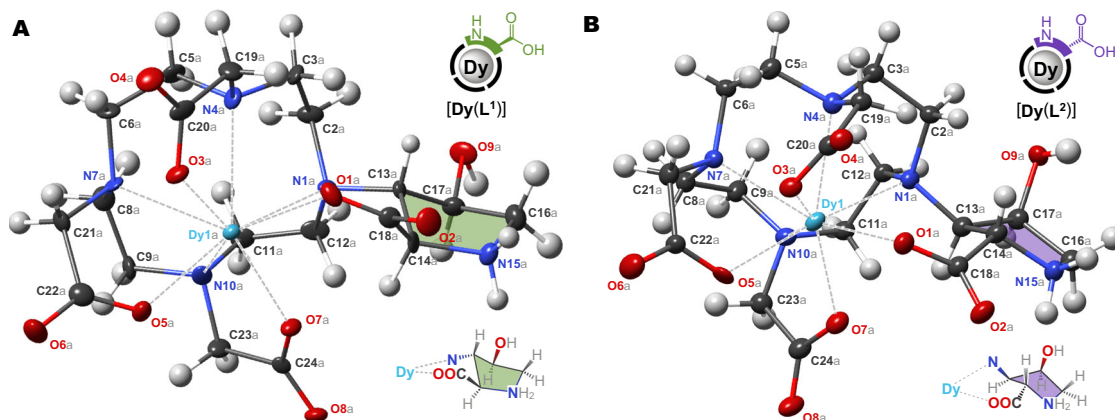
coordination isomers typical for macrocyclic chelators and thus manage the problem of isomeric speciation<sup>36,52,53</sup>. Indeed, only one dominant species was found with NMR spectroscopy in solution for all tested lanthanides (see the next section).

**Encodable molecular framework.** Having the suitable building blocks, we designed a prototype of an encodable molecular framework. Tripeptide 1 (TP1) consists of two  $L^1$  units, one at each terminus, and a middle amino acid bearing a reporter  $\text{CF}_3$  group. Analogous tripeptide 2 (TP2) molecule contains  $L^2$  unit at the C-terminus. This construction allows the reporter to perceive magnetic fields from two metal ions simultaneously in  $M^1M^2\text{-TP1}$  or  $M^1M^2\text{-TP2}$  molecules (Fig. 5). We selected the  $\text{CF}_3$  reporter for the high sensitivity and negligible natural background of the  $^{19}\text{F}$  nucleus, which facilitate observation and interpretation of signals in the RF NMR channel<sup>40,41,53,54</sup>.

The chemical shifts induced by paramagnetic lanthanide ions are difficult to predict from theory with confidence<sup>35</sup>, especially if the structure in solution is not precisely known. Therefore, to



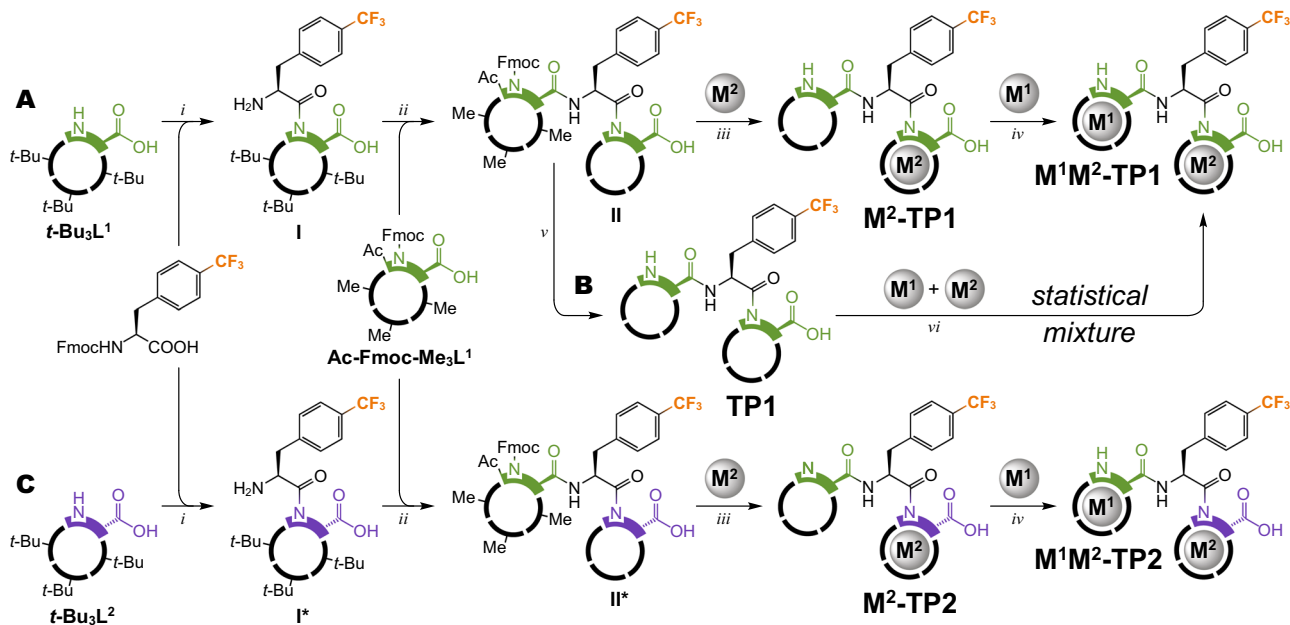
**Fig. 3 Synthetic scheme for building blocks  $L^1$ ,  $L^2$ , and their protected variants.** Conditions: (i) BnBr (benzyl bromide), TEA (triethylamine), THF (tetrahydrofuran); (ii) PPh<sub>3</sub>, THF, 0 °C followed by DIAD (diisopropyl azodicarboxylate), MeI, 0 °C → RT; (iii) (PhSe)<sub>2</sub>, EtOH, NaBH<sub>4</sub>, 0 °C; (iv) H<sub>2</sub>O<sub>2</sub>, THF, 0 °C → RT; (v) MCPBA (*meta*-chloroperoxybenzoic acid), CHCl<sub>3</sub>, 85 °C; (vi) cyclen, *t*-BuOH, 105 °C; (vii) *t*-BuO<sub>2</sub>CCH<sub>2</sub>Br, K<sub>2</sub>CO<sub>3</sub>, MeCN; (viii) H<sub>2</sub>, Pd/C, AcOH, MeOH; (ix) TFA; (x) FmocCl, aq. borate/NaOH buffer (pH 9.0), MeCN; (xi) MeO<sub>2</sub>CCH<sub>2</sub>Br, K<sub>2</sub>CO<sub>3</sub>, MeCN; (xii) Ac<sub>2</sub>O, TEA, DMAP (4-dimethylaminopyridine), MeCN. Intermediates in brackets were not isolated. Colors define stereochemistry of the hydroxyproline moiety: 2R,3S,4S (green) and 2S,3R,4S (violet). Detailed synthetic procedures are provided in Supplementary Figs. 22–36.



**Fig. 4 Molecular structures of [Dy(L<sup>1</sup>)] and [Dy(L<sup>2</sup>)] in the solid state.** In both structures, the ligand coordinated to the central Dy<sup>III</sup> cation with four ring nitrogen donors and four carboxylate donors. The view highlights the absolute configuration of the proline ring. Thermal ellipsoids were set at 50% probability. **A** [Dy(L<sup>1</sup>)] contains 2R,3S,4S configuration of the proline ring (green shading) and the chelate adopts  $\Delta\lambda\lambda\lambda$  square antiprismatic (SA) conformation. For detailed information, see Supplementary Figs. 1–3. **B** [Dy(L<sup>2</sup>)] contains 2S,3R,4S configuration of the proline ring (violet shading) and the chelate adopts  $\Delta\delta\delta\delta\delta$  twisted square antiprismatic (TSA) conformation. For detailed information, see Supplementary Figs. 5–7.

gain a quick overview, we used a metal-free TP1 for post-synthetic complexation of equimolar binary lanthanide mixtures to obtain statistical mixtures of four M<sup>1</sup>M<sup>2</sup>-TP1 compounds in 1:1:1:1 molar ratio (Fig. 5B). In all studied cases, <sup>19</sup>F NMR spectra showed four major singlets, mostly free of overlaps, shifted and broadened by the paramagnetic action of the respective Ln<sup>3+</sup> ions (Supplementary Fig. 9). We selected the Dy<sup>3+</sup>/Ho<sup>3+</sup> combination for further detailed study because of a good balance between chemical shift and NMR signal broadening (Fig. 6A). Specific

M<sup>1</sup>M<sup>2</sup> sequences were encoded into M<sup>1</sup>M<sup>2</sup>-TP1 compounds via controlled synthesis using several different strategies, some of which also included direct peptide coupling with metal-preloaded derivatives of L<sup>1</sup> (Supplementary Figs. 64–86). Each of the well-defined M<sup>1</sup>M<sup>2</sup>-TP1 compounds matched one of the four major <sup>19</sup>F NMR peaks observed in the statistical Dy<sup>3+</sup>/Ho<sup>3+</sup> mixture, confirming that each M<sup>1</sup>M<sup>2</sup> sequence generates a unique signal (Fig. 6B). Some minor peaks detected in the <sup>19</sup>F NMR spectra of (purified) M<sup>1</sup>M<sup>2</sup>-TP1 compounds could be attributed to low-



**Fig. 5** Synthesis of  $M^1M^2$ -TP1 and  $M^1M^2$ -TP2 compounds. **A** Main synthetic pathway to well-defined  $M^1M^2$ -TP1 products. **B** Post-synthetic complexation pathway to statistical mixtures of four major  $M^1M^2$ -TP1 products. **C** Synthetic pathway to well-defined  $M^1M^2$ -TP2 products (analogous to A). Conditions: (i) Fmoc-Phe(*p*-CF<sub>3</sub>)-OH, PyAOP ((7-azabenzotriazol-7-yloxy)tripyrrolidinophosphonium hexafluorophosphate), DIPEA (*N,N*-diisopropylethylamine), DMSO (dimethyl sulfoxide) followed by DBU (1,8-diazabicyclo[5.4.0]undec-7-ene), DMF (dimethylformamide); (ii) Ac-Fmoc-Me<sub>3</sub>L<sup>1</sup>, PyAOP, DIPEA, DMSO followed by TFA (trifluoroacetic acid); (iii) M<sup>2</sup>Cl<sub>3</sub>, aq. MOPS (3-(*N*-morpholino)propanesulfonic acid)/NaOH (pH 7.0) followed by LiOH, H<sub>2</sub>O, MeOH; (iv) M<sup>1</sup>Cl<sub>3</sub>, aq. MOPS/NaOH (pH 7.0); (v) LiOH, H<sub>2</sub>O, MeOH; (vi) M<sup>1</sup>Cl<sub>3</sub>, M<sup>2</sup>Cl<sub>3</sub>, aq. MOPS/NaOH (pH 7.0). For the full structures of the building blocks, see Fig. 3. Detailed synthetic procedures including alternative synthetic pathways are provided in Supplementary Figs. 37–95.

abundance forms of the same compound (conformers or isomers), because most of them coalesced with the main peak at elevated temperatures (Supplementary Fig. 10). Nevertheless, these minor peaks presented no obstacle in further analysis of the NMR data. As expected, the induced chemical shifts strongly depended also on the precise coordination environment and mutual arrangement of the two metal centers. The molecular system of  $M^1M^2$ -TP2, despite its (deceptive) similarity to  $M^1M^2$ -TP1, provided dramatically different chemical shifts with the same metal ion sequences (Fig. 6C and Supplementary Fig. 11). Furthermore, to investigate the importance of rigid and compact molecular design on paramagnetic encoding, we performed a control experiment using an alternative architecture. The same Dy<sup>3+</sup>/Ho<sup>3+</sup> sequences were encoded into a tripeptide TP3 that was constructed with the previously described DOTA-K building block<sup>49,50</sup> containing a flexible sidechain linker (Supplementary Fig. 12). Conformational flexibility, in this case, resulted in metal ion position averaging, diminishing the paramagnetic shifts, and consequently the TP3 system failed to produce distinguishable signals for different  $M^1M^2$  sequences (Fig. 6D and Supplementary Fig. 13). Moreover, the signals appeared split, likely due to the *cis*–*trans* isomerism of the coordinated amide moiety noted previously for similar compounds<sup>36</sup>. This problem was not observed with the TP1 and TP2 systems using the DO3A-Hyp family building blocks, demonstrating that a rigid and stereochemically well-defined molecular framework is essential for the purpose of paramagnetic encoding.

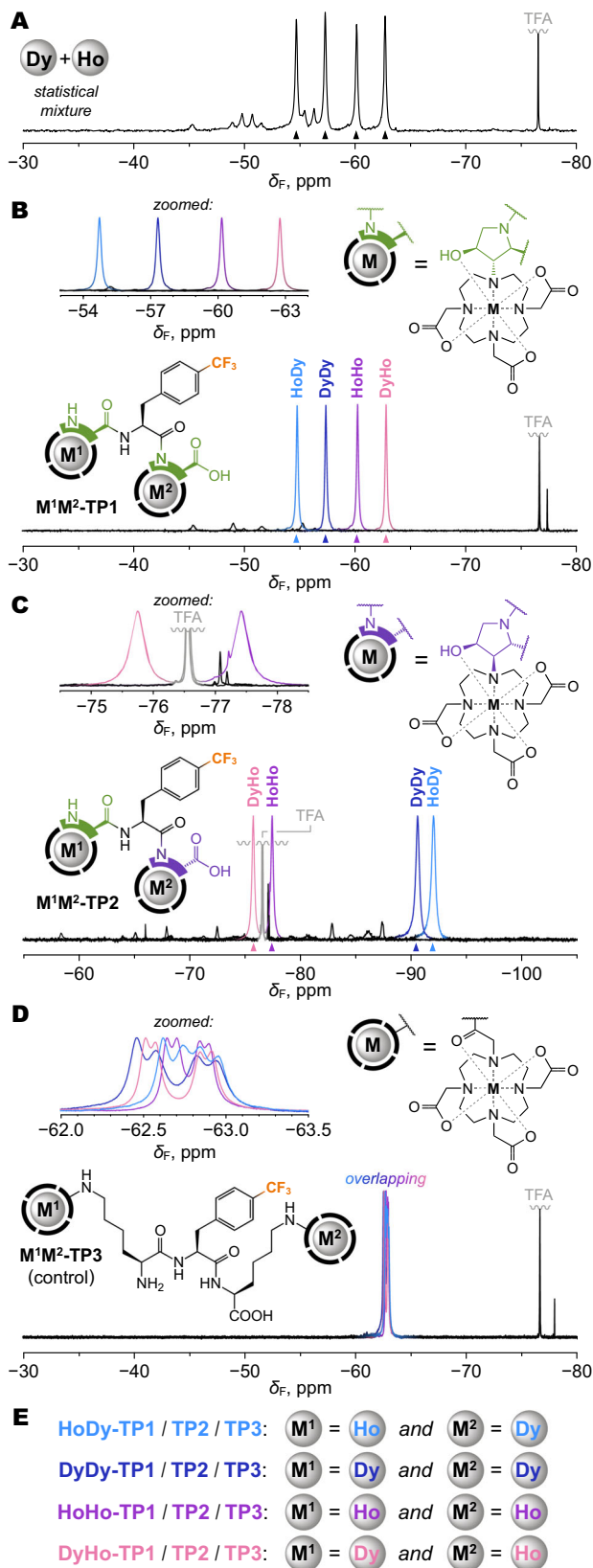
**Rules of paramagnetic encoding.** Various molecular systems have been previously explored that used a single lanthanide ion per molecule to generate discernible NMR shifts, primarily for imaging applications<sup>43–45</sup>. However, the way that magnetic susceptibility tensors of multiple lanthanide ions combine within one molecule is a rarely studied phenomenon<sup>46,47</sup>. We, therefore,

examined the rules of this in the  $M^1M^2$ -TP1 system in detail. Our hypothesis was that the observed <sup>19</sup>F chemical shift ( $\delta_F$ ) was a sum of three independent parameters:

$$\delta_F = \delta_p^1 + \delta_p^2 + \delta_d, \quad (1)$$

where  $\delta_p^1$  and  $\delta_p^2$  are individual pseudocontact paramagnetic shifts generated by metal M in a given position (1 or 2, as defined in Fig. 1), and  $\delta_d$  is a constant diamagnetic contribution.

To test this hypothesis, we first analyzed a limited dataset comprised of Dy<sup>3+</sup>/Ho<sup>3+</sup>, Dy<sup>3+</sup>/Y<sup>3+</sup> and Ho<sup>3+</sup>/Y<sup>3+</sup> statistical mixtures of  $M^1M^2$ -TP1 compounds. The diamagnetic Y<sup>3+</sup> ion was included to anchor the  $\delta_d$  component. The assignment of peaks to specific  $M^1M^2$  sequences was already partially known (Fig. 6B) and the rest was inferred or determined by trial and error. A multiparametric least-square fit according to Eq. (1) converged to a single set of  $\delta_p^1$ ,  $\delta_p^2$ ,  $\delta_d$  values. An independent dataset of Dy<sup>3+</sup>/Tm<sup>3+</sup>, Dy<sup>3+</sup>/Lu<sup>3+</sup>, Tm<sup>3+</sup>/Lu<sup>3+</sup> mixtures provided similar  $\delta_p^1$ ,  $\delta_p^2$  values for Dy<sup>3+</sup> that was present in both datasets and nearly identical  $\delta_d$  (Supplementary Fig. 14). These results demonstrated that the contributions of individual paramagnetic centers were independent and additive, confirming the hypothesis. Knowing the parameters for Dy<sup>3+</sup>, Ho<sup>3+</sup>, and Tm<sup>3+</sup> allowed us to decipher  $\delta_p^1$ ,  $\delta_p^2$  values for the remaining tested lanthanides. Final global fit provided very good agreement between the calculated and measured shifts, with residual differences not exceeding 0.34 ppm within a 38.8 ppm overall range (Supplementary Fig. 15). Because paramagnetic shifts are very sensitive to changes in molecular structure, this agreement demonstrates that the  $M^1M^2$ -TP1 compounds were practically isostructural regardless of the inserted lanthanide ions. For each element, the  $\delta_p^1$ ,  $\delta_p^2$  values were of opposite signs (Supplementary Fig. 15B), indicating substantial mutual rotation of the two magnetic susceptibility tensors in  $M^1M^2$ -TP1 molecule<sup>35</sup>. The  $\delta_p^1$  and  $\delta_p^2$  values roughly followed tabulated Bleaney constants<sup>33</sup>,



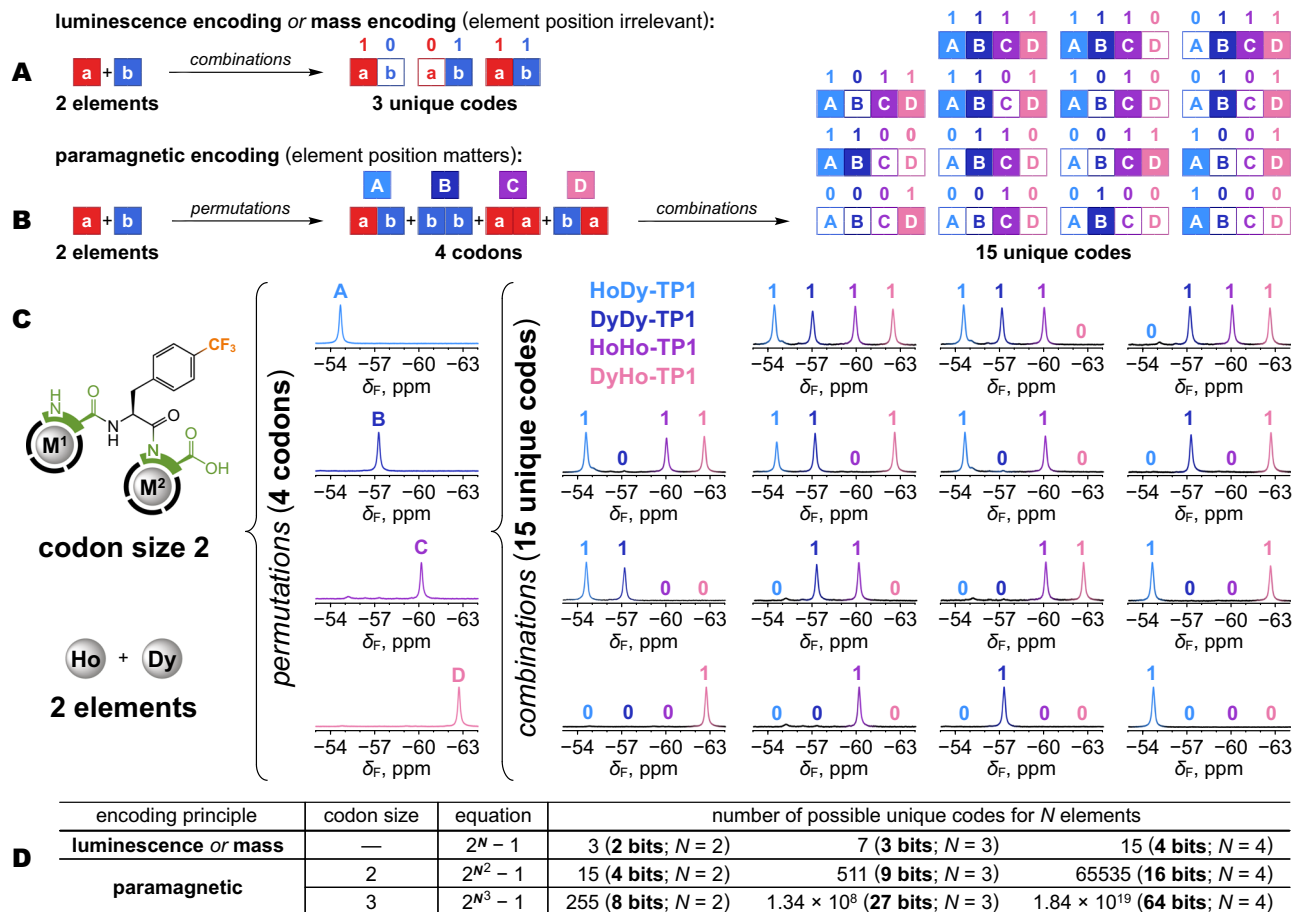
**Fig. 6**  $^{19}\text{F}$  NMR spectroscopy (470.4 MHz,  $T = 298.1\text{K}$ ) of molecules encoded with  $\text{Dy}^{3+}/\text{Ho}^{3+}$  ions. Spectra were measured in aq. MOPS/NaOH buffer (pH = 7) with external  $\text{D}_2\text{O}$  for frequency lock. TFA (trifluoroacetic acid) was used as internal NMR reference ( $-76.55$  ppm). Charges of  $\text{Ln}^{3+}$  ions were omitted for clarity. **A** Statistical mixture resulting from a post-synthetic complexation of  $\text{Dy}^{3+}/\text{Ho}^{3+}$  ions provided four major peaks (marked with triangles) with clearly distinguishable shifts representing all  $M^1M^2$  permutations of the metal ions. The mixture was measured without purification. **B**  $M^1M^2$ -TP1 compounds encoded with specific sequences of  $\text{Dy}^{3+}/\text{Ho}^{3+}$  ions, prepared via controlled synthesis, each show one major peak matching with the statistical mixture in the panel above. **C**  $M^1M^2$ -TP2 compounds encoded with the same sequences of  $\text{Dy}^{3+}/\text{Ho}^{3+}$  ions, prepared via controlled synthesis, each show one major peak, but the shifts (both absolute and relative) are very different from the analogous  $M^1M^2$ -TP1 compounds in the panels above. **D**  $M^1M^2$ -TP3 control compounds encoded with the same sequences of  $\text{Dy}^{3+}/\text{Ho}^{3+}$  ions fail to produce distinguishable peaks. **E** Legend for color-coding of  $M^1M^2$  sequences.

Alterations to the structure of the building blocks present an opportunity to optimize future molecular designs to achieve better spaced-out NMR signals with fewer overlaps.

**Sequence decoding.** Current strategies to decode sequences of nongenomic oligomers mostly rely on their sequential degradation (e.g., Edman degradation) or fragmentation with tandem MS techniques<sup>8–10</sup>, resulting in the destruction of the molecules. Non-destructive spectroscopic methods typically cannot read sequences of monomer units. This is different for paramagnetic metal ions, where the magnetic susceptibility tensors bring directionality. Metal ion sequences in the  $M^1M^2$ -TP1 system can be decoded from one-dimensional NMR spectra, provided that the  $\delta_p^1$ ,  $\delta_p^2$  components are known and the  $\delta_F$  shifts are uniquely distinguishable. The latter condition is somewhat limiting, but, nevertheless, suitable sets of elements can be identified. We found by simulation that  $\text{Tb}^{3+}$ ,  $\text{Dy}^{3+}$ ,  $\text{Ho}^{3+}$ , and  $\text{Yb}^{3+}$  can be freely combined within the  $M^1M^2$ -TP1 system, with a minimum peak distance of 0.36 ppm (Supplementary Fig. 17). Moreover, a particularly important property of paramagnetic encoding is that the basic parameters of the system (i.e.,  $\delta_p^1$ ,  $\delta_p^2$ ,  $\delta_d$  values in the case of  $M^1M^2$ -TP1) can be obtained from limited experimental data and then used to reliably simulate and identify sequences that were not yet encountered. This is significant for the design of advanced future systems, where the number of combinations may be too high to be comprehensively explored by experimentation. We tested the reliability of the predictions and sequence decoding on the sub-system of  $M^1M^2$ -TP1 compounds encoded with  $\text{Tb}^{3+}/\text{Dy}^{3+}/\text{Ho}^{3+}/\text{Yb}^{3+}$  ions. To ensure that we erased prior knowledge, we excluded from the experimental data all cases where these elements were combined in statistical mixtures. An unbiased set of  $\delta_p^1$ ,  $\delta_p^2$ ,  $\delta_d$  values was fitted from this reduced dataset (Supplementary Fig. 18). Using these parameters, we simulated shifts for all 16 permutations. The best match between the experimental  $\delta_F$  and simulated  $^{\text{sim}}\delta_F$  shifts confidently identified all 12 heteronuclear  $M^1M^2$  sequences (Supplementary Fig. 19), none of which was used to obtain the  $\delta_p^1$ ,  $\delta_p^2$ ,  $\delta_d$  values that allowed prediction of the  $^{\text{sim}}\delta_F$ .

thus showing a reasonable agreement with theory (Supplementary Fig. 16), although the accuracy of Bleaney's theoretical model has recently been questioned<sup>35</sup>. Limited synthetic yields prevented us from performing a similar extensive analysis on the  $M^1M^2$ -TP2 system, but from the comparison in Fig. 6, it is clear that the  $\delta_p^1$ ,  $\delta_p^2$  values must be very different from  $M^1M^2$ -TP1.

**Multiplexing and combinatorial implications.** Multiplexing is a way to encode information by combining unique signals into composite patterns that can be received through a single channel and decoded back to the original components. Chemical compounds are usable for this purpose by virtue of their signals being detectable by analytical methods. For example,



**Fig. 7 Molecular codes achievable with multiplexing.** **A** Methods that rely on direct detection of the elements (e.g., by luminescence or mass) provide a maximum of three unique codes from two elements. **B** Paramagnetic encoding with two elements provides four permutations (codons) with unique signals that are further multiplexed into 15 unique codes. **C** Physical realization of paramagnetic encoding and multiplexing with  $\text{Ho}^{3+}/\text{Dy}^{3+}$  ions in  $\text{M}^1\text{M}^2\text{-TP1}$  system leading to 15 unique codes. The color-coding of the NMR peaks corresponds to specific  $\text{M}^1\text{M}^2$  permutations as shown in the top middle of this panel. **D** Enumeration of codes achievable with methods of direct element reading vs. paramagnetic encoding for 2–4 elements and considering codon size of two or three elements for paramagnetic encoding. The null case (all zeros) is ignored, hence the  $-1$  term in the equations.

digital files have been encoded into mixtures of organic compounds readable by mass spectrometry<sup>19,20</sup> or fluorescence spectroscopy<sup>23</sup>. Lanthanides were previously multiplexed to generate molecular barcodes based on their unique isotope mass<sup>55</sup>, luminescence<sup>24</sup>, or paramagnetic NMR shifts<sup>45</sup>. However, decoding of such chemical systems typically proceeds with Boolean logic, considering only the presence/absence of each component, and ignoring their repetition and order. This severely limits the number of codes that can be generated, especially when the number of basic components itself is limited (e.g., chemical elements). For example, two elements can provide only three unique codes readable by luminescence or mass (Fig. 7A). Notably, this applies also to multiplexing NMR signals of molecules that carry a single lanthanide ion<sup>45</sup>. Paramagnetic encoding based on multiple lanthanides within each molecule improves the combinatorics by inserting permutations into the process. First, two elements combine into a codon, where repetitions and order are recognized, resulting in 4 permutations with unique signals. In the second step, these codons are multiplexed, providing 15 unique codes (Fig. 7B). A comprehensive physical realization of this two-level encoding is demonstrated in Fig. 7C with the  $\text{M}^1\text{M}^2\text{-TP1}$  system and  $\text{Dy}^{3+}/\text{Ho}^{3+}$  ions. Controlled synthesis was essential for the whole encoding process to work. If uncontrolled post-synthetic

complexation was used instead, the system would revert to only 3 out of the 15 codes displayed in Fig. 7C: 0010 ( $\text{Dy}^{3+}$ ), 0100 ( $\text{Ho}^{3+}$ ) and 1111 ( $\text{Dy}^{3+}/\text{Ho}^{3+}$  mixture). The combinatorial implications of the two-level encoding are very significant, as the number of unique codes grows with a double-exponential dependency on the number of elements. Paramagnetic encoding thus vastly outperforms encoding based on luminescence, mass, or other methods that read signals directly from the basic components (Fig. 7D). The current  $\text{M}^1\text{M}^2\text{-TP1}$  prototype is capable of at least 16-bit encoding (65,535 codes) with the  $\text{Tb}^{3+}/\text{Dy}^{3+}/\text{Ho}^{3+}/\text{Yb}^{3+}$  set discussed above that provides 16 distinguishable signals. Theoretically, the combinatorial power can be further improved by increasing the codon size to 3 elements (Fig. 7D). The major limitation is the number of spectroscopically distinguishable codons. However, the NMR spectral window is large in comparison e.g., to optical methods<sup>12,23,24</sup> and can accommodate hundreds of signal peaks at high magnetic fields. Moreover, even overlapping signals may be distinguishable by computational methods such as deconvolution, or thanks to different  $T_1$  or  $T_2$  relaxation times induced by the paramagnetic ions<sup>53,54</sup>. With an optimized molecular system and method of reading, 64-bit ( $\sim 10^{19}$  codes) or higher encoding is conceivable. This would be enough to cover realistic needs for molecular barcodes<sup>1,2,13</sup>. For

comparison, the current MAC (media access control) system uses 48-bit addresses for the unique identification of all devices connected to the Internet.

**Digital information encoding and reading.** To demonstrate a practical example of information encoding, we employed the system of  $M^1M^2$ -TP1 and  $Tb^{3+}/Dy^{3+}/Ho^{3+}/Yb^{3+}$  ions. All 16 possible compounds were individually synthesized and provided unique and distinguishable NMR signals (Supplementary Fig. 20). With each signal representing one bit (present = 1, absent = 0), it is possible to encode 16 bits of digital data by multiplexing the compounds. Standard ASCII (American Standard Code for Information Interchange) characters use 7 bits, so a single mixture can carry two characters. That is not enough for practical purposes. Therefore, we organized five different mixtures within a capillary into layers that were spatially separated by an immiscible signal-free solvent  $CCl_4$  (Fig. 8A). Thus, a composite NMR sample was created with the capacity to store ten ASCII characters, enough to encode a reasonably strong password. To read back the information, we used *z*-resolved NMR spectroscopy to acquire spectra of all layers simultaneously (Fig. 8B). After referencing the chemical shifts in each layer to an internal standard (trifluoroacetic acid, TFA), the presence/absence of the signals was evaluated (Fig. 8C) and converted to a binary code (Fig. 8D). Due to the lower spectral resolution of *z*-resolved NMR, we avoided the confusion of the two least-resolved peaks ( $DyDy/YbTb$ , Fig. 8C) by shifting the beginning of the binary code to one of them ( $DyDy$ ), thus eliminating it by definition (binary codes of all standard ASCII characters start with 0, Fig. 8E). Conversion of the binary code to characters revealed the encoded 10-character password (Fig. 8F, G). It is important to note that multiple pairs of positional isomers (e.g.,  $YbHo/HoYb$ ) present in the mixtures make the encoded information practically unreadable with methods other than magnetic resonance since the isomers are very difficult to distinguish by other means. This is a very attractive feature especially for potential counterfeit applications of such molecular codes.

**Parallel reading with MRI.** A key aspect of reading and communicating information is whether it proceeds in a sequential or parallel way. Parallel systems offer higher and more scalable throughput, which is also suitable for imaging applications. For example, optical reading methods capture information simultaneously from many objects in the field of view, while MS can only process one sample at a time. Magnetic resonance can provide both spectroscopic and spatial information simultaneously, thus allowing parallel reading and imaging. Spectrally resolved MRI has been previously realized with various systems based on chemically different diamagnetic compounds<sup>38–41</sup>, host-guest interactions<sup>42,45</sup>, and lanthanide-induced NMR shifts<sup>43–45</sup>. The latter approach offers the advantage to select the NMR shifts without changing the chemical properties, which is particularly useful for in-vivo imaging applications<sup>43,44</sup>. However, previous systems based on a single lanthanide per molecule offered only a very limited choice. The programmable  $M^1M^2$ -TP1 system provides an advanced level of control to generate a higher number of practically usable NMR signals. To demonstrate this, we used just two lanthanide ions  $Dy^{3+}/Ho^{3+}$  to encode the  $M^1M^2$ -TP1 molecules, which were then pipetted into a  $7 \times 5$ -well plate in patterns to write one letter with each compound. Each well thus contained a homogenous mixture of 0–4 compounds (Fig. 9A). The whole well plate was imaged on a preclinical 4.7 T MRI scanner with a  $^{19}F$  CSI (Chemical Shift Imaging) pulse sequence, where two dimensions represent space coordinates and the third is a frequency domain. The information contained in the 35 wells

was obtained simultaneously, the encoded molecules were distinguished owing to their different resonance frequencies (Fig. 9B), and the four letters could be resolved and independently displayed (Fig. 9C–F). A fly-through view of the MRI data is provided in Supplementary Movie 1.

**Limit of detection.** NMR spectroscopy is typically not considered as a method for reading molecular codes because of its relatively low sensitivity compared to e.g., mass spectrometry. This prompted us to investigate the limit of detection achievable with  $M^1M^2$ -TP1 molecules encoded with  $Dy^{3+}/Ho^{3+}$  ions. Previous works have demonstrated that shortening of longitudinal relaxation times ( $T_1$ ) induced by paramagnetic ions allows to use a rapid scan rate and boost sensitivity by acquiring more signal per time<sup>53,54</sup>. We took advantage of this effect in  $M^1M^2$ -TP1 compounds (Supplementary Fig. 21). The signal-to-noise ratio (SNR) in NMR grows with the square root of the number of scans, but there are practical limits to the experiment time. We extrapolated the detection limit at 11.7 T and  $SNR = 3$  to be 5.8 nmol with a short measurement time (1.6 s, 32 scans), or 123 pmol with a reasonably long measurement time (60 min, 72,000 scans). Paramagnetic codes are therefore conceivably usable as identification tags in bead-based bioassays (a single 90- $\mu$ m TentaGel bead can hold 100 pmol). The detection limit could be further lowered by technical as well as chemical means, e.g., by optimization of the NMR hardware and pulse sequence, optimizing the molecular system to provide narrower NMR signals (higher SNR), or including a higher number of magnetically equivalent  $^{19}F$  nuclei. Although the conventional NMR/MRI technology is large and expensive, in principle it can be re-designed for the purpose of reading paramagnetic molecular codes. Downsizing using cost-efficient permanent magnets can dramatically lower the cost and potentially improve the sensitivity to match a particular application<sup>56,57</sup>.

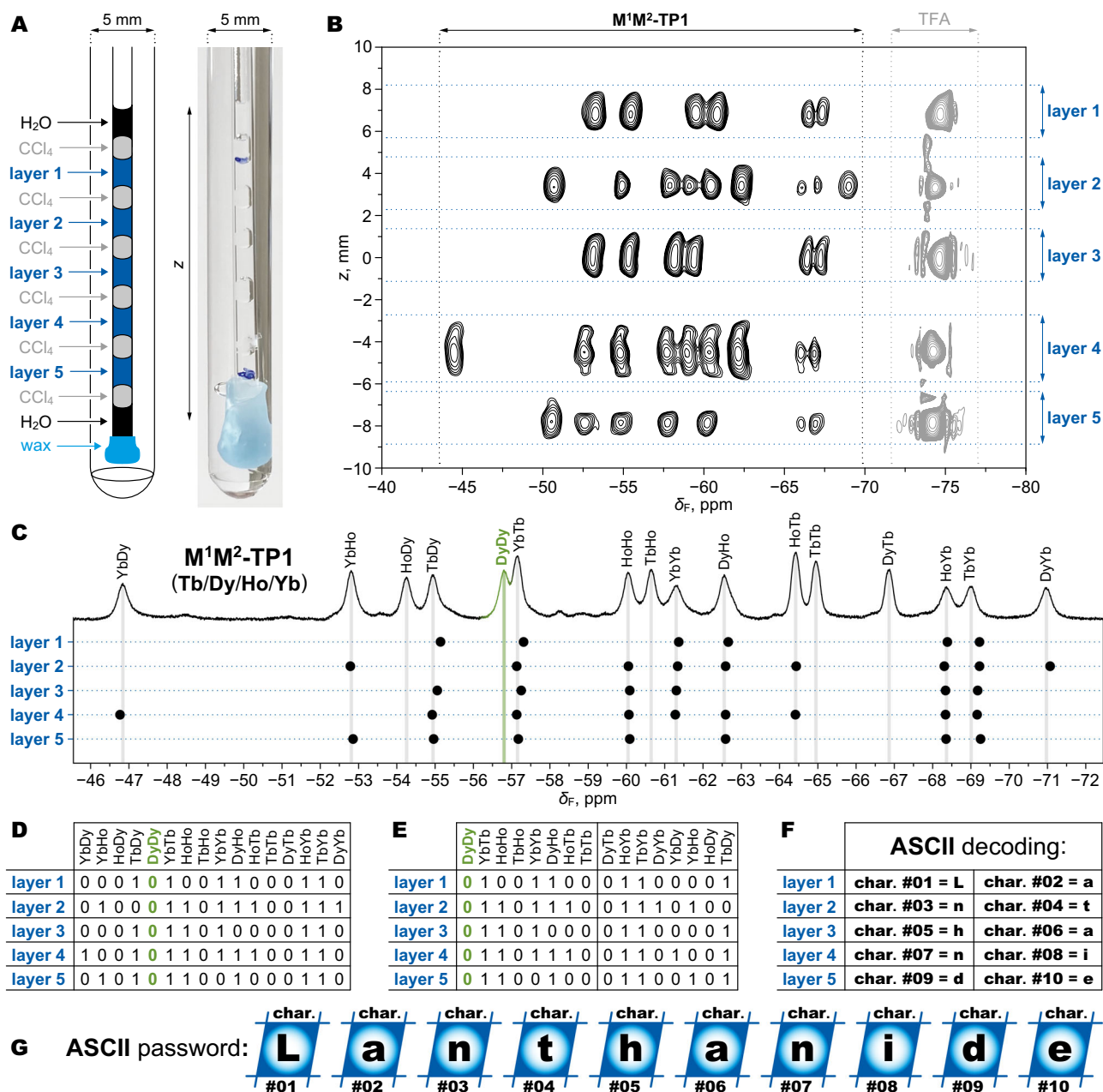
## Discussion

Properties of paramagnetically encoded molecules depart from typical chemical systems and approach macroscopic hardware and software. As the encoded information is hidden in chemically inaccessible *f*-electrons of lanthanides, the hardware (molecule) remains physically unchanged (isostructural) when loaded with different software (metal ion sequence). Simple and clear programming rules, which can be calibrated from a small number of synthesized examples, provide predictable behavior of the molecular code. Reading in the RF spectrum draws parallels with the technology of radiofrequency identification (RFID) tags by being repeatable and non-destructive, and by working through space without a direct line of sight. The paramagnetic codes presented here are nearly chemically identical, based on elements with negligible biological background (lanthanides, fluorine), and structurally related to clinically approved MRI contrast agents. These are favorable qualities for tagging biological objects or artificial objects in the biological milieu. We refer to the  $M^1M^2$ -TP1 and  $M^1M^2$ -TP2 systems as prototypes, because many aspects of their modular construction can be further optimized to build more advanced systems with enhanced functionality. The major roadblock to immediate practical use is the low sensitivity of reading with conventional NMR technology. However, rapid advancements in quantum sensors may soon make the detection of single molecules a routine reality<sup>58,59</sup>, opening the door to as-yet unforeseen applications.

## Methods

**Synthesis of building blocks, tripeptides, and their metal chelates.** Detailed synthetic strategies and procedures, including the conditions for the separation of isomers and characterization of all products, are provided in Supplementary Information. All syntheses proceeded in solution. Some lanthanide ions were excluded





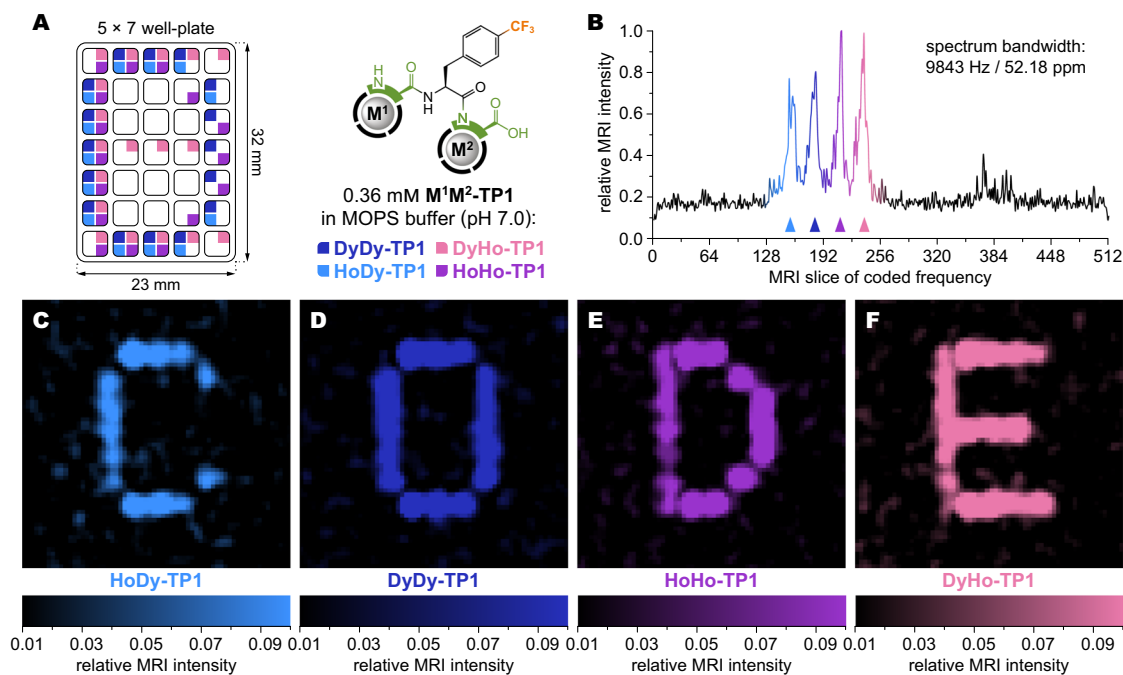
**Fig. 8 Password decoding from a single sample with z-resolved NMR.** **A** Physical realization of a single NMR sample containing 5 different mixtures of M<sup>1</sup>M<sup>2</sup>-TP1 compounds encoded with Tb<sup>3+</sup>/Dy<sup>3+</sup>/Ho<sup>3+</sup>/Yb<sup>3+</sup> ions. Aqueous solutions (in MOPS/NaOH buffer, pH = 7) of the mixtures (layers 1–5) are separated by layers of CCl<sub>4</sub>. The capillary is submerged in D<sub>2</sub>O. **B** z-resolved <sup>19</sup>F NMR spectrum (470 MHz, T = 298.1 K) of the sample from panel A shows signals of the M<sup>1</sup>M<sup>2</sup>-TP1 compounds and internal standard (TFA, trifluoroacetic acid) resolved for the layers along the z-axis. Shifts from different layers are not perfectly aligned due to bulk magnetic susceptibility (BMS) effects that varied for each layer. **C** Evaluation of the signals from panel B by comparison with an independently obtained 1D NMR spectrum of a mixture containing all 16 M<sup>1</sup>M<sup>2</sup>-TP1 compounds. Note that peaks of DyDy and YbTb are close but distinguished in the 1D spectrum. Black dots mark shifts from the z-resolved spectrum after correcting for BMS effects in each layer (referencing to TFA). Good match is achieved between expected shifts (vertical lines) and determined shifts (black dots). **D** Binary representation of the information from C. **E** Final table of the binary codes for 10 ASCII (American Standard Code for Information Interchange) characters encoded in the sample. Relative to D, the binary codes are rearranged to start at the position of DyDy (green shading), which is not used to avoid confusion with YbTb. **F** ASCII characters obtained by conversion from the binary codes in E with indication of their order in a password. **G** The final decoded 10-character password.

from testing for various reasons (La<sup>3+</sup>: redundant diamagnetic, Ce<sup>3+</sup>: 3 + /4+ redox chemistry, Pm<sup>3+</sup>: radioactive, Gd<sup>3+</sup>: inducing no paramagnetic shift). Concentrations of Ln<sup>3+</sup> ions in stock solutions for synthesis and in solutions of M<sup>1</sup>M<sup>2</sup>-TP1, M<sup>1</sup>M<sup>2</sup>-TP2, and M<sup>1</sup>M<sup>2</sup>-TP3 compounds were determined with ICP-AES.

**X-ray crystallography.** Detailed procedures for preparation of the single crystals of [Dy(L<sup>1</sup>)]·3.5H<sub>2</sub>O and [Dy(L<sup>2</sup>)]·3H<sub>2</sub>O, details of structure solving and

refinement, as well as additional details of the obtained structures are provided in Supplementary Information (Supplementary Methods and Supplementary Figs. 1–3, 5–7, 65, 97).

**<sup>19</sup>F NMR spectroscopy.** Measurements of all <sup>19</sup>F NMR spectra were done on Avance III<sup>TM</sup> HD 500 MHz spectrometer (Bruker, 470.4 MHz for <sup>19</sup>F) equipped with a broad-band cryo-probe with ATM module (5 mm CPBBO BB-<sup>1</sup>H/<sup>19</sup>F/<sup>15</sup>N/D Z-



**Fig. 9** Image encoding and reading with  $^{19}\text{F}$  MRI (4.7 T). **A** Distribution pattern of  $M^1M^2$ -TP1 compounds encoded with  $\text{Dy}^{3+}/\text{Ho}^{3+}$  ions within a 7 × 5-well plate (final concentrations  $\sim 0.36$  mM each compound, well volumes completed to 75  $\mu\text{L}$  with water). **B**  $^{19}\text{F}$  MR spectrum obtained from the entire well plate shows four distinct peaks of the four encoded compounds (marked with triangles, color-coding as explained in A). The spectrum was created by projecting the maximum signal intensity from each slice of (CSI) MRI data onto the frequency dimension. The spectral resolution was 19.2 Hz (0.10 ppm) per slice. **C-F** Selective display of  $\text{HoDy-TP1}$  (C),  $\text{DyDy-TP1}$  (D),  $\text{HoHo-TP1}$  (E) and  $\text{DyHo-TP1}$  (F) compounds reveals the patterns encoded with them into the well plate as legible letters (average of 15 adjacent slices for each compound peak). Displayed field of view: 50 × 50 mm, voxel size 0.78 × 0.78 × 10 mm. Image intensities were mapped to colors as shown in the legends. See also Supplementary Movie 1 for a fly-through view of the MRI data.

GRD). Solutions of samples were prepared in MOPS/NaOH buffer (pH = 7) and placed into an insert capillary that was inserted into a 5-mm NMR tube containing  $\text{D}_2\text{O}$  (used for NMR frequency lock). The volume of solution in the insert observable by the spectrometer probe was  $\sim 50$   $\mu\text{L}$ . Longitudinal relaxation times ( $T_1$ ) were obtained by inversion recovery sequence. Spectra of statistical mixtures or individually synthesized  $M^1M^2$ -TP1 and  $M^1M^2$ -TP2 compounds were measured with these parameters: 90° observation pulse, pre-acquisition delay = 1 s, acquisition time = 0.58 s and 512 or 4096 scans. Measurements of SNR were performed on a sample containing a mixture of  $\text{HoDy-TP1}$ ,  $\text{DyDy-TP1}$ ,  $\text{HoHo-TP1}$ , and  $\text{DyHo-TP1}$  compounds, each at molar amount of 0.18  $\mu\text{mol}$  in the volume observable by the NMR probe (50  $\mu\text{L}$ ). Parameters for SNR measurements were: pre-acquisition delay = 0, acquisition time = 50 ms, receiver gain maximum value = 203. The SNR values were obtained using *MestReNova* software (version 12.0.3) after the application of an exponential window function of 1 Hz. The limit of detection (SNR = 3) was extrapolated from SNR measured with 32 scans (1.6 s total time), considering that SNR scales linearly with the analyte concentration and with the square root of the number of scans. Further details for SNR measurements are provided in Supplementary Fig. 21.  $^{19}\text{F}$  z-resolved NMR spectrum was acquired using a phase encoding pulse sequence with resolution 2048 × 64 points at 20 mm field of view, number of scans = 32 K, acquisition time = 18 ms, relaxation delay = 100  $\mu\text{s}$  (for technical reasons). The excitation pulse was  $\sim 31^\circ$ , which was the result of optimization aiming to enhance the SNR of weaker (broader) signals, sacrificing the SNR of sharper peaks. This was possible due to the differences in  $T_1$  relaxation times. The total experimental time was  $\sim 12$  h. For subsequent processing, only 700 × 64 raw points were used in order to reduce the noise. Data were apodized using QSINE window functions in both F2 (SSB = 10) and F1 (SSB = 1) dimensions and zero-filled to 8192 × 128 points. Processing was done in *Bruker TopSpin 3.5* software. Spectra were visualized using the *NumPy* and *Matplotlib Python* libraries.

**Magnetic resonance imaging.** Stock solutions of individually synthesized  $\text{HoDy-TP1}$ ,  $\text{DyDy-TP1}$ ,  $\text{HoHo-TP1}$ , and  $\text{DyHo-TP1}$  compounds in 0.5 M MOPS/NaOH buffer (pH = 7.0) were pipetted into a 7 × 5-well plate (cut out of standard 384-well plate) and the volume in each well was completed with water to 75  $\mu\text{L}$  to reach compound concentrations  $\sim 0.36$  mM each. The well plate was covered with plastic tape and imaged on 4.7 T scanner *Bruker Biospec 47/20 (Bruker BioSpin, Ettlingen, Germany)* with a custom-built dual  $^1\text{H}/^{19}\text{F}$  RF surface coil. First,  $^1\text{H}$  MRI was performed in all three planes (axial, coronal, and sagittal) for localization of the sample. Then,  $^{19}\text{F}$  NMR spectra were acquired by 90° single pulse sequence to precisely determine resonance frequencies of the

compounds.  $^{19}\text{F}$  MRI images were obtained using a CSI sequence (120,000 scans, repetition time = 700 ms, field of view = 50 × 50 mm; slice thickness = 10 mm, matrix of acquired image = 16 × 16 × 512, matrix of reconstructed image = 64 × 64 × 512, acquisition time = 23 h 20 min, resonance frequency = 188,630,130 Hz, bandwidth = 9843 Hz/52.18 ppm). Images for display were prepared in *Matlab* software. First, the data in the Bruker format were imported to *Matlab* using *read\_2dseq* function (by Cecil Yen (2021), [https://www.mathworks.com/matlabcentral/fileexchange/69177-read\\_2dseq-quickly-reads-brucker-s-2dseq-mri-images](https://www.mathworks.com/matlabcentral/fileexchange/69177-read_2dseq-quickly-reads-brucker-s-2dseq-mri-images)). Maximum intensity projection along the frequency axis was used to find the peaks of the respective samples. Background signal (outside of main peaks) was calculated as an average of 41 slices (# 10–50 perpendicular to the frequency axis) and subtracted from each slice of the original data. Then, the intensity was normalized to 1 for the highest intensity. The final images displayed in Fig. 9 were obtained by averaging 15 slices (perpendicular to the frequency axis) around the maximum for each peak, and mapping the image intensity to specific color maps. Supplementary Movie 1 was created analogously, with each frame showing a floating average of 15 slices.

**Data fitting.** Fitting of experimental shifts  $\delta_{\text{F}}$  from statistical mixtures of  $M^1M^2$ -TP1 compounds to obtain  $\delta_{\text{p}}^1$ ,  $\delta_{\text{p}}^2$ ,  $\delta_{\text{d}}$  contributions was done according to Eq. (1) with the method of least squares using the *Solver* function in *Microsoft Excel*. Sequences of  $\text{Dy}^{3+}/\text{Ho}^{3+}$  synthesized as individual compounds were the starting point for the assignment of  $\delta_{\text{F}}$  to particular  $M^1M^2$  sequences. Next, homonuclear sequences were identified based on their repeated occurrences between mixtures containing the particular element. The assignment of heteronuclear sequences then proceeded simultaneously with the fitting, one statistical mixture at a time. If a satisfactory fit could not be obtained, it indicated that the assignment was wrong and the positions in  $M^1M^2$  sequences were exchanged. This was continued until consensus was reached across all data in a final global fit. The paramagnetic shift contributions of the  $\text{Sm}^{3+}$  ion were too small to be obtained simultaneously with the other lanthanide ions. Therefore,  $\delta_{\text{p}}^1$ ,  $\delta_{\text{p}}^2$  parameters for  $\text{Sm}^{3+}$  were fitted separately from the  $\text{Sm}^{3+}/\text{Y}^{3+}$  statistical mixture of  $M^1M^2$ -TP1 compounds and were then used as fixed parameters in the global fit.

#### Data availability

The crystallographic data for the structures reported in this study have been deposited at the Cambridge Crystallographic Data Centre (CCDC) under deposition number 2072548 for  $[\text{Dy}(\text{L}^1)] \cdot 3.5\text{H}_2\text{O}$  and 2143481 for  $[\text{Dy}(\text{L}^2)] \cdot 3\text{H}_2\text{O}$  and can be obtained free of charge

from the Centre via its website ([www.ccdc.cam.ac.uk/getstructures](http://www.ccdc.cam.ac.uk/getstructures)). All other data supporting the findings in this study are available within the article and its Supplementary Information, as well as from the corresponding author upon request.

### Code availability

No custom computer code was used in this study.

Received: 18 June 2021; Accepted: 10 May 2022;

Published online: 08 June 2022

### References

- Birtwell, S. & Morgan, H. Microparticle encoding technologies for high-throughput multiplexed suspension assays. *Integr. Biol.* **1**, 345–362 (2009).
- Binan, L., Drobtzky, E. A. & Costantino, S. Exploiting molecular barcodes in high-throughput cellular assays. *SLAS Technol. Transl. Life Sci. Innov.* **24**, 298–307 (2019).
- Grass, R. N., Heckel, R., Puddu, M., Paunescu, D. & Stark, W. J. Robust chemical preservation of digital information on DNA in silica with error-correcting codes. *Angew. Chem. Int. Ed.* **54**, 2552–2555 (2015).
- Koch, J. et al. A DNA-of-things storage architecture to create materials with embedded memory. *Nat. Biotechnol.* **38**, 39–43 (2020).
- Qian, J. et al. Barcoded microbial system for high-resolution object provenance. *Science* **368**, 1135–1140 (2020).
- Doroschak, K. et al. Rapid and robust assembly and decoding of molecular tags with DNA-based nanopore signatures. *Nat. Commun.* **11**, 5454 (2020).
- Yang, K., McCloskey, C. M. & Chaput, J. C. Reading and writing digital information in TNA. *ACS Synth. Biol.* **9**, 2936–2942 (2020).
- Lee, J. M. et al. High-density information storage in an absolutely defined aperiodic sequence of monodisperse copolyester. *Nat. Commun.* **11**, 56 (2020).
- Laurent, E. et al. High-capacity digital polymers: storing images in single molecules. *Macromolecules* **53**, 4022–4029 (2020).
- König, N. F. et al. Photo-editable macromolecular information. *Nat. Commun.* **10**, 3774 (2019).
- Al Ouahabi, A., Charles, L. & Lutz, J.-F. Synthesis of non-natural sequence-encoded polymers using phosphoramidite chemistry. *J. Am. Chem. Soc.* **137**, 5629–5635 (2015).
- Zhang, J. et al. Rapidly sequence-controlled electrosynthesis of organometallic polymers. *Nat. Commun.* **11**, 2530 (2020).
- Rassay, S., Ramezani, M., Shomaji, S., Bhunia, S. & Tabrizian, R. Clandestine nanoelectromechanical tags for identification and authentication. *Microsyst. Nanoeng.* **6**, 1–8 (2020).
- Baumbauer, C. L. et al. Printed, flexible, compact UHF-RFID sensor tags enabled by hybrid electronics. *Sci. Rep.* **10**, 16543 (2020).
- Zaeimbashi, M. et al. Ultra-compact dual-band smart NEMS magnetoelectric antennas for simultaneous wireless energy harvesting and magnetic field sensing. *Nat. Commun.* **12**, 3141 (2021).
- Shi, C. et al. Application of a sub-0.1-mm3 implantable mote for in vivo real-time wireless temperature sensing. *Sci. Adv.* **7**, eabf6312 (2021).
- Won, S. M., Cai, L., Gutruf, P. & Rogers, J. A. Wireless and battery-free technologies for neuroengineering. *Nat. Biomed. Eng.* **1**–19 <https://doi.org/10.1038/s41551-021-00683-3> (2021).
- Rivnay, J., Wang, H., Fenno, L., Deisseroth, K. & Malliaras, G. G. Next-generation probes, particles, and proteins for neural interfacing. *Sci. Adv.* **3**, e1601649 (2017).
- Arcadia, C. E. et al. Multicomponent molecular memory. *Nat. Commun.* **11**, 691 (2020).
- Cafferty, B. J. et al. Storage of information using small organic molecules. *ACS Cent. Sci.* **5**, 911–916 (2019).
- Rosenstein, J. K. et al. Principles of information storage in small-molecule mixtures. *IEEE Trans. NanoBioscience* **19**, 378–384 (2020).
- Ratner, T., Reany, O. & Keinan, E. Encoding and processing of alphanumeric information by chemical mixtures. *ChemPhysChem* **10**, 3303–3309 (2009).
- Nagarkar, A. A. et al. Storing and reading information in mixtures of fluorescent molecules. *ACS Cent. Sci.* **7**, 1728–1735 (2021).
- Kreidt, E., Leis, W. & Seitz, M. Direct solid-phase synthesis of molecular heterooligonuclear lanthanoid-complexes. *Nat. Commun.* **11**, 1346 (2020).
- Valm, A. M. et al. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017).
- Shi, H. et al. Highly multiplexed spatial mapping of microbial communities. *Nature* **588**, 676–681 (2020).
- Geiss, G. K. et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **26**, 317–325 (2008).
- Han, M., Gao, X., Su, J. Z. & Nie, S. Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules. *Nat. Biotechnol.* **19**, 631–635 (2001).
- Zhang, F. et al. Rare-earth upconverting nanobarcodes for multiplexed biological detection. *Small* **7**, 1972–1976 (2011).
- Gorris, H. H. & Wolfbeis, O. S. Photon-upconverting nanoparticles for optical encoding and multiplexing of cells, biomolecules, and microspheres. *Angew. Chem. Int. Ed.* **52**, 3584–3600 (2013).
- Lee, J. et al. Universal process-inert encoding architecture for polymer microparticles. *Nat. Mater.* **13**, 524–529 (2014).
- Coronado, E. Molecular magnetism: from chemical design to spin control in molecules, materials and devices. *Nat. Rev. Mater.* **5**, 87–104 (2020).
- Bleaney, B. et al. Origin of lanthanide nuclear magnetic resonance shifts and their uses. *J. Chem. Soc. Chem. Commun.* 791–793 <https://doi.org/10.1039/C3972000791B> (1972).
- Parigi, G., Ravera, E. & Luchinat, C. Magnetic susceptibility and paramagnetism-based NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **114–115**, 211–236 (2019).
- Parker, D., Sutturina, E. A., Kuprov, I. & Chilton, N. F. How the ligand field in lanthanide coordination complexes determines magnetic susceptibility anisotropy, paramagnetic NMR shift, and relaxation behavior. *Acc. Chem. Res.* **53**, 1520–1534 (2020).
- Häussinger, D., Huang, J. & Grzesiek, S. DOTA-M8: an extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 14761–14767 (2009).
- Wahsner, J., Gale, E. M., Rodríguez-Rodríguez, A. & Caravan, P. Chemistry of MRI contrast agents: current challenges and new frontiers. *Chem. Rev.* **119**, 957–1057 (2019).
- McMahon, M. T. et al. New “multicolor” polypeptide diamagnetic chemical exchange saturation transfer (DIACEST) contrast agents for MRI. *Magn. Reson. Med.* **60**, 803–812 (2008).
- Liu, G. et al. In vivo multicolor molecular MR imaging using diamagnetic chemical exchange saturation transfer liposomes. *Magn. Reson. Med.* **67**, 1106–1113 (2012).
- Lim, Y. T. et al. Multiplexed imaging of therapeutic cells with multispectrally encoded magnetofluorescent nanocomposite emulsions. *J. Am. Chem. Soc.* **131**, 17145–17154 (2009).
- Akazawa, K. et al. Perfluorocarbon-based 19F MRI nanoprobe for in vivo multicolor imaging. *Angew. Chem. Int. Ed.* **57**, 16742–16747 (2018).
- Yang, Y. et al. Coloring ultrasensitive MRI with tunable metal-organic frameworks. *Chem. Sci.* **12**, 4300–4308 (2021).
- Ferrauto, G., Castelli, D. D., Terreno, E. & Aime, S. In vivo MRI visualization of different cell populations labeled with PARACEST agents. *Magn. Reson. Med.* **69**, 1703–1711 (2013).
- Harvey, P. et al. Moving the goal posts: enhancing the sensitivity of PARASHIFT proton magnetic resonance imaging and spectroscopy. *Chem. Sci.* **4**, 4251–4258 (2013).
- Goren, E., Avram, L. & Bar-Shir, A. Versatile non-luminescent color palette based on guest exchange dynamics in paramagnetic cavitands. *Nat. Commun.* **12**, 3072 (2021).
- Ishikawa, N., Iino, T. & Kaizu, Y. Study of 1H NMR spectra of dinuclear complexes of heavy lanthanides with phthalocyanines based on separation of the effects of two paramagnetic centers. *J. Phys. Chem. A* **107**, 7879–7884 (2003).
- Polovkova, M. A. et al. Determination of the structural parameters of heteronuclear (phthalocyaninato)bis(crownphthalocyaninato)lanthanide(III) triple-deckers in solution by simultaneous analysis of NMR and single-crystal X-ray data. *Inorg. Chem.* **55**, 9258–9269 (2016).
- Sørensen, T. J. & Faulkner, S. Multimetallic lanthanide complexes: using kinetic control to define complex multimetallic arrays. *ACC Chem. Res.* **51**, 2493–2501 (2018).
- Brückner, K., Zitterbart, R., Seitz, O., Beck, S. & Linscheid, M. W. Solid phase synthesis of short peptide-based multimetal tags for biomolecule labeling. *Bioconjug. Chem.* **25**, 1069–1077 (2014).
- De León-Rodríguez, L. M., Kovacs, Z., Dieckmann, G. R. & Sherry, A. D. Solid-phase synthesis of DOTA-peptides. *Chem. - Eur. J.* **10**, 1149–1155 (2004).
- Boros, E., Polasek, M., Zhang, Z. & Caravan, P. Gd(DOTA)la: a single amino acid Gd-complex as a modular tool for high relaxivity MR contrast agent development. *J. Am. Chem. Soc.* **134**, 19858–19868 (2012).
- Woods, M., Kovacs, Z., Zhang, S. & Sherry, A. D. Towards the rational design of magnetic resonance imaging contrast agents: isolation of the two coordination isomers of lanthanide DOTA-type complexes. *Angew. Chem. Int. Ed.* **42**, 5889–5892 (2003).
- Chalmers, K. H., Botta, M. & Parker, D. Strategies to enhance signal intensity with paramagnetic fluorine-labelled lanthanide complexes as probes for 19F magnetic resonance. *Dalton Trans.* **40**, 904–913 (2011).

54. Kislukhin, A. A. et al. Paramagnetic fluorinated nanoemulsions for sensitive cellular fluorine-19 magnetic resonance imaging. *Nat. Mater.* **15**, 662–668 (2016).
55. Bodenmiller, B. et al. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat. Biotechnol.* **30**, 858–867 (2012).
56. Michal, C. A. Low-cost low-field NMR and MRI: instrumentation and applications. *J. Magn. Reson.* **319**, 106800 (2020).
57. Cooley, C. Z. et al. Design and implementation of a low-cost, tabletop MRI scanner for education and research prototyping. *J. Magn. Reson.* **310**, 106625 (2020).
58. Cujia, K. S., Boss, J. M., Herb, K., Zopes, J. & Degen, C. L. Tracking the precession of single nuclear spins by weak measurements. *Nature* **571**, 230–233 (2019).
59. Schwartz, I. et al. Blueprint for nanoscale NMR. *Sci. Rep.* **9**, 6938 (2019).

## Acknowledgements

This work was supported by the Czech Science Foundation (17-22834Y, M.P.). Additional support is acknowledged from the Student Grant Scheme at the Technical University of Liberec (SGS-2022-3002, M.V.), and from the Ministry of Health of the Czech Republic - conceptual development of research organization (Institute for Clinical and Experimental Medicine-İKEM, IN 00023001, D.J.). The authors thank to Kelsea Grace Jones for language editing, Veronika Urbanová for proof-reading and Adam Jaroš for molecular modelling.

## Author contributions

M.P. conceived the project and supervised the work. J.K. synthesized the compounds. R.J. and M.K. designed and conducted the synthesis of the epoxide precursor. M.D. and O.S. conducted NMR measurements. D.J. and M.V. conducted MRI experiments and constructed a radiofrequency coil. I.C. performed X-ray diffraction and solved the structures. T.D. created the artwork. M.P., J.K., and T.D. wrote the manuscript. All authors proofread and approved the manuscript.

## Competing interests

M.P. and J.K. are co-inventors on a pending patent application no. PCT/CZ2020/050032 filed in the name of applicant Ustav Organické Chemie a Biochemie AV ČR V.V.I. The application covers some of the compounds discussed in this work. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-30811-9>.

**Correspondence** and requests for materials should be addressed to Miloslav Polasek.

**Peer review information** *Nature Communications* thanks Carlos Platas-Iglesias and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022