Chirality- and sequence-selective successive self-sorting via specific homo- and complementary-duplex formations

Wataru Makiguchi¹, Junki Tanabe¹, Hidekazu Yamada¹, Hiroki Iida¹, Daisuke Taura¹, Naoki Ousaka¹ & Eiji Yashima¹

Self-recognition and self-discrimination within complex mixtures are of fundamental importance in biological systems, which entirely rely on the preprogrammed monomer sequences and homochirality of biological macromolecules. Here we report artificial chirality- and sequence-selective successive self-sorting of chiral dimeric strands bearing carboxylic acid or amidine groups joined by chiral amide linkers with different sequences through homo- and complementary-duplex formations. A mixture of carboxylic acid dimers linked by racemic-1,2-cyclohexane bis-amides with different amide sequences (NHCO or CONH) self-associate to form homoduplexes in a completely sequence-selective way, the structures of which are different from each other depending on the linker amide sequences. The further addition of an enantiopure amide-linked amidine dimer to a mixture of the racemic carboxylic acid dimers resulted in the formation of a single optically pure complementary duplex with a 100% diastereoselectivity and complete sequence specificity stabilized by the amidinium-carboxylate salt bridges, leading to the perfect chirality- and sequence-selective duplex formation.

¹Department of Molecular Design and Engineering, Graduate School of Engineering, Nagoya University, Chikusa-ku, Nagoya 464-8603, Japan. Correspondence and requests for materials should be addressed to E.Y. (email: yashima@apchem.nagoya-u.ac.jp).
High-fidelity self-recognition and self-discrimination are of principal importance in living systems, which enable biological macromolecules, such as proteins and DNA, to self-organize into uniform quaternary and double helical structures with a controlled handedness, respectively, through noncovalent interactions even within complex mixtures of subunits or molecular strands with a similar shape, size and sequence, thereby providing sophisticated functions that are essential for human life. Such an incredible self-sorting performance observed in biological systems ultimately relies on the preprogrammed monomer sequences and homochirality of their building blocks and components. Recent advances in the total chemical synthesis of unnatural ¿-proteins and ¿-DNA clearly revealed the indispensable role of the homochirality along with the monomer sequences of natural ¿-proteins and ¿-DNA towards chiral specificities in entanoselective reactions and complementary double-helix formations, respectively.

In organic and supramolecular chemistry, the control of the size and shape, and topology, as well as the handedness of supramolecular assemblies from mixtures of a variety of components via spontaneous self-sorting has become one of the urgent and emerging topics because it will not only enrich our understanding of the principles underlying the precise recognition behaviour of biological macromolecules but also contribute to the development of novel supramolecular catalysts, optoelectrical devices and sensors. In most cases, however, self-sorting has been achieved in a size- and shape-, and topology-selective way through the formation of macrocycles, capsules or cages, and helices, while a limited number of supramolecular systems undergoes chiral self-sorting producing either homochiral or heterochiral assemblies, which is mostly relevant to the biological processes as well as the state-of-the-art asymmetric catalysis and optical resolution on crystallization.

Here we report artificial chirality- and sequence-selective chiral self-sorting of dimeric strands consisting of carboxylic acid or amidine groups joined by chiral amide linkers with different sequences (NHCO or CONH) through specific homoduplex and subsequent hetero- (complementary-) duplex formations with implications for biological double helices in DNA. DNA-like double helices have a significant advantage in the precise recognition of the monomer sequence and chain length between the complementary strands. Although hydrogen-bond-driven peptide nucleic acids and metal-coordinate-bonded helicates are known to show self-sorting with respect to their sequences and/or chain lengths, chirality- and sequence-selective successive self-sorting is currently unknown.

Results
Design and synthesis of carboxylic acid and amidine dimers.
Our molecular design is mainly based on the previously reported heterodouble helices composed of complementary dimer strands with m-terphenyl backbones intertwined through amidinium–carboxylate salt bridges (Fig. 1a, AA·CC). Owing to the high tolerance of the salt bridges towards various functional groups, a variety of linkers (L) could be introduced while maintaining the double-stranded helical structures with a one-handed helical sense biased by the chirality introduced on the amidine residues. During the course of our study, we also found that an achiral carboxylic acid dimer joined by a p-diethynylbenzene linker with n-octyl (R®) substituents self-associated into a racemic homodouble helix through interstrand hydrogen bonds between the carboxy groups (Fig. 1a, (CC®)).

We anticipated that either a right- or left-handed homodouble helix would be induced in the carboxylic acid dimers when chiral linkers, such as (S,S)- or (R,R)-trans-1,2-cyclohexane-based (cHex) bis-amide derivatives, were introduced between the monomer units (1 and 2, Fig. 1b). Incidentally, we have found a complete sequence-selective chiral self-sorting in a mixture of racemic (rac)-1a and rac-2a mediated by unique interstrand multihydrogen-bond-driven homoduplex formations, the structures of which are significantly different from our expected self-associated homodouble helices, and are further different from each other depending on the linker amide sequences (NHCO–cHex (1L) or CONH–cHex (1L), Fig. 1c). The further addition of an enantiopure amide-linked amidine dimer (3 or 4) with a particular linker amide sequence to a mixture of rac-1a and rac-2a results in the formation of a single optically pure complementary duplex with a 100% diastereoselectivity and sequence specificity stabilized by the amidinium–carboxylate salt bridges (Fig. 1c).

A series of chiral dimers of carboxylic acids (1, 2) and amidines (3, 4) linked through chiral or meso amide residues with different sequences was prepared according to the reported methods (see Supplementary Methods).
(Δδ/ΔT (ppb per °C)) in DMSO-d6 (−5.8 and −4.9) than those in CDCl3 (−2.0 and −3.1; Supplementary Figs 5–8)34.

The X-ray crystallographic analyses of analogous (S,S)-1a′ and -2a′ whose structures are assumed to be almost identical to those of (S,S)-1a and -2a, respectively, except for the pendant and/or terminal substituents (see Fig. 1b), revealed a unique deeply intertwined duplex structure (Fig. 2c,d). Unexpectedly, these structures are completely different from a self-associated homodouble helix-like CC2 (see Fig. 1a). Each strand adopts a similar ‘U-shape’ structure resulting from the (S,S)-trans-1,2-cyclohexyl linker residue, binding together through the eight interstrand hydrogen bonds, in which each amide group at the linker moiety is sandwiched between the two carboxy groups of the other strand. The remarkable structural difference between (S,S)-1a′ and -2a′ in the solid state is that the homoduplex ((S,S)-2a′)2 has a well-packed structure due to the almost parallel orientation of the strands (Fig. 2d), while the intertwined strands of (S,S)-1a′ are oriented perpendicular to each other, resulting in a less sterically hindered structure (Fig. 2c). This difference in the steric hindrance between ((S,S)-1a′)2 and ((S,S)-2a′)2 is reasonably correlated with the difference in their self-association affinities; the Kd value of (S,S)-1a is ~105 times greater than that

Figure 1 | Chirality- and sequence-selective successive self-sorting. (a) Structures of heterodouble helix (AA·CC) and homodouble helix (CC). (b) Structures of carboxylic acid and amidine dimer strands linked through chiral or meso amide residues (X = L1 or L2). (c) Schematic representation of chirality- and sequence-selective successive self-sorting via specific homo- and complementary-duplex formations. Carboxylic acid dimers of rac-1a and rac-2a with different linker amide sequences form homo/hetero- (i) and homochiral (ii) duplexes, respectively. A mixture of rac-1a and rac-2a completely selfsorts in sequence-selective way to give duplexes of 1a and 2a (iii). Only one enantiomer among rac-1a and rac-2a (S,S)- or (R,R)-2a forms a complementary duplex with an amidine dimer linked through the same linker amide chirality and sequence to that of 1a or 2a via salt bridges, leading to perfect chirality- and sequence-selective duplex formation (iv).
of (S,S)-2a. The 1H two-dimensional (2D) nuclear Overhauser effect spectroscopy (NOESY) analysis of (R,R)-1a and (S,S)-2a in CDCl3 (Supplementary Figs 9–13) showed characteristic interstrand NOE cross-peaks including those between the linker cHex protons and the terminal aromatic protons of the m-terphenyl moieties along with the interstrand aromatic protons, indicating that such unique intertwined structures in the solid state were retained in solution. The observed large downfield shifts of the amide NH and CO2H protons of (S,S)-1a and (R,R)-2a in their 1H NMR spectra (Fig. 2b) also support the structures.

We next investigated the chiral self-sorting behaviour of rac-1a and rac-2a using 1H NMR in CDCl3 at 25 °C. The 1H NMR spectrum of rac-1a showed two sets of signals with an integral ratio of 2:1 (0% enantiomeric excess (e.e.), Fig. 3a), which did not coalesce even at 50 °C (Supplementary Fig. 14), whereas a single set of signals was observed in DMSO (Supplementary Fig. 15), in which the diastereomeric duplex formation is sensitive to temperature and its molar ratio changed on heating or cooling (Supplementary Fig. 14c). The amide NH resonances were assigned using 1H 2D NMR spectroscopy (Supplementary Figs 11–13). (c,d) X-ray crystal structures of (S,S)-1a (c) and (S,S)-2a (d). The diastereomeric excess (d.e.) values (%) of 1a (homochiral duplex versus heterochiral duplex) were then plotted versus the % e.e. of 1a (Fig. 3a). Therefore, the minor signals are unambiguously assigned to the heterochiral duplex (R,R)-1a · (S,S)-1a.

The diastereomeric excess (d.e.) values (%) of 1a (homochiral duplex versus heterochiral duplex) were then plotted versus the % e.e. of 1a, which matched well with the simulated curve using the equation \( Y = 100\{ -2 \times 4 - 3(X/100)^2 \}^{0.5} + 5/3 \) (inset in Fig. 3b; equation A) obtained by a modification of the reported equation \( Y = 100\{ -2 \times 4 - 3(X/100)^2 \}^{0.5} \) (for details, see Supplementary Methods), suggesting no preference in the diastereomeric duplex formation, thus forming the homochiral and heterochiral 1a2 duplexes in a 2:1 molar ratio in CDCl3 at 25 °C. As anticipated, however, the diastereomeric duplex formation is sensitive to temperature and its molar ratio changed on heating or cooling (Supplementary Fig. 14d), being attributed to a subtle change in the relative interstrand hydrogen-bonding strength between the homo- and heterochiral 1a2, which was in accordance with the observed differences in their amide NH temperature coefficients (Supplementary Fig. 14c).

In sharp contrast, rac-2a was completely chiral self-sorted to form only the homochiral duplexes (R,R)-2a2 and (S,S)-2a2, giving one set of 1H NMR signals in CDCl3 independent of the % e.e. of 2a (Fig. 3c) and temperature from −20 to 50 °C.
that amide linkers with a different sequence, in which the average derived from (equation A), where X and Y are % e.e. of 1a and % d.e. of 1a, respectively (for details, see Supplementary Methods). (c) 1H NMR spectra of carboxylic acid CO$_2$H and amide NH proton resonances of 2a (0.50 mM) with different % e.e. in CDCl$_3$ at 25°C. The plots were fitted using $Y = 100 - (24 - 3\times100)^{2.055} + 5.07/3$ (equation A), where X and Y are % e.e. of 1a and % d.e. of 1a, respectively (for details, see Supplementary Methods). (d) Partial 1H NMR spectra (0.50 mM) of (S,S)-2a, (S,S)-2c, an equimolar mixture (1.0 mM) of (S,S)-2a and (S,S)-2c, and an equimolar mixture (1.0 mM) of (R,R)-2a and (S,S)-2c in CDCl$_3$ at 25°C. Red circles denote the signals because of the duplex (S,S)-2a (S,S)-2c. Full-scale spectra are shown in Supplementary Fig. 20.

The difference in the chiral self-sorting behaviour between 1a and 2a may be due to the difference in the twist angles between the two rigid diphenylethylene groups connecting to the chiral amide linkers with a different sequence, in which the average twist angle in ((S,S)-2a)$_2$ (ca. 48°) is much larger than that in ((S,S)-1a)$_2$ (ca. 18°) in the solid state (Fig. 2c,d). The density functional theory (DFT) calculations revealed that the homochiral duplex ((S,S)-1a) is only 17.0 kJ mol$^{-1}$ more stable than the heterochiral duplex (R,R)-1a · (S,S)-1a, whereas, for 2a, the homochiral duplex formation ((S,S)-2a) is much more favourably than the heterochiral duplex formation (R,R)-2a · (S,S)-2a by 51.7 kJ mol$^{-1}$ (Supplementary Fig. 21a–d), being reasonably consistent with the previously discussed experimental results. The observed large energy difference (AE) between the homochiral and heterochiral duplexes for 2a (−51.7 kJ mol$^{-1}$) as compared with that for 1a (−17.0 kJ mol$^{-1}$) is attributed to a distorted structure of the heterochiral homoduplex (R,R)-2a · (S,S)-2a that lacks one interstrand hydrogen bond between the amide NH and carboxy C=O groups (Supplementary Fig. 21d), while all of the possible interstrand hydrogen bonds are retained in both the homo- and heterochiral 1a, during the calculations because of the perpendicular orientation of the strands as seen in the solid state (Fig. 2c), resulting in a similar duplex structure (Supplementary Fig. 21a,b). Therefore, predominant homochiral self-sorting could not take place for rac-1a (Fig. 3a).

Of further interest is that an equimolar mixture of rac-1a and rac-2a showed complete self-sorting in terms of their linker amide sequences, giving only the enantiomeric pairs of the 1a and 2a, as explicitly observed in their 1H NMR spectra in CDCl$_3$ (Fig. 4a). Each duplex of 1a and 2a is unable to exchange, that is further supported by the 1H NMR and CD measurements of an equimolar mixture of (S,S)-1a and (S,S)-2a in CDCl$_3$ (Fig. 4a,b); the 1H NMR spectrum of an equimolar mixture of (S,S)-1a and (S,S)-2a in CDCl$_3$ displayed only two sets of the signals corresponding to ((S,S)-1a) and ((S,S)-2a) (Fig. 4a). Moreover, the mixture showed a CD spectrum precisely identical to the simulated one (Fig. 4b), indicating that the complete sequence-selective chiral self-sorting took place within mixtures of rac-1a and rac-2a. The DFT calculations demonstrate the important role of the interstrand hydrogen bonds; the energy-minimized duplex structure of (S,S)-1a · (S,S)-2a with the different linker amide sequences has a mismatched arrangement of the interstrand hydrogen bonds (Supplementary Fig. 21e),
and is 11.1 and 32.6 kJ mol$^{-1}$ less stable than those of the chiral self-sorted homoduplexes ((S,S)-2a), and ((S,S)-1a)$_2$, respectively (Supplementary Fig. 21a,c). On the basis of the total energies of the calculated structures (Supplementary Fig. 21), the stabilities of the duplexes decrease in the following order: ((S,S)-1a)$_2$ > (R,R)-1a · (S,S)-1a > ((S,S)-2a)$_2$ > (S,S)-1a · (S,S)-2a > (R,R)-2a · (S,S)-2a, which are in good agreement with the experimental results, and the energetically disfavoured duplexes (S,S)-1a · (S,S)-2a and (R,R)-2a · (S,S)-2a were not detected at all under the present experimental conditions.

Chirality- and sequence-selective heteroduplex formation. With all the above results taken together with our previous findings of the complementary double-helix formations through amidinium–carboxylate salt bridges$^{44,47–50}$, we envisaged that the optically active amidine dimers linked through the chiral or active amidine dimers linked through the chiral or meso amide linkers with a different sequence (NHCO-cHex (L1) or CONH-cHex (L2)) (3a–c and 4a–c, Fig. 1b) could selectively recognize the carboxylic acid dimers (1a and 2a) according to the linker chirality and sequences, thereby leading to a unprecedented diastereoo- and sequence-selective complementary duplex formation.

Figure 4 | Sequence-selective chiral self-sorting behaviour of rac-1a and rac-2a. (a) $^1$H NMR spectra of carboxylic acid CO$_2$H and amide NH proton resonances of rac-1a (1.0 mM), rac-2a (1.0 mM), an equimolar mixture (1.0 mM) of rac-1a and rac 2a (S,S)-1a (0.50 mM), (S,S)-2a (0.50 mM) and an equimolar mixture (1.0 mM) of (S,S)-1a and (S,S)-2a in CDCl$_3$ at 25°C. Full-scale spectra are shown in Supplementary Fig. 22. (b) CD and absorption spectra of (S,S)-1a (0.50 mM), (S,S)-2a (0.50 mM) and those observed and simulated for the mixture of an equimolar amount of (S,S)-1a and (S,S)-2a (0.50 mM) in CDCl$_3$ at 25°C.
On mixing rac-2a and (R,R,R,R,R)-3a composed of the dimeric (R,R,R) amine linked through the (R,R)-L2 linker in a 2:1 molar ratio in CDCl$_3$, the $^1$H NMR spectrum immediately changed to that consisting of complementary diastereomeric duplexes of 3a-(R,R)-2a and 3a-(S,S)-2a together with the remaining 2a (Supplementary Fig. 25). The nonequivalent N–H proton signals appeared at a low magnetic field (13–15.5 p.p.m.), suggesting the preferred-handed duplex formation stabilized by salt bridges$^{44,47–50}$. Because the chain exchange rate between the complementary dimer strands is slower than the NMR timescale$^{48}$, the diastereoselectivity between (R,R,R,R,R)-3a and rac-2a was estimated to be d.e. = 58% ([(R,R)-2a]-rich, run 1, Table 1) from their integral ratio, indicating the homochiral (R,R) selectivity with respect to their cHex linker chirality. This diastereoselectivity was further confirmed by CD; the observed CD spectrum is almost identical to the simulated CD (Supplementary Fig. 26).

Interestingly, the amine dimer (R,R,S,S,R,R)-3b bearing the opposite (S,S)-L2 linker showed a perfect diastereoselectivity towards rac-2a, producing the complementary duplex of 3b-(S,S)-2a with d.e.>$99\%$ as well as (R,R)-2a (e.e.>$99\%$) remaining as a homochiral duplex, as evidenced by the $^1$H NMR and identical experimental and simulated CD spectra (run 2, Table 1, Fig. 5a–c and Supplementary Figs 27 and 28). It should be noted that highly disfavoured 3b-(R,R)-2a was hardly formed even in the presence of a large excess of (R,R)-2a (12 equivalents), indicating that the binding affinity of (S,S)-2a to 3b is at least 10$^5$ times higher than that of (R,R)-2a (Supplementary Fig. 29). This surprisingly high diastereoselectivity$^{57,58}$ of 3b towards rac-2a enabled the separation of rac-2a into the (S,S)- and (R,R)-2a enantiomers by facile silica gel column chromatography, giving the corresponding optically pure enantiomers in 85% and 63% yield, respectively (Supplementary Fig. 30). (R,R,S,S,R,R)-3b could be recovered for further separation of rac-2a. Moreover, the resolved optically pure (R,R)-2a (e.e.>$99\%$) could also be used to separate rac-3b in principle. More practically, the chemical bonding of (R,R,S,S,R,R)-3b to chromatographic supports including silica gel will enable more efficient and preparative separation of rac-2a as well as other enantiomers as a novel chiral stationary phase (CSP) for chromatographic enantioseparation. The complementary 1:1 duplex formation of 3b-(S,S)-2a was further evidenced using an electron-spray ionization mass spectrometric (ESI-MS) measurement and vapour pressure osmometry experiment (Supplementary Figs 31–33).

More interestingly, (R,R)-3c and (R,R)-3d, in which the chiral amidine residues were replaced by achiral isopropyl (3c) and cyclohexyl (3d) substituents, respectively, also performed a perfect diastereoselectivity towards rac-2a to form duplexes only with (R,R)-2a (d.e.>$99\%$) having the same linker chirality, which also resulted in the optically pure (S,S)-2a (e.e.>$99\%$) as the remaining homochiral duplex (runs 3, 4, Table 1 and Supplementary Figs 34–37), suggesting that the diastereoselective duplex formation is governed by the linker chirality, and the chiral amidine residues may be no longer required. In other words, an anticipated complementary double-stranded helix formation stabilized by chiral amidinium–carboxylate salt bridges that would occur on both sides of the linker may not be prerequisite in order to achieve diastereoselective duplex formations. Therefore, the amine dimer having the chiral amidine residues linked through achiral meso-linker (R,R,meso,R,R)-3c completely lost its diastereoselectivity (d.e. = 0%; run 5, Table 1 and Supplementary Figs 38 and 39).

Various attempts to obtain crystals of the complementary duplex dimers suitable for an X-ray analysis produced only amorphous solids. Therefore, the energy-minimized structures of the duplexes of (R,R,S,S,R,R)-3b-(S,S)-2a and (R,R,S,S,R,R)-3b-(R,R)-2a were constructed using the semiempirical molecular orbital calculations followed by the DFT calculations based on an analogous crystal structure$^{47}$. The initial model structures of (R,R,S,S,R,R)-3b-(S,S)-2a and its diastereomer (R,R,S,S,R,R)-3b-(R,R)-2a with a right-handed twisted conformation, in which the pendant 1-octyl groups are replaced by hydrogen atoms, were constructed so as to satisfy the following experimental results: (1) as shown in their $^1$H NMR spectra (Fig. 5d and Supplementary Fig. 62), all of the amide NH resonances of (R,R,S,S,R,R)-3b-(S,S)-2a showed significant downfield shifts compared with those of the amide strand (R,R,S,S,R,R)-3b (Δδ = 0.53 or 0.73 p.p.m.; Table 2, entry 2), indicating that all of the amide protons participate in hydrogen bonds. In contrast, the amide NH resonances of (R,R,S,S,R,R)-3b-(R,R)-2a were slightly shifted downfield (Δδ = 0.12 or 0.09 p.p.m.), suggesting weak hydrogen bonds. (2) An analogous complementary dimeric duplex composed of the identical (R,R)-amine and carboxylac acid dimer strands linked by diacetylene residues bound together through salt bridges was determined to have a right-handed helical structure using the single-crystal X-ray analysis$^{47}$.

The resultant energy-minimized structures of (R,R,S,S,R,R)-3b-(S,S)-2a and (R,R,S,S,R,R)-3b-(R,R)-2a with their total energies are depicted in Fig. 6a,b, respectively, which revealed that they take a largely bent-shaped, right-handed double-helix-like structure, in which (R,R,S,S,R,R)-3b-(S,S)-2a is 57.4 kJ mol$^{-1}$ more stable than the other. All of the linker amide protons in 3b-(S,S)-2a form inter- or intramolecular hydrogen bonds with the amide C=O groups (average NH···O

| Table 1 | Results of diastereoselective duplex formations between chiral amidine dimers (3a–4c) and racemic carboxylic acid dimers (rac-1a and rac-2a). |
|-----------------|-----------------|-----------------|
| Amidine dimer (amide sequence) | rac-2a (CONH-chex) | rac-1a (NHCO-chex) |
| Run | d.e. (%)$^*$ | Suppl.$^*$ Fig. no. | Run | d.e. (%)$^*$ | Suppl.$^*$ Fig. no. |
| (R,R,R,R,R)-3a (CONH-cHex) | 1 | 58 (R,R) | 6 | 14 (S,S) | 40 |
| (R,R,S,S,R)-3b (CONH-cHex) | 2 | >99 (S,S) | 7 | 64 (R,R) | 42 |
| (R,R)-3c (CONH-chex) | 3 | >99 (R,R) | 8 | 64 (S,S) | 44 |
| (R,R)-3d (CONH-chex) | 4 | >99 (R,R) | 9 | 68 (S,S) | 46 |
| (R,R,meso,R,R)-3e (CONH-chex) | 5 | 70 (S,S) | 10 | 0 (R,R) | 48 |
| (R,R,S,S,R)-4a (NHCO-chex) | 11 | 70 (S,S) | 14 | 64 (R,R) | 56 |
| (R,R,S,S,R)-4b (NHCO-chex) | 12 | 34 (R,R) | 15 | 74 (S,S) | 58 |
| (R,R)-4c (NHCO-chex) | 13 | 80 (S,S) | 16 | 70 (R,R) | 60 |

$^*$Estimated using $^1$H NMR (CDCl$_3$, 25°C).

$^*$Supplementary Fig. no.
The hydrogen-bonding networks between average NH/C1/C1/C1 (3b) structures reasonably explain the present unexpectedly high distance (2.30 Å; Fig. 6b). The observed 2D NOESY spectra of 3b (0.50 mM) showed interstrand NOE cross-peaks including those between the terminal TMS protons and the phenyl moieties of the amidine residues along with the interstrand aromatic protons, while interstrand NOEs between the linker cHex protons were not identified because of the same linker amide sequence (NHCO-cHex) whose chemical shifts were quite similar to each other (Supplementary Figs 68–70). On the other hand, an interstrand NOE was clearly observed between the linker cHex protons of 4b - (R,R)-2a because of the different linker amide sequences (Supplementary Figs 71 and 72). Considering all the results including the 2D NOESY, ESI-MS, vapour pressure osmometry, salt-bridge formations and large downfield shifts of TMS protons were different from that of rac-2a because of the different linker amide sequences (Supplementary Figs 40–47).

The 2D NOESY spectra of 3b (0.50 mM) and two equivalents of rac-2a (1.0 mM) in CDCl3 at ambient temperature. For the simulated CD and absorption spectra, see Supplementary Methods. (d) Partial 1H NMR spectra (0.5 mM) of 3b, 3b - (R,R)-2a and 3b - (S,S)-2a in CDCl3 at 25 °C. The linker amide NH resonances were assigned using 1H 2D NMR spectroscopy (Supplementary Fig. 62).
Table 2 | Relationships between the diastereoselective complementary duplex formations and downfield chemical shifts ($\Delta \delta$) of the linker amide NH resonances of the amidine strands of 3b·1a–4b·2a duplexes from the monomeric amidine strands (3b and 4b).

| Entry | Combination | Duplex | $\Delta \delta_{\text{NH}}$ (p.p.m.) | d.e. (%) | Supplementary Fig no. |
|-------|-------------|--------|-----------------------------------|----------|-----------------------|
| 1     | (R,R,S,S,R),3b (CONH-cHex) and 1a (NHCO-cHex) | 3b·(R,R)-1a favoured, 3b·(S,S)-1a disfavoured | 0.56 | 64 ((R,R)-1a-rich) | 64 |
|       |             | 4b·(R,R)-2a disfavoured | 0.03 | | |
| 2     | (R,R,S,S,R),3b (CONH-cHex) and 2a (CONH-cHex) | 3b·(R,R)-2a disfavoured, 3b·(S,S)-2a disfavoured | 0.12 or 0.09* | > 99 ((S,S)-2a-rich) | 62 |
|       |             | 4b·(R,R)-1a disfavoured, 4b·(S,S)-1a disfavoured | 0.73 or 0.53* | | |
| 3     | (R,R,S,S,R),4b (NHCO-cHex) and 1a (NHCO-cHex) | 4b·(R,R)-1a disfavoured, 4b·(S,S)-1a disfavoured | 0.52 | 74 ((S,S)-1a-rich) | 67 |
|       |             | 4b·(R,R)-2a disfavoured, 4b·(S,S)-2a disfavoured | 1.16 or 1.08* | | |
| 4     | (R,R,S,S,R),4b (NHCO-cHex) and 2a (CONH-cHex) | 4b·(R,R)-2a disfavoured, 4b·(S,S)-2a disfavoured | 0.90 | 34 ((R,R)-2a-rich) | 66 |

*cHex, (RR)-trans-1,2-cyclohexane.

*The amide NH resonances of the amidine strands of duplexes could not be distinguished from those of the carboxylic acid strands because of overlapping with the phenyl and aliphatic proton signals. For the assignments of the amide NH proton signals, see Supplementary Figs 62–67.

Figure 6 | Structures of complementary duplexes. Capped-stick drawings of the structures for the homochiral duplex (R,R,S,S,R),3b·(S,S)-2a (a) and the heterochiral duplex (R,R,S,S,R),3b·(R,R)-2a (b) with respect to the cHex linker chirality optimized by DFT calculations. DFT-calculated energies are also shown in the bottom.
duplex with (S,S)-1a in 64% d.e., which is in significant contrast to the homochiral selectivity observed between 3a–3d and rac-2a (runs 1–4, Table 1). Again, the linker chirality plays a critical role so that (R,R,meso,R,R)-3e showed no diastereoselectivity (run 10, Table 1 and Supplementary Figs 48 and 49). The reason for this heterochiral preference between 3a-d and rac-1a is not totally understood, but may be due to the difference in the interstrand hydrogen-bonding networks that could be more efficiently formed between the heterochiral linker amide residues than between the homochiral counterparts. This speculation is supported by the fact that the amide NH resonances of, for instance, the heterochiral (R,R,SS,S,R)-3b·(R,R)-1a showed higher downfield shifts (Δδ = 0.56 p.p.m.) as compared with those of the homochiral (R,R,SS,S,R)-3b·(S,S)-1a (Δδ = 0.03 p.p.m.), respectively (Supplementary Fig. 64 and Table 2, entry 1).

We also investigated the diastereoselective duplex formations using a series of amidine dimers (4a–4c) with the NHCO-CH₃ linker amide sequence towards rac-1a and rac-2a. In all the duplex formations, moderate diastereoselectivities (64–80% d.e.) were observed except for that between 4b and rac-2a (34% d.e.; runs 11–16, Table 1 and Supplementary Figs 50–61), which were in contrast to the perfect diastereoselectivities achieved between 3b–3d and rac-2a. As for the selectivities with regard to the cHex amide linker chirality, there is the same tendency; 4a–4c favourably formed duplexes with 1a and 2a with the same and opposite configurations, respectively. As a typical example, (R,R)-4c preferentially formed a duplex with (R,R)-1a, but with (S,S)-2a in 70 and 80% d.e., respectively. The observed diastereoselectivities also rely on the difference in the interstrand hydrogen-bond strengths between the diastereomeric duplexes as revealed by more downfield shifts of the amide NH protons for the major duplex diastereoselectively formed during the complexations (Supplementary Figs 65–67 and Table 2, entries 3 and 4).

On the basis of these results, we anticipated that particular chiral amidine dimers would simultaneously recognize the chirality and sequence within a mixture of complementary carboxylic acid dimers via specific duplex formations. In fact, the mixing of two equivalents of each of rac-1a and rac-2a with (R,R,SS,S,R)-3b in CDCl₃ resulted in the formation of only the (R,R,SS,S,R)-3b·(S,S)-1a duplex (d.e. >99%), while no duplex formation was observed towards rac-1a. As a result, an optically pure (R,R)-2a (e.e. >99%) quantitatively remained as a

**Figure 7 | Chirality- and sequence-selective complementary duplex formation.** (a) Partial ¹H NMR spectra of rac-2a (0.50 mM), rac-1a (0.50 mM), 3b·(R,R)-2a (0.50 mM), 3b·(S,S)-2a (0.50 mM), 3b·(R,R)-1a (0.50 mM), 3b·(S,S)-1a (0.50 mM) and a mixture of 3b (0.50 mM) and two equivalents of rac-2a (1.0 mM) and rac-1a (1.0 mM) in CDCl₃ at 25°C. Full-scale spectra are shown in Supplementary Fig. 73. (b) CD (upper) and absorption (bottom) spectra (0.50 mM) of 3b·(S,S)-2a, (R,R)-2a and rac-1a in CDCl₃ under ambient temperature. (c) Experimental and simulated CD (d.e. = 100%) and absorption spectra for a mixture of 3b (0.50 mM) and two equivalents of each of rac-1a (1.0 mM) and rac-2a (1.0 mM) in CDCl₃ under ambient temperature.
homochiral duplex together with free rac-1a (Fig. 7); thus, the perfect sequence- (CONH-CHex or NHCO-CHex) and chirality- (R,R) or (S,S) selective duplex formation was achieved as unambiguously evidenced using their 1H NMR analysis (Fig. 7a) and the observed and simulated CD spectra (Fig. 7b,c). The linker chirality rather than the amidine chirality also plays a vital role in the sequence- and chirality-selective complementary duplex formation; therefore, (R,R)-3d composed of achiral amidine residues formed a duplex only with (R,R)-2a with a complete diastereoselectivity (d.e. >99%) and sequence specificity in a mixture of rac-1a and rac-2a (Supplementary Figs 74 and 75). Discussion

We demonstrate here an unprecedented successive chiral self-sorting during the unique homo- and subsequent heteroduplex formations through interstrand multihydrogen bonds that take place in a perfect sequence-selective way accompanied with an extraordinary high diastereoselectivity in the latter case, which enables the separation of the racemic strands into the extraordinary high diastereoselectivity in the latter case, therefore, the sequence- and chirality-selective complementary duplex formation; and the present chirality- and sequence-selective successive self-sorting is most likely achieved because of the rigid chiral geometries generated by the chiral amide-linked dimeric strands that arrange functional groups in a suitable way for specific duplex formations.

In our previous studies, we showed that a carboxylic acid dimer such as CC formed an intertwined homodouble helix via interstrand association of the carboxylic acids61, which further formed a double-stranded helix (AA·CC) with the complementary amide dimer such as AA (Fig. 1a)44,47–49. The helical handedness of the complementary double helix is fully controlled by the chirality introduced on the amidine residues49. On the basis of the present studies, however, it appears that the chiral and sequence (NHCO-CHex or CONH-CHex) of the amide linkers of dimeric carboxylic acid and amide strands are of primary importance and dictate the overall chirality- and sequence-selective self-assemblies of the dimer strands, resulting from the unique multihydrogen-bonding networks formed between the linker amide residues. Importantly, the amidine chirality is not almost involved in the present diastereoselective complementary duplex formations that mostly rely on the sequence and chirality of the amide linkers of the amidine dimers relative to those of the carboxylic acid dimers.

The dimers of amides and carboxylic acids possess reactive trimethylylthyl vinyl groups at the ends, which allows longer oligomers joined by chiral amide linkers with specific sequence and chirality to be synthesized, which would provide a unique strategy for asymmetric template synthesis and an artificial replication system39,60 based on chirality- and sequence-selective multihydrogen-bond-assisted duplex formations, and also provide a clue towards a better understanding of the biological chiral self-sorting process.

Methods

General procedures for the complementary duplex formations. A typical procedure for the diastereoselective duplex formations between amide and carboxylic acid dimers is described below. Stock solutions of (R,R,R,R,R)-3a (2.0 mM; solution I) and rac-2a (2.0 mM; solution II) were prepared in dry CDCl3. Aliquots of I (0.40 mol, 200 µl), II (0.80 mol, 400 µl) and dry CDCl3 (200 µl) were added to an NMR tube, and its 1H NMR spectrum was measured at 25 °C to determine the % d.e. of the 3a·2a duplexes (Supplementary Fig. 25). The solution was also used for measuring the CD (Supplementary Fig. 26). In a similar way, other diastereoselective duplex formations were preformed and their % d.e. values were estimated (Table 1).

Optical resolution of rac-2a via the hetoduodeplex formation. (R,R,R,R,R,R)-3b (5.44 mg, 3.58 µmol) and two equivalents of rac-2a (9.46 mg, 17.6 µmol) were dissolved in CHCl3 (2 ml), which produced an equimolar mixture of 3b·(S,S)-2a and (R,R)-2a judging from its 1H NMR spectrum (Fig. 5a) and were separated into the first and second fractions, respectively, using flash column chromatography (SiO2, 2 cm (i.d.) × 12 cm; eluent: CHCl3/Methanol = 1/0–50/1, v/v). The % d.e. of the 3b·(S,S)-2a duplex obtained (6.3 mg) was >99% as estimated using its CD and 1H NMR spectra (Supplementary Fig. 30a,c), while the second fraction mainly contained (R,R)-2a along with a small amount of 3b·(S,S)-2a and an unknown compound probably generated during the column chromatography. Thus, the second fraction was purified again using flash column chromatography, yielding 3b·(S,S)-2a (2.3 mg, >99% d.e.; total yield: 8.6 mg; 85%) and (R,R)-2a containing a small amount of the unknown compound. Further purification using recyle size-exclusion chromatography (eluent: CHCl3) afforded pure (R,R)-2a (2.99 mg) in 63% yield; its e.e. value was estimated to be >99% on the basis of the CD and 1H NMR measurements (Supplementary Fig. 30b,c).

References

1. Berg, J. M., Tymoczko, J. L. & Stryer, L. Biochemistry 7th ed (W. H. Freeman, 2007).
2. Kuriyan, J., Kontforti, B. & Wemmer, D. The Molecules of Life: Physical and Chemical Principles (Garland Science - Taylor & Francis Group, 2012).
3. Rowan, A. E. & Nolte, R. J. M. Helical molecular programming. Angew. Chem. Int. Ed. 37, 63–68 (1998).
4. Milton, R. C. D., Milton, S. C. F. & Kent, S. B. H. Total chemical synthesis of a D-enzyme: the enantiomers of HIV-1 protease show demonstration of reciprocal chiral substrate-specificity. Science 256, 1445–1448 (1992).
5. Manzini, K. et al. Chemical synthesis and X-ray structure of a helical D-protein antagonist plus vascular endothelial growth factor) protein complex by racemic crystallography. Proc. Natl Acad. Sci. USA 109, 14779–14784 (2012).
6. Garbisi, A. et al. L-DNAs as potential antimessenger oligonucleotides: a reassessment. Nucleic Acids Res. 21, 4139–4163 (1993).
7. Kretzmer, R., Lehlt, J.-M. & Schneider, M. A. Self-recognition in helicate self-assembly: spontaneous formation of helical metal complexes from mixtures of ligands and metal ions. Proc. Natl Acad. Sci. USA 90, 5394–5398 (1993).
8. Caulder, D. L. & Raymond, K. N. Supramolecular Self-recognition and self-assembly in gallium(III) catecholamine triple helices. Angew. Chem. Int. Ed. Engl. 36, 1440–1442 (1997).
9. Albrecht, M., Schneider, M. & Rötttele, H. Template-directed self-recognition of alkyl-bridged bis(catechol) ligands in the formation of helicate-type complexes. Angew. Chem. Int. Ed. 38, 557–559 (1999).
10. Wu, A. & Isaacs, L. Self-sorting: the exception or the rule? J. Am. Chem. Soc. 125, 4831–4835 (2003).
11. Addicott, C., Das, N. & Stang, P. J. Self-recognition in the coordination driven self-assembly of 2-D polygoms. Inorg. Chem. 43, 5335–5338 (2004).
12. Cao, L.-P., Wang, J.-G., Ding, J.-Y., Wu, A.-X. & Isaacs, L. Reassembly self-sorting triggered by heterodimerization. Chem. Commun. 47, 8548–8550 (2011).
13. Jimenez, A. et al. Selective encapsulation and sequential release of guests within a self-sorting mixture of three tetrahedral caged. Angew. Chem. Int. Ed. 53, 4556–4560 (2014).
14. Talotta, C., Gaeta, C., Qi, Z., Schalley, C. A. & Neri, P. Pseudorotaxanes with self-sorted sequence and stereochemical orientation. Angew. Chem. Int. Ed. 52, 7437–7441 (2013).
15. Ayme, J.-F., Beves, J. E., Campbell, C. J. & Leigh, D. A. The self-sorting behavior of C/2 helicases and molecular knots and links. Angew. Chem. Int. Ed. 53, 7823–7827 (2014).
16. Masood, M. A., Enemark, E. J. & Stack, T. D. P. Ligand self-recognition in the self-assembly of a [(CuI)]2+ complex: the role of chirality. Angew. Chem. Int. Ed. 37, 928–932 (1998).
17. Prins, L. J., Huiskens, I., de Jong, F., Timmerman, P. & Reinhoudt, D. N. Complete asymmetric induction of supramolecular chirality in a hydrogen-bonded assembly. Nature 398, 498–502 (1999).
18. Telfer, S. G., Bernardinelli, G. & Williams, A. F. Diastereospecific synthesis of amino-acid substituted 2,2’-bipyridyl complexes. Chem. Commun. 1498–1499 (2001).
19. Shi, K., Fettinger, J. C. & Davis, I. T. Homochiral G-quadruplexes with Ba2+ but not with K+: the cation programs enantiomer specific recognition. J. Am. Chem. Soc. 123, 6738–6739 (2001).
20. Ishida, Y. & Aida, T. Homochiral supramolecular polymerization of an “S”-shaped chiral monomer: translation of optical purity into molecular weight distribution. J. Am. Chem. Soc. 124, 14017–14019 (2002).
21. Chung, D. M. & Nowick, J. S. Enzyme-like recognition between β-sheets. J. Am. Chem. Soc. 126, 3062–3063 (2004).
22. Hwang, I.-W. et al. Porphyrin boxes constructed by homochiral self-sorting assembly: optical separation, exciton coupling, and efficient excitation energy migration. J. Am. Chem. Soc. 126, 16187–16198 (2004).
23. Telfer, S. G., Sato, T. & Kuroda, R. Noncovalent ligand strands for transition-metal helicates: the straightforward and stereoselective self-assembly of dinuclear double-stranded helicates using hydrogen bonding. Angew. Chem. Int. Ed. 43, 581–584 (2004).
24. Hutin, M. et al. Self-sorting chiral subcomponent rearrangement during crystallization. J. Am. Chem. Soc. 129, 8774–8780 (2007).

25. Ličkoski, J. Enantioselective self-recognition in homo- and heteronuclear macrocyclic lanthanide(III) complexes. Inorg. Chem. 50, 5567–5576 (2011).

26. Safont-Sempere, M. M. et al. Impact of molecular flexibility on binding strength and self-sorting of chiral π-surfaces. J. Am. Chem. Soc. 133, 9580–9591 (2011).

27. Helmich, F., Smulders, M. M. I., Lee, C. C., Schenning, A. P. H. J. & Meijer, E. W. Effect of stereochemical centers on the self-sorting, depolymerization, and atropisomerization kinetics of porphyrin-based aggregates. J. Am. Chem. Soc. 133, 12338–12346 (2011).

28. Ponnuswamy, N., Coughon, F. B. L., Clough, J. M., Pantos, G. D. & Sanders, J. K. M. Discovery of an organic trefoil knot. Science 338, 783–785 (2012).

29. Lin, N.-T. et al. Enantioselective self-sorting on planar, π-acidic surfaces of chiral amine-π transporters. Chem. Sci. 3, 1121–1127 (2012).

30. Roche, C. et al. Homochiral columns constructed by chiral self-sorting during supramolecular helical organization of hat-shaped molecules. J. Am. Chem. Soc. 136, 7169–7185 (2014).

31. Gütz, C. et al. Enantioselectively pure trinuclear helicates via diastereoselective self-recognition and characterization of their redox chemistry. J. Am. Chem. Soc. 136, 11830–11838 (2014).

32. Safont-Sempere, M. M., Fernández, G. & Würtzner, F. Self-sorting phenomena in complex supramolecular systems. Chem. Rev. 111, 5784–5814 (2011).

33. Saha, M. L. & Schmittel, M. Degree of molecular self-sorting in macromolecular lanthanide(III) complexes. Angew. Chemie Int. Ed. 52, 5275–5279 (2013).

34. Ito, H., Furusho, Y., Hasegawa, T. & Yashima, E. Sequence- and chain-length-dependent crystallization. J. Polym. Sci. 53, 990–999 (2015).

35. Yamada, H., Furusho, Y. & Yashima, E. Diastereoselective imine-bond formation through complementary double-helix formation. J. Am. Chem. Soc. 134, 7250–7253 (2012).

36. Cierpicki, T. & Otlewski, J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. J. Biomed. NMR 21, 249–261 (2001).

37. Luchinat, C. & Roelens, S. Enantiomeric purity determination of 1,2-diols through NMR spectroscopy without chiral auxiliaries. J. Am. Chem. Soc. 108, 4873–4878 (1986).

38. Safont-Sempere, M. M., Oswald, P., Radacki, K. & Würtzner, F. Chiral self-recognition and self-discrimination of strapped perylene bisimides by π-stacking dimerization. Chem. Eur. J. 16, 7380–7384 (2010).

39. Kubo, Y., Maeda, S., Tokita, S. & Kubo, M. Colorimetric chiral recognition by a molecular sensor. Nature 382, 522–524 (1996).

40. Levkin, P. A., Maier, N. M., Schurig, V. & Lindner, W. Strong detrimental effect of a minute enantiomeric impurity of a chiral selector on the enantioselectivity factor. Angew. Chem. Int. Ed. 49, 7742–7744 (2010).

41. Lindsey, J. Self-assembly in synthetic routes to molecular devices. Biological principles and chemical perspectives: a review. New J. Chem. 15, 153–180 (1991).

42. Robertson, A., Sinclair, A. J. & Philp, D. Minimal self-replicating systems. Chem. Soc. Rev. 29, 141–152 (2000).

Acknowledgements

This work was supported in part by a Grant-in-Aids for Scientific Research (S) from the Japan Society for the Promotion of Science (NPS no. 25228004) and by the Nanotechnology Platform Program (Molecule and Material Synthesis) of the Ministry of Education, Culture, Sports, Science and Technology, Japan. W.M. and J.T. express thanks for a JSPS Research Fellowship for Young Scientists (no. 480 for W.M. and no. 8886 for J.T.).

Author contributions

E.Y. designed and directed the project, and W.M., J.T. and H.Y. performed the experiments. H.I. confirmed the validity of the X-ray crystallographic analysis. N.O. calculated the duplex structures by DFT. All authors analysed the data and discussed the results. N.O. and E.Y. co-wrote the manuscript.

Additional information

Accession codes: The X-ray crystallographic coordinates for structures ((SS)-1a and (SS)-2a) reported in this Article have been deposited at the Cambridge Crystallographic Data Centre (CCDC), under deposition number CCDC 1036590 and 1036591. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Makiguchi, W. et al. Chirality- and sequence-selective successive self-sorting via specific homo- and complementary-duplex formations. Nat. Commun. 6:7236 doi: 10.1038/ncomms8236 (2015).