Chemistry–A European Journal
Research Article
doi.org/10.1002/chem.202402637
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

Unraveling the Source of Self-Induced Diastereomeric Anisochronism in Chiral Dipeptides

Fabio Spiaggia,[a] Federica Aiello,*[b] Luca Sementa,[b] Jean-Marc Campagne,[c] Renata Marcia de Figueiredo,[c] Gloria Uccello Barretta,[a] and Federica Balzano[a]

Mastering of analytical methods for accurate quantitative determinations of enantiomeric excess is a crucial aspect in asymmetric catalysis, chiral synthesis, and pharmaceutical applications. In this context, the phenomenon of Self-Induced Diastereomeric Anisochronism (SIDA) can be exploited in NMR spectroscopy for accurate determinations of enantiomeric composition, without using a chiral auxiliary that could interfere with the spectroscopic investigation. This phenomenon can be particularly useful for improving the quantitative analysis of mixtures with low enantiomeric excesses, where direct integration of signals can be tricky. Here, we describe a novel analysis protocol to correctly determine the enantiomeric composition of scalemic mixtures and investigate the thermodynamic and stereochemical features at the basis of SIDA. Dipeptide derivatives were chosen as substrates for this study, given their central role in drug design. By integrating the experiments with a conformational stochastic search that includes entropic contributions, we provide valuable information on the dimerization thermodynamics, the nature of non-covalent interactions leading to self-association, and the differences in the chemical environment responsible for the anisochronism, highlighting the importance of different stereochemical arrangement and tight association for the distinction between homochiral and heterochiral adducts. An important role played by the counterion was pointed out by computational studies.

Introduction

Self-Induced Diastereomeric Anisochronism (SIDA) constitutes one of the most intriguing phenomena among detectable chiral recognition processes in NMR spectroscopy, relying on the ability of enantiomers to self-assemble in solution and to form diastereomeric adducts.[1,2] Due to the presence of functional groups that can promote self-aggregation through intermolecular interactions like π-π stacking and, more importantly, hydrogen bonding, homochiral and heterochiral species can associate to originate dimers, which behave in all respects as diastereomers.[1–3] In that way, anisochronism is induced in the NMR spectra of intrinsically isochronous enantiomers, thus avoiding the use of an external chiral auxiliary[4,5] that could create interferences, such as unwanted signal superimpositions, in the quantitative analysis of the enantiomeric mixtures.

In spite of the enormous potentiality of SIDA, in most cases this phenomenon has been neglected, albeit it could be related to the occurrence of nonlinear effects in catalytic asymmetric synthesis, chiral separations and in the optical properties of chiral substrates.[6–8] As a matter of fact, a full rationalization of SIDA is still far from the current state of knowledge: sometimes it has been accidentally encountered and recognized,[9–11] sometimes it constituted the way to explain the presence of unexpected species in investigations dealing with chirality.[12–14]

In general, a confirmation of the occurrence of SIDA can be obtained by comparing the NMR spectra of enantiomerically pure and racemic compounds, for which NMR signals are unexpectedly detected at different chemical shifts.

The first case of SIDA was recognized in dihydroquinine derivatives in 1969, when Williams *et al.*[15] found out that the ^{1}H NMR spectrum of enantiomerically pure sample of dihydroquinine is significantly different from the ^{1}H NMR spectrum of the racemic mixture when both are recorded at the same experimental conditions (concentration, temperature, and solvent). Additionally, the authors discovered that when a scalemic mixture is prepared and the ^{1}H NMR spectrum is recorded, some signals can undergo splitting, and the split signals areas correspond to the enantiomeric composition of the mixture.

Since the mid-70's, several articles were published where SIDA was observed in structurally different compounds. Different combinations of functional groups can favor the occurrence of SIDA, such as amide,[16] carboxylic[17] and hydroxy groups.[18] Among the compounds that are capable of giving SIDA, many of them are amino acid derivatives[19,20] and peptide-like molecules.[16]

These systems are of particular relevance: besides being the building block of proteins, peptides are involved in many

[a] F. Spiaggia, G. Uccello Barretta, F. Balzano
Department of Chemistry and Industrial Chemistry, University of Pisa, Via G. Moruzzi 13, 56124 Pisa, Italy
[b] F. Aiello, L. Sementa
Institute for Chemical and Physical Processes (IPCF), National Research Council (CNR), Via G. Moruzzi 1, 56124 Pisa, Italy
E-mail: federica.aiello@cnr.it
[c] J.-M. Campagne, R. Marcia de Figueiredo
ICGM, Univ Montpellier, CNRS, ENSCM, 34296 Montpellier, France

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202402637

physiological processes, and for this reason, they have a wide application as therapeutic agents in various areas. Nanotubes or hydrogels can be obtained by self-assembling of oligopeptides containing two or three amino acids,[21–23] and structures like micelles, vesicles, 2D-sheets, 3D networks, can be obtained from amphiphilic peptides.[24] They constitute the starting material for the design of biocompatible and low toxic drug-delivery systems[25] to apply in nanobiotechnology and nanomedicine.[26] Even dipeptides can be used for the development of drug delivery systems and therapeutic agents thanks to their low molecular weight, the relatively low costs related to their synthesis, and the possibility of being administered orally.[27] These features favor the preparation of derivatives endowed with pharmacological properties and of supramolecular structures that can be exploited in multiple ways by fine-tuning their aggregation mechanism.

A clear understanding of the phenomenon of self-assembly in peptide-based systems, therefore, allows a better comprehension of the behavior of the related supramolecular systems. The investigation of the thermodynamics and the kinetics controlling the process, hence, is paramount. Thermodynamics is dependent on non-covalent interactions (π-π stacking, intermolecular hydrogen bonding and electrostatic interactions), and factors influencing the kinetic include pH, temperature, concentration, and solvent.[28] In general, the propensity of peptides to self-aggregate is related to several factors, such as the constituent amino acids, pH, ionic strength, concentration, and the nature (polarity) of the solvent, that can deeply influence the intermolecular peptide-peptide interactions and hence favor different supramolecular architectures.[29,30]

With the aim of shedding some light on the peculiar self-assembly phenomenon known as SIDA, in the present work an intertwined (synthesis - physicochemical characterization - computational investigation) study has been undertaken to understand the origins of SIDA in chiral dipeptides. LL- and DD-valine-leucine (Val-Leu) dipeptides (Figure 1) were synthesized, and their homochiral (LL/LL and DD/DD) and heterochiral (LL/DD) mixtures were analyzed *via* NMR.

This pair of amino acids was chosen in consideration of their peculiarities. Leucine administration is related to important anti-catabolic actions,[31] and valine is one of the essential amino acids responsible for growth and development,[32] known for its pharmaceutical, cosmetics, and nutritional properties.[33] Moreover, the wide applications of valine in chiral stationary phases for gas chromatography are well-known. As an example, the chiral column Chirasil-Val® is largely used for the chiral analysis of amino acids, hydroxy acids, alcohols, amines, and biphenyl derivatives.[34]

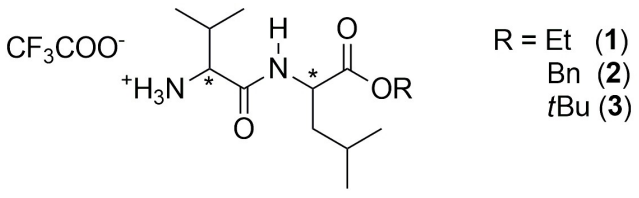


Figure 1. Chemical structures of the TFA•H_2N-Val-Leu-OR dipeptides **1**–**3**.

NMR spectroscopy was chosen in this work as analytical technique thanks to the possibility of delving into the phenomenon for obtaining both thermodynamic and kinetic information. The substrates analyzed are characterized as trifluoroacetate salts at the *N*-terminus of the valine residue and by an ester derivatization at the *C*-terminus of the leucine fragment (TFA·H_2N-Val-Leu-OR). Three different types of ester derivatives were investigated, namely ethyl (Et), benzyl (Bn) and *tert*-butyl (*t*Bu) (**1**–**3**, Figure 1), so that the effect of the derivatization could be evaluated as well.

NMR DOSY (Diffusion Ordered SpectroscopY) and ROESY (Rotating-frame Overhauser Enhancement SpectroscopY) techniques were exploited for evaluating the occurrence of aggregation in solution, and hence determining whether dimers or higher-order structures were formed, and the conformational features of the aggregates. The thermodynamic parameters (dimerization constants, Gibbs free energy, enthalpy and entropy) were determined for both the homo- and heterochiral dipeptide aggregates for all the three substrates analyzed by applying the dilution method.

A computational investigation has been conducted on derivative **1** to predict the thermodynamics of self-aggregation. This analysis provided atomistic insight into the equilibrium conformations and the nature of the interactions promoting stable adducts, supporting the data obtained from NMR spectroscopy.

Results and Discussion

Synthesis of Dipeptides 1–3

LL- and DD-valine-leucine *tert*-butoxycarbonyl (Boc)-dipeptides bearing different ester groups were synthesized *via* a solution-phase strategy using activated α-amino esters (Scheme S1).[35] This method relies on the alternative amine activation instead of traditional carboxylic acid activation.[36,37] Boc-dipeptides with Et and Bn ester groups were transformed into the corresponding TFA salts **1** and **2** through treatment with TFA, whereas Boc-dipeptides bearing a *t*Bu ester group required a two-step procedure (Scheme S1). For this particular case (absence of orthogonal protecting groups), methanesulfonic acid was used in combination with *t*BuOAc in CH_2Cl_2 to selectively remove the *N*-urethane Boc protecting group in the presence of a *tert*-butyl ester.[38] Then, after basic workup, the free amines so obtained were transformed into TFA salts **3** by treatment with TFA (1 equiv) in cold Et_2O.

NMR Investigation of TFA•H_2N-Val-Leu-OEt (1)

The self-aggregation propensity of the ethyl derivative was investigated by comparing the NMR spectra of enantiomerically pure LL-**1** at different concentrations. This process, in fact, should determine a shift of the resonances, enhanced for the nuclei more involved in the intermolecular interaction processes. Indeed, the comparison of the proton (^{1}H) NMR spectra

of LL-**1** in a range of concentration between 0.5 mM and 15 mM showed a significant variation in the chemical shift of the amide protons (δ_{NH}) from 7.14 ppm to 7.72 ppm (Figure 2). This trend suggested the occurrence of self-aggregation in solution and supported the hypothesis of hydrogen bonding between the monomeric species of **1**.

To further verify the supramolecular aggregation, the diffusion coefficient (D) of LL-**1** was measured (Table S1). This parameter can be obtained by exploiting the NMR DOSY technique,[39,40] and is related to the molecular sizes by the Stokes-Einstein equation (Eq. (1)), strictly holding in the spherical approximation:

$$D = \frac{k_B T}{6\pi\eta R_H} \quad (1)$$

where k_B is the Boltzmann constant, T is the absolute temperature, η is the solution viscosity that, in diluted solutions, can be approximated to that of the solvent, and R_H is the hydrodynamic radius. As Eq. (1) highlights, the higher the molecular sizes of a species, the lower is its diffusion coefficient.

In the most diluted solution (0.5 mM), where the dipeptide can be assumed to be present mainly as a monomer, a D of $(6.5 \pm 0.8)\times10^{-10}$ m^2s^{-1} was measured, to be compared to a value of $(4.5 \pm 0.3)\times10^{-10}$ m^2s^{-1} measured in the most concentrated solution (15 mM, Figure S1).

The decrease of the diffusion parameter with increasing of the concentration confirmed the supramolecular aggregation already hypothesized based on the chemical shift variations reported in Figure 2. The ratio between the highest and the lowest diffusion coefficient, equal to 1.44, suggested the presence in the most concentrated solution of dimeric species at the most. The diffusion coefficient of the counterion, measured by means of ^{19}F NMR DOSY, indicated that the ionic couple is tight, as a good correspondence between the ^1H and the ^{19}F diffusion coefficients was found (Table S1).

Solutions at different concentrations of racemic **1** were then prepared, and a variation in chemical shift of the amide proton was observed also in this set of samples (Figure S2). More notably, a different value of δ_{NH} was measured in the proton spectra of LL-**1** and its racemate at the same concentration. As an example, at 10 mM δ_{NH} was equal to 7.65 ppm for LL-**1** (Figure 3a) but shifted to 7.56 ppm in racemic **1** (Figure 3b). This feature is a distinctive and one of the first ever described traits of SIDA,[15] referred as deviation in early theoretical treatments[41] and attributable to the formation of homospecies or of both homo- and heteroaggregates.

Briefly, SIDA occurs when enantiomeric substrates self-associate in solution and give origin to homo- and heteroaggregates.[1,2] In fast-exchange conditions with respect to the NMR timescale, the chemical shifts observed are a weighted average of the monomer and the homo- and heteroaggregates. If the self-assembly determines the formation of dimers and no higher order species are present, as in the case of **1**, the chemical shifts observed for each enantiomer (δ_L and δ_D) can be expressed according to Eq. (2) and Eq. (3):

$$\delta_L = x_{L,m}\delta_{L,m} + x_{LL}\delta_{homo} + x^L_{LD}\,\delta_{hetero} \quad (2)$$

$$\delta_D = x_{D,m}\delta_{L,m} + x_{DD}\delta_{homo} + x^D_{LD}\delta_{hetero} \quad (3)$$

where δ_m, δ_{homo} and δ_{hetero} are the chemical shifts of the monomer (L or D), the homo- (LL or DD), and the heterodimers (LD/DL), respectively; x_L/x_D, x_{LL}/x_{DD}, and x^L_{LD}/x^D_{LD} are the fractions of the enantiomer in the monomer, in the homodimer and in the heterodimer with respect to the total enantiomer concentration.

When suitable experimental conditions for favoring SIDA (*e.g.*, solvent, temperature, total concentration) are chosen, the NMR spectrum of an enantiopure substrate will differ from that of the corresponding racemate because, in the latter case, the solute-solute interactions involve not only species with the same chirality, but also species with the opposite configuration, thus determining the formation of both homo- and heterodimers. Due to the additional contribution of δ_{hetero} to the observed δ, the chemical shifts of the two samples (enantiopure *vs.* racemate) will not be coincident.

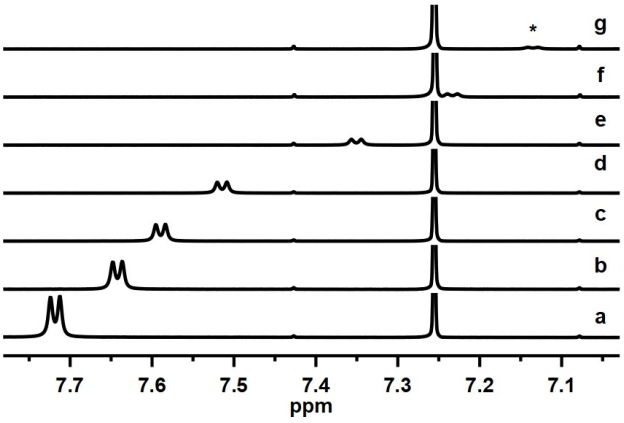


Figure 2. ^1H NMR (600 MHz, $CDCl_3$, 25 °C) spectral region of the amide proton of LL-**1** at: 15 mM (a), 10 mM (b), 7.5 mM (c), 5 mM (d), 2 mM (e), 1 mM (f), 0.5 mM (g). * indicates the amide signal in 0.5 mM sample.

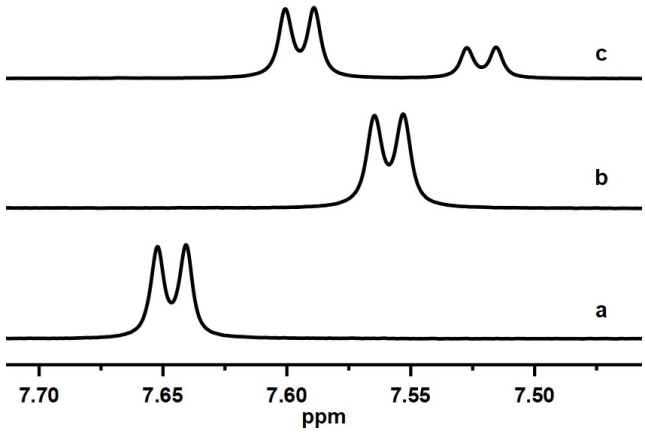


Figure 3. ^1H NMR (600 MHz, $CDCl_3$, 25 °C, 10 mM) spectral region of the amide proton of LL-**1** (a), racemic **1** (b) and enantiomerically enriched (70% LL – 30% DD) **1** (c).

When instead an enantiomerically enriched mixture is analyzed, the main enantiomer will show propensity toward homo-aggregation, whereas the minor one will associate mainly to give heterochiral species. Therefore, $\delta_L \neq \delta_D$, and the NMR spectrum will be characterized by split signals, with their integrated areas directly representing the enantiomeric excess (ee = %L − %D, with %L and %D being the percentages of the L and D enantiomers, respectively) of the mixture. The difference between the chemical shift of the split signals will give the nonequivalence ($\Delta\delta = \delta_L - \delta_D$), which will depend on both stereochemical (the chemical shifts of the dimeric species) and thermodynamic factors (the molar fractions of the species and, consequently, the dimerization constants).

Indeed, the analysis of an enantiomerically enriched sample of **1** revealed the presence of two split signals for the amide proton (Figure 3c), and their integrated area reflected the enantiomeric composition of the mixture.

With the aim of evaluating the feasibility of the NMR analytical method in determining the enantiomeric purity and in investigating the SIDA phenomenon, a deep investigation was carried out by performing an enantiomeric titration, *i.e.*, by recording proton spectra of solutions of **1** at different ee, in which the total concentration was kept constant (10 mM). Starting from a nominal ee of ±10%, splitting of signal was observed for the amide proton (Figure 4). The magnitude of the nonequivalence ($\Delta\delta_{NH}$) increased with the increasing of ee, by reaching a significant value of 0.187 ppm for the nominal ee of 98% (Table S2). A linear relationship was found between ee determined by NMR integration, for all the cases with baseline separation, and nominal ee (Figure S3) with an excellent correlation coefficient (0.9999). The absolute errors were all within ±1% (Table S2). For the mixtures with nominal ee of ±10%, for which no separation at the baseline was observed, no accurate integration was possible.

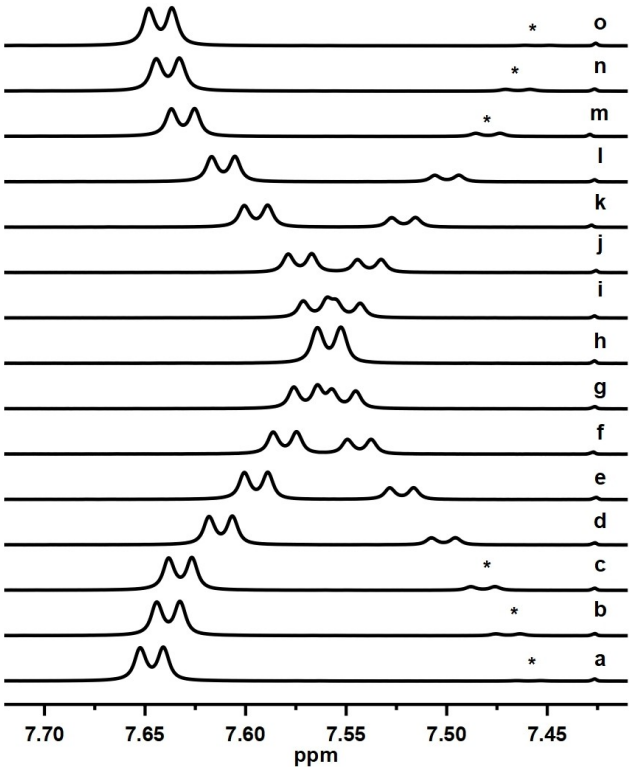


Figure 4. ^{1}H NMR (600 MHz, $CDCl_3$, 25 °C, 10 mM) spectral region of the amide proton of **1** at different nominal ee: −98% (a), −90% (b), −80% (c), −60% (d), −40% (e), −20% (f), −10% (g), 0% (h), 10% (i), 20% (j), 40% (k), 60% (l), 80% (m), 90% (n), 98% (o). * indicate the amide signal belonging to the minor enantiomer in the most enantiomerically enriched samples.

Figure 5 depicts the trend of ee *vs.* $\Delta\delta_{NH}$ (in absolute value) obtained by applying least-square fitting. The presence of a linear relationship prompted us to investigate whether the enantiomeric composition could be alternatively calculated by measuring the nonequivalence.

As detailed in Section 2 of Supporting Information, a calibration curve can be constructed by plotting $\Delta\delta_{NH}$ *vs.* enantiomer percentage (for instance %LL) for all enantiomerically enriched mixtures with baseline separation observed in the proton spectrum. Based on the linear relationship reported in Figure S19, the two mixtures with nominal ee of ±20%, where integration is still possible, were used as validation points, obtaining a calculated content of LL-**1** with an error of 0.5%. Therefore, the enantiomeric composition could be obtained for mixtures with significantly overlapped signals, such as in those with a nominal ee of ±10%, for which integration is challenging. As reported in Table S9, a content of LL-**1** of 44.8% (nominal ee = −10%) and of 54.4% (nominal ee = 10%) were calculated, within 1% error with respect to data from deconvolution.

The strong agreement between values calculated from the calibration curve and those determined *via* integration or deconvolution procedures suggests that this novel approach is effective for analyzing enantiomeric mixtures with overlapped signals, to determine the enantiomeric composition directly from chemical shift data.

Having delved into the analytical application of SIDA phenomena, the attention focused on the determination of the thermodynamic parameters of the self-assembly processes of **1** in solution: dimerization constants, Gibbs free energy (ΔG°), enthalpy (ΔH°) and entropy (ΔS°).

The preference toward formation of either homo- or heterodimer was investigated by analyzing the relationship between the percentage of LL-**1** enantiomer in solution and the corresponding chemical shift of its amide proton. The plot

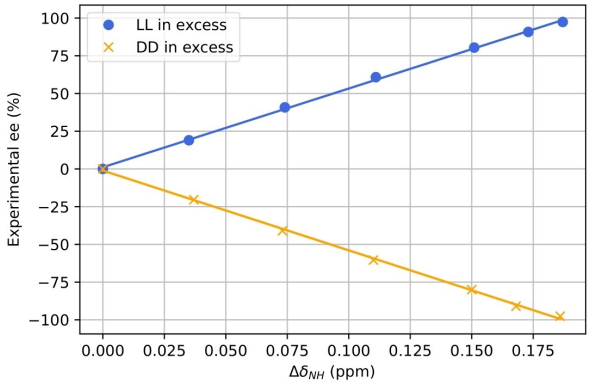


Figure 5. ee *vs.* $\Delta\delta_{NH}$ for **1**.

reported in Figure S4 shows a linear relationship between the two parameters that, according to the literature,[1,16] reflects the absence of thermodynamic preference toward one of the two dimeric aggregates. Moreover, it allows connecting the two dimerization constants as reported in Eq. (4):

$$K_{d(hetero)} = 2K_{d(homo)} \quad (4)$$

where $K_{d(hetero)}$ and $K_{d(homo)}$ are the dimerization constants for the heterodimer and the homodimer, respectively.[2]

Enantiomerically pure samples of **1** were then exploited to determine $K_{d(homo)}$, from which $K_{d(hetero)}$ can be directly obtained (Eq. (4)).

An alternative approach to the determination of $K_{d(hetero)}$ required the analysis of racemic mixtures to determine a "mean" dimerization constant ($K_{d(mean)}$, Eq. (5)), from which $K_{d(hetero)}$ can be extracted (Eq. (6)).[42]

$$K_{d(mean)} = \frac{K_{d(hetero)} + 2K_{d(homo)}}{4} \quad (5)$$

$$K_{d(hetero)} = 4K_{d(mean)} - 2K_{d(homo)} \quad (6)$$

For determining $K_{d(homo)}$, a dilution curve was built for LL-**1**: δ_{NH} was plotted against the total concentration C_0 (Figure S5 and Table S3) and Eq. (7) was used to fit the data:[43]

$$\delta = \delta_m + (\delta_{homo} - \delta_m)\frac{\sqrt{(1 + 8K_{d(homo)}C_0) - 1}}{\sqrt{(1 + 8K_{d(homo)}C_0) + 1}} \quad (7)$$

A $K_{d(homo)}$ of $(164 \pm 18)\ M^{-1}$ at 25 °C was calculated by feeding the data into the "NMR Dimer Aggregation" model (see Experimental Section). From this value, a $K_{d(hetero)}$ of $328\ M^{-1}$ was obtained from Eq. (4).

The enthalpy and the entropy of dimerization were determined by building a van't Hoff plot. The dilution curve was built for LL-**1** at 25 °C, 30 °C, 35 °C, 40 °C, and 45 °C, and $K_{d(homo)}$ was calculated for each temperature (Table S4). The van't Hoff plot was obtained by reporting the natural logarithm of $K_{d(homo)}$ against the inverse of the absolute temperature (Figure S6); the data was fitted based on Eq. (8) via linear least squares:

$$lnK_{d(homo)} = -\frac{\Delta H^0}{RT} + \frac{\Delta S^0}{R} = \frac{2830.9}{T} - 4.4 \quad (8)$$

The van't Hoff plot was consistent with the case of an exothermic process like self-aggregation, in which the enthalpy is negative. Furthermore, the negative intercept of the equation was consistent with a process related to a reduction of entropy in the system, like in the case of a dimerization process of the type $2\ A \rightleftharpoons A_2$.

From the slope and intercept of Eq. (8), an enthalpy of dimerization equal to $(-23.5 \pm 1.0)\ kJmol^{-1}$ and an entropy of dimerization equal to $(-36.6 \pm 3.3)\ JK^{-1}mol^{-1}$ were obtained for LL-**1**. It is important to underline that the calculated value of self-association enthalpy fits with the typical value of one NH···O=C hydrogen bond in nonpolar solvents, which usually falls in the range of −20 to −35 $kJmol^{-1}$.[44]

Following the same approach used for calculating $K_{d(homo)}$, a $K_{d(mean)}$ of $(168 \pm 23)\ M^{-1}$ was calculated from racemic samples of **1**. From this value, a $K_{d(hetero)}$ equal to $(344 \pm 99)\ M^{-1}$ was calculated based on Eq. (6), comparable to the expected value of $328\ M^{-1}$ obtained from Eq. (4), thus confirming the absence of a thermodynamic preference toward the formation of homodimers or heterodimers.

Once ascertained the thermodynamic parameters, the stereochemistry of the self-assembly of the two aggregates (homo- and heterodimer) was investigated by analyzing and comparing the dipolar effects detected in ROESY experiments for both the enantiomerically pure and the racemic sample of **1**. First, the conformation in solution of monomeric **1** was ascertained by studying the dependance of the dihedral angles on the vicinal coupling constant through the Karplus' equation[45] and by analyzing through-space dipolar interactions by means of 1D ROESY experiments. Then, the supramolecular self-assembly was investigated by focusing on ROE effects that could not be ascribed solely to intramolecular dipolar effects.

The prevailing conformation of monomeric **1** in solution was established by focusing on each fragment composing the dipeptide (the valine and the leucine residues, and the ethyl ester). The ROE effects observed between the NH proton of leucine (H5) and methine protons on the stereocenter of leucine (H6) and valine (H3) (Figure 6b) pointed out that H5 is in a cisoid arrangement with H3 and transoid with respect to its methine proton H6. The absence of dipolar interactions between H5 and ethyl protons (H11/H11′ and H12) supported a conformational arrangement in which the amide proton H5 is cisoid with respect to the carbonyl oxygen of the ester group, thus favoring the occurrence of an intramolecular hydrogen bond. Regarding the valine fragment, the comparable dipolar effects observed for the proton on the chiral carbon (H3) with its adjacent methine of isopropyl moiety (H2) and with H5 (Figure 6c) suggested that H2 and H5 are approximately at the same distance with respect to H3. A similar analysis could not be performed for the leucine fragment, due to partial superimposition of the resonances belonging to isobutyl moiety (H7/H7′ and H8) (Figure 6a). Finally, beside the main dipolar effect observed for the methyl protons (H12) belonging to the ethyl moiety on the adjacent methylene group H11/H11′, an effect on the diastereotopic methyl protons H9/H9′ of the leucine moiety was detected, as expected, due to the spatial proximity of the two fragments in the monomer (Figure 6d).

Interestingly, a ROE effect between methyl protons H12 and the diastereotopic methyl protons H1/H1′ of valine was detected, more intense than that on methyl protons H9/H9′ of leucine (Figure 6d). Such a strong dipolar interaction cannot be ascribed uniquely to an intramolecular ROE effect, but rather to a combination of both intra- and intermolecular interaction. The intensity of this effect, therefore, was attributed to the occurrence of an intermolecular interaction between two monomeric units, confirming once again the self-assembly of **1**. Considering that the self-association enthalpy of the homo-

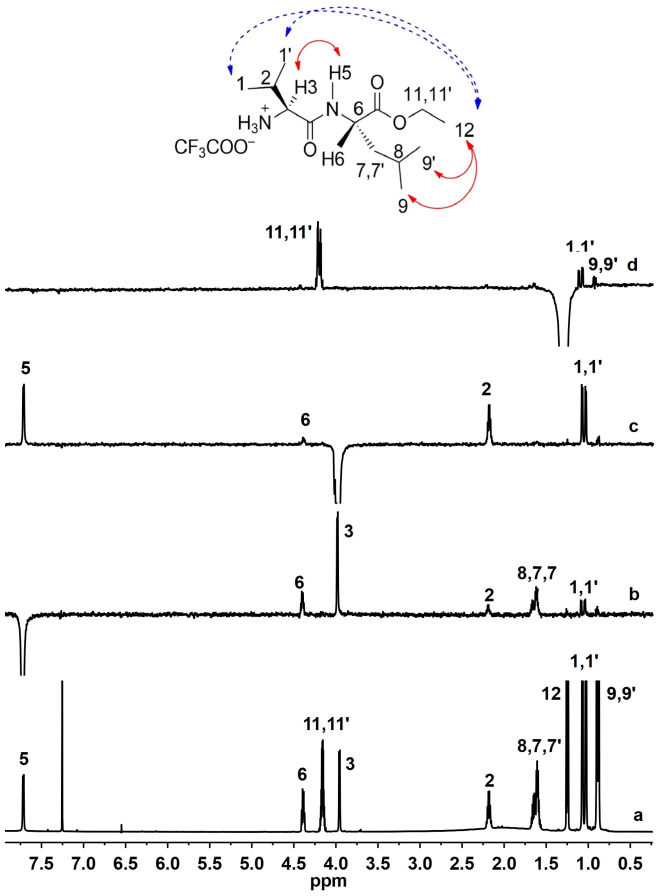


Figure 6. ^{1}H NMR (600 MHz, $CDCl_3$, 25 °C) spectrum (a) and 1D ROESY (600 MHz, $CDCl_3$, 25 °C, mixing time = 300 ms) spectra of H5 (b), H3 (c), and H12 (d) protons of enantiopure **1** (15 mM). The arrows highlight the most significant intramolecular (red, plain) and intermolecular (blue, dashed) effects observed.

dimer previously calculated was compatible with the formation of one net intermolecular hydrogen bond, a conformation in which the two dipeptide molecules are arranged in a parallel way in β-sheet fashion could be ruled out, in favor of an intermolecular interaction where the two units are almost perpendicular.

The dipolar effects detected for the racemic mixture (Figure S7) were comparable to those observed for the enantiomerically pure sample, thus suggesting that the interaction model proposed for the homodimer could be also applied for the heterodimer.

NMR Investigation of TFA•H_2N-Val-Leu-OBn (2) and TFA·H_2N-Val-Leu-O*t*Bu (3)

The occurrence of self-assembly was confirmed also for the benzyl (**2**) and the *tert*-butyl (**3**) ester derivatives.

The comparison between the NMR spectra of the diluted and the concentrated solution of **2** pointed out that the amide moiety is, once again, the most involved in the formation of the aggregates in solution since a significant variation in chemical shift (from 7.12 ppm at 0.4 mM to 7.63 ppm at 13.3 mM) was observed (Figure S8). Similarly, a shift from 6.90 ppm (0.5 mM) to 7.56 ppm (15 mM) was observed for substrate **3** (Figure S8).

The trend of the diffusion coefficients measured for **2** and **3** at different concentrations was comparable to what observed for the ethyl derivative **1** (Tables S5 and S6). Also in these cases, hence, the self-assembly gave dimeric species.

The enantiomeric titrations performed (Figures S9 and S10) indicated linearity between ee and $\Delta\delta_{NH}$ (in absolute value) for both **2** and **3** (Figures S11 and S12).

The thermodynamic parameters were then determined by building up a dilution curve for the two dipeptides (Figure S5). For **2**, a $K_{d(homo)}$ of $(221 \pm 26)\ M^{-1}$ was calculated at 25 °C (Table S7). From the van't Hoff plot (Figure S6), an enthalpy of (-19.7 ± 0.6) kJmol^{-1}, and an entropy of (-20.8 ± 1.8) JK^{-1}mol^{-1} were calculated. A $K_{d(mean)}$ of $(244 \pm 39)\ M^{-1}$ was obtained for racemic **2**, comparable to $K_{d(homo)}$.

For *tert*-butyl derivative **3**, the fitting of the dilution curve provided a value for $K_{d(homo)}$ of $(154 \pm 24)\ M^{-1}$ at 25 °C (Table S8), slightly lower with respect to **1** and **2**. The dimerization enthalpy and entropy obtained from the van't Hoff plot (Figure S6) were equal to (-28.0 ± 2.5) kJmol^{-1} and (-51.5 ± 8.2) JK^{-1}mol^{-1}, respectively.

A value of $(137 \pm 20)\ M^{-1}$ was calculated for $K_{d(mean)}$ of **3**, superimposable to that of $K_{d(homo)}$. Therefore, the condition $K_{d(mean)} = K_{d(homo)}$ was fulfilled for both derivatives **2** and **3**.

The values of $K_{d(homo)}$ found for **1**–**3** at 25 °C were similar, as summarized in Table 1. This result reflected into a comparable Gibbs free energy for the three systems; however, the self-association enthalpies and entropies showed a marked enthalpy-entropy compensation behavior (Figure S13). By comparing the dipeptides investigated, the magnitude of nonequivalence can be correlated to the different enthalpy of self-association. For **3**, which showed the highest nonequivalences, the highest value of enthalpy (in absolute value) was calculated (−28.0 kJmol^{-1}), whereas the lowest value was found for **2** (−19.7 kJmol^{-1}) that also showed the lowest nonequivalence (Table 1); the self-association entropy followed the same trend,

Table 1. Nonequivalences (total concentration 10 mM, ee = 40 %) and thermodynamic data for the TFA·H_2N-Val-Leu-OR dipeptide **1**–**3**.

	$\Delta\delta_{NH}$ (ppm)	$K_{d(homo)}$ (M^{-1})	ΔG (kJ mol^{-1})	ΔH (kJ mol^{-1})	ΔS (J K^{-1} mol^{-1})
1	0.074	164 ± 18	−12.6 ± 0.3	−23.5 ± 1.0	−36.6 ± 3.3
2	0.058	221 ± 26	−13.4 ± 0.3	−19.7 ± 0.6	−20.8 ± 1.8
3	0.105	154 ± 24	−12.5 ± 0.4	−28.0 ± 2.5	−51.5 ± 8.2

indicating a loss of freedom due to a higher association propensity.

Regarding the dependence of the nonequivalence from the ee, a direct proportionality was found when the total concentration of the scalemic mixtures is kept constant. The dependence of the nonequivalence from both thermodynamic ($K_{d(homo)}$) and stereochemical (δ_{homo} - δ_{hetero}) factors is shown in Supporting Information (Section 3). In the cases where no thermodynamic preference is detected, the differentiation of the chemical shifts between homodimer and heterodimer (stereochemical factor) plays a fundamental role for SIDA to occur, as in the present case.

NMR Investigation of HCl•H_2N-Val-Leu-OEt (1 b)

With the aim to investigate the role of the counterion in the self-assembly of dipeptide salts and in SIDA, the corresponding chloride derivative of substrate **1**, here indicated as **1 b**, was synthesized (Scheme S2) to be analyzed *via* NMR by adopting the same protocol used for the analysis of the TFA derivatives. Nevertheless, this substrate showed severe solubility issues that prevented from carrying out an accurate spectroscopic investigation. Precipitation was observed for solutions of pure LL- and DD-**1 b**, even at 5 mM concentration. For this reason, the preparation of enantiomerically enriched samples with known total concentration and ee was not feasible, and the enantiomeric titration could not be performed.

The comparison of the ^{1}H NMR spectra of LL-**1** and LL-**1 b** (Figure S14), however, pointed out that the proton signals of the chloride derivative are characterized by a significant linewidth with respect to the TFA one, particularly enhanced for the amide proton and the methine protons of valine and leucine fragment. In contrast, a narrower signal is observed for the proton resonance belonging to the NH_3^+ group, suggesting a minor involvement of the amino site (and its counterion) in the intermolecular interaction. These differences could be imputable to a different mechanism of self-assembly.

Moreover, a diffusion coefficient of $(3.7\pm0.1)\times10^{-10}$ m^2s^{-1} was measured for this substrate (Figure S15) at 5 mM concentration, to be compared to a value of $(4.9\pm0.1)\times10^{-10}$ m^2s^{-1} measured for LL-**1** at the same concentration (Table S1). The significant decrease of this parameter suggested the presence in solution of higher order aggregates (trimers or even bigger supramolecular adducts) already at low concentration. This result was attributed to the nature of the counterion: chloride is not endowed with the coordinating features of TFA and, therefore, is not capable to play a role in the formation of dimeric species. In these conditions, the dipeptides are likely more prone to aggregate and form bigger adducts, thus explaining the propensity for precipitation.

Computational Results on 1

To support the experimental findings, molecular simulations were performed revealing the interactions leading to dimer formation and elucidating the protons local chemical environment and the effects on the NMR response thereof.

Dimerization energies were estimated by identifying the most stable conformational arrangements for the individual monomer, the homodimer, and the heterodimer. A Basin Hopping (BH) algorithm was employed on a system comprising two monomers and the corresponding TFA ions, which compensate for the positive charge of the protonated valine amine moiety. The choice of tetra-molecular models was driven by the experimental evidence (DOSY measurements) of a strong interaction between the dipeptide and its TFA counterion. On this basis, preliminary molecular dynamics lasting 0.5 ns yielded a ratio between the hydrodynamic radii of dimer and monomer of 1.40, which closely matches the experimental value of 1.44 calculated from NMR diffusion measurements.

Figure 7 depicts ball-and-stick models of the Global Minimum (GM) for the monomer, homodimer, and heterodimer identified by the sampling algorithm. For the monomer, the predicted optimal TFA arrangement is in proximity to the valine NH_3^+ group and its nearest H3 proton. Interestingly, the algorithm found a quasi-symmetric C2 conformation for the homochiral model. Even though this symmetric conformation is not possible when the dimer is formed by different stereoisomers, the energy difference between the GM diastereomers is less than 2 $kJmol^{-1}$. Homodimer and heterodimer GM structures revealed that the interaction between the monomers is mediated by the TFA anions, which adopt a bridged configuration in which both oxygen atoms connect the positively charged NH_3^+ groups of the protonated valine residues of different monomers. There is a key difference in how the leucine residues interact with the TFA anions. In the homodimer, both leucine amide protons point toward the TFA oxygens, while in the heterodimer, one of the leucine amide protons is forced to point away and toward the solvent.

This difference in the local atomic environment is likely the reason for the observed shift in the NMR signal measured for the amide protons of Leu when the solution contains both diastereomers. The simulations predicted chemical shift values

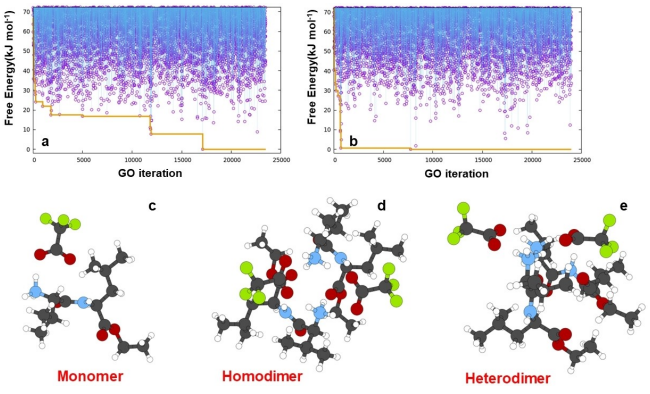


Figure 7. Free Energy (pink circle) for the homodimer (a) and heterodimer (b) system as a function of the GO iteration. The orange curve shows the free-energy evolution for the transitory GM along the GO run. Ball and Stick representation of the putative GM for the monomer (c), the homodimer (d) and heterodimer (e); different colors are associated with different atomic types: H/white, C/grey, N/light-blue, O/red, F/green.

of 7.74 ppm, 6.96 ppm and 5.93 ppm for the amide protons in the homodimer, in the heterodimer and in the monomer, respectively. For the two dimers, the reported values average the simulated chemical shifts for these species. It is worth noting that this trend is in agreement with the experimental one, in which a higher chemical shift was observed for enantiomerically pure LL-**1** at the highest concentration with respect to the most diluted solution (Figure S2), where monomer contribution is preponderant, and an intermediate value was measured in the racemate, where both homodimer and heterodimer are present.

As for the dimerization free energy, the value calculated by comparing the energy of the dimer to the GM of the individual monomer/TFA pair was -46.1 kJmol^{-1}. This value is significantly higher in absolute value than the experimentally measured free energy. This discrepancy likely arises because the computational approach does not account for the entropic contribution coming from the existence of multiple conformations of the chemical system considered.

Well-tempered metadynamics were then employed to accurately evaluate all the entropic contributions to the dimerization process.[46] The distance between the leucine N atoms of different monomers served as the collective variable describing adduct formation. Sampling times were accelerated by constraining the centers of mass of the monomers within a cylindrical space with a radius and height of 10 nm and 25 nm, respectively. Figure S16 illustrates the metadynamics approach that yields a dimerization free energy of approximately -23 kJmol^{-1}, which, although reduced compared to the previous value, still overestimates experimental results. This discrepancy might stem from the limited accuracy of both the semiempirical method employed to describe the electronic structure of our models and the implicit solvent model that critically affects dipolar interactions and intermolecular forces that hold the monomers together in the dimer models.

To further support the experimental findings, we analyzed the trajectories generated by 10 ns of dynamics within the canonical ensemble at T 300 K starting from the GM dimer structures. The reported values are collected after system equilibration, *i.e.*, along the last 7.5 ns of the simulations. The angular distributions (Figure S17) revealed similar dynamics for both the diastereomers with a preferred transoid conformation for H6-H5, consistent with the observed dipolar effects observed in ROESY experiments. Interestingly, the H3-H5 dihedral angle showed two peaks, indicating that thermal energy allows rotation around the peptide bond enabling stereoisomers to explore two different conformations during the dynamics: the former with H5-H3 in cisoid arrangement (angle of about 50°) and the latter with H3-H5 in a transoid arrangement (angle of about 125°). The population of the conformation with the cisoid arrangement, wherein H3 and H5 protons are in close proximity, likely accounts for the intense signal in the ROE spectra (Figure 6b).

Figure 8 depicts the distribution, generated from the dynamics, of the shortest distance between the methyl protons H12 of OEt and the methyl groups H1/H1′ of valine residues. The analysis considered whether the residues are in the same monomer (intramolecular) or separate monomer (intermolecular) within the homochiral dimer model. In roughly 30% of the configurations sampled during the simulation, protons H12-H1/H1′ belonging to the same monomer are closer than 5 Å, potentially generating a signal detectable by ROE experiments (red bars in Figure 8b). Similarly, approximately 25% of the configurations exhibit protons with intermolecular distances falling within 5 Å (Figure 8a), supporting the concept that the intense dipolar effects detected with ROESY were the results of both intermolecular and intramolecular interactions. A similar result was found for the heterodimer (Figure S18).

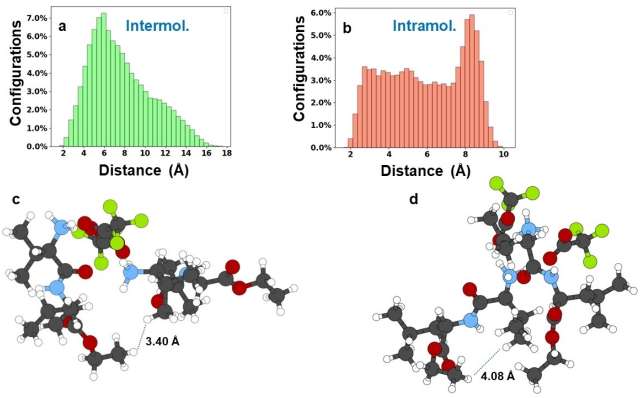


Figure 8. Intermolecular (a) and intramolecular (b) distribution of the shortest distance between the methyl protons H12 of ethyl moiety and the methyl protons of valine (H1/H1′). Ball and stick representations of configurations extracted from the MD with proton intermolecular (c) and intramolecular distance (d) less than 5 Å for the homodimer of **1**.

Conclusions

The combined spectroscopic and computational investigation conducted in this study delves into the origins of self-discrimination processes in chiral dipeptides. This approach overcomes the limitations of chiral analysis that typically requires the selection of an external chiral auxiliary to generate anisochronism in the enantiomeric components.

The NMR data indicate that, when a linear dependency between the observed chemical shift of a selected nucleus and the enantiomeric fraction is found, the homodimer and the heterodimer have the same thermodynamic (*i.e.*, enthalpic) preference for formation. The driving force of SIDA within the three Val-Leu systems investigated, therefore, lies in the different stereochemical arrangements characterizing the homodimer and the heterodimer. This difference underlies the variations in chemical shifts observed between δ_{homo} and δ_{hetero}. However, for these differences to manifest, a tight association is necessary. If the dimers are mildly associated (more positive values of enthalpy), their nuclei will not experience high differences in their chemical environment. Therefore, a tight association is a necessary condition for the nonequivalence to arise. Nevertheless, it is not a sufficient condition: even tightly associated homodimers and heterodimers may not exhibit nonequivalence if their nuclei do not experience different

chemical environments. This can occur in scenarios such as a hypothetical coplanar bidentate interaction where minimal stereochemical differences exist between the homodimer and the heterodimer.

By integrating stochastic sampling and accelerated dynamics, we demonstrated that the tight association between the monomers is mediated by TFA anions, which, being achiral, keep pairs of similar or different enantiomers together with the same strength, at a distance that allows for differentiation of the protons chemical environments in the adducts due to the stereochemical conformation assumed by the dipeptides. Further studies using different counterions, capable of expressing less specific interactions with the amino acid moieties, would help to strengthen our hypothesis on the importance of NMR secondary species in SIDA.

Finally, from an analytical perspective, it has been shown that a proper analysis of the dependence of the nonequivalence on the enantiomeric composition can overcome the limitations associated with the need to accurately integrate the signals corresponding to the two enantiomers. This facilitates NMR analysis in cases of partial signal superimposition, as the enantiomeric composition can be directly obtained from chemical shift detection (nonequivalence).

Experimental Section

Materials

CDI (≥90%), copper (II) bromide (99%), hydroxybenzotriazole hydrate (97%), HCl (ACS reagent 37%), $MgSO_4$ (Anhydrous, ReagentPlus®, ≥99.5%), pentane (98%), Et_2O (ACS Reagent, ≥99.8%), *t*BuOAc (≥99%), and methanesulfonic acid (99%) were purchased from Sigma-Aldrich. Boc-(L)-Val-OH (>95%), Boc-(D)-Val-OH (98%), HCl.H_2N-(D)-Leu-O*t*Bu (>95%), TosOH.H_2N-(L)-Leu-OBn (>97%), TosOH.H_2N-(D)-Leu-OBn: (>95%), HCl.H_2 N-(L)-Leu-OEt (>95%), and HCl.H_2N-(D)-Leu-OEt (>95%) were purchased from Biosynth-Carbosynth. HCl.H_2N-(L)-Leu-OtBu (>95%) was purchased from Fluorochem. $NaHCO_3$ (99–100.5%) was purchased from Fluka. Ethyl acetate (Emplura®, 99%) was purchased from Supelco. CH_2Cl_2 (pure, stabilized with amylene) was purchased from Carlo Erba. TFA (99%) was purchased from ThermoFisher. Deuterated chloroform ($CDCl_3$) used for the NMR experiments was purchased from Deutero GmbH.

General Procedures for Dipeptides 1–3 Synthesis

The synthetic schemes and the numbering schemes for proton and carbon nuclei are reported in Supporting Information.

*Boc-dipeptides LL- and DD-Val-LeuOR (**5**).*[35] To a solution of in situ prepared activated α-aminoesters Imi-Leu-OR (**4**) (1.0 equiv) in CH_2Cl_2 (1.0 M) were successively added copper bromide (10 mol%), 1-hydroxybenzotriazole hydrate (10 mol%), and Boc-Val-OH (1.5 equiv) at 25 °C. The obtained blue suspension was thus stirred at this temperature for 20 h. After reaction completion (TLC monitoring), the mixture was washed once with aqueous HCl 0.5 N. Then, the aqueous phase was extracted with CH_2Cl_2 (2x). The organic layers were combined, washed with saturated aqueous $NaHCO_3$ and brine. The organic phase was thus dried over $MgSO_4$ and evaporated. Purification was performed by filtration through a short pad of silica gel (eluent Pentane/EtOAc) to afford the Boc-dipeptide products (**5**).

*TFA salts of dipeptides LL- and DD-Val-LeuOEt (**1**) and TFA salts of dipeptides LL- and DD-Val-LeuOBn (**2**).* To a solution of dipeptides Boc-Val-Leu-OR (R=Et, Bn) in CH_2Cl_2 at 0 °C was added TFA (20 equiv). The reaction was then stirred at this temperature for 30 minutes and at 25 °C for 2 – 4 h. After reaction completion (TLC monitoring), the solvents were removed in vacuum. The corresponding TFA dipeptide salts **1** and **2** were precipitated in Et_2O (50–75% overall yields).

TFA•H_2N-Val-Leu-OEt (**1**). ^{1}H NMR (600 MHz, $CDCl_3$, ppm) δ = 7.78 (d, $^3J_{H5,H6}$ = 6.8 Hz, 1H; H5), 4.37 (dt, $^3J_{H6,H7-7'}$ = 8.2 Hz, $^3J_{H6,H5}$ = 6.8 Hz, 1H; H6), 4.16 (m, 2H; H11/H11′), 3.98 (d, $^3J_{H3,H2}$ = 6.7 Hz, 1H; H3), 2.16 (m, 1H; H2), 1.65 (m, 1H; H8), 1.60 (m, 2H; H7/H7′), 1.24 (t, $^3J_{H12,H11/H11'}$ = 7.1 Hz, 3H; H12), 1.07 (d, $^3J_{H1,H2}$ = 6.7 Hz, 3H; H1), 1.02 (d, $^3J_{H1',H2}$ = 6.7 Hz, 3H; H1′), 0.88 (d, $^3J_{H9,H8}$ = 6.3 Hz, 3H; H9), 0.86 (d, $^3J_{H9',H8}$ = 6.3 Hz, 3H; H9′).

^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ = 172.1 (C10), 168.7 (C4), 61.4 (C11), 58.6 (C3), 51.8 (C6), 40.2 (C7), 30.4 (C2), 24.6 (C8), 22.4 (C9), 21.5 (C9′), 17.9 (C1), 17.8 (C1′), 14.0 (C12).

TFA•H_2N-Val-Leu-OBn (**2**). ^{1}H NMR (600 MHz, $CDCl_3$, ppm) δ = 7.63 (d, $^3J_{H5,H6}$ = 6.7 Hz, 1H; H5), 5.15 (d, $^2J_{H11,H11'}$ = 12.0 Hz, 1H; H11), 5.09 (d, $^2J_{H11',H11}$ = 12.0 Hz, 1H; H11′), 4.44 (m, 1H; H6), 3.96 (d, $^3J_{H3,H2}$ = 6.2 Hz, 1H; H3), 2.11 (m, 1H; H2), 1.61 (m, 3H; H7/H7′, H8), 0.98 (d, $^3J_{H1,H2}$ = 6.8 Hz, 3H; H1), 0.93 (d, $^3J_{H1',H2}$ = 6.8 Hz, 3H; H1′), 0.87 (d, $^3J_{H9,H8}$ = 5.7 Hz, 3H; H9), 0.85 (d, $^3J_{H9',H8}$ = 5.7 Hz, 3H; H9′).

^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ = 172.0 (C10), 168.5 (C4), 135.1 (Ph), 128.7 (Ph), 128.5 (Ph), 128.4 (Ph), 128.3 (Ph), 128.2 (Ph), 67.2 (C11), 58.6 (C3), 40.2 (C7), 30.1 (C2), 24.5 (C8), 22.4 (C9), 21.6 (C9′), 17.8 (C1), 17.7 (C1′).

*TFA salts of dipeptides LL- and DD-Val-LeuOtBu (**3**).*[38] Boc-dipeptides *tert*-butyl esters (1.0 equiv) were dissolved in a mixture of *t*BuOAc/CH_2Cl_2 4:1 (0.5 M) at 25 °C. Then, methanesulfonic acid (1.5 equiv) was added to this solution. After 2 h of stirring at 25 °C, TLC monitoring showed that the reaction was uncompleted. Then, additional methanesulfonic acid (1.5 equiv) was added, and the reaction was stirred until complete consumption of the starting materials (additional 1 – 3 h). Afterwards, the reaction mixture was neutralized by addition of an aqueous saturated solution of $NaHCO_3$, and the products were extracted with EtOAc (3x). The solvents were removed in vacuum to provide the free amines (>95% purity by ^{1}H NMR), which were used without further purification in the next step. Free amines were dissolved into Et_2O and placed at 0 °C under strong stirring. Thus, TFA (1.0 equiv) was added and the precipitation of the TFA salts **3** occurred instantly. After filtration, the solids were vacuum dried to provide dipeptides **3** in good yields (68–75% overall yields).

^{1}H NMR (600 MHz, $CDCl_3$, ppm) δ = 7.60 (d, $^3J_{H5,H6}$ = 6.9 Hz, 1H; H5), 4.28 (dt, $^3J_{H6,H7/H7'}$ = 8.5 Hz, $^3J_{H6,H5}$ = 6.9 Hz, 1H; H6), 3.98 (d, $^3J_{H3,H2}$ = 6.5 Hz, 1H; H3), 2.19 (m, 1H; H2), 1.65 (m, 1H; H8), 1.56 (m, 2H; H7/H7′), 1.44 (s, 9H; H12), 1.09 (d, $^3J_{H1,H2}$ = 6.8 Hz, 3H; H1), 1.04 (d, $^3J_{H1',H2}$ = 6.8 Hz, 3H; H1′), 0.89 (d, $^3J_{H9,H8}$ = 6.6 Hz, 3H; H9), 0.87 (d, $^3J_{H9',H8}$ = 6.6 Hz, 3H; H9′).

^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ = 171.4 (C10), 168.4 (C4), 82.0 (C11), 58.6 (C3), 52.6 (C6), 40.4 (C7), 30.3 (C2), 27.9 (C12), 24.6 (C8), 22.4 (C9), 21.7 (C9′), 18.0 (C1), 17.9 (C1′).

*HCl salts of dipeptides LL- and DD-Val-LeuOEt (**1b**).*[47] A solution of HCl in dioxane (14 mL, 4 M) was added to a 50 mL round-bottom flask containing Boc-Val-Leu-OEt dipeptide (0.7 mmol) under argon at 0 °C. After stirring at this temperature for 5 min, the ice-bath was removed and the mixture was stirred at room temperature for 1 h.

After reaction completion (TLC monitoring), CH_2Cl_2 and Et_2O were added to the reaction mixture and the solvents were removed in vacuum. The corresponding HCl dipeptide salts (**1 b**) were precipitated in Et_2O (88—98% overall yields).

^{1}H NMR (600 MHz, $CDCl_3$, ppm) $\delta = 8.29$ (br s, 3H; NH_3^+), 7.90 (br s, 1H; H5), 4.43 (br s, 1H; H6), 4.23 (br s, 1H; H3), 4.17 (m, 2H; H11/H11′), 2.38 (br s, 1H; H2), 1.79 (m, 1H; H8), 1.72 (m, 1H; H7), 1.66 (m, 1H; H7′), 1.25 (t, $^3J_{H12,H11/H11'} = 7.1$ Hz, 6H; H12), 1.15 (d, $^3J_{H1,H2} = 6.3$ Hz, 3H; H1), 1.09 (d, $^3J_{H1',H2} = 6.2$ Hz, 3H; H1′), 0.94 (d, $^3J_{H9,H8} = 6.1$ Hz, 3H; H9); 0.92 (d, $^3J_{H9',H8} = 6.6$ Hz, 3H; H9′).

^{13}C NMR (150 MHz, $CDCl_3$, ppm) $\delta = 172.5$ (C10), 168.1 (C4), 61.4 (C11), 58.6 (C3), 52.1 (C6), 40.3 (C7), 30.4 (C2), 24.7 (C8), 22.7 (C9), 22.0 (C9′), 18.5 (C1), 18.1 (C1′), 14.1 (C12).

NMR Methods

^{1}H, ^{19}F and ^{13}C{^{1}H} NMR measurements were carried out on a spectrometer operating at 600 MHz, 564 MHz, and 150 MHz for ^{1}H, ^{19}F and ^{13}C nuclei, respectively. ^{1}H and ^{13}C chemical shifts are referred to tetramethylsilane (TMS) as the secondary reference standard, ^{19}F chemical shifts are referred to trifluorotoluene as the external standard and the temperature was controlled (± 0.1 °C). The proton spectra were recorded by using a $\pi/2$ pulse that was calibrated for each investigated dipeptide system, a relaxation delay of 5 s and a number of scans sufficient to provide a signal to noise (S/N) ratio of at least 150. For all the 2D NMR spectra employed for the characterization of the substrates, the spectral width was the minimum required in both dimensions. The gCOSY (gradient COrrelation SpectroscopY) and TOCSY (TOtal Correlation SpectroscopY) maps were recorded by using a relaxation delay of 1 s, 128 increments of 4 transients, each with 2 K data points, and a mixing time of 80 ms (TOCSY). The 1D TOCSY spectra were recorded using a selective pulse, transients ranging from 64 to 256, a relaxation delay of 1 s, and a mixing time of 80 ms. The 1D ROESY (Rotating-frame Overhauser Enhancement SpectroscopY) spectra were recorded using a selective inversion pulse with transients ranging from 1024 to 6144, a relaxation delay of 1 s and a mixing time of 300 ms. The gHSQC (gradient Heteronuclear Single Quantum Coherence) and gHMBC (gradient Heteronuclear Multiple Bond Correlation) experiments were recorded with a relaxation delay of 1.2 s, 128 and 200 increments, respectively, with 4 transients, each of 2 K data points. The gHMBC experiments were optimized for a long-range coupling constant of 8 Hz. DOSY (Diffusion-Ordered SpectroscopY) experiments were carried out by using a stimulated echo sequence with self-compensating gradient schemes and with 60 K data points. Typically, g was varied in 15 steps (4–256 transients) and Δ and δ were optimized to obtain an approximately 85–90% decrease in the resonance intensity at the largest gradient amplitude. The baselines of all arrayed spectra were corrected prior to processing the data. After data acquisition, each FID was apodised with 1.0 Hz line broadening and Fourier-transformed. The data were processed with the DOSY macro (involving the determination of the resonance heights of all the signals above a pre-established threshold and the fitting of the decay curve for each resonance to a Gaussian function) to obtain pseudo two-dimensional spectra with NMR chemical shifts along one axis and calculated diffusion coefficients along the other.

Sample Preparation for the NMR Analysis

The solutions used for the enantiomeric titration experiments were prepared by mixing the proper volume of stock solutions of LL- and DD-enantiomer of dipeptides **1**, **2** and **3**; the final volume was adjusted to 0.7 mL.

The samples used for the dilution experiments were prepared by properly diluting a stock solution of dipeptide, either in enantiomerically pure or in racemic form.

Determination of Dimerization and Mean Constants

Dimerization constants were obtained using Eq. (7) as the fitting function and employing the curve_fit function of the scipy.optimize module in Python to perform the non-linear least squares fitting. The algorithm employed for fitting was the Levenberg–Marquardt algorithm. The curve_fit function returned an array of the optimal values for the function parameters (δ_{mono}, δ_{homo} and $K_{d(homo)}$) and the associated standard deviation errors. Mean constants and associated standard deviation errors were obtained using the same equation as indicated by the literature.[1]

Computational Methods

Electronic structures and atomic forces were generated at the semi-empirical level by using the Extended Semiempirical Tight Binding (GFN2-xTB) model proposed by Grimme's group.[48–51] We performed the conformational stochastic sampling with our proprietary Python code. The sampling algorithm proceeds by: i) generating new conformers by stochastically varying the dihedral angles and the positions of the center of mass of monomers and TFA ions, ii) calculating the free energy (ΔF) of the relaxed conformer by evaluating the entropic contribution from translational, rotational, and vibrational degrees of freedom, and iii) accepting or rejecting the new structure based on the Metropolis criterion,[52,53] which is determined by the probability min(1,exp($-\Delta$F/kBT)), where T (3000 K in this work) represents a fictitious temperature heuristically chosen. The sampling terminates if the algorithm does not identify a new putative GM within a pre-defined number of iterations.

The NMR responses of the diastereomers, evaluated at the equilibrium geometry predicted by xTB, were obtained using the ORCA code[54] with the B3LYP functional and def2-TZV2P basis sets. Solvent effects on both structural relaxation and NMR response prediction were accounted for using the implicit models implemented in XTB (ALPB)[55] and ORCA (CPCM)[56] respectively. The reported chemical shifts are referenced to the calculated NMR response of TMS, whose structure was generated at the semi-empirical GFN2-xTB level.

Molecular dynamics simulations were performed by using the ASE library,[57,58] within the canonical ensemble with the Langevin algorithm.[59] Plumed software[60–62] accelerated the Langevin simulations according to the well-tempered metadynamics scheme.[63] We sampled the systems' dynamics at 300 K, integrating the equation of motion with a timestep of 0.5 fs. When performing the metadynamics, the Gaussian terms added with a pace of 250 fs to the GFN2-xTB potential energy surface had a width of 0.5 Å and starting height of 0.25 kJmol^{-1} regulated along the simulation by using a biasfactor of 5. Simulations were analyzed using our proprietary tools that employ some Python classes encoded in the ASE library.

GM geometries of the monomer, homodimer, and heterodimer are available in the Zenodo repository.[64]

Author Contributions

Conceptualization: F.A., G.U.B., F.B.; Data curation: F.S., L.S.; Formal analysis: F.S., L.S.; Funding Acquisition: G.U.B., R.M.F.;

Investigation: F.S., L.S., R.M.F.; Methodology: F.S., L.S., R.M.F., J.M.C., F.B.; Project administration: F.A., G.U.B., F.B.; Resources: J.M.C., R.M.F., G.U.B., F.B.; Software: F.S., L.S.; Supervision: F.A., G.U.B. (experimental studies), L.S. (molecular modeling); Validation: F.S., F.A., F.B.; Visualization: F.S., F.A., L.S.; Writing-original draft: F.S., F.A., L.S., J.M.C., R.M.F., G.U.B., F.B.; Writing-review and editing: F.S., F.A., G.U.B., F.B.

Acknowledgements

This work was supported by Università di Pisa under the "PRA – Progetti di Ricerca di Ateneo" (Institutional Research Grants) – Project no. PRA_2022–2023_"New challenges of transition metal and lanthanide complexes in the perspective of green chemistry". L.S. gratefully acknowledges computational support from the CINECA supercomputing center. R.M.F and J.M.C. thank the Agence Nationale de la Recherche (project NOPS, ANR-18-CE07-0038-01) for financial support. Open Access publishing facilitated by Consiglio Nazionale delle Ricerche, as part of the Wiley - CRUI-CARE agreement.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

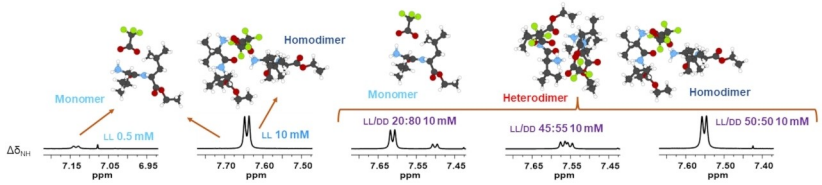
The data that support the findings of this study are available in the supplementary material of this article.

Keywords: Computational chemistry · Dimerization · Molecular recognition · NMR spectroscopy · Self-assembly

[1] Z. Szakács, Z. Sánta, A. Lomoschitz, C. Szántay, *TrAC Trends Anal. Chem.* **2018**, *109*, 180–197.
[2] F. Aiello, G. Uccello Barretta, F. Balzano, F. Spiaggia, *Molecules* **2023**, *28*, 6854.
[3] V. Dašková, D. Padín, B. L. Feringa, *J. Am. Chem. Soc.* **2022**, *144*, 23603–23613.
[4] F. Balzano, G. Uccello-Barretta, F. Aiello, *in Chiral Anal. Second Ed* (Ed: P. L. Polavarapu), Elsevier, Amsterdam **2018**, 367–427.
[5] G. Uccello Barretta, T. J. Wenzel, F. Balzano, *in Compr. Chirality Second Ed.* (Ed: J. Cossy), Academic Press, Oxford **2024**, 560–592.
[6] J. Han, O. Kitagawa, A. Wzorek, K. D. Klika, V. A. Soloshonok, *Chem. Sci.* **2018**, *9*, 1718–1739.
[7] J. Han, A. Wzorek, K. D. Klika, V. A. Soloshonok, *in Late-Stage Fluorination Bioact. Mol. Biol.-Relev. Substrates* (Ed.: A. Postigo), Elsevier, Amsterdam **2019**, 321–355.
[8] J. Han, A. Wzorek, M. Kwiatkowska, V. A. Soloshonok, K. D. Klika, *Amino Acids* **2019**, *51*, 865–889.
[9] M. Mikoajczyk, R. Urawiski, P. Kiebasiski, M. W. Wieczorek, J. Baszczyk, W. R. Majzner, *Synthesis* **1997**, *1997*, 356–365.
[10] S. K. Ghosh, *J. Pept. Res. Off. J. Am. Pept. Soc.* **1999**, *53*, 261–274.
[11] Z. Xu, Q. Wang, J. Zhu, *J. Am. Chem. Soc.* **2015**, *137*, 6712–6724.
[12] A. B. Ouryupin, M. I. Kadyko, P. V. Petrovskii, E. I. Fedin, *Chirality* **1994**, *6*, 1–4.
[13] A. B. Ouryupin, M. I. Kadyko, P. V. Petrovskii, E. I. Fedin, A. Okruszek, R. Kinas, W. J. Stec, *Tetrahedron: Asymmetry* **1995**, *6*, 1813–1824.
[14] K. D. Klika, M. Budovská, P. Kutschy, *Tetrahedron: Asymmetry* **2010**, *21*, 647–658.
[15] T. Williams, R. G. Pitcher, P. Bommer, J. Gutzwiller, M. Uskokovic, *J. Am. Chem. Soc.* **1969**, *91*, 1871–1872.
[16] M. T. Cung, M. Marraud, J. Neel, *Biopolymers* **1977**, *16*, 715–729.
[17] M. Kwiatkowska, A. Wzorek, A. Kolbus, M. Urbaniak, J. Han, V. A. Soloshonok, K. D. Klika, *Symmetry* **2021**, *13*, 543.
[18] V. Nieminen, D. Y. Murzin, K. D. Klika, *Org. Biomol. Chem.* **2009**, *7*, 537–542.
[19] G. Storch, M. Haas, O. Trapp, *Chem. Eur. J.* **2017**, *23*, 5414–5418.
[20] J.-M. Menke, O. Trapp, *Chem. Eur. J.* **2024**, *30*, e202400623.
[21] L. Adler-Abramovich, E. Gazit, *Chem. Soc. Rev.* **2014**, *43*, 6881–6893.
[22] M. S. Levine, M. Ghosh, M. Hesser, N. Hennessy, D. M. DiGuiseppi, L. Adler-Abramovich, R. Schweitzer-Stenner, *Soft Matter* **2020**, *16*, 7860–7868.
[23] L. Li, L. Xie, R. Zheng, R. Sun, *Front. Chem.* **2021**, *9*, 739791–739805.
[24] K. Pandurangan, B. Roy, K. Rajasekhar, Y. V. Suseela, P. Nagendra, A. Chaturvedi, U. R. Satwik, N. A. Murugan, U. Ramamurty, T. Govindaraju, *ACS Appl. Bio Mater.* **2020**, *3*, 3413–3422.
[25] D. Berillo, A. Yeskendir, Z. Zharkinbekov, K. Raziyeva, A. Saparov, *Medicina (Mex.)* **2021**, *57*, 1209.
[26] G. Colombo, P. Soto, E. Gazit, *Trends Biotechnol.* **2007**, *25*, 211–218.
[27] S. Santos, I. Torcato, M. A. R. B. Castanho, *Pept. Sci.* **2012**, *98*, 288–293.
[28] J. Wang, K. Liu, R. Xing, X. Yan, *Chem. Soc. Rev.* **2016**, *45*, 5589–5604.
[29] P. Kumaraswamy, R. Lakshmanan, S. Sethuraman, U. M. Krishnan, *Soft Matter* **2011**, *7*, 2744–2754.
[30] L. Majumder, K. Bera, K. Khamaru, U. Pal, N. C. Maiti, B. Banerji, *J. Mol. Struct.* **2022**, *1266*, 133455.
[31] N. E. Zanchi, H. Nicastro, A. H. Lancha, *Nutr. Metab.* **2008**, *5*, 20.
[32] S. Chatterjee, A. Mavani, J. Bhatttacharyya, *in Nutr. Funct. Foods Boost. Dig. Metab. Immune Health* (Eds.: D. Bagchi, S. E. Ohia), Academic Press, Oxford **2022**, 3–14.
[33] P. E. V. Paul, V. Sangeetha, R. G. Deepika, *in Recent Dev. Appl. Microbiol. Biochem.* (Ed.: V. Buddolla), Academic Press, Oxford **2019**, 107–125.
[34] G. Betzenbichler, L. Huber, S. Kräh, M.-L. K. Morkos, A. F. Siegle, O. Trapp, *Chirality* **2022**, *34*, 732–759.
[35] R. M. de Figueiredo, J.-S. Suppo, C. Midrier, J.-M. Campagne, *Adv. Synth. Catal.* **2017**, *359*, 1963–1968.
[36] J.-S. Suppo, G. Subra, M. Bergès, R. M. de Figueiredo, J.-M. Campagne, *Angew. Chem. Int. Ed.* **2014**, *53*, 5389–5393.
[37] E. Tosi, J.-M. Campagne, R. M. de Figueiredo, *J. Org. Chem.* **2022**, *87*, 12148–12163.
[38] L. S. Lin, T. Lanza, S. E. de Laszlo, Q. Truong, T. Kamenecka, W. K. Hagmann, *Tetrahedron Lett.* **2000**, *41*, 7013–7016.
[39] R. Evans, G. Dal Poggetto, M. Nilsson, G. A. Morris, *Anal. Chem.* **2018**, *90*, 3987–3994.
[40] R. Evans, *Prog. Nucl. Magn. Reson. Spectrosc.* **2020**, *117*, 33–69.
[41] M. I. Kabachnik, T. A. Mastryukova, E. I. Fedin, M. S. Vaisberg, L. L. Morozov, P. V. Petrovskii, A. E. Shipov, *Russ. Chem. Rev.* **1978**, *47*, 821.
[42] D. Gut, A. Rudi, J. Kopilov, I. Goldberg, M. Kol, *J. Am. Chem. Soc.* **2002**, *124*, 5449–5456.
[43] J. S. Chen, R. B. Shirts, *J. Phys. Chem.* **1985**, *89*, 1643–1646.
[44] V. R. Mundlapati, S. Gautam, D. K. Sahoo, A. Ghosh, H. S. Biswal, *J. Phys. Chem. Lett.* **2017**, *8*, 4573–4579.
[45] M. Karplus, H. J. Kolker, *J. Chem. Phys.* **2004**, *39*, 1493–1506.
[46] P. Pracht, S. Grimme, *Chem. Sci.* **2021**, *12*, 6551–6568.
[47] G. Han, M. Tamaki, V. J. Hruby, *J. Pept. Res.* **2001**, *58*, 338–341.
[48] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *WIREs Comput. Mol. Sci.* **2021**, *11*, e1493.
[49] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.
[50] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
[51] P. Pracht, E. Caldeweyher, S. Ehlert, S. Grimme, **2019**, DOI: 10.26434/chemrxiv.8326202.v1.
[52] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **1953**, *21*, 1087–1092.
[53] W. K. Hastings, *Biometrika* **1970**, *57*, 97–109.
[54] F. Neese, F. Wennmohs, U. Becker, C. Riplinger, *J. Chem. Phys.* **2020**, *152*, 224108.
[55] S. Ehlert, M. Stahn, S. Spicher, S. Grimme, *J. Chem. Theory Comput.* **2021**, *17*, 4250–4261.
[56] M. Garcia-Ratés, F. Neese, *J. Comput. Chem.* **2019**, *40*, 1816–1828.
[57] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C.

Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, *J. Phys. Condens. Matter* **2017**, *29*, 273002.
[58] S. R. Bahn, K. W. Jacobsen, *Comput. Sci. Eng.* **2002**, *4*, 56–66.
[59] R. W. Pastor, *in Mol. Dyn. Liq. Cryst.* (Eds: G. R. Luckhurst, C. A. Veracini), Springer Netherlands, Dordrecht **1994**, 85–138.
[60] M. Bonomi, G. Bussi, C. Camilloni, G. A. Tribello, P. Banáš, A. Barducci, M. Bernetti, P. G. Bolhuis, S. Bottaro, D. Branduardi, R. Capelli, P. Carloni, M. Ceriotti, A. Cesari, H. Chen, W. Chen, F. Colizzi, S. De, M. De La Pierre, D. Donadio, V. Drobot, B. Ensing, A. L. Ferguson, M. Filizola, J. S. Fraser, H. Fu, P. Gasparotto, F. L. Gervasio, F. Giberti, A. Gil-Ley, T. Giorgino, G. T. Heller, G. M. Hocky, M. Iannuzzi, M. Invernizzi, K. E. Jelfs, A. Jussupow, E. Kirilin, A. Laio, V. Limongelli, K. Lindorff-Larsen, T. Löhr, F. Marinelli, L. Martin-Samos, M. Masetti, R. Meyer, A. Michaelides, C. Molteni, T. Morishita, M. Nava, C. Paissoni, E. Papaleo, M. Parrinello, J. Pfaendtner, P. Piaggi, G. Piccini, A. Pietropaolo, F. Pietrucci, S. Pipolo, D. Provasi, D. Quigley, P. Raiteri, S. Raniolo, J. Rydzewski, M. Salvalaglio, G. C. Sosso, V. Spiwok, J. Šponer, D. W. H. Swenson, P. Tiwary, O. Valsson, M. Vendruscolo, G. A. Voth, A. White, *Nat. Methods* **2019**, *16*, 670–673.
[61] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, *Comput. Phys. Commun.* **2014**, *185*, 604–613.
[62] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, M. Parrinello, *Comput. Phys. Commun.* **2009**, *180*, 1961–1972.
[63] A. Barducci, *Phys. Rev. Lett.* **2008**, *100*, 020603.
[64] L. Sementa, **2024**, DOI: 10.5281/zenodo.12198517.

Manuscript received: July 11, 2024
Accepted manuscript online: August 11, 2024
Version of record online: ■■, ■■



NMR spectroscopy and computational analysis shed light on the SIDA (Self-Induced Diastereomeric Anisochronism) phenomenon occurring in non-racemic mixtures of chiral dipeptide derivatives. Self-assembly in solution gives rise to diastereomeric homochiral and heterochiral adducts tightly associated, which can be differentiated by proton NMR analysis based on their different chemical shift, due to a different stereochemical arrangement.

F. Spiaggia, F. Aiello, L. Sementa, J.-M. Campagne, R. Marcia de Figueiredo, G. Uccello Barretta, F. Balzano*