DOI: 10.1002/chem.200601289

A Combined Computational and Experimental Approach for the Analysis of the Enantioselective Potential of a New Macrocyclic Receptor for N-Protected α-Amino Acids

Andrea Ragusa,[a, d] Joseph M. Hayes,*[b, e] Mark E. Light,[c] and Jeremy D. Kilburn*[a]

Abstract: A new macrocyclic receptor incorporating a thiourea moiety has been synthesised. Crystal structures of the macrocycle showed that the receptor has a rigid backbone but the thiourea moiety can orientate itself to bind to a DMSO solvent molecule. Forcefield (MMFFs) calculations were performed to model the macrocycle and its binding properties with respect to N-protected amino acids, which were measured experimentally by NMR titration. Binding free energies were calculated by using the mode integration algorithm (MINTA) or free-energy perturbation (FEP). Excellent qualitative agreement with experiment was obtained. To further exploit the accuracy of the free-energy predictions for this system, the faster free-energy algorithm MINTA was used as a prediction tool to test the binding affinity of the macrocycle towards a series of several other amino acid derivatives, which speeded up considerably the screening process and reduced laboratory costs.

Keywords: enantioselectivity · host–guest systems · molecular modeling · molecular recognition · receptors

Introduction

A significant issue for supramolecular chemists is effective and time-efficient analysis of the selectivity of new chiral receptors. X-ray crystallography and NMR data are necessary to determine key structural features and predict possible binding mechanisms in receptor–enantiomer complexes. Small changes in receptor structure can in some cases lead to large changes in enantioselectivity.[1] However, it is not practical to examine every putative receptor of interest due to the costs and the time-consuming nature of the experimental procedures. Computation is seen by many as a useful tool for prediction of binding modes to support the design and development of new enantioselective "synthetic receptors". However, depending on the complexity of the systems, the energetics and sometimes the geometrical features of binding from forcefield calculations alone must be treated with caution. A key question is how suitable are the forcefield parameters for the system under study? To determine this, benchmarking of test cases against results from experiment or high-level theoretical calculations (quantum mechanics) is required. In many cases, forcefield reparametrisation is necessary.[2,3] Furthermore, calculation of differences in binding free energy for chiral separation can be troublesome for a number of reasons and thus a major obstacle to successful computational predictions. Many authors in the past have reported differences in binding as differences in

[a] A. Ragusa, Prof. J. D. Kilburn
School of Chemistry
University of Southampton
Southampton, SO171BJ (UK)
Fax: (+44) 2380-596-805
E-mail: jdk1@soton.ac.uk

[b] Dr. J. M. Hayes
Anterio Consult and Research GmbH
Augustaanlage 26
68165 Mannheim (Germany)
E-mail: joe@klingon.uab.es

[c] Dr. M. E. Light
EPSRC National Crystallography Service
School of Chemistry
University of Southampton
Southampton, SO171BJ (UK)

[d] A. Ragusa
Current address: Laboratory of Glycotechnology, IIQ, CSIC
Calle Américo Vespucio 49, 41092 Seville (Spain)

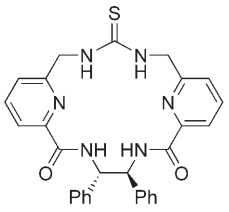
[e] Dr. J. M. Hayes
Current address: Departament de Química
Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona) (Spain)
Fax: (+34) 935-812-920

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

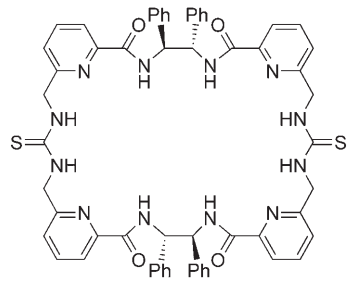
enthalpies, but results can differ significantly depending on the entropy contribution.[2,4,5] A range of methods for calculation of binding free energies are available, but no one method is considered completely satisfactory. These vary from the fast but least accurate empirical scoring function (docking) methods[6–9] to the most accurate free-energy perturbation (FEP) and thermodynamic integration (TI) methods.[10–12] One of us has discussed free-energy methods in greater detail elsewhere,[2] so we limit our discussion here to the two free-energy methods used in this work, FEP and the mode integration algorithm (MINTA).[13,14]

The mode integration algorithm (MINTA) is one of the recently emerged so-called "direct methods" which involve direct computation of the configuration integral as the sum of the contributions of low-energy conformational states. While the "mining minima" method evaluates the configuration integral over soft modes or torsion angles,[15] MINTA includes all degrees of freedom for more accurate free energies. Although it is too early to predict exactly how accurate the MINTA is with respect to FEP, MINTA has already, in a short period of time, proven to be a highly efficient method for prediction of relative binding free energies in studies on chiral separation.[2,4,16] Also, MINTA is intrinsically faster than FEP. The latter can be run either in explicit solvent or in combination with a continuum solvation model. However, depending on the size of the system, FEP often requires unrealistically long run times, especially when explicit solvent models are employed.[3] On the other hand, MINTA relies on the use of a continuum solvation model, and for the calculation of the *A* value of methylcyclohexane was at least 43 times faster than FEP.[13] Success of FEP and MINTA are dependent on adequate sampling of the conformational potential energy surface (PES). The FEP simulations rely on sampling of the PES in either molecular (MD) or stochastic (SD) dynamics, Monte Carlo (MC) or MC/SD simulations,[11] while MINTA relies on location of all relevant low-energy minima in the conformational space, which is best achieved by using an MC algorithm. Sampling is problematic when there are large barriers to conformational interconversion or when the conformational space is sparsely populated.[17–19] There is a certain degree of uncertainty over the effectiveness of MINTA when calculating differences in binding free energy (BFE) for systems that are more sterically diverse than enantiomers, due to neglect of rotational and translational contributions to the configuration integral.[20] However, in this study this should not be a problem, as in previous studies on chiral separation with MINTA.[2,4,16,21] The ideal computational technique relates computed binding free energy differences between L and D receptor-bound enantiomers ($\Delta\Delta G^{L-D}$) to experimental enantioselectivities (*ee*). If this is achieved then computation can serve as a stand-alone tool to probe enantioselectivity.

Here we present X-ray crystal structures, NMR titration experiments and computational calculations which complement each other by showing excellent agreement in their predictions. None of the previously mentioned computational difficulties were encountered. Such is the accuracy of the computational results using the forcefield employed for our system (MMFFs) that it is used alone as the prediction tool for the chiral separation potential of macrocycle **1** for a



1



2

series of amino acids. This can have wide implications in the field of chiral separation. Results are therefore not only analyzed and discussed with respect to the enantioselectivity of macrocycle **1** studied here but also from a wider perspective: how computational chemistry with accurate forcefield parameters and the correctly chosen free-energy calculation method (particularly MINTA) can be used as an efficient stand-alone tool for probing enantioselectivity. The success of the initial screening process using MINTA has already been reported.[21] The extended work reported here highlights more comprehensively the quality of agreement between computation and experiment, the results of the full screening of a virtual library of amino acid ligands using MINTA, results of the binding properties of macrocycle **1** in different solvents and the accuracy of results which can also be obtained using the more time-consuming FEP method. We note that the rigidity of the macrocycle minimises forcefield errors and the possibility of inadequate sampling of the conformational space. However, with work on the development of polarisable forcefields a current focus for some,[22] accurate forcefield description of a more diverse range of systems in the near future is a distinct possibility.

Results

Receptor **1** was designed on the basis of the already reported macrocycle **2**.[1,23] Macrocycle **2** proved to be highly enantioselective for 1:1 binding of N-protected glutamate and aspartate and was found to have remarkable solvent-dependent behaviour in that the receptor adopts two different

conformations depending on the polarity of the solvent used. Receptor **1** represents the monomer analogue of receptor **2** and thus has the same structural characteristics but a smaller cavity and reduced degrees of freedom, and was designed as a putative receptor for mono- rather than dicarboxylates.

Design and synthesis: As in the dimer analogue, the main carboxylate binding site in macrocycle **1** is the thiourea moiety, while the two amide groups should provide additional hydrogen-bonding functionalities. The pyridine units should help to preorganise the macrocycle into a more favourable conformation through intramolecular hydrogen bonds,[24] and the two chiral centres in the diamine spacer should enable enantioselective recognition. Figure 1 shows the proposed/expected binding configuration.

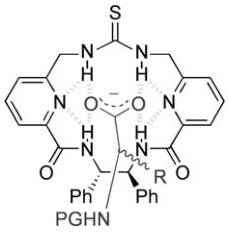
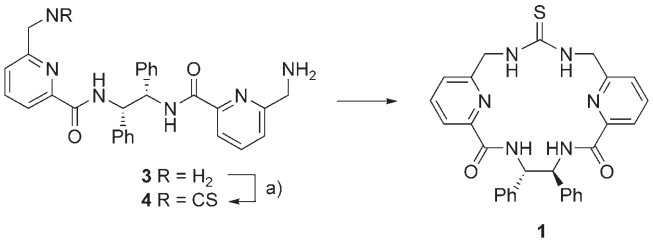


Figure 1. Proposed complex of a generic amino acid with macrocycle **1**. PG = protecting group.

For the synthesis of macrocycle **1** (Scheme 1) the same route used to prepare receptor **2** was used to give amine **3**.[1] Treatment with one equivalent of thiophosgene gave monoisothiocyanate **4**, and subsequent intramolecular reaction between the remaining amino group and the isothiocyanate group gave macrocycle **1** in 17% yield.



Scheme 1. a) Addition by syringe pump over 3 h of a solution of thiophosgene (1 equiv) in CH_2Cl_2 to a solution of diamine **3** and Et_3N in CH_2Cl_2, 17%.

Structural studies: Macrocycle **1** was recrystallised from CH_3CN and from DMSO and X-ray crystal structures were obtained in both cases (Figure 2).

Crystals of **1** from CH_3CN are orthorhombic with two host molecules and a molecule of solvent in the asymmetric unit, but no significant interactions between them.[25] In the crystal structure, the thiourea group does not form hydrogen bonds to the pyridine nitrogen atom, and instead is orthogonal to the plane drawn by the macrocyclic backbone. Thus the thiourea hydrogen atoms point out of the host cavity, in a conformation that is not ideal for maximising electronic and/or steric interactions when binding a carboxylate guest.

A much better preorganised molecular arrangement is found in the crystal structure of **1** recrystallised from DMSO.[21] From this solvent a monoclinic crystal system is formed and there are two host/solvent pairs in the asymmetric unit cell. The macrocycle backbone is now slightly concave, with the thiourea and the amidic protons pointing inwards and hydrogen bonding with both the pyridine moieties and the sulfoxide oxygen atom of a DMSO molecule. Hydrogen-bond distances between the sulfoxide oxygen atom and the four NH groups range from 2.03 to 2.37 Å (Table 1) and indicate a strong interaction between the two molecules, as also predicted by computation (vide infra). Intramolecular hydrogen bonds between the four NH groups and the pyridine nitrogen atoms are also observed. This type of binding conformation was already observed the crystal structure of a similar compound.[26] Although the angles for these interactions are quite acute compared to optimal values, they probably contribute to the overall conformation adopted by the macrocycle.

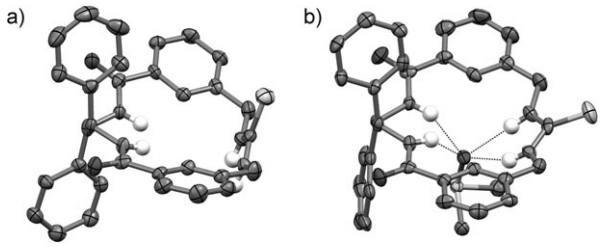


Figure 2. X-ray crystal structures of macrocycle **1** from a) CH_3CN and b) DMSO showing conformational reorganization of the thiourea moiety upon complexation of a solvent molecule.

Table 1. X-ray crystal and computed hydrogen-bond distances [Å] and angles [°] for macrocycle **1** and bound DMSO molecule.[a]

NH[b]	d(NH···O)[c]	∢(NH···O)[c]	d(NH···N)[d]	∢(NH···N)[d]
R1	2.37 (2.65)	139.9 (128.7)	2.31 (2.22)	106.2 (107.4)
R2	2.03 (2.06)	153.2 (161.3)	2.30 (2.37)	107.1 (101.3)
R3	2.13 (2.16)	153.8 (147.2)	2.34 (2.24)	99.6 (106.5)
R4	2.12 (2.03)	152.4 (157.3)	2.42 (2.31)	96.7 (101.2)

[a] Computed values were optimised at the B3LYP/6-31+G* level of theory and are given in parentheses. [b] As defined in Figure 5. [c] Hydrogen bond to guest oxygen atom. [d] Intramolecular hydrogen bond to pyridine nitrogen atom.

Macrocycle **1**, as distinct from receptor **2**, proved to be soluble in neat $CDCl_3$, as well as in CD_3CN or [D_6]DMSO. The ^{1}H NMR spectra of **1** in each of these solvents did not differ significantly, apart from the shifts for the NH signals, and this suggests that the conformation of the macrocycle in solution is not particularly effected by solvent polarity, but the spectrum was better resolved in the more polar solvent (Figure 3). This is in sharp contrast to macrocycle **2**,[1] which has a significantly different conformation in $CDCl_3$ than in [D_6]DMSO or CD_3CN, but is not surprising considering the flexibility of that receptor compared to the rigidity of the monomer analogue.

The main difference in the three spectra is the change in the chemical shift of the signals for the thiourea and amide

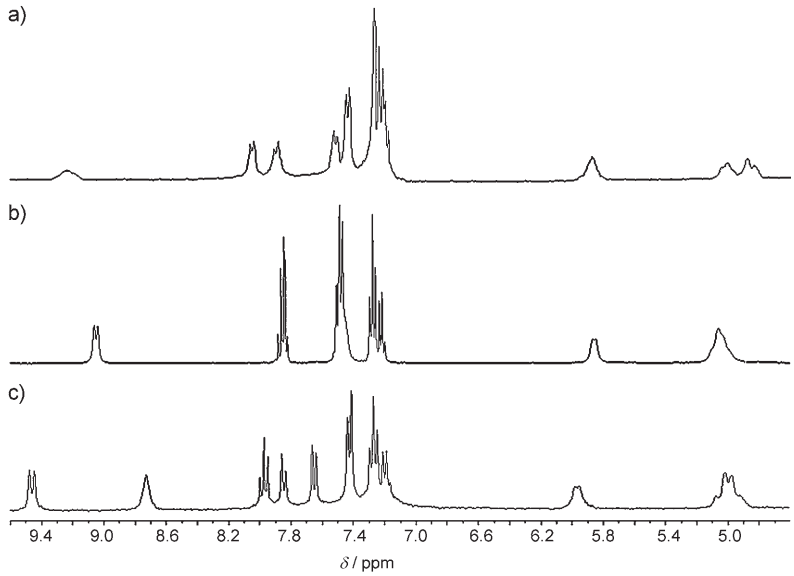


Figure 3. ^{1}H NMR spectra (300 MHz, room temperature) of macrocycle **1** in a) $CDCl_3$, b) CD_3CN and c) [D_6]DMSO.

NH protons. On moving from $CDCl_3$ or CD_3CN to [D_6]DMSO the thiourea proton is shifted downfield by more than 1 ppm (from ca. 7.5 ppm in $CDCl_3$ or CD_3CN to 8.72 ppm in [D_6]DMSO) as a consequence of hydrogen bonding to the solvent molecules. Similarly, the signal for the amide proton appears at $\delta = 9.25$ ppm in $CDCl_3$ and at $\delta = 9.05$ ppm in CD_3CN, shifted downfield in each case by hydrogen bonding to the pyridyl units,[26] but strong hydrogen bonding with the more polar [D_6]DMSO solvent molecules causes a further downfield shift to $\delta = 9.45$ ppm. The NMR data suggests that in [D_6]DMSO the macrocycle adopts a similar conformation to that observed in the crystal structure (Figure 2b) with strong hydrogen bonding between the thiourea and amide NH groups and the sulfoxide oxygen atom of DMSO.

Binding studies: Binding studies were carried out by ^{1}H NMR titration experiments monitoring the shift of both the thiourea and amide NH protons upon addition of the carboxylate guests as their tetrabutylammonium salts. Data from the ^{1}H NMR titrations were fitted to a 1:1 (host:guest) binding isotherm by using NMRTit HG software.[27] Titration experiments were carried out in a range of solvents of increasing polarity, from $CDCl_3$ to CD_3CN and [D_6]DMSO.

Binding studies in $CDCl_3$ (Table 2): Initial investigation of the binding properties of receptor **1** was carried out with *N*-Boc-phenylalanine (Boc = *tert*-butoxycarbonyl). Addition of *N*-Boc-L-phenylalanine tetrabutylammonium salt to a solution of **1** in $CDCl_3$ led to a significant downfield shift of the NH signals, consistent with the formation of strong hydrogen bonds ($\Delta\delta = 1.4$ ppm for thiourea NH and $\Delta\delta = 0.65$ ppm for amide NH). As expected, the binding data could be fitted to a 1:1 binding isotherm ($K_a = 890\,M^{-1}$, $\Delta G = -16.8$ kJ mol^{-1}), consistent with the proposed mode of binding. However, addition of *N*-Boc-D-phenylalanine led to very similar changes in the ^{1}H NMR spectrum of the macrocycle, and the titration data gave $K_a = 800\,M^{-1}$ ($\Delta G = -16.5$ kJ mol^{-1}) and hence a small degree of enantioselectivity in favour of L binding ($\Delta\Delta G^{L-D} = -0.3$ kJ mol^{-1}, $K_a^{L/D} = 1.1$).

To check the influence of the bulky *tert*-butyloxycarbonyl moiety in the recognition process, binding studies were carried out with the same amino acid but with a smaller protecting group. Binding data obtained for *N*-Ac-L-phenylalanine were only slightly different from the values obtained with the Boc-protected guest ($K_a = 770\,M^{-1}$, $\Delta G = -16.5$ kJ mol^{-1}), while *N*-Ac-D-phenylalanine gave a smaller association constant ($K_a = 490\,M^{-1}$, $\Delta G = -15.3$ kJ mol^{-1}) and better enantioselectivity ($\Delta\Delta G^{L-D} = -1.2$ kJ mol^{-1}, $K_a^{L/D} = 1.6$).

Table 2. Binding constants K_a and free energies of complexation $-\Delta G$ for the 1:1 complexes formed between macrocycle **1** and tetrabutylammonium salts of N-protected amino acids in $CDCl_3$ at 298 K.

Guest	K_a [M^{-1}]	$-\Delta G$ [kJ mol^{-1}]	$\Delta\Delta G^{L-D}$ [kJ mol^{-1}]	$K_a^{L/D}$
N-Boc-L-Phe	8.87×10^2	16.8	−0.3	1.1
N-Boc-D-Phe	7.96×10^2	16.5		
N-Ac-L-Phe	7.74×10^2	16.5	−1.2	1.6
N-Ac-D-Phe	4.86×10^2	15.3		
N-Ac-L-Gln	1.47×10^3	18.1	+0.6	0.8
N-Ac-D-Gln	1.94×10^3	18.7		

Binding studies were also carried out with *N*-Ac-glutamine, since the amide group of the amino acid side chain of this guest could form additional hydrogen bonds with the macrocycle, and if these interactions were specific for just one enantiomer, then chiral recognition would be improved. Furthermore, computational calculations (vide infra), which were able to reproduce the measured enantioselectivities with the phenylalanine substrates, suggested that the selectivity with *N*-Ac-glutamine as guest would be reversed in favour of the D enantiomer.

Addition of one equivalent of *N*-Ac-L-glutamine to a solution of **1** in $CDCl_3$ caused considerable downfield shifts of the NH signals ($\Delta\delta = 1.5$ ppm for the thiourea NH signal and $\Delta\delta = 0.45$ ppm for the amide NH signal), and the binding data could again be fitted to a 1:1 binding isotherm ($K_a = 1500\,M^{-1}$, $\Delta G = -18.1$ kJ mol^{-1}). The addition of one equivalent of *N*-Ac-D-glutamine also led to similar downfield shifts of the thiourea NH signal ($\Delta\delta = 1.5$ ppm) and to

a slightly larger shift of the amide NH signal ($\Delta\delta =$ 0.75 ppm). The binding data could be fitted to a 1:1 binding isotherm and gave $K_a = 1900\,M^{-1}$ ($\Delta G = -18.7\,kJ\,mol^{-1}$). Thus this guest is indeed bound more strongly than the phenylalanine derivatives, and the enantioselectivity, although still modest, is reversed with this guest ($\Delta\Delta G^{L-D} = +0.6\,kJ\,mol^{-1}$, $K_a^{L/D} = 0.8$), as predicted computationally (vide infra).

The obvious assumption is that the glutamine primary amide can establish additional interactions (hydrogen bonds) that drive the host:guest system into a binding conformation which is slightly favoured for the D enantiomer. However, the computational results show this assumption to be incorrect (vide infra).

Binding studies in CD_3CN (Table 3): Binding studies with the *N*-Boc-phenylalanine and the *N*-Ac-phenylalanine salts in CD_3CN gave a similar picture to that obtained in the less polar $CDCl_3$. However, the shifts of the NH signals on addition of guest carboxylate are greater in CD_3CN (e.g., $\Delta\delta =$ 2 ppm for the thiourea NH signal of macrocycle **1** in CD_3CN after addition of one equivalent of *N*-Boc-L-phenylalanine) and association constants are correspondingly larger.

Table 3. Binding constants K_a and free energies of complexation $-\Delta G$ for the 1:1 complexes formed between macrocycle **1** and tetrabutylammonium salts of N-protected amino acids in CD_3CN at 298 K.

Guest	$K_a\,[M^{-1}]$	$-\Delta G\,[kJ\,mol^{-1}]$	$\Delta\Delta G^{L-D}\,[kJ\,mol^{-1}]$	$K_a^{L/D}$
N-Boc-L-Phe	4.80×10^3	21.0	−1.8	2.1
N-Boc-D-Phe	2.30×10^3	19.2		
N-Ac-L-Phe	4.69×10^3	20.9	−1.6	1.9
N-Ac-D-Phe	2.41×10^3	19.3		
N-Ac-L-Ala	2.93×10^3	19.8	+1.8	0.5
N-Ac-D-Ala	6.18×10^3	21.6		

The stronger binding in the more polar solvent indicates that the entropic gain from desolvating the host–guest surfaces is significant in driving binding in polar solvents.[28] To examine this hypothesis binding studies with *N*-Ac-alanine were carried out. This guest lacks the aromatic ring compared to phenylalanine, and thus the different solvation effects in the free host and guest and the complex might influence greatly the binding affinity in the polar aprotic CD_3CN. After addition of one equivalent of either enantiomer of *N*-Ac-alanine the amide NH signal of the macrocycle shifted downfield by about 1 ppm, while the thiourea NH signal shifted by 2.00 ppm upon addition of the L enantiomer and 2.34 ppm upon addition of the D enantiomer. The association constant for *N*-Ac-L-Ala ($K_a = 2900\,M^{-1}$, $\Delta G = -19.8\,kJ\,mol^{-1}$) is smaller than that for *N*-Ac-L-Phe ($K_a = 4700\,M^{-1}$, $\Delta G = -20.9\,kJ\,mol^{-1}$) while the association constant for *N*-Ac-D-Ala ($K_a = 6200\,M^{-1}$, $\Delta G = -21.6\,kJ\,mol^{-1}$) is larger than that for *N*-Ac-D-Phe ($K_a = 2400\,M^{-1}$, $\Delta G = -19.3\,kJ\,mol^{-1}$), and thus the enantioselectivity is reversed ($\Delta\Delta G^{L-D} = +1.8\,kJ\,mol^{-1}$, $K_a^{L/D} = 0.5$). Comparing the two L enantiomers, the binding affinity indeed decreases, as a possible result of the different solvation effects. However, this may just be speculation, since the association constant for *N*-Ac-D-Ala is the biggest observed in CD_3CN, and this suggests that other factors influence complex formation, for example, steric effects.

Binding studies in [D_6]DMSO: In the more competitive solvent [D_6]DMSO, addition of either enantiomer of the *N*-Boc-phenylalanine salt did not lead to any noteworthy change in the ^1H NMR spectrum of the macrocycle. Computations revealed, as expected, that the energetic cost of breaking the hydrogen bonds with the solvent molecules is not compensated by the binding of a guest molecule (vide infra).

Computational Studies

Macrocycle conformational preferences: Molecular modeling was used to probe the degree of influence of solvent polarity on the conformational preferences of macrocycle **1**. Conformational searches were performed using MacroModel 8.1[29,30] and random adjustment of defined torsion angles using the Monte Carlo multiple minima (MCMM) algorithm.[31] Solvent effects on macrocyclic conformation for two solvents on opposite ends of the polarity scale were studied (H_2O and $CHCl_3$) by using the generalised Born/surface area (GB/SA) continuum model.[32–34] The MMFFs forcefield was used.[35,36] Gas-phase conformational preferences were calculated as a reference point for measuring solvation effects. The same result was obtained in each case (gas phase, in $CHCl_3$ and H_2O): a nearly planar global minimum macrocycle, as shown in Figure 4.

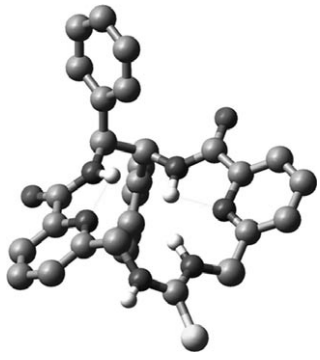


Figure 4. Global minimum conformation of macrocycle **1** in H_2O as determined from a Monte Carlo (MC) conformational search using the MMFFs forcefield and GB/SA model for solvation.

Results for the superimposition of the global minima are shown in Table 4. The effects of solvation on the macrocycle are minimal. The root mean square (RMS) distance between gas-phase and chloroform-solvated macrocycle heavy atoms is just 0.0743 Å. This difference is greater between the gas-

Table 4. Effect of solvation on the gas-phase structure of macrocycle **1**: superimposition of computed global minimum conformations.[a]

Phase	RMS[b]	RMS (no Ph groups)[b]
gas	0.0	–
$CHCl_3$	0.0743	0.0427
H_2O	0.5498	0.2688

[a] As calculated using the MMFFs forcefield.[35,36] Global minimum conformations from MCMM conformational searches. [b] Root mean square (RMS) distance between heavy atoms in Å.

phase and water-solvated macrocycles (0.5498 Å), but this is mainly because of different orientations of the phenyl rings. Excluding the heavy atoms of the phenyl rings reduces the RMS to 0.2688 Å. Computations therefore consistently predict a rigid macrocycle with minor deviations in the global minimum conformation with changes in the solvent environment. Located slightly higher in energy for the three phases (gas, $CHCl_3$ and H_2O) is a conformation similar to the X-ray crystal structure with DMSO. Table 5 lists the results of superimposition of these conformations on the X-ray crystal coordinates. The more polar the macrocyclic environment, the closer in energy this conformation is to the global minimum. Most notably, the conformation in H_2O is the second lowest energy conformation and just 2.4 kJ mol^{-1} above the global minimum. The computational conformations obtained in all cases have RMS distances between heavy atoms of about 0.5 Å in each case. This reduces to about 0.3 Å on excluding the heavy atoms of the phenyl ring where one of the rings adopts a tilted conformation relative to the crystal structure in each case.

Table 5. Results for superimposition of computed MMFFs gas- and solution-phase conformations of macrocycle **1** on the X-ray crystal structure in DMSO.

Phase	Conformer[a]	Energy[a]	RMS[b]	RMS (no Ph groups)[b]
gas	3	+15.7	0.4703	0.2766
$CHCl_3$	4	+13.0	0.5120	0.3908
H_2O	2	+2.4	0.5390	0.2894
DMSO[c]	"1"	"0"	0.5696	0.3590

[a] Energy with respect to global minimum conformation in kJ mol^{-1}. [b] Root mean square distance between heavy atoms in Å. [c] The DMSO conformer "1" at energy "0" is the average minimised conformation from MD simulations, and for comparison purposes is taken here to be the global minimum.

Although DMSO is not parametrised for use with the continuum GB/SA model in MacroModel 8.1, the conformational preferences of macrocycle **1** in DMSO were extracted from our molecular dynamics (MD) study for binding (or lack of binding) of amino acids in explicit DMSO (vide infra). The results reveal that binding of a single DMSO molecule in the macrocycle cavity is favoured over binding of the amino acid, as was experimentally observed. The average minimised macrocycle conformation from the MD simulations is shown in Figure 5 with the DMSO molecule fitting well within the macrocycle cavity to maximise the hydrogen-bond enthalpic contributions. Binding of the DMSO molecule results in a more planar macrocyclic geometry than the global minima observed in gas, $CHCl_3$ and H_2O phases. The superimposition of the computed DMSO-binding macrocyclic conformation from the MD simulations and the X-ray structure of the crystal from DMSO are shown in Figure 6. The RMS distances between the structures are 0.5696 Å and 0.3590 Å including and excluding phenyl groups in the superimposition, respectively (Table 5). The hydrogen-bond distances between the sulfoxide oxygen atom and the four NH groups are predicted to be tighter than in the DMSO crystal structure (2.03–2.37 Å) and range from 1.88 to 2.05 Å. Such discrepancy in H-bond distances is not unexpected when using forcefields. However, on optimization of the MMFFs conformation by higher level DFT calculations at the B3LYP/6-31+G* level, much closer agreement with experiment was obtained. Key structural features of the DFT optimised DMSO/macrocycle conformation compared with the corresponding crystal structure features are listed in Table 1. The hydrogen-bond distances at this level of theory range from 2.03 to 2.65 Å, while the angles (128.7–161.3°) are also in reasonably good agreement with the X-ray crystal values (139.9–153.8°), although only a medium-sized 6-31+G* basis set was used. Finally, CH–π interactions between one of the DMSO methyl groups and a macrocyclic phenyl group, which were present in the

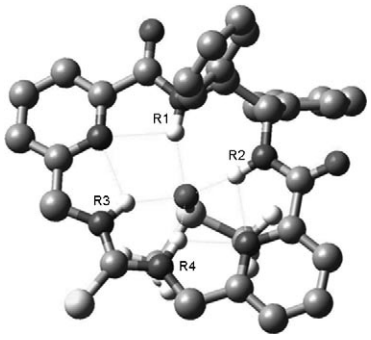


Figure 5. Average macrocycle **1** conformation from MD simulations in explicit DMSO showing efficient binding to the oxygen atom of one DMSO solvent molecule.

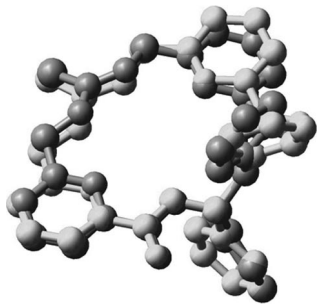


Figure 6. Superimposition of an average minimised macrocyclic conformation from MD simulations in DMSO and the X-ray crystal structure in DMSO.

MMFFs global minimum, are accentuated by the DFT optimisation. The overall excellent agreement of the computed conformational preferences with experiment highlights the effectiveness of the MMFFs forcefield in modeling the macrocycle.

Binding studies in $CHCl_3$: Modeling of macrocycle **1** for binding of N-protected α-amino acids was first performed for *N*-Ac-phenylalanine and *N*-Boc-phenylalanine to check for agreement with experiment. Conformational searches for binding studies were performed by using the mixed-mode MCMM/LMCS algorithm and, as before, the MMFFs forcefield and continuum GB/SA model for $CHCl_3$ solvation. With this conformational search algorithm, as well as following the torsional variations in the system using MCMM, the "soft" vibrational modes (low-frequency eigenvectors) were also used to direct the search by using the low mode conformational search (LMCS) algorithm.[37] MCMM/LMCS has already been used successfully for exhaustive conformational searches of organic compounds,[38] receptor–ligand binding studies in proteins[39] and studies on chiral separation.[1,2,4,21] Conformations generated from the searches were reminimised to increase the ratio between conformations found and conformations with good convergence, and to discard previously unidentified duplicate conformations. Enthalpy differences (0 K) for binding of L and D enantiomers to the macrocycle ($\Delta\Delta H^{L-D}$) were calculated from the global minimum conformations, and $\Delta\Delta G^{L-D}$ values obtained using the L and D configurations in separate MINTA computations of G^L and G^D.

The modeling results indicate that the carboxylate is bound by four hydrogen bonds, one from each NH group of the macrocycle to just one of the carboxylate oxygen atoms, with additional interactions in most cases between the amino acid NH group and the pyridine nitrogen atom. An intramolecular hydrogen bond from the amino acid NH group to the bound carboxylate oxygen atom is often present. To illustrate this, the lowest energy conformations for binding of *N*-Ac-phenylalanine are displayed in Figure 7. The hydrogen bonds are predicted to be in general tighter for binding of the L enantiomers and range from 1.79 to 2.11 Å compared to 1.87–2.08 Å for the D enantiomers. In the case of *N*-Ac-alanine (Figure 8), however, the hydrogen bonds are tighter (1.77–2.04 Å) due to reduced steric hindrance of this guest, which allows deeper binding within the macrocycle cavity. The binding of the D enantiomer is less favourable, however, as the host:guest complex does not have the additional interaction between the amino acid NH group and pyridine nitrogen atom.

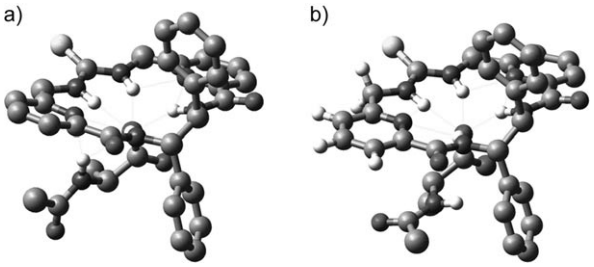


Figure 7. Molecular modeling global minimum conformations obtained for complexes of a) the L and b) the D enantiomer of *N*-Ac-phenylalanine with macrocycle **1** in $CHCl_3$.

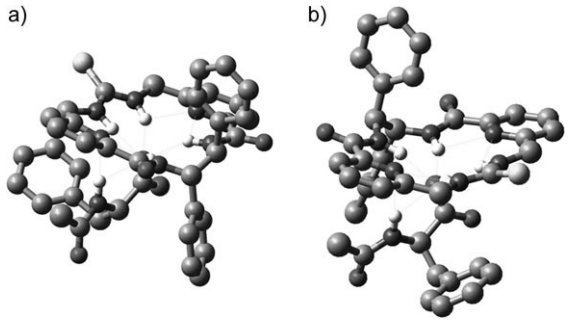


Figure 8. Molecular modeling global minimum conformations obtained for complexes of a) the L and b) the D enantiomer of *N*-Ac-alanine with macrocycle **1** in $CHCl_3$.

To summarise, it is clear that in all cases the main enthalpic contribution to chiral discrimination is limited by steric hindrance to entry of the carboxylate into the macrocyclic cavity, which will be significantly less for binding of one carboxylate oxygen atom rather than both, as in the original proposed complex model (Figure 1).

The thermodynamic results (computational and experimental) for binding of N-protected amino acids to macrocycle **1** are shown in Table 6. For the test ligands *N*-Ac-phenylalanine and *N*-Boc-phenylalanine the calculations accurately

Table 6. Thermodynamic properties for the chiral separation of amino acids by macrocycle **1** in $CHCl_3$.[a]

Guest	Experiment	Computation			Agreement
	$\Delta\Delta G^{L-D}$	$\Delta\Delta H^{L-D}$[b]	$\Delta\Delta G^{L-D}$[c]	$\Delta\Delta S^{L-D}$[d]	
N-Ac-Phe	−1.3	−1.4	−4.7	−3.3 [↑]	yes
N-Boc-Phe	−0.2	−1.3	−1.4	−0.1 [↑]	yes
N-Ac-Gln	+0.4	+2.7	+2.3	−0.4 [↓]	yes
N-Ac-Ala	–	−4.6	−4.9	−0.3 [↑]	
N-Boc-Ala	–	−3.0	−1.4	+1.6 [↓]	
N-Ac-Ser	–	−0.3	−0.7	−0.4 [↑]	
N-Boc-Ser	–	−0.9	−2.0	−1.1 [↑]	
N-Ac-Val	–	−0.8	−0.6	+0.2 [↓]	
N-Boc-Val	–	+0.3	+0.3	0.0 [–]	
N-Ac-Asp	–	−4.1	−4.8	−0.7 [↑]	
N-Boc-Asp	–	−2.8	−1.9	+0.9 [↓]	
N-Ac-Tyr	–	−2.0	−1.4	+0.6 [↓]	
N-Boc-Tyr	–	+5.9	+3.9	+2.0 [↓]	
N-Ac-Iso	–	−4.3	−1.6	+2.7 [↓]	
N-Ac-Try	–	+1.2	−1.2	−2.4 [–]	
N-Ac-Cys	–	+2.9	+2.7	−0.2 [↓]	

[a] All values are in kJ mol^{-1}. [b] At 0 K. Calculated from difference in energies between global-energy L and D ligand complexes from the MCMM/LMCS conformational searches. [c] As calculated using MINTA at 300 K. Error in $\Delta\Delta G^{L-D}$ values is ±1.2–1.4 kJ mol^{-1}. [d] Entropy estimates, calculated as the difference between $\Delta\Delta H^{L-D}$ (0 K) and $\Delta\Delta G^{L-D}$ (300 K). Arrows indicate whether entropy increases or decreases chiral separation relative to value of $\Delta\Delta H^{L-D}$.

predict the experimental direction of enantioselectivity but also significantly the relative degree of enantioselectivity.

Following the success of the computational model for the test ligands, a "virtual library" of amino acids was screened. The calculations indicated that for some substrates the sense of enantioselectivity is reversed, so that binding of the D enantiomer is favoured. Therefore, to further check the quality of the computational screening method, the enantioselective binding of *N*-Ac-glutamine was measured experimentally ($\Delta\Delta G^{L-D} = +0.4$ kJ mol^{-1}, vide supra) and verified the sense of enantioselectivity predicted computationally ($\Delta\Delta G^{L-D} = +2.3$ kJ mol^{-1}). Furthermore, the relative degree of enantioselectivity between experiment and computation for all three test ligands is completely consistent. We note that the MINTA errors for $\Delta\Delta G^{L-D}$ in Table 6 are not negligible (± 1.2–1.4 kJ mol^{-1}) and could be reduced further by using a greater number of sampling points per block (see Experimental Section). However, even within the bounds of the current errors the experimental trend is still followed by the test ligands. Further reduction of MINTA errors will not alter this key result. With respect to the binding in the *N*-Ac-glutamine global minima, an additional intramolecular hydrogen bond is present in the complexes of both enantiomers in which the NH group of the glutamine amide side chain interacts with the unbound carboxylate oxygen atom. The global minimum complexes are shown in Figure 9. Although the host:guest complex of the D enantiomer does not have the interaction between the amino acid NH group and pyridine nitrogen atom, it is slightly more enthalpically favourable than the L complex. A breakdown of the energy components for the global minimum complexes cannot assign the difference to any one component. However, the largest energy difference is found for the torsion-energy components. The diversity can be assigned to differences in the seven-membered ring formed by the additional intramolecular hydrogen bond, where in the case of the L complex extra torsional strain is involved in forming a "boat"-type conformation as opposed to the "chair"-type conformation in the L complex (Figure 9). From the full set of results in Table 6 it is clear that enantioselectivity is modest for all guests tested. The best enantioselectivities were obtained for *N*-Ac-alanine, *N*-Ac-phenylalanine and *N*-Ac-aspartate, with computational $\Delta\Delta G^{L-D}$ values of -4.7 to -4.9 kJ mol^{-1}. However, with a rigid macrocycle such as **1** only a few structural features contribute to discrimination between binding of L and D enantiomers except for the two phenyl groups, one pointing above the plane and the other below. A few trends can be observed in Table 6. The largest computed $\Delta\Delta G^{L-D}$ values obtained (>4 kJ mol^{-1}) are for the *N*-Ac-protected amino acids. These all have a favourable entropic contribution to separation/binding of the same enantiomer (L), favoured both by enthalpy and entropy. Hence, the degree of enantioselectivity is considerably dependent on entropy.

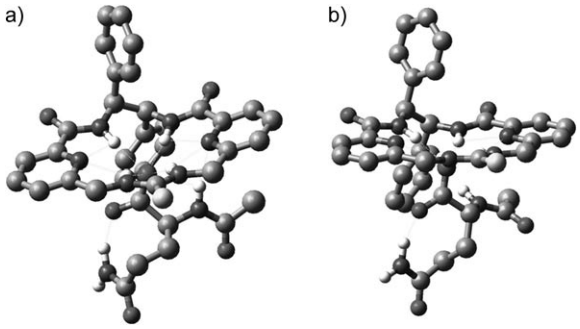


Figure 9. Molecular modeling global minimum conformations obtained for complexes of a) the L and b) the D enantiomer of *N*-Ac-glutamine with macrocycle **1** in $CHCl_3$.

Binding studies in CH_3CN: The origin of the enantioselectivity experimentally observed for macrocycle **1** in the more polar solvent CH_3CN was probed computationally for the three guests: *N*-Ac-alanine, and *N*-Ac- and *N*-Boc-phenylalanine. As the GB/SA solvation model has not as yet been parametrised for acetonitrile with MacroModel, a different modeling approach was used. Molecular dynamics simulations at 300 K using the MMFFs forcefield and explicit CH_3CN solvation were performed. MacroModel substructures were used to define the system with a frozen outer solvation shell, semi-frozen middle solvation shell and only the centre shell containing macrocycle, ligand and inner solvent molecules completely free. Following the initial minimization and equilibration period, sample conformations from the data-collection phase were collected and clustered into groups with similar geometry. Enthalpy differences for binding of L and D enantiomers to the macrocycle at 300 K were calculated ($\Delta\Delta H^{L-D}$), and the effects of entropy on these values analyzed by performing cluster analysis and intermolecular hydrogen-bond analysis of the conformations saved during the data-collection phase.

The results from the MD simulations are shown in Tables 7 and 8. The computed enthalpies ($\Delta\Delta H^{L-D}$, Table 7) predict the experimental enantioselectivity for L over D binding for both *N*-Ac- and *N*-Boc-phenylalanine (-2.6 and -3.7 kJ mol^{-1}, respectively). Entropy estimates from clustering of the saved MD trajectory conformations and measurement of their average H-bond populations indicates that en-

Table 7. Comparison between computational and experimental results for enantioselectivity predictions in CH_3CN.

Guest	Experiment[a]		Computation[a]			Agreement
	$\Delta\Delta G^{L-D}$	$\Delta\Delta H^{L-D}$	entropy direction[b]	$\Delta\Delta G^{L-D}$ est[b]	$\Delta\Delta G^{L-D}$ calcd[c]	
N-Boc-Phe	-1.8	-3.7	L	<-3.7	–	yes
N-Ac-Phe	-1.6	-2.6	L	<-2.6	–	yes
N-Ac-Ala	$+0.8$	-2.4	D	>-2.4	$+2.8$	yes

[a] In kJ mol^{-1}. [b] Entropy estimates based on conformation clustering and percentage H-bonding populations (see Table 8 and text). [c] As calculated using FEP. The error in the $\Delta\Delta G^{L-D}$ value is ± 0.6 kJ mol^{-1}.

Table 8. Analysis of molecular dynamics trajectory conformations for binding of amino acids to macrocycle **1** in CH_3CN as an indicator of the direction of entropy contributions to chiral separation.

Guest	No. of clusters (members of cluster 1)		Entropy favours	Intermolecular H-bond populations [%]		Entropy favours
	L complex	D complex		L complex	D complex	
N-Boc-Phe	8 (47)	8 (59)	L	43.7	47.8	L
N-Ac-Phe	10 (53)	6 (76)	L	40.6	45.9	L
N-Ac-Ala	9 (26)	18 (17)	D	47.5	41.9	D

tropy also favours L binding (Table 8). For binding of *N*-Boc-phenylalanine eight different types of conformations among the trajectory conformations were saved for each enantiomer complex. However, the average percentage of possible intermolecular hydrogen bonds in the L (43.7%) compared to the D conformations (47.8%) clearly indicates that the complexes of the L enantiomer are less tightly bound and hence will be favoured by entropy. The situation is more clear-cut for *N*-Ac-phenylalanine, for which both the number of clusters and the average percentage of hydrogen bonds in the trajectory conformations indicate that entropy favours L binding. Hence in both cases (for *N*-Boc- and *N*-Ac-Phe binding) computation predicts $\Delta\Delta G^{L-D} > \Delta\Delta H^{L-D}$ in favour of L binding and in agreement with experiment (Table 7). For binding of *N*-Ac-alanine the situation is quite different: $\Delta\Delta H^{L-D}$ is in favour of L binding ($-2.4\,kJ\,mol^{-1}$), in contradiction to experiment where binding of the D enantiomer is preferred ($\Delta\Delta G^{L-D} = +0.8\,kJ\,mol^{-1}$). However, the entropy estimates from the clustering of MD conformations and percentage hydrogen-bond analysis indicates a strong entropy preference for D binding. The 18 types of conformations for the D complex compare to only nine for the L complex, while the average percentage hydrogen bonding for D complexation (41.9%) is much less than for L (47.5%).

As discussed in the Experimental Section, entropy effects in the more polar solvents play a key role in binding and hence in differences in binding free energies. The question remains whether the entropy effect in favour of binding the D enantiomer is sufficient to override the enthalpic preference for the L enantiomer and give an overall free-energy preference for D binding in agreement with experiment. To answer this, $\Delta\Delta G^{L-D}$ was calculated directly by using MMFFs. MacroModel free-energy perturbation (FEP) calculations using the Monte Carlo/stochastic dynamics (MC/SD) conformational search algorithm were performed. The success of FEP is largely dependent on a thorough scan of the conformational PES. MC/SD is a very efficient conformational search method[40,41] and is the recommended algorithm for use with MacroModel FEP calculations.[42] For reasons of computational efficiency we therefore chose MC/SD. However, as MMFFs is not available for use with MacroModel 8.1 FEP computations, we had to make some approximations: we identified the torsion and partial charges as the key forcefield parameters in conformation determination and exchanged AMBER* forcefield parameters with their MMFFs equivalents. FEP calculations were then performed with AMBER* but with MMFFs torsion and partial-charge parameters. By this method, calculation of $\Delta\Delta G^{L-D}$ shows that entropy effects are so great that the trend is indeed reversed for $\Delta\Delta H^{L-D}$: MMFFs predicts enantioselectivity in favour of D binding ($\Delta\Delta G^{L-D} = +2.8\,kJ\,mol^{-1}$) and again in agreement with experiment.

To summarise, as in $CHCl_3$, the enantioselectivity in CH_3CN is dependent on a positive entropy effect on the separation for *N*-Ac- and *N*-Boc-phenylalanine. However, in the case of *N*-Ac-alanine, enthalpy effects in favour of L separation are outweighed by entropy effects in favour of D binding, so that $\Delta\Delta G^{L-D}$ becomes positive. Although more time consuming, the MD and MCSD/FEP simulations in explicit solvent gave excellent agreement between computation and experiment.

Binding studies in DMSO: Molecular dynamics simulations at 300 K with explicit DMSO solvation were performed to probe the lack of binding observed experimentally in this solvent. A similar system setup to that used for the CH_3CN MD simulations was used. The MMFFs forcefield was used and substructures employed as before. Two different starting configurations were used: macrocycle with already bound *N*-Ac-L-alanine, and macrocycle with a bound DMSO molecule within the macrocycle cavity. The average enthalpy for the unbound *N*-Ac-alanine simulation ($\langle H_{unbound}\rangle = 3730\,kJ\,mol^{-1}$) was found to be $46\,kJ\,mol^{-1}$ lower than that for the bound case ($\langle H_{bound}\rangle = 3776\,kJ\,mol^{-1}$). The energetic cost of breaking the hydrogen bonds with the solvent molecule is therefore not compensated by binding of a guest molecule. As observed in the crystal structure, the small DMSO molecule fits particularly well inside the cavity of the receptor, with CH–π interactions between a DMSO methyl group and one of the aromatic rings of the macrocycle (Figure 5).

Discussion

Design of macrocycle **1** was based on the highly enantioselective receptor **2**[1] but with a smaller cavity and just one carboxylate binding site. X-ray structural studies and ^1H NMR spectra in $CDCl_3$, CD_3CN or $[D_6]$DMSO show that the macrocycle has a rigid backbone and adopts essentially the same conformation in solvents of different polarity, although the thiourea moiety is able to twist in or out of the cavity. In DMSO the thiourea moiety twists into the cavity to allow the four NH groups of the macrocycle to interact with the sulfoxide oxygen atom of a DMSO molecule. This is confirmed by computation.

Macrocycle **1** forms strong 1:1 (host:guest) complexes with several N-protected amino acids in $CDCl_3$ and CD_3CN, but enantioselectivities were modest. Changing the guest's protecting group (from the small acetyl to the bulky *tert*-butoxycarbonyl) or moving from aliphatic (alanine) to aromatic (phenylalanine) amino acids did not improve the enantiose-

lectivity of the receptor. Using glutamine, an amino acid with additional binding functionality, did lead to increased binding affinity and reversed enantioselectivity. However, computation revealed that the extra functionality does not interact with the receptor but serves to create an extra intramolecular hydrogen bond within the bound enantiomer.

No binding was observed in the more competitive solvent DMSO. As observed in the crystal structure, a DMSO molecule can fit into the host cavity with tight hydrogen bonding of the sulfoxide oxygen atom to two thiourea and two amide protons, as shown by the short interatomic distances. Molecular dynamics simulations confirmed our assumption that binding of a single DMSO molecule is energetically more favourable due to the stronger intermolecular interactions.

Molecular modeling simulations suggest a different mode of binding to that originally proposed. Instead of a typical bidentate motif between the thiourea moiety and the carboxylate group, modeling suggests that the four NH groups of the macrocycle form hydrogen bonds with only one of the carboxylate oxygen atoms,[43] similar to the interaction observed in the structure of the crystal obtained from DMSO between the amide and a pyridine nitrogen atom of the guest. The binding process proved to be largely entropy driven.

The MMFFs forcefield parameters proved to be very effective in modeling the macrocycle, and excellent agreement was observed between experimental and computational results, which replicate not only the binding preference between the two enantiomers of each single guest in free-energy calculations (both FEP and MINTA), but also the relative degree of enantioselectivity between different ligands. To exploit such efficiency a "virtual library" of amino acid guests was created and chiral separation with macrocycle **1** in $CHCl_3$ was screened computationally by using the relatively fast and efficient MINTA algorithm.

Conclusion

Herein we present the success of a complementary experimental and computational approach for the analysis of the separation efficiency of a new chiral receptor that could represent a future direction for enantioselective separations. The work described here highlights the potential of MINTA and FEP for calculation of free energies when used in combination with high-quality forcefield parameters. Although the FEP calculations in CH_3CN were very time consuming, MINTA was particularly efficient in screening the enantioselectivities for the binding of a "virtual library" of amino acids. A similar screening in CH_3CN was not practical with FEP for reasons of computational expense, the computational mark-up in using explicit solvent for CH_3CN being a significant factor.[3] As mentioned in the introduction, with the advent of polarisable forcefields fast approaching,[22] the potential to substitute expensive and time-consuming experimental procedures for screening databases of receptors/ligands with state-of-the-art molecular modeling methods such as MINTA provides a valuable alternative to the supramolecular chemist.

Experimental Section

General experimental and instrumentation: Reactions were carried out in solvents of commercial grade which, where necessary, were distilled prior to use (for solvent-distilling procedures, see W. L. F. Armarego, D. D. Perrin, *Purification of Laboratory Chemicals* (4th ed.), Pergamon, Oxford, 1996). CH_2Cl_2 was distilled from calcium hydride. TLC was conducted on foil-backed sheets coated with silica gel (0.25 mm) which contained the fluorescent indicator UV_{254}. Column chromatography was performed on Sorbsil C60, 40–60 mesh silica.

^{1}H NMR spectra were obtained at 400 MHz on a Bruker DPX 400 spectrometer. ^{13}C NMR spectra were recorded at 100 MHz on a Brüker DPX 400 spectrometer. Spectra were referenced to the residual solvent peak of the deuterated solvent. Infrared spectra were recorded on BIORAD Golden Gate FTS 135. Spectra were obtained on neat solids. Melting points were determined in open capillary tubes using a Gallenkamp Electrothermal melting point apparatus and are uncorrected. Optical rotations were measured on a PolAr2001 polarimeter in the stated solvent; the concentration is given in g per 100 mL. Electrospray mass spectra were obtained on a Micromass platform with a quadrupole mass analyzer. High-resolution accurate mass measurements were carried out at 10000 resolution on a Brüker Apex III FT-ICR mass spectrometer.

(14*S*,15*S*)-14,15-Diphenyl-4-thioxo-3,5,13,16,22,23-hexaaza-tricyclo-[16.3.1.^{17,11}]tricosa-1(22),7,9,11(23),18,20-hexaene-12,17-dione (1): Diamine **3** (200 mg, 0.42 mmol) was dissolved in dry CH_2Cl_2 (500 mL) and Et_3N (63 μL, 0.45 mmol) added. Thiophosgene (34 μL, 0.45 mmol) was dissolved in dry CH_2Cl_2 (2 mL) and slowly added by syringe pump over 3 h to the diamine solution. The mixture was stirred for 17 h under a slow stream of nitrogen. The solvent was removed in vacuo and the residue was purified by flash column chromatography (70% ethyl acetate/petroleum ether up to neat ethyl acetate) to yield macrocycle **1** as a white solid (37 mg, 17%). R_f = 0.30 (80% ethyl acetate/petroleum ether); m.p. 140–142 °C; $[\alpha]_D^{21}$ = 125.0 (c = 1 in $CHCl_3$); ^{1}H NMR (400 MHz, CD_3CN): δ = 9.06 (d, J = 8.5 Hz, 2H; NHCO), 7.85 (m, 2H; pyrH), 7.48–7.22 (m, 16H; pyrH, CH Ar and NHCS), 5.86 (br s, 2H; C*H*NHCO), 5.04 ppm (br s, 4H; CH_2NHCS); ^{13}C NMR (100 MHz, CD_3CN): δ = 164.6 (0), 158.1 (0), 149.7 (0), 141.2 (0), 139.3 (0), 129.3 (1), 128.3 (1), 127.9 (1), 126.0 (1), 121.2 (1), 57.2 (2), 48.2 (2) ppm; IR (neat): $\tilde{\nu}_{max}$ = 3298 (br), 2360 (w), 1658 (s), 1527 (m), 1455 cm^{-1} (w); MS (ES$^+$): *m/z*: 1067.2 [2*M*+Na]$^+$, 545.2 [*M*+Na]$^+$, 523.2 [*M*+H]$^+$; HRMS (ES$^+$) calcd for $C_{29}H_{27}N_6O_2S_1^+$: 522.1838; found: 522.1834.

General procedure for NMR titration experiments: All ^{1}H NMR titration experiments were conducted on a Bruker AM 300 spectrometer at 298 K. $CDCl_3$ was passed over a pad of basic alumina prior to use and collected over molecular sieves (4) Å). A sample of host was dissolved in the deuterated solvent. A portion of this solution was used as the host NMR sample and the remainder used to dissolve a sample of the guest, so that the concentration of the host remained constant throughout the titration. Successive aliquots of the guest solution were added to the host NMR sample and ^{1}H NMR spectra recorded after each addition. The hydrogen atoms monitored during binding studies were the thiourea or amide protons in the host molecule. The changes in chemical shifts of all the host signals as a function of guest concentration were analyzed with purpose-written software, kindly provided by C. A. Hunter,[27] assuming a 1:1 binding mode. This program fits the data to the appropriate binding model to yield the association constant, the bound chemical shift and the free chemical shift. For a greater degree of accuracy, association constants are quoted as an average of all the association constants obtained from each proton monitored in the host molecule, and in all cases errors are less than 10%.

Computational details

Macrocycle phase-dependent conformational preferences: To probe the rigidity of the macrocycle in the gas phase and in solvents of different po-

larity, MCMM conformational searches[31] were performed. The MMFFs forcefield[35,36] was used with solvation effects included with the GB/SA continuum model.[32–34] 1500 MCMM steps were used in each case with 2–5 of the defined macrocyclic torsions randomly adjusted each step. The definition of ring opening/closures allowed for the torsional variations within the macrocycle. A "use-directed" Monte Carlo structure selection (MCSS) option was employed so that the least investigated structure within a 50 kJ mol^{-1} energy window of the current "global" minimum was used as the starting point for the next MC step. Computations were performed using MacroModel 8.1.[29,30] The conformational preferences of the macrocycle in DMSO were determined from MD simulations in explicit solvent, with more accurate geometric parameters obtained using DFT quantum chemical calculations (vide infra).

Macrocycle binding studies: For macrocycle **1** and binding of *N*-protected α-amino acid enantiomers in $CHCl_3$, conformational searches were performed using the mixed-mode MCMM/LMCS algorithm.[31,37] Global ligand translations were included in the searches, and these are seen as critical to success of the conformational search. In total, 2–5 torsions were adjusted in each MCMM step, with a 1:1 ratio of MC torsional moves and/or ligand translation to the LMCS low-mode moves. The MCSS use-directed structure selection option was again used. 3000 MCMM/LMCS steps were performed for binding of each enantiomer in each case, with an energy window of 50 kJ mol^{-1} above the global minimum used for saving conformations. The MMFFs forcefield and GB/SA model for $CHCl_3$ solvation were used as before.

For the free-energy calculations using MINTA, all conformations generated from the conformational searches were first reminimised (truncated Newton conjugate gradient (TNCG)[44]) using the MacroModel "multiple minimization" algorithm to attempt to fully minimise any unconverged conformations and to discard any unidentified duplicate conformations from the original MCMM/LMCS conformational search. The resulting conformations (generally <500) were used in separate calculations of G^{R-L} and G^{R-D} at 300 K for L and D enantiomer:receptor (R) complexes, respectively. MINTA integrals were calculated as block averages with 10×1000 energy evaluations per conformation. Numerical integration in all degrees of freedom was used throughout, with hard and soft limits of 1 Å and three units of standard deviation, respectively, for sampling along the normal modes. Note that MINTA gives a statistical error calculated from the block averages.[13] The error decreases with increasing number N of sampling points for single block sampling as $\sim1/\sqrt{N}$. With multiple blocks the error is therefore smaller, so that MINTA allows for calculation of free energies with very high accuracy if required. Finally, as the free energies of the isolated enantiomer are the same ($G^{L}=G^{D}$), the binding free energy difference $\Delta\Delta G^{L-D}$ was calculated simply as [Eq. (1)].

$$\Delta\Delta G^{L-D} = G^{R-L} - G^{R-D} \quad (1)$$

The conformational searches and MINTA computations were performed using MacroModel 8.1.

For the binding studies in CH_3CN, since there is no GB/SA parametrisation available within MacroModel for CH_3CN, instead explicit solvation models were used and MacroModel 8.1 molecular dynamics simulations were used to probe the conformational energy surface. Global minimum receptor:enantiomer complexes from the conformational searches in $CHCl_3$ were taken as starting points for the simulations, which were imbedded in a previously equilibrated $35.42\times35.42\times35.42$ Å cubic box of 512 CH_3CN solvent molecules at experimental density (0.7857 g cm^{-3} at 298.15 K and 1 atm).[45] The equilibration of the CH_3CN solvent box is described in detail elsewhere.[1] Overlapping solvent molecules and those giving rise to repulsive interactions with macrocycle:ligand systems were first deleted. Next, substructures consisting of three shells were created. The first shell was the unconstrained free shell and contained macrocycle:ligand and between 20 and 38 (depending on ligand size) of the nearest solvating CH_3CN molecules; the second shell was a semi-frozen shell containing 13–27 solvent molecules; while the third outer shell was completely frozen and contained 180–316 CH_3CN molecules. Remaining solvent molecules outside the outer shell were deleted. Short minimizations (2500 steps, PR conjugate gradient (PRCG) method[46]) were first performed to remove any bad contacts. MD simulations using a standard constant-temperature velocity-Verlet algorithm were subsequently performed using a simulation temperature and time step of 300 K and 1 fs. For the MD runs, 500–750 ps of equilibration was followed by lengthy 10–17.5 ns data-collection phases. The larger *N*-Ac-phenylalanine (500 ps equilibration, 15 ns data collection) and *N*-Boc-phenylalanine (750 ps equilibration, 17.5 ns data collection) systems were given longer simulation times compared to the *N*-Ac-alanine complex (500 ps equilibration, 10 ns data collection). Sample conformations (100) from the resulting MD trajectories (data-collection phases) were subsequently collected and analyzed for binding properties such as percentage intermolecular H-bond populations in the complexes, and clustered into groups having similar conformations (based on the RMS distances between heavy atoms). A representative conformation was obtained for each cluster, the representative conformation from the largest cluster being the most important. Clustering calculations were performed using NMRclust 1.2.[47] Finally, for the binding enthalpies, $\Delta\Delta H^{L-D}$ (300 K) was calculated using an analogous equation to Equation (1) for enthalpies.

For the FEP calculations of $\Delta\Delta G^{L-D}$ for binding of *N*-Ac-alanine enantiomers in CH_3CN, the mixed-mode MC/SD conformational search algorithm was used. The system setup was similar to that used for the MD simulations, with substructures again employed. The first shell contained the macrocycle, bound L- or D-*N*-Ac-alanine enantiomer and 20 CH_3CN solvent molecules; the second shell was semi-frozen with 13 CH_3CN molecules; and the third, outer shell was completely frozen with 180 CH_3CN molecules. A target temperature of 298.15 K was used. The FEP calculations were performed starting from the D ligand complex, and the chiral C–H group of the enantiomer was gradually mutated over 16 windows into C–CH_3 (and vice versa) to form the L ligand complex. Double wide sampling was employed. Within each window, 2500-step TNCG minimization was followed by equilibration for 50 ps and data collection over 3 ns. To be consistent with our other calculations, we needed to employ the MMFFs forcefield. However, with MacroModel 8.1 and FEP calculations only the AMBER* forcefield is fully supported and implemented. For this reason, simulations were performed using AMBER* with MMFFs partial charges and torsion parameters, the parameters most critical to conformational analysis. Parameter changes were effected by editing the input maestro (.mae) coordinates file for charges and the AMBER* forcefield file (amber.fld) for torsions.

Finally, the lack of experimentally observed binding properties in DMSO was probed by MD simulations in explicit solvent. Two different starting conformations were used: 1) macrocycle and unbound *N*-Ac-L-alanine enantiomer system (this also allowed us to investigate the macrocycle conformation preferences in bulk DMSO for comparison with the DMSO crystal structure) and 2) macrocycle with bound *N*-Ac-L-alanine. Systems 1 and 2 were prepared by introducing the macrocycle and enantiomer to a previously equilibrated 512 molecule DMSO cubic box ($39.28\times39.28\times39.28$ Å) at experimental density (1.096 g cm^{-3}).[1,45] Overlapping solvent molecules and those giving rise to repulsive interactions were then deleted. Next, three substructures or shells were created, as for the MD simulations in CH_3CN. The shells for each system 1 and 2 contained the same number of solvent molecules so that the thermodynamic (enthalpy) data from the MD simulations could be directly compared. The first shell contained unconstrained macrocycle, guest and 65 DMSO solvent molecules; the second, semi-frozen shell 32 DMSO molecules; and the third, outer, frozen shell 277 DMSO molecules. The semi- and fully frozen shells were identical for both systems. Remaining solvent outside the three shells was deleted. The systems and shells were much larger than those used in the CH_3CN simulations, as system 1 had to solvate a separated macrocycle and enantiomer system. Short minimizations (5000 steps, PRCG algorithm) were performed to remove remaining bad contacts.[46] Full MD simulations followed which consisted of 2 ns equilibration and 30 ns data-collection phases. The two MD simulations 1 and 2 allowed us to analyze the binding (or lack of binding) of amino acids (*N*-Ac-alanine) in DMSO by comparing the average enthalpies from the simulations, $\langle H_{unbound}\rangle$ from 1 versus $\langle H_{bound}\rangle$ from 2. Further, as stated above, simulation 1 allowed us to monitor DMSO solvation effects on

the macrocycle conformation alone. To do so, 100 snapshots from the MD trajectory (data-collection phase) were saved and the macrocycle with bound DMSO molecule clustered into groups having similar conformations (again based on RMS distances between heavy atoms). A representative conformation was obtained for each cluster. The representative conformation from the largest cluster was minimised (1000 steps, TNCG algorithm) and compared with the DMSO crystal structure. Clustering calculations were performed using NMRclust 1.2.[47] Finally, the average minimised conformation from the MD simulations was also optimised at the B3LYP/6-31+G* level. DFT calculations were performed using Jaguar 6.0.[48]

Acknowledgements

We thank the European Commission (TMR Network grant "Enantioselective Separations" ERB FMRX-CT-98–0233) for financial support, postgraduate fellowships (A.R.) and post-doctoral fellowships (J.M.H.). We are also grateful to Schrödinger Inc. for technical support with the FEP Macromodel calculations.

[1] A. Ragusa, S. Rossi, J. M. Hayes, M. Stein, J. D. Kilburn, *Chem. Eur. J.* **2005**, *11*, 5674–5688.
[2] J. M. Hayes, M. Stein, J. Weiser, *J. Phys. Chem. A* **2004**, *108*, 3572–3580.
[3] R. H. Henchman, J. D. Kilburn, D. L. Turner, J. W. Essex, *J. Phys. Chem. B* **2004**, *108*, 17571–17582.
[4] A. del Rio, J. M. Hayes, M. Stein, P. Piras, C. Roussel, *Chirality* **2004**, *16*, S1-S11.
[5] K. B. Lipkowitz, *J. Chromatogr. A* **2001**, *906*, 417–442.
[6] R. L. DesJarlais, R. P. Sheridan, G. L. Seibel, J. S. Dixon, I. D. Kuntz, R. Venkataraghavan, *J. Med. Chem.* **1988**, *31*, 722–729.
[7] Glide 2.7, Schrödinger, LLC, New York, NY, **2003**.
[8] C. McMartin, R. S. Bohacek, *Comput.-Aided Mol. Des.* **1997**, *11*, 333–344.
[9] D. S. Goodsell, A. Olsen, *J. Protein Chem.* **1990**, *9*, 195–202.
[10] R. W. Zwanzig, *J. Chem. Phys.* **1954**, *22*, 1420–1426.
[11] P. Kollman, *Chem. Rev.* **1993**, *93*, 2395–2417.
[12] M. R. Mruzik, F. F. Abraham, D. E. Schreiber, G. M. Pound, *J. Chem. Phys.* **1976**, *64*, 481–491.
[13] I. Kolossváry, *J. Phys. Chem. A* **1997**, *101*, 9900–9905.
[14] I. Kolossváry, US Patent Serial No. 08/940, 145, Mode Integration (MINTA): Method and Apparatus for Selecting a Molecule Based on Conformational Free Energy, filed September 29, **1997**.
[15] M. S. Head, J. A. Given, M. K. Gilson, *J. Phys. Chem. A* **1997**, *101*, 1609–1618.
[16] I. Kolossváry, *J. Am. Chem. Soc.* **1997**, *119*, 10233–10234.
[17] D. Q. McDonald, W. C. Still, *Tetrahedron Lett.* **1992**, *33*, 7747–7750.
[18] P. A. Kollman, *Acc. Chem. Res.* **1996**, *29*, 461–469.
[19] A. E. Mark, S. P. van Helden, P. E. Smith, L. H. M. Janssen, W. F. van Gunsteren, *J. Am. Chem. Soc.* **1994**, *116*, 6293–6302.
[20] a) M. J. Potter, M. K. Gilson, *J. Phys. Chem. A* **2002**, *106*, 563–566; b) C. -E. Chang, M. J. Potter, M. K. Gilson, *J. Phys. Chem. B* **2003**, *107*, 1048–1055.
[21] A. Ragusa, J. M. Hayes, M. E. Light, J. D. Kilburn, *Eur. J. Org. Chem.* **2006**, *16*, 3545–3549.
[22] a) J. R. Maple, Y. Cao, W. Damm, T. A. Halgren, G. A. Kaminski, L. Y. Zhang, R. A. Friesner, *J. Chem. Theory Comput.* **2005**, *1*, 694–715; b) V. M. Anisimov, G. Lamoureux, I. V. Vorobyov, N. Huang, B. Roux, A. D. MacKerrell, *J. Chem. Theory Comput.* **2005**, *1*, 153–168; c) S. Patel, C. L. Brooks, *J. Comput. Chem.* **2004**, *25*, 1504–1514; d) L. R. Olano, S. W. Rick, *J. Comput. Chem.* **2005**, *26*, 699–707; e) S. Patel, C. L. Brooks, *Molec. Simul.* **2006**, *32*, 231–249.
[23] S. Rossi, G. M. Kyne, D. L. Turner, N. J. Wells, J. D. Kilburn, *Angew. Chem.* **2002**, *114*, 4407–4409; *Angew. Chem. Int. Ed.* **2002**, *41*, 4233–4236.
[24] C. A. Hunter, D. H. Purvis, *Angew. Chem.* **1992**, *104*, 779–782; *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 792–795.
[25] Single-crystal X-ray diffraction study: Data were collected on a Bruker Nonius KappaCCD mounted at the window of an Mo rotating anode. Standard data collection and reduction procedures were followed. Crystal data for **1**: $C_{30}H_{27.5}N_{6.5}O_2S$, $M_r = 543.15$, $T = 120(2)$ K, orthorhombic, space group $P2_12_12_1$, $a = 11.7527(2)$, $b = 13.0264(2)$, $c = 35.5767(9)$ Å, $V = 5446.63(19)$ Å^3, $\rho_{calcd} = 1.325$ g cm^{-3}, $\mu = 0.16$ mm^{-1}, $Z = 8$, 27591 reflections collected, 11764 independent reflections ($R_{int} = 0.1051$), final R indices [$I > 2\sigma I$]: $R_1 = 0.0750$, $wR_2 = 0.1196$, R indices (all data): $R_1 = 0.2370$. $wR_2 = 0.1607$. CCDC-609671 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.
[26] G. M. Kyne, M. E. Light, M. B. Hursthouse, J. de Mendoza, J. D. Kilburn, *J. Chem. Soc. Perkin Trans. 1* **2001**, 1258–1263.
[27] A. P. Bisson, C. A. Hunter, J. C. Morales, K. Young, *Chem. Eur. J.* **1998**, *4*, 485–851; all binding data from NMR titration experiments are provided in the Supporting Information.
[28] D. B. Smithrud, F. Diederich, *J. Am. Chem. Soc.* **1990**, *112*, 339–343.
[29] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440–467.
[30] MacroModel version 8.1, Schrödinger, LLC, New York, NY, **2003**.
[31] G. Chang, W. C. Guida, W. C. Still, *J. Am. Chem. Soc.* **1989**, *111*, 4379–4386; M. Saunders, K. N. Houk, Y. D. Wu, W. C. Still, M. Lipton, G. Chang, W. C. Guida, W. C. Still, *J. Am. Chem. Soc.* **1990**, *112*, 1419–1427.
[32] W. C. Still, A. Tempczyk, R. C. Cawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
[33] J. Weiser, A. A. Weiser, P. S. Shenkin, W. C. Still, *J. Comput. Chem.* **1998**, *19*, 797–808.
[34] J. Weiser, P. S. Shenkin, W. C. Still, *J. Comput. Chem.* **1999**, *20*, 217–230, 586–596.
[35] T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490–519, 520–552, 553–586, 616–641; T. A. Halgren, R. B. Nachbar, *J. Comput. Chem.* **1996**, *17*, 587–615.
[36] T. A. Halgren, *J. Comput. Chem.* **1999**, *20*, 720–729, 730–748.
[37] I. Kolossváry, W. C. Guida, *J. Am. Chem. Soc.* **1996**, *118*, 5011–5019.
[38] P. A. Greenidge, S. A. M. Merette, R. Beck, G. Dodson, C. A. Goodwin, M. F. Scully, J. Spencer, J. Weiser, J. J. Deadman, *J. Med. Chem.* **2003**, *46*, 1293–1305.
[39] I. Kolossváry, W. C. Guida, *J. Comput. Chem.* **1999**, *20*, 1671–1684.
[40] F. Guarnieri, W. C. Still, *J. Comput. Chem.* **1994**, *15*, 1302–1310.
[41] D. Q. McDonald, W. C. Still, *J. Am. Chem. Soc.* **1994**, *116*, 11550–11553.
[42] MacroModel 8.5 Reference Manual, Schrödinger, LLC, New York, NY, **2003**, p. 152.
[43] J. L. Sessler, A. Andrievsky, V. Král, V. Linch, *J. Am. Chem. Soc.* **1997**, *119*, 9385–9392.
[44] J. W. Ponder, F. M. Richards, *J. Comput. Chem.* **1987**, *8*, 1016–1024.
[45] www.chemfinder.com
[46] E. Polak, G. Ribiere, *Rev. Fr. Inform. Rech. Oper.* **1969**, *16-R1*, 35–43.
[47] L. A. Kelley, S. P. Gardner, M. J. Sutcliffe, *Protein Eng.* **1996**, *9*, 1063–1065; http://neon.chem.le.ac.uk/nmr.clust/.
[48] Jaguar 6.0, Schrödinger, LLC, New York, NY, **2003**; www.schrodinger.com.

Received: September 7, 2006
Published online: January 2, 2007

