Molecular recognition of oxoanions based on guanidinium receptors

Pascal Blondeau,[a] Margarita Segura,[a] Ruth Pérez-Fernández†[a] and Javier de Mendoza*[ab]

Received 1st September 2006
First published as an Advance Article on the web 9th November 2006
DOI: 10.1039/b603089k

Guanidinium is a versatile functional group with unique properties. In biological systems, hydrogen-bonding and electrostatic interactions involving the arginine side chains of proteins are critical to stabilise complexes between proteins and nucleic acids, carbohydrates or other proteins. Leading examples of artificial receptors for carboxylates, phosphates and other oxoanions, such as sulfate or nitrate are highlighted in this *tutorial review*, addressed to readers interested in biology, chemistry and supramolecular chemistry.

1. Introduction

Nature frequently uses guanidinium moieties to coordinate different anion groups. Present in the side chain of the amino acid arginine, the guanidinium group forms strong ion-pairs with oxoanions such as carboxylates or phosphates in enzymes and antibodies, and it also contributes to the stabilisation of protein tertiary structures *via* internal salt bridges, mainly with carboxylates.[1] Not surprisingly, guanidinium-based compounds are found in many drugs and have been extensively used in molecular recognition studies, leading to the design and synthesis of various receptors for anions.[2]

The capacity of the guanidinium group to bind oxoanions is due to its geometrical Y-shaped, planar orientation, which directs the hydrogen bonding, and to its high pK_a value (around 12–13),[3] which ensures protonation over a wide pH range. The positive charge is delocalized over the three nitrogen atoms, and four out of the five hydrogen bond donors present in the guanidinium group of arginine can complement bidentate oxoanion acceptors, along the two edges available (Fig. 1). This accounts for the geometrical versatility of the binding modes. From the energy point of view, binding to oxoanions results from both ion-pairing and hydrogen bonding, and this turns out to be a difficult challenge in highly polar solvents or in water. In fact, the binding energy arises from the difference of the energy released by the host–guest interactions and the energy penalty necessary to remove the solvation shell around the host, which is quite high in water.

In proteins, the guanidinium–oxoanion interaction usually occurs inside hydrophobic pockets or in areas of low dielectric constant. On the contrary, in artificial synthetic systems designed to work in water or polar solvents, complexation takes place in an environment more exposed to solvation effects which compete with the donor and acceptor sites, causing a substantial decrease of the binding. This is usually overcome by increasing the number of charges or hydrogen bond donors or by the design of more sophisticated receptors

[a]*Institute of Chemical Research of Catalonia (ICIQ), 43007 Tarragona, Spain*
[b]*Department of Organic Chemistry, Universidad Autónoma de Madrid, 28049 Madrid, Spain*
† Current address: Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom.

Pascal Blondeau

Pascal Blondeau was born in Le Mans, France in 1978. He completed his MSc degree at University of Montpellier, France, in 2002 working on hybrid self-organized materials applied to transport biomimetism. He is currently doing his PhD under the supervision of Prof. Javier de Mendoza at ICIQ in Tarragona. His research interests involve the design and synthesis of chiral guanidinium receptors for molecular recognition of anions, enantioselective recognition of carboxylic acid derivatives as well as transport of amino acids.

Margarita Segura was born in Leiden, The Netherlands in 1969. She received her MSc degree in chemistry (1993) from the University of Granada, Spain and her PhD degree (1998) from the Universidad Autónoma (Madrid), supervised by Prof. Javier de Mendoza, working on molecular recognition of oxoanions from phosphodiesters, uronic acids and dipeptides. After postdoctoral research at the University of Parma, Italy (1998–2000) working on carbohydrate recognition in water in the group of Prof. Rocco Ungaro, she again joined de Mendoza's group in Madrid as a postdoctoral researcher (2000–2003) leading a project on supramolecular donor–acceptor electroactive systems linked by multiple hydrogen-bonding. Since 2004 she has been the Group Coordinator at Prof. de Mendoza's group at ICIQ in Tarragona.

Margarita Segura

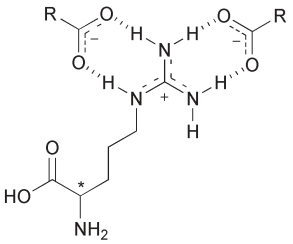


Fig. 1 The guanidinium group of arginine and its two possible binding modes with carboxylates.

where the access to the solvent is restricted. In this review, several examples on how this has been achieved in natural systems and in artificial guanidinium receptors are provided.

2. Guanidinium–oxoanion ion pairs in proteins and nucleic acids

2.1. Proteins

Protein structure has been at the forefront of research studies with the goal of better understanding the function of these biomolecules in the chemistry, physiology and pathology of the cell. Proteins are remarkably flexible and susceptible to the influence of the environment. Both intramolecular and intermolecular interactions involving the protein and the solvent define the native conformation.[4] To perform its function, a protein has to fold properly, a task where the various intra-protein or inter-protein interactions, as well as the interactions of the protein with metals or other molecules (such as co-factors, lipids or carbohydrates), are essential elements of control. An illustrative example of misfolding is the prion protein, which results in aggregated copies of the protein causing the ''Mad Cow Disease'' deadly condition.[5] Hydrogen-bonded salt bridges, such as those involving guanidinium–carboxylate, are relevant contributors to α-helical stabilization and sometimes destabilization of peptides and proteins.[6]

Guanidinium salt bridges play also important roles in enzyme active sites. Typical examples are carboxypeptidase A,[7a] creatine kinase,[7b] fumarate reductase,[7c] and malate dehydrogenase.[7d]

Protein–protein hetero-dimerization processes are often mediated by salt bridges involving arginine on one molecule and phosphorylated amino acids on the other. For example, phosphorylation of the OH group of a serine residue in a receptor enables the simultaneous interaction with two adjacent arginine residues of another receptor. On the other hand, phosphorylation of serines (or threonines) adjacent to the arginines of the same molecule slows down the attraction between the receptors.[8]

A related case is the involvement of the guanidinium group in cell–cell and cell–matrix adhesion motifs such as the tripeptide sequence RGD (arginine-glycine-aspartate). Adhesive proteins like fibronectin, osteopontin, vitronectin and collagens display the RGD sequence at their cell recognition site in extracellular matrices,[9] which is recognized by at least one member of the structural related integrins, a family of α,β hetero-dimeric transmembrane cellular receptors (Fig. 2).[10] On the cytoplasmatic side of the plasma membrane, the receptors connect the extracellular matrix to the cytoskeleton.

Thus, osteopontin (OPN), a multifunctional phosphorylated glycoprotein recognized as a key molecule in a multitude of biological processes such as bone mineralization or cancer metastasis, contains an integrin-binding RGD sequence. A significant regulation of OPN function is mediated through post-translational phosphorylation and glycosylation, a process that is essential for osteoclast attachment.[11] Osteoclasts are cells that actively reabsorb old bones so that a new bone may be replaced. Osteoporosis (bone loss) occurs when osteoclasts reabsorb bone faster than the osteoblasts cells are producing it.

Ruth Pérez-Fernández

Ruth Pérez was born in Madrid, Spain in 1975. She obtained her degree in chemistry (organic chemistry) at the Universidad Autónoma (Madrid). She was awarded a European PhD at the same university in October 2005 under the supervision of Prof. J. de Mendoza and Prof. P. Prados working in the field of protein–ligand interactions, self-assembly and dynamic evolution based on guanidinium–oxoanion interactions. In November 2005 she joined Prof. Jeremy Sanders' group at the University of Cambridge as a postdoctoral researcher and her current work concerns molecular recognition using a dynamic combinatorial approach.

Javier de Mendoza (Barcelona, Spain 1944) is Professor of Organic Chemistry at Universidad Autónoma (Madrid) and

Javier de Mendoza

Group Leader at the ICIQ (Tarragona). He obtained MSc and PhD degrees in Pharmacy (Barcelona, 1967 and 1971). He has been Assistant Professor (1971–1975), Associate Professor (1975–1980) and Full Professor at the Universities of Barcelona, Bilbao, Alcalá de Henares and Autónoma (Madrid). He pursued a postdoctoral stay (1971–1972) in Montpellier (France) under the supervision of Prof. Robert Jacquier and Dr José Elguero. He is Chevalier de l'Ordre du Mérite of France since 1994 and he was awarded in 1999 with the Research National Prize and Medal of the Spanish Royal Chemical Society. Javier de Mendoza pioneered the introduction of Supramolecular Chemistry in Spain and his current research interests range from molecular recognition to calixarene chemistry, self-assembly and catalysts design.

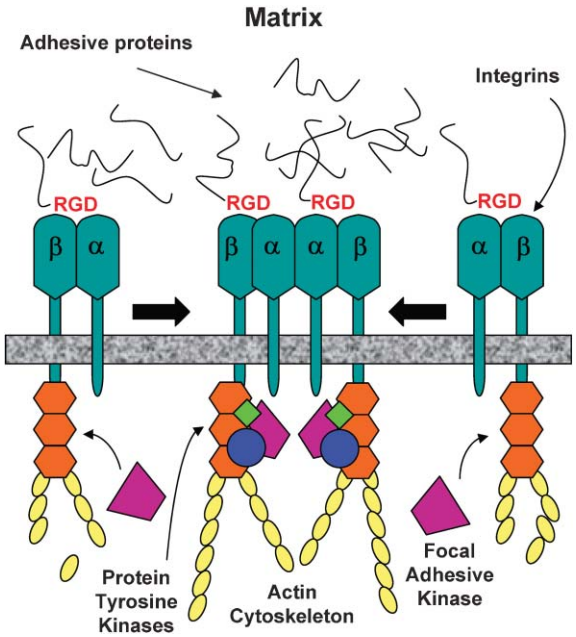


Fig. 2 Scheme of the binding of the RGD sequence to integrins in cell–cell and cell–matrix adhesion processes.

2.2. Nucleic acids

Proteins that interact with nucleic acids have a key role in biological processes. They are necessary for the control of the genetic information, replication, packaging and protection. Arginine is again essential for the interaction of proteins with DNA. In the nucleosome, in which the DNA winds around the arginine-rich histone, the amino acid side chains clearly show a direct interaction with the DNA phosphodiester chains (Fig. 3).[12] Due to the diversity of binding modes in this system, the 39 arginine residues present in the four histone proteins forming the nucleosome core may be divided into three groups: a first group of 20 arginine residues involved in histone–histone interactions not contacting DNA, followed by 7 arginines which enter the minor groove of DNA and are essential for histone–DNA binding; and a final group of 12 arginines which show direct guanidinium–phosphate salt bridge interaction. Methylation of arginine residues in the histone core leads to a conformational change allowing DNA transcription.[13] In this way, the transcription of genes can be regulated.

3. Guanidinium-based artificial receptors for oxoanions

Lehn and co-workers first reported in the late 1970's guanidinium-containing macrocycles for the recognition of phosphate PO_4^{3-} in water.[14] The weak association constants (K_a = 50 (**1**), 158 (**2**) and 251 (**3**) M^{-1}, pH titrations) can be explained in terms of the more delocalised charge of guanidinium over ammonium and accounts for the electrostatic prevailing interaction.

The guanidinium can be incorporated into a bicyclic framework (Fig. 4a) in order to improve its solubility in apolar solvents, where the hydrogen bonds are stronger, and to avoid the *anti* conformation, not suitable for hydrogen-bonding to oxoanions (Fig. 4b). As a result, the hydration of the cation is reduced and the conformational freedom restricted. Inserted into a decaline framework, the guanidinium cation becomes therefore an almost ideal complement for oxoanions, since both NH protons are docking sites for the two *syn* lone pairs of the oxoanion. The resulting ionic

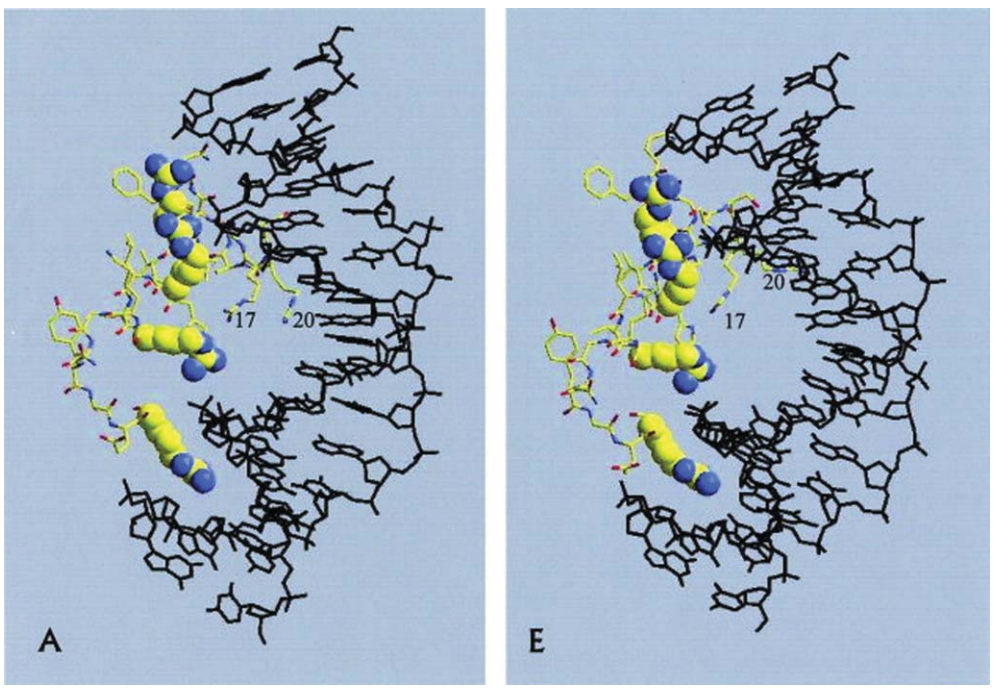
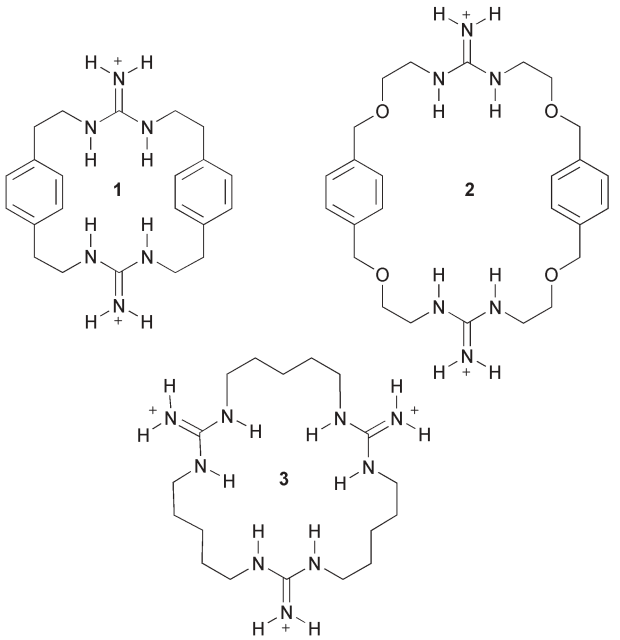


Fig. 3 Stereoview of a double stranded DNA interacting with arginines of H2A histones along the major groove. (Reprinted with permission from Subirana *et al.*[12] *Biopolymers*, 2003, **69**, 432–439. Copyright (2003) Wiley Periodicals, Inc.)



DD-AA (donor-donor–acceptor-acceptor) hydrogen-bonded complex is particularly stable and geometrically well defined. Due to the large pK_a difference between guanidinium and carboxylic acids (*ca.* 9 pK_a units in water) a trans-protonation that would destroy the salt bridge and give a less robust AD-DA hydrogen bond interaction[15] is unlikely, although it could occur in non-polar solvents, where the differences in pK_a are substantially reduced. Finally, C_2 symmetry can be introduced into the molecule by two stereogenic centres at the vicinal atoms, allowing chiral recognition of the oxoanion guest. Such a chiral bicyclic guanidinium binding subunit can be conveniently prepared in multigram quantities in nine steps from chiral amino acids (asparagine and methionine).[16]

The association constant between bicyclic guanidinium derivatives and carboxylates are quite high in chloroform or apolar solvents. Thus, UV titrations between **4** (tetraphenylborate salt) and tetrabutylammonium (TBA) *p*-nitrobenzoate gave $K_a = 7 \times 10^6$ M^{-1}.[17] The crystal structure of an acetate salt confirmed the formation of two strong symmetric hydrogen bonds between the host and the guest (N···O 2.850 Å). This first binding study confirmed the good match of oxoanions by guanidinium receptors through ion pair and a linear array of hydrogen bonds in apolar solvents.

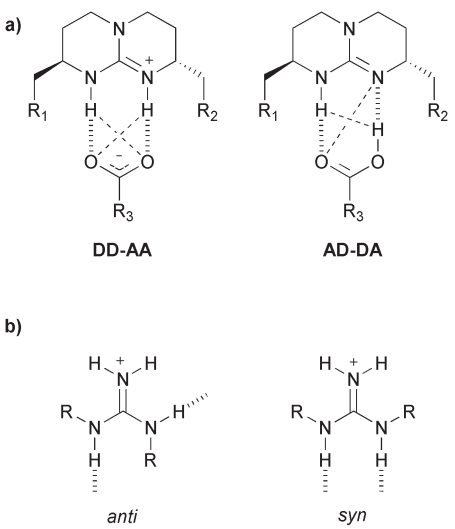
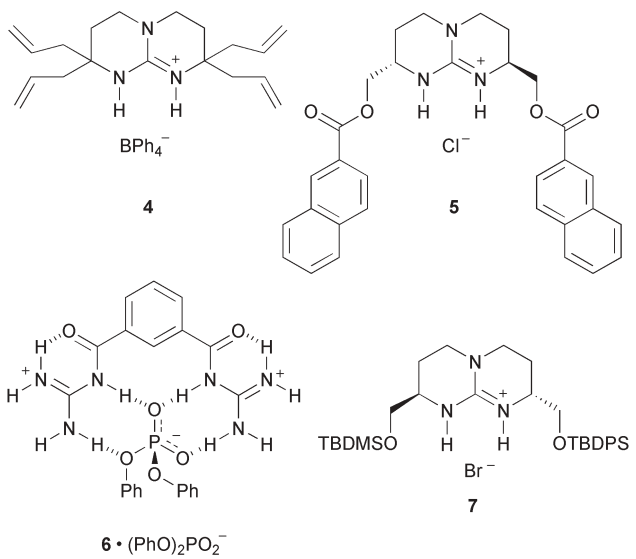


Fig. 4 a) Chiral bicyclic guanidinium receptor. b) *Anti* and *syn* conformations of guanidinium group.

We developed receptor **5** (chloride as counterion) for aromatic carboxylates, but the stability constant with TBA *p*-nitrobenzoate was much lower ($K_a = 1.6 \times 10^3$ M^{-1}, ^{1}H NMR titrations in $CDCl_3$).[18] This example illustrates the competition with the initial counterion and the importance of the counterion in binding strength: in this case the tetraphenylborate counterion results in significantly weaker binding than chloride. Thus, poorly coordinating counterions such as hexafluorophosphate or tetraphenylborate are necessary if strong binding constants are desired.

The strong deshielding of the NH signals in the ^{1}H NMR spectrum of **5**·*p*-nitrobenzoate indicates the presence of hydrogen bonds. Moreover, stacking interactions between the naphthoyl side arms and the *p*-nitrophenyl moiety are evidenced by the shifting of the aromatic signals. Despite their ionic character, hosts **4** and **5** are insoluble in water but soluble in chlorinated solvents. Thus, liquid–liquid extractions of water solutions of carboxylate salts give quantitatively the ion pair in the organic solvent, free from any competing ion.



Hamilton and co-workers synthesised bis-acylguanidinium salt **6** as a receptor for phosphodiesters. The binding constant with TBA diphenylphosphate ($K_a = 4.6 \times 10^4$ M^{-1}, measured by UV in CH_3CN), was one order of magnitude higher than with a simpler benzoylguanidinium tetraphenylborate.[19] The carbonyl groups contribute to the binding in two ways: they increase the acidity of the guanidinium NHs (but not to such an extent that trans-protonation can occur) and they pre-organise the host by intramolecular hydrogen bonds (chelation effect). The combination of these two factors and the additional hydrogen bonding from the guanidinium groups allows strong complexation in more polar solvents, such as acetonitrile.

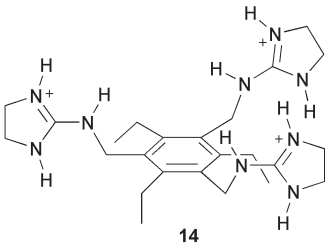
Schmidtchen studied guanidinium–carboxylate interactions by isothermal titration calorimetry (ITC).[20] The isotherm

binding curve between **7** (bromide) and tetraethylammonium acetate in acetonitrile ($K_a = 2.0 \times 10^5\ M^{-1}$) revealed that the process was both entropically and enthalpically favourable for a 1 : 1 complex. Although thermodynamic parameters could be determined in both CH_3CN and DMSO, the reaction in MeOH produced too little heat to allow quantification of the association constant. This result shows that the stabilization of the guanidinium–carboxylate is not only due to the strong electrostatic interactions (ΔH°) but also to a favourable release of solvent molecules (ΔS°), which strongly emphasises the importance of solvation in host–guest interactions, a factor often neglected in receptor design.

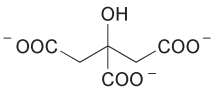
The thermodynamic aspects of dicarboxylate recognition by artificial receptors with increasingly acidic hydrogen bond donor groups such as two ureas (**8**), thioureas (**9**), or guanidiniums (**10** and **11**) in polar solvents (from DMSO to water) were studied by Hamilton (Fig. 5).[21]

As expected, association constants with carboxylate groups (**12** and **13**) increase with hydrogen acidity but are decreased in more polar solvents. While guanidinium–carboxylate association in DMSO is enthalpically driven, in more polar solvents such as methanol or water the association becomes an entropically driven process due to the liberation of solvent molecules upon binding.

Anslyn and co-workers developed receptor **14**, with three guanidinium moieties into a 1,3,5-triethyl-2,4,6-trimethylbenzene preorganized tripod platform,[22] showing selective binding towards citrate **15** in pure water ($K_a = 6.9 \times 10^3\ M^{-1}$, ^1H NMR titrations). The host was able to complex citrate even from a crude extract of orange juice, which highlights its selectivity relative to other carboxylates. This receptor shows how the solvent competition can be overcome by accumulation of hydrogen bond donors (three guanidinium subunits) in a suitable fashion.

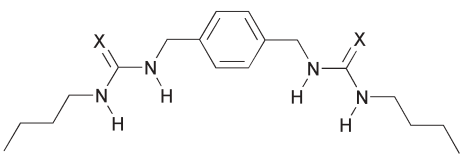


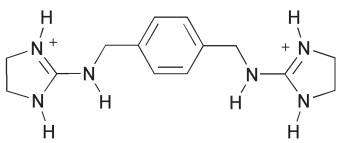
14

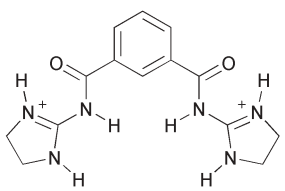


15

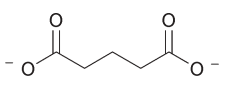
The same principles inspired Schmuck's 2-(guanidiniocarbonyl)-1*H*-pyrroles (Fig. 6), designed to complex carboxylate groups in highly competitive media, such as water.[23] Whereas the simple guanidinium cation **16** ($pK_a = 13$) does not show any sign of complexation with carboxylates in aqueous DMSO, the increased acidity of the acylguanidinium **17** ($pK_a = 7$–8), rises the binding affinity ($K_a = 50\ M^{-1}$). An additional hydrogen bond from the pyrrole NH (as in **18**) increases the association significantly ($K_a = 130\ M^{-1}$) and the additional amide group (**19**) adds a further hydrogen donor well oriented to reach the *anti* oxygen lone pair ($K_a = 770\ M^{-1}$). The predicted geometries have been confirmed by X-ray



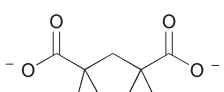




		ΔG kcal mol^{-1}	ΔH kcal mol^{-1}	ΔS cal mol^{-1} K^{-1}
			DMSO	
8 X = O	**12**	-3.1	-2.5	+5.9
	13	-3.3	-4.5	+0.1
9 X = S	**12**	-3.9	-4.1	+4.3
	13	-4.1	-5.9	+0.6
			MeOH	
10	**12**	-3.4	+3.7	+28
	13	-3.9	+4.0	+32
			MeOH	
11	**12**	-3.8	+4.0	+31
	13	-4.1	+4.4	+34



12



13

Fig. 5 Urea- (**8**), thiourea- (**9**) and guanidinium-based (**10–11**) receptors and association data for dicarboxylates **12** and **13** by isothermal titration calorimetry.

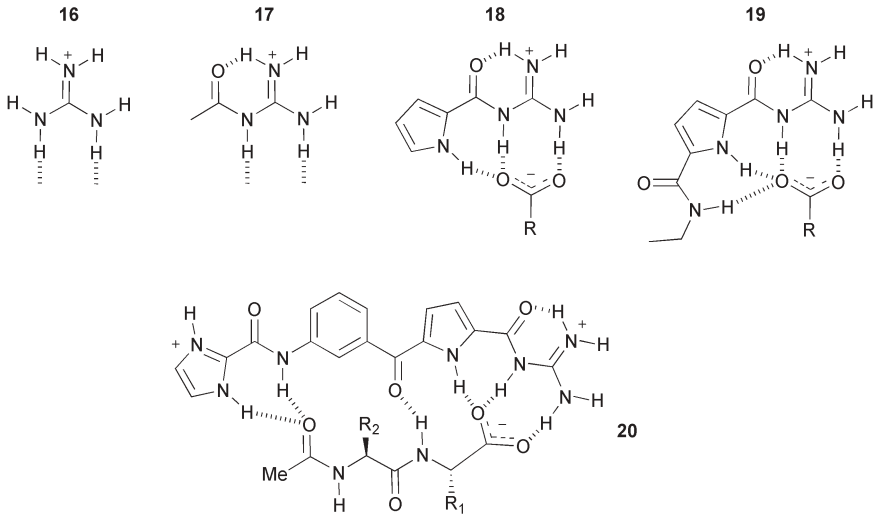


Fig. 6 Guanidinium cations **16–20** and their carboxylate complexes.

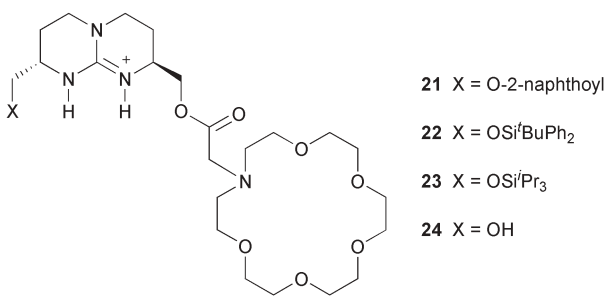
crystal structures. Even dipeptides are bound efficiently in water by a receptor such as **20** (K_a = 54300 M^{-1} for Val-Val).[24] A similar scaffold has been used in a combinatorial approach showing the importance of additional interactions caused by the side arm to improve selectivity.

4. Chiral guanidines for the enantioselective recognition of carboxylates

Chiral discrimination of anions based on abiotic receptors is still an underdeveloped area of supramolecular chemistry. Enantiomerically pure compounds are usually obtained by asymmetric synthesis, crystallisation of diastereomeric salts, kinetic resolution of racemic mixtures or chiral chromatography. An interesting alternative to these methods is the separation of enantiomers based on the complementarity of a receptor. Those processes based on the translocation of a guest between immiscible phases (chromatography, extraction, membrane transport) are particularly attractive. If the receptor is chiral, one of the enantiomers can be complexed preferentially and a kinetic resolution could be achieved. Moreover, the process needs only a catalytic amount of receptor since it can transfer several substrate molecules across the phases, without being removed from its own (stationary or liquid) phase.

In this context, a useful concept, developed for chiral chromatography, is the *three-point binding rule*, which states that a minimum of three simultaneous interactions between the chiral stationary phase and for instance one of the enantiomers are necessary to achieve enantioselection, with at least one of these interactions being stereochemically dependent.[25] For anions, receptors based on ammonium groups, amides, ureas, thioureas and guanidinium moieties, as well as porphyrins, saphyrins, or metal-containing ligands have been employed. Only chiral guanidines aimed at the discrimination of the enantiomers of amino acids will be reviewed here.

The first example of chiral recognition of a carboxylate by a guanidinium-based receptor was reported by de Mendoza in 1989.[18] Indeed, compound **5** was shown to extract enantiomeric salts of *N*-protected amino acids, such as tryptophan, with modest selectivities (up to 17% excess of *N*-Ac-L-Trp or *N*-Boc-L-Trp were extracted by (*S,S*)-**5** from water to chloroform). ^{1}H NMR titrations of the triethylammonium salts of *N*-acetyltryptophan in $CDCl_3$ gave K_a = 1000 and 500 M^{-1} for the L- and D-enantiomers, respectively. For the non protected, strongly solvated zwitterionic amino acids, receptor (*S*,*S*)-**21** was designed.[26] The compound features non self-complementary binding sites for carboxylate (the guanidinium function) and ammonium (a crown ether moiety), preventing the receptor from internal collapse, and an aromatic planar surface (a naphthalene ring) as a third point for additional stacking interactions (Fig. 7a). Up to 40% of racemic tryptophan or phenylalanine were extracted by (*S*,*S*)-**21** from saturated aqueous neutral solutions into dichloromethane, with a *ca.* 80% content of the L-enantiomer. Reciprocally, chiral host (*R*, *R*)-**21** extracts mainly D-Trp.



Further guanidinium receptors were then synthesised in order to optimise the binding and extraction properties, and were tested as membrane carriers (U-tube tests with dichloromethane between two water phases).[27] Interestingly, both single liquid–liquid extractions and U-tube transport experiments revealed that **22** and **23** transported Trp with degrees of selectivity comparable with **21**. This suggests that the aromatic naphthoyl group does not play a significant role in the discrimination process. Even compound **24**, lacking the potential π–π interaction, was enantioselective, although to a lesser extent. Another binding mode was then proposed, without participation of the naphthoyl arm, as the outcome of

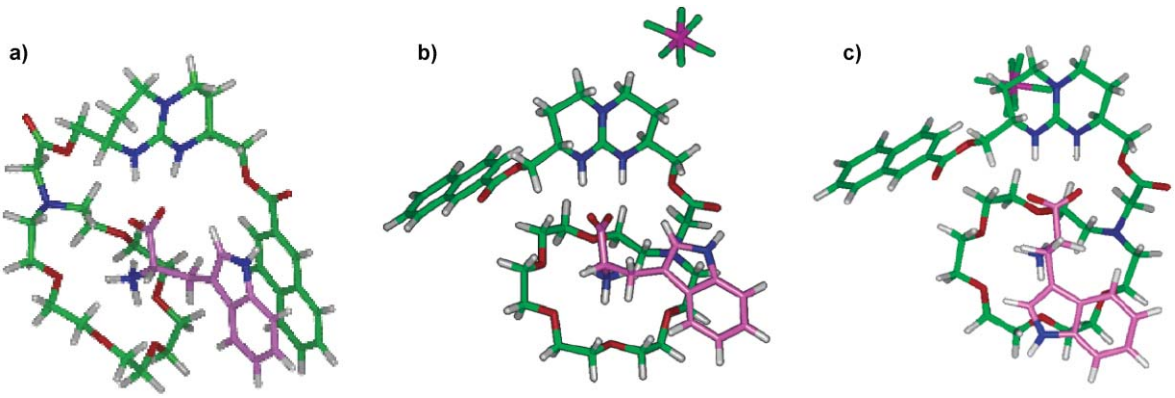
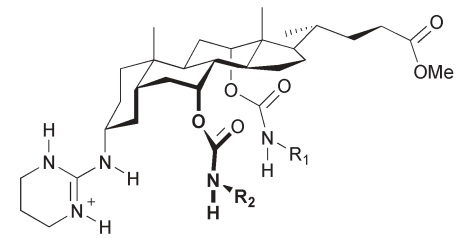
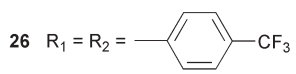


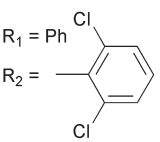
Fig. 7 a) Three-point binding mode for receptor (*S*,*S*)-**21** and L-Trp.[26b] b) Two-point binding mode for L-Trp.[27] c) Two-point binding mode for D-Trp.[27]



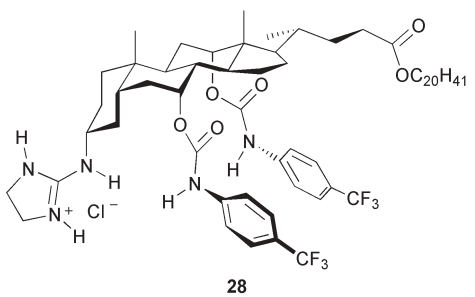
25 $R_1 = R_2 = Ph$



26 $R_1 = R_2 =$



27 $R_1 = Ph$, $R_2 =$



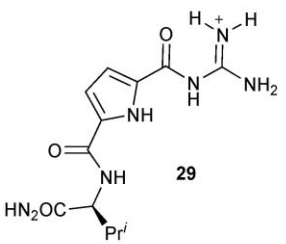
28

molecular dynamics calculations with explicit solvent molecules. In this model, binding of D-Trp exposes a highly polar area of the receptor (around the crown ether nitrogen) to the apolar solvent, causing the overall energy to increase (Fig. 7b,c).[27]

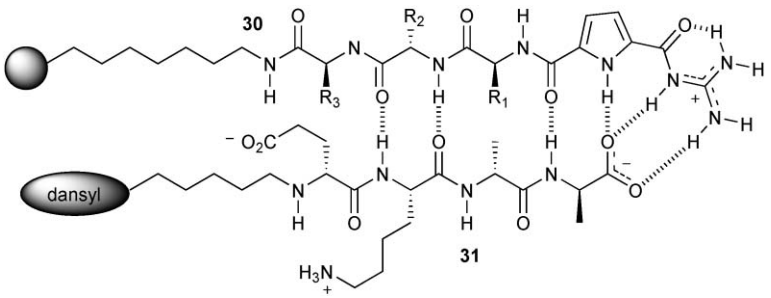
A series of receptors for *N*-protected amino acids, bearing guanidinium and carbamate moieties anchored to the curved and lipophilic surface of cholic acid (compounds **25–27**) have been reported by Davis and co-workers.[28] The chirality is provided by the steroidal framework, the guanidinium as well as the carbamate groups establishing the ion pair and hydrogen bonds with the substrate. All these hosts efficiently extract (52–87%) *N*-acyl α-amino acids from an aqueous phosphate buffer solution (pH 7.4) into chloroform. Compound **25** showed high enantioselectivity (up to 7 : 1, ^{1}H NMR measurements) for several *N*-acyl α-amino acids although this selectivity decreased dramatically for the more hindered *N*-Boc derivatives. On the contrary, chiral discrimination increased (9 : 1) with derivatives **26** and **27**, carrying the more acidic carbamoyl groups.

The highly lipophilic receptor **28** was then synthesised in gram amounts for transport studies with *N*-acetylphenylalanine either in U-tube bulk liquid membranes (dichloromethane) or with hollow-fibre membrane contactors (2.5% octanol in hexane).[29] High enantioselectivity and transport rates were observed in the U-tubes (27% of *N*-Ac-Phe transported in 24 h with 56% e.e.), as well as with the large scale hollow fibre system (*ca.* 70 equiv. of substrate transported after 48 hours) although in this case the initial selectivity (*ca.* 30%) decreased over time.

Guanidiniocarbonyl pyrrole systems have also been tested for enantioselection. Schmuck reported host **29** which was able to bind strongly carboxylates in water.[23] Despite its flexible structure and the fact that it bears only one chiral centre, this



29



30

31

receptor showed enantioselectivity towards *N*-acetylalanine (K_a = 1610 and 910 M^{-1} for *N*-Ac-L-Ala and *N*-Ac-D-Ala, respectively), a remarkable result considering the small size of alanine's side chain. Curiously, other amino acids with bulkier side chains (such as *N*-acetylphenylalanine or *N*-acetyltryptophan) showed only slight differences in binding for both enantiomers.

More recently, tris-cationic receptors based on the guanidiniocarbonyl pyrrole scaffold were developed by combinatorial chemistry. One compound (**30**, $R_1 = R_2$ = Lys; R_3 = Phe) showed efficient binding to the sequence D-Glu-L-Lys-D-Ala-D-Ala-OH (**31**) with $K_a > 10^4$ M^{-1} in buffered water.[30] This peptide sequence is related to the bacterial peptidoglycan that is recognised by the vancomycin family of antibiotics, preventing formation of the cell wall.

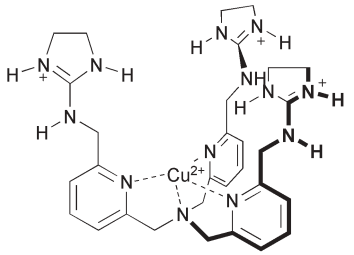
5. Phosphate, sulfate and nitrate recognition

In addition to carboxylates or phosphodiesters, other oxo-anions such as phosphate, sulfate and nitrate are biologically relevant[31] and chemically challenging to recognise, due to their weak basicity. At neutral pH, phosphate (as HPO_4^{2-}) and sulfate present a tetrahedral binding mode with two negative charges, although nitrate has a trigonal planar binding motif with just one negative charge. Thus, two guanidines are required for phosphate and sulfate but only one is needed for nitrate, among other hydrogen bond donor atoms. The design of suitable linkers between the hydrogen donors with optimal orientation and maximum participation of host's lone pairs constitutes the major issue in the field.

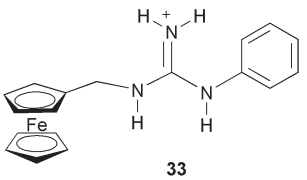
5.1. Phosphates

Anslyn and co-workers showed metallo-receptor **32** selective binding for monoprotonated phosphate (HPO_4^{2-}) and arsenate ($HAsO_4^{2-}$) over other anions (such as AcO^-, NO_3^-, HCO_3^- or Cl^-) at biological pH ($K_a = 10^4$ M^{-1} in 98 : 2 water/methanol, UV/vis and ITC titrations).[32] Hosts bearing only the Cu(II) centre were less effective ($K_a = 10^2$ M^{-1}), highlighting the role of the cavity and the presence of the guanidinium groups. Thermodynamic data showed that association of HPO_4^{2-} with guanidinium derivative **32** was both enthalpically and entropically driven, whereas complexation with an ammonium analogue was mainly governed by entropy. The different mode of binding was rationalized in terms of the different solvation energies of both binding groups.

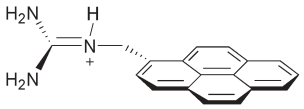
Ferrocenyl-based receptor **33** gives moderately strong complexes with pyrophosphate $P_2O_7^{4-}$ (K_a = 4600 M^{-2}, 50% methanol–water) showing a 2 : 1 host–guest stoichiometry.[33] The presence of a redox active subunit (the ferrocene) allows its use as an electrochemical sensor for this biologically relevant anion. Another receptor binding pyrophosphate in a 2 : 1 fashion ($K_a = 1.2 \times 10^8$ M^{-2} and 1.0×10^4 M^{-1} for 2 : 1 and 1 : 1 complexes, respectively) is guanidinium **34**, containing a fluorescent pyrene subunit, which appeared to be highly selective for $P_2O_7^{4-}$ over a variety of anions.[34] Moreover, ^1H NMR suggests that the two hosts in the 2 : 1 complex are self-assembled through pyrene–pyrene stacking interactions.



32

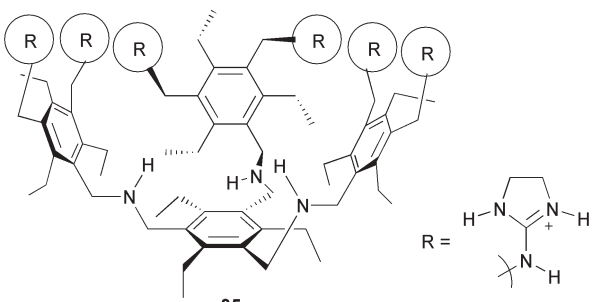


33

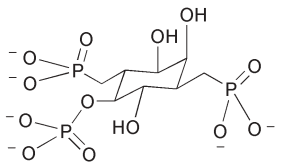


34

Polyanionic messenger inositol-1,4,5-triphosphate (**36**) was recognized by receptor **35** bearing up to six guanidinium subunits on top of a pre-organized 2,4,6-triethylbenzene platform.[35] Steric gearing causes the guanidinium groups to converge toward the cavity. As a result, a cleft-like cavity is formed. Since **35** has no chromophore, the binding constant ($K_a = 2.2 \times 10^4$ M^{-1} in a buffered solution, 1.0×10^8 M^{-1} in MeOH) was measured by competition with a fluorescent guest (5-carboxyfluorescein), which is released in the presence of the preferred guest **36**.



35

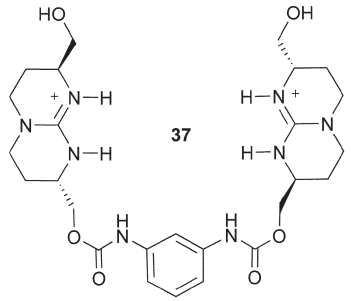


36

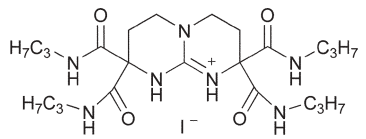
Schmidtchen designed a urethane-linked bis-guanidinium receptor **37** for the binding of ditopic tetrahedral anions.[36] A binding constant of 10^6 M^{-1} in water for *p*-nitrophenyl phosphate and cytidine-5′-phosphate was determined by ^1H NMR. Such a high value in a polar solvent was explained by the simultaneous complexation of both guanidinium groups to the tetrahedral guest.

Recently, binding studies between **38** and **4** and phosphates of different sizes were measured by both ^1H NMR and ITC in acetonitrile.[37] For **38**, ^1H NMR gave a curve fitting for a 1 : 1 stoichiometry whereas ITC predicted a 1 : 2 host–guest binding model and revealed that the binding was not caused by a large enthalpic contribution but to a strong entropic factor instead.

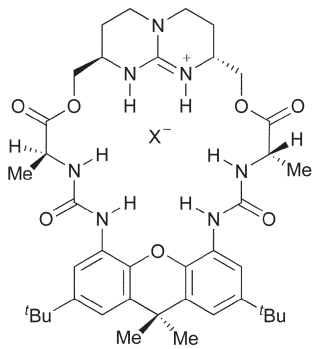
Calorimetry indeed prevents misleading conclusions from NMR in cases where rapid interconverting species are in equilibrium. Thus, introduction of several hydrogen bond donors in the receptor scaffold counteracts rather than enhances the enthalpic stabilization of the host–guest complex.



37



38



39

Macrocycle **39**, based on the chiral bicyclic guanidinium subunit, has been designed by de Mendoza and co-workers to afford six strong hydrogen bonds oriented towards its cavity to facilitate wrapping around tetrahedral oxoanions.[38] Although diphenylphosphate was readily extracted from water, the binding constant could not be measured from the tetraphenylborate salt ($K_a > 10^5$ M^{-1}) by NMR in $CDCl_3$. However, from the chloride salt the constant was 10^3 M^{-1}, which indicates that the receptor is selective for diphenylphosphate over chloride. Contrary to the expectation that the guest would be threaded across the cavity, splitting of most signals at low temperature indicated the rapid counterion scrambling between both sides of the macrocycle. Consequently, attempts to make a rotaxane by using the bulkier 3,5-di-*tert*-butyldiphenylphosphate as a template during the cyclisation (clipping) were unsuccessful.

5.2. Sulfates

Sulfate recognition by guanidinium receptors has been much less explored than carboxylate or phosphate binding. Therefore, no detailed studies comparing binding constants with sulfates and other anions are available. However, despite both phosphate and sulfate being tetrahedral, the latter is less basic, thus the affinity for guanidinium receptors is shifted towards phosphate.

In 1996 de Mendoza and co-workers reported on chiral bicyclic bis-guanidinium (**40** and **41**) and tetrakis-guanidinium (**42**) salts whose sulfate counterions, unlike the corresponding chloride salts, required hydrogen donors from two different molecules to balance the charges and to fully wrap around the anion, since the spacer CH_2SCH_2 is simply too short to use guanidines from the same chain (Fig. 8).[39] Therefore, two subunits are forced to self-assemble orthogonally around the tetrahedral anion in a double-helical structure (sulfate helicates). ^{1}H NMR spectra showed large downfield shifts of guanidinium NH's as dimers or tetramers complexed sulfate anion. Moreover, ROESY spectra confirmed intermolecular contacts due to the folded conformation.

A recent computational study concluded that for simple sulfate–guanidinium interactions several minima of similar

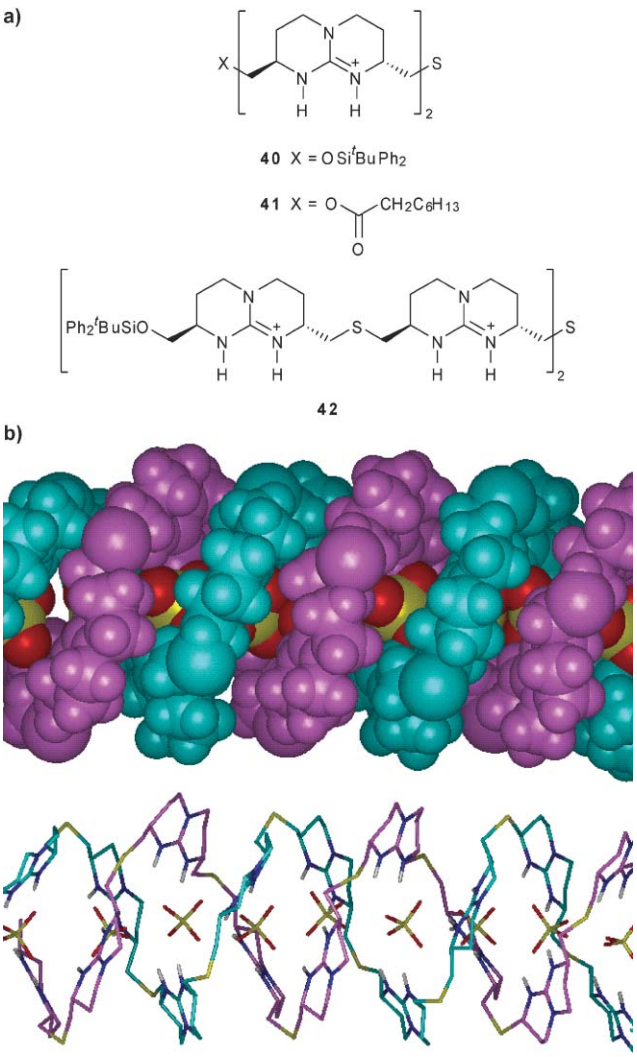


Fig. 8 a) Chiral bicyclic bis-guanidinium salts **40** and **41** and tetrakis-guanidinium salt **42**. b) Optimised model of a sulfate helicate from (*S,S*)-guanidines.

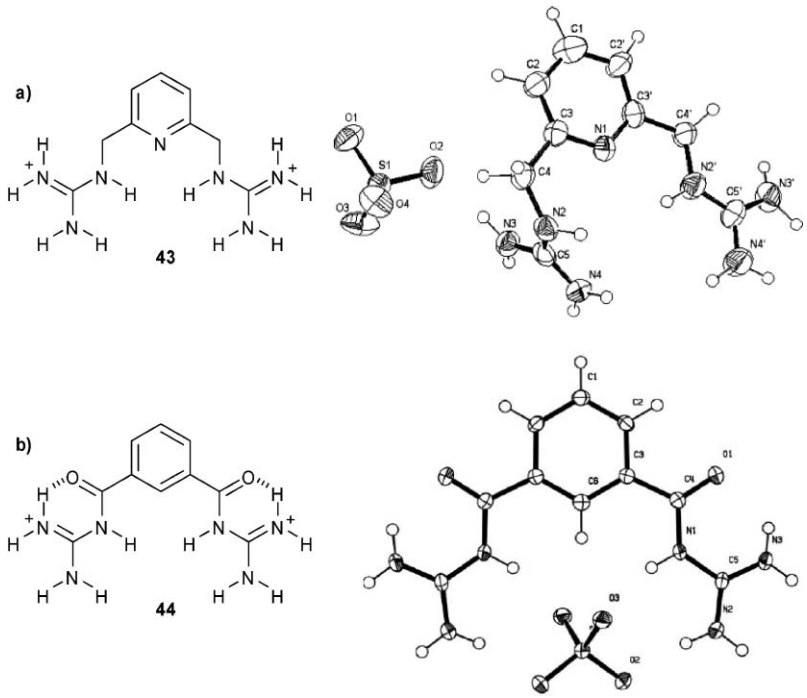
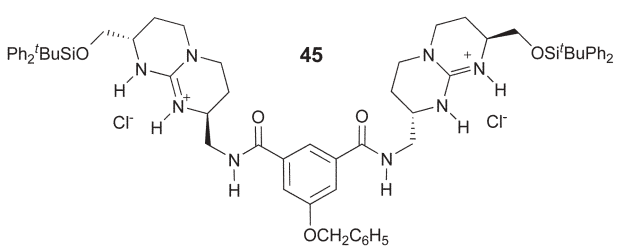


Fig. 9 Crystal structure of **43**·H_2SO_4 (a) and of **44**·H_2SO_4 (b). (Reprinted with permission from Grossel *et al.*[41] *CrystEngComm*, 2003, **5**, 77–81. Copyright (2003) The Royal Society of Chemistry.)

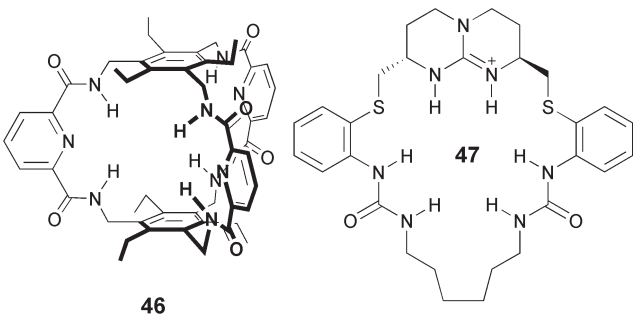
energies could be found.[40] For more complex guanidines, such as the closely related ligands **43** and **44**, the crystal structure shows 1 : 1 sulfate complexes, with a good docking of the anion into the cavity of **44**, but this was not the case for **43** (Fig. 9).[41] Interestingly, the pyridine nitrogen of **43** induces pre-organisation by intramolecular hydrogen bonding but causes repulsion of the anion due to the increased charge density in the pocket around the heteroatoms.

The energetics of the guanidinium–sulfate system have been analysed by Schmidtchen using two bicyclic guanidinium subunits linked through a suitable spacer, such as **45**.[42] ITC measurements reveal that guest complexation is strongly endothermic with entropy as the driving force. The interactions are strong enough in methanol ($K_a = 6.8 \times 10^6$ M^{-1}) to overcome the positive enthalpy change. Assuming that free host and free guest are more highly solvated than the complex, the positive enthalpy reflects the reorganisation of the solvent shell upon complexation. A comparison between receptors which combine two guanidinium groups (such as **45**) and monotopic bicyclic guanidinium ones, which showed little or no interaction with sulfate in methanol, accounts for the importance of the bridging spacer between both cationic subunits.



5.3. Nitrates

Despite its highly symmetrical trigonal planar binding mode, nitrate is a rather weak base with little tendency to establish robust hydrogen-bonded frameworks in solution. Hence, it can hardly compete with other anions such as chloride or carboxylates for the hydrogen donor sites of the receptors. For instance, with cyclophane **46**, an amide-based receptor, nitrate has a higher affinity for the receptor than chloride ($K_a = 300$ *vs.* 40 M^{-1} respectively, ^1H NMR titration in 25% CD_2Cl_2 in CD_3CN) but acetate binds better ($K_a = 770$ M^{-1}).[43]



The nitrate anion offers six optimal sites for proton location according to the number and orientation of the lone pair orbitals. These features are invariably observed in the solid state and have been supported by theoretical calculations.[44] Some X-ray structures involve ion-pairing with guanidinium donors.[45]

De Mendoza and co-workers have recently designed macrocycles, such as **47**, combining a guanidinium and two urea moieties for an optimal complement of charge, size and binding sites to nitrate.[46] Both enthalpy and entropy are

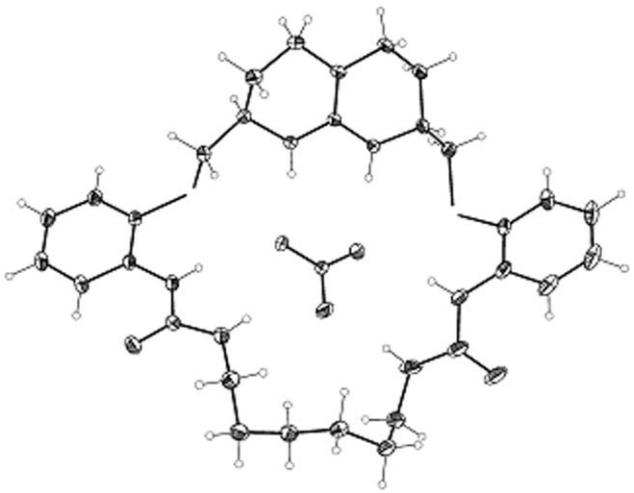


Fig. 10 Nitrate complex with guanidinium receptor **47**.

driving forces for the association (K_a = 7.4 × 10^4 M^{-1}, ITC with TBA nitrate in acetonitrile), and the selectivity ratio Cl^-/NO_3^- is rather moderate (1.3). The solid state structure (Fig. 10) nicely shows the encapsulation of the anion in the predicted orientation for optimal binding.

6. Peptide and protein surface recognition

The design of small molecules that interact with proteins is a challenging topic for the discovery of novel drugs. The catalytic sites of enzymes are well-defined cavities more or less isolated from solvent. Binding sites in the pockets are typically hydrogen bonds, salt bridges or electrostatic forces. Small molecules inspired in the structure of the substrates are often good enzyme inhibitors. On the other hand, protein–protein interactions are essential in numerous biological processes, such as cell proliferation, growth and differentiation. Unfortunately, unlike the case for the enzyme pockets, the design of small molecules aimed at disrupting such interactions between proteins, is a formidable task, essentially because the large and flat interfacial areas (typically of *ca.* 1600 $Å^2$) make it very difficult for any small molecule to be competitive. Also, protein surfaces contain complex arrangements of highly solvated functional groups, such as aspartate, glutamate, or arginine.[47]

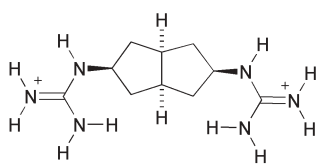
Since the distribution of groups depends on the conformation of the peptide backbone at the protein surface (α-helix, β-sheet, β-turn, *etc.*), initial efforts in the area concentrated on the design of small molecules that recognise functional side chains located at positions typical of these secondary structures. For guanidinium-based receptors, the complementary residues are obviously the carboxylate side chains of aspartate and glutamate. Hamilton synthesised compound **48**, a rigid, concave scaffold that orients two guanidinium subunits to interact with two aspartates of model peptide **49** in $i(i + 3)$ positions (K_a = 2200 M^{-1} in 10% water–methanol, NMR titrations).[48] Peptides **50** and **51**, having the aspartates in $i(i+4)$ and $i(i+11)$ positions, gave weaker complexes (K_a = 770 and 390 M^{-1}, respectively). Peptide **49** was designed to posses a significant α-helical character, a necessary condition to establish ion-pairs with **48** because in this conformation the carboxylates are spaced by 4–5 Å in an approximate parallel arrangement, so they can interact with the guanidinium groups of the receptor.

In a fruitful collaborative effort, Hamilton, Giralt, de Mendoza and their co-workers studied tetraguanidinium **42** for the same purpose.[49] Remarkably, peptide **52**, endowed with four aspartates at $i(i+3)$ positions, not only gave a strong complex with **42** in a highly competitive medium (K_a = 1.6 × 10^5 M^{-1} in 10% water/methanol, circular dicroism titrations), but also increased its α-helical content from an initial 21% to 61%.[49a] Likely, the stabilization and helical increase arises from the fact that in this conformation, each of the carboxylates lies in front of one of the guanidinium residues of **42** (Fig. 11a). It is interesting to note that the peptide sequence influences binding also by residues other than aspartates (or glutamates) in $i(i + 3)$. For instance, addition of tryptophan residues so that stacking interactions are possible between the α-helix and the bicyclic guanidinium ring (such as in **53**, Fig. 11b), results in strong binding (K_a = 1.1 × 10^8 M^{-1} in 1 : 1 water–trifluoroethanol, circular dicroism) and induces a four-fold increase of the peptide α-helical content.[49b]

Thus, oligomers of chiral bicyclic guanidines such as **42** stabilize α-helical secondary structures with anionic residues (aspartates or glutamates) oriented toward the same side of the helical backbone [$i(i + 3)$ or $i(i + 4)$]. Expectedly, **42** should bind to protein surfaces displaying such structural features. One example is the tetramerisation domain of p53, a tumour suppressor protein which is a key therapeutic target for cancer treatment. Each subunit of this domain contains an anionic sequence formed by residues Glu336, Glu339, Glu343, Glu346, Glu349, and Asp352, and tetraguanidinium **42** binds to two overlapping domains of this sequence with high affinity, as demonstrated by chemical shift perturbation (CSP) and saturation transfer difference (STD) NMR techniques, thus contributing to the stabilization of the protein.[49c]

7. Outlook

Guanidinium is a versatile functional group with unique properties. It remains protonated over a wide pH range and forms well structured hydrogen bonded ion-pairs with oxoanion substrates, in some cases even in rather polar solvents or water. Guanidines have been often used to bind to DNA or RNA as well as enzyme models for phosphodiesterases to

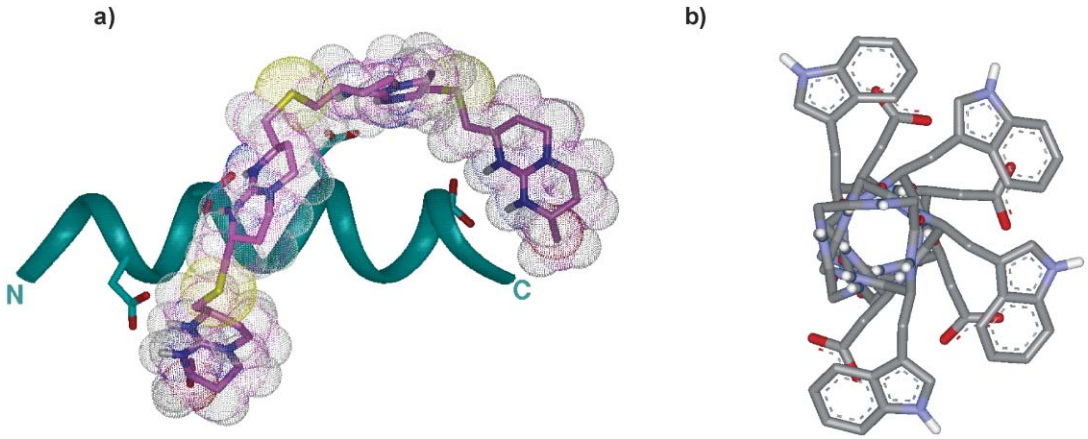


48

49 Ac-Ala-Ala-Gln-**Asp**-Ala-Ala-**Asp**-Ala-Ala-Ala-Ala-Ala-Gln-Ala-Ala-Tyr-NH_2

50 Ac-Ala-Ala-Gln-**Asp**-Ala-Ala-Ala-**Asp**-Ala-Ala-Ala-Ala-Gln-Ala-Ala-Tyr-NH_2

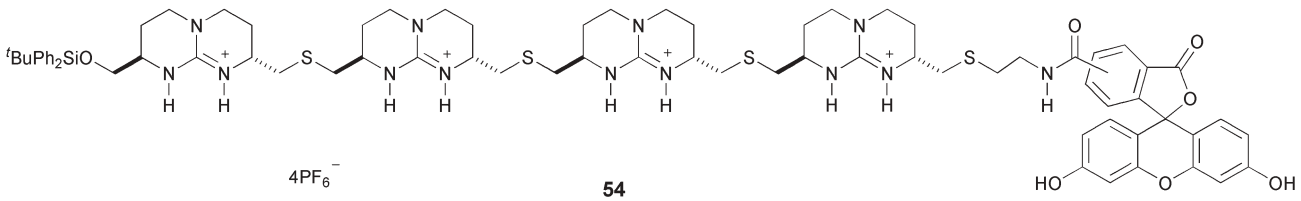
51 Ac-Ala-Ala-Gln-**Asp**-Ala-Ala-Ala-Ala-Ala-Ala-Ala-Ala-Gln-Ala-**Asp**-Tyr-NH_2



52 Ac-Ala-Ala-Ala-**Asp**-Gln-Leu-**Asp**-Ala-Leu-**Asp**-Ala-Gln-**Asp**-Ala-Ala-Tyr-NH_2

53 Ac-Ala-Ala-Ala-Trp-Gln-Leu-Trp-**Asp**-Leu-Trp-**Asp**-Ala-Trp-**Asp**-Ala-Gln-**Asp**-Ala-Ala-Ala-NH_2

Fig. 11 a) Optimised model of an α-helical peptide backbone with four aspartates at $i(i + 3)$ and tetraguanidinium **42**. b) Orientation of tryptophan and aspartate residues in peptide **53**.



54

cleave RNA.[50] Also, guanidines are among the most promising organocatalysts for reactions that proceed through anionic transition states.[51] When inserted into a bicyclic framework, the guanidinium function can be efficiently used for extractions or membrane transport. Over the last decade, arginine-rich peptides able to cross the cell membranes and to transport non-permeant molecules inside the cytoplasm have been identified and thoroughly studied as cell penetrating agents. Synthetic, non-peptidic oligoguanidines such as **54** also efficiently internalize into human cancer cells. This opens new perspectives in drug delivery.[52]

References

1 For an excellent recent review on non-covalent binding interactions involving guanidinium and phosph[on]ate/sulf[on]ate groups see: K. A. Schug and W. Lindner, *Chem. Rev.*, 2005, **105**, 67–113 and references therein.
2 For a recent review on guanidinium-containing receptors for anionic species see: M. D. Best, S. L. Tobey and E. V. Anslyn, *Coord. Chem. Rev.*, 2003, **240**, 3–15.
3 The pK_a value of the guanidinium moiety may vary depending on adjacent group effects; in general it will be lowered by acyl ⩾ phenyl ⩾ alkyl substitution (see ref. 1).
4 M. F. Jarrold, *Annu. Rev. Phys. Chem.*, 2000, **51**, 179–207.
5 C. Soto, G. P. Saborio and L. Anderes, *Trends Neurosci.*, 2002, **25**, 390–394.
6 R. Luo, L. David, H. Hung, J. Devaney and M. K. Gilson, *J. Phys. Chem. B*, 1999, **103**, 727–736.
7 (*a*) D. H. Kim and J. Park, *Bioorg. Med. Chem. Lett.*, 1996, **6**, 2967–2970; (*b*) P. F. Wang, M. J. McLeish, M. M. Kneen, G. Lee and G. L. Kenyon, *Biochemistry*, 2001, **40**, 11698–11705; (*c*) C. G. Mowat, R. Moysey, C. S. Miles, D. Leys, M. K. Doherty, P. Taylor, M. D. Walkinshaw, G. A. Reid and S. K. Chapman, *Biochemistry*, 2001, **40**, 12292–12298; (*d*) J. K. Bell, H. P. Yennawar, S. K. Wright, J. R. Thompson, R. E. Viola and L. J. Banaszak, *J. Biol. Chem.*, 2001, **276**, 31156–31162.
8 A. S. Woods and S. Ferre, *J. Proteome Res.*, 2005, **4**, 1397–1402.
9 E. Ruoslahti, *Annu. Rev. Cell Dev. Biol.*, 1996, **12**, 697–715.
10 M. J. Humphries, *Biochem. Soc. Trans.*, 2000, **4**, 311–339.
11 B. Christensen, M. S. Nielsen, K. F. Haselmann, T. E. Petersen and S. Sørensen, *Biochem. J.*, 2005, **390**, 285–292 and references therein.
12 M. García-Pérez, M. Pinto and J. A. Subirana, *Biopolymers*, 2003, **69**, 432–439.
13 Y. Zhang and D. Reinberg, *Genes Dev.*, 2001, **15**, 2343–2360.
14 B. Dietrich, T. M. Fyles, J.-M. Lehn, L. G. Pease and D. L. Fyles, *J. Chem. Soc., Chem. Commun.*, 1978, 934–936.
15 The secondary interactions between the H-donor and the neighbouring acceptor atoms are attractive in a DD-AA scaffold, but repulsive in a DA-AD one: W. L. Jorgensen and J. Pranata, *J. Am. Chem. Soc.*, 1990, **112**, 2008–2010.
16 (*a*) A. M. Echavarren, A. Galán, J. de Mendoza, A. Salmerón and J.-M. Lehn, *Helv. Chim. Acta*, 1988, **71**, 685–692; (*b*) H. Kurzmeier and F. P. Schmidtchen, *J. Org. Chem.*, 1990, **55**, 3749–3755.
17 G. Müller, J. Riede and F. P. Schmidtchen, *Angew. Chem., Int. Ed. Engl.*, 1988, **27**, 1516–1518.
18 A. Echavarren, A. Galán, J.-M. Lehn and J. de Mendoza, *J. Am. Chem. Soc.*, 1989, **111**, 4994–4995.
19 R. P. Dixon, S. J. Geib and A. D. Hamilton, *J. Am. Chem. Soc.*, 1992, **114**, 365–366.
20 M. Berger and F. P. Schmidtchen, *J. Am. Chem. Soc.*, 1999, **121**, 9986–9993.
21 B. R. Linton, M. S. Goodman, E. Fan, S. A. van Arman and A. D. Hamilton, *J. Org. Chem.*, 2001, **66**, 7313–7319.
22 A. Metzger, V. M. Lynch and E. V. Anslyn, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 862–864.
23 C. Schmuck, *Chem.–Eur. J.*, 2000, **6**, 709–718.

24 C. Schmuck and L. Geiger, *J. Am. Chem. Soc.*, 2004, **126**, 8898–8899.
25 W. H. Pirkle and T. C. Pochapsky, *Chem. Ber.*, 1989, **89**, 347–362.
26 (*a*) A. Galán, D. Andreu, A. M. Echavarren, P. Prados and J. de Mendoza, *J. Am. Chem. Soc.*, 1992, **114**, 1511–1512; (*b*) J. de Mendoza and F. Gago, *Computational Approaches in Supramolecular Chemistry, NATO ASI Ser., Series C*, 1994, **426**, 79–99.
27 P. Breccia, M. Van Gool, R. Pérez-Fernández, S. Martin-Santamaria, F. Gago, P. Prados and J. de Mendoza, *J. Am. Chem. Soc.*, 2003, **125**, 8270–8284.
28 L. J. Lawless, A. G. Blackburn, A. J. Ayling, M. N. Perez-Payan and A. P. Davis, *J. Chem. Soc., Perkin Trans. 1*, 2001, **11**, 1329–1341.
29 B. Baragan, A. G. Blackburn, P. Breccia, A. P. Davis, J. de Mendoza, J. M. Padron-Carrillo, P. Prados, J. Riedner and J. G. de Vries, *Chem.–Eur. J.*, 2002, **8**, 2931–2936.
30 (*a*) C. Schmuck and P. Wich, *Angew. Chem., Int. Ed.*, 2006, **45**, 4277–4281; (*b*) C. Schmuck and M. Heil, *Chem.–Eur. J.*, 2006, **12**, 1339–1348.
31 (*a*) B. L. Jacobson and F. A. Quiocho, *J. Mol. Biol.*, 1988, **204**, 783–787; (*b*) H. Luecke and F. A. Quiocho, *Nature*, 1990, **347**, 402–406.
32 S. L. Tobey and E. V. Anslyn, *J. Am. Chem. Soc.*, 2003, **125**, 14807–14815.
33 P. D. Beer, M. G. B. Drew and D. K. Smith, *J. Organomet. Chem.*, 1997, **543**, 259–261.
34 S. Nishizawa, Y. Kato and N. Teramae, *J. Am. Chem. Soc.*, 1999, **121**, 9463–9464.
35 K. Niikura, A. Metzger and E. V. Anslyn, *J. Am. Chem. Soc.*, 1998, **120**, 8533–8534.
36 F. P. Schmidtchen, *Tetrahedron Lett.*, 1989, **30**, 4493–4496.
37 V. D. Jadhav and F. P. Schmidtchen, *Org. Lett.*, 2005, **7**, 3311–3314.
38 V. Alcázar, M. Segura, P. Prados and J. de Mendoza, *Tetrahedron Lett.*, 1998, **39**, 1033–1036.
39 J. Sánchez-Quesada, C. Seel, P. Prados and J. de Mendoza, *J. Am. Chem. Soc.*, 1996, **118**, 277–278.
40 I. Rozas and P. K. Kruger, *J. Chem. Theory Comput.*, 2005, **1**, 1055–1062.
41 M. C. Grossel, D. A. S. Merckel and M. G. Hutchings, *CrystEngComm*, 2003, **5**, 77–81.
42 M. Berger and F. P. Schmidtchen, *Angew. Chem., Int. Ed.*, 1998, **37**, 2694–2696.
43 A. P. Bisson, V. M. Lynch, M.-K. C. Monahan and E. V. Anslyn, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 2340–2342.
44 (*a*) B. P. Hay, M. Gutowski, D. A. Dixon, J. Garza, R. Vargas and B. A. Moyer, *J. Am. Chem. Soc.*, 2004, **126**, 7925–7934; (*b*) B. P. Hay, T. K. Firman and B. A. Moyer, *J. Am. Chem. Soc.*, 2005, **127**, 1810–1819.
45 (*a*) A. Gleich, F. P. Schmidtchen, P. Mikulcik and G. Müller, *J. Chem. Soc., Chem. Commun.*, 1990, 55–57; (*b*) L.-P. Lu, M.-L. Zhu and P. Yang, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, 2004, **C60**, m18–m20.
46 P. Blondeau and J. de Mendoza, submitted for publication.
47 For a leading reference on disruption of protein–protein interactions, see: H. Yin and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2005, **44**, 4130–4163.
48 J. S. Albert, M. S. Goodman and A. D. Hamilton, *J. Am. Chem. Soc.*, 1995, **117**, 1143–1144.
49 (*a*) M. W. Peczuch, A. D. Hamilton, J. Sánchez-Quesada, J. de Mendoza, T. Haack and E. Giralt, *J. Am. Chem. Soc.*, 1997, **119**, 9327–9328; (*b*) B. P. Orner, X. Salvatella, J. Sánchez-Quesada, J. de Mendoza, E. Giralt and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2002, **41**, 117–119; (*c*) X. Salvatella, M. Martinell, M. Gairí, M. G. Mateu, M. Feliz, A. D. Hamilton, J. de Mendoza and E. Giralt, *Angew. Chem., Int. Ed.*, 2004, **43**, 196–198.
50 D. M. Perreault, L. A. Cabell and E. V. Anslyn, *Bioorg. Med. Chem.*, 1997, **5**, 1209–1220 and references therein.
51 T. Ishikawa and T. Isobe, *Chem.–Eur. J.*, 2002, **8**, 552–557.
52 J. Fernández-Carneado, M. Van Gool, V. Martos, S. Castel, P. Prados, J. de Mendoza and E. Giralt, *J. Am. Chem. Soc.*, 2005, **127**, 869–874.

