Noncovalent Binding between Guanidinium and Anionic Groups: Focus on Biological- and Synthetic-Based Arginine/Guanidinium Interactions with Phosph[on]ate and Sulf[on]ate Residues

Kevin A. Schug* and Wolfgang Lindner

University of Vienna, Department of Analytical Chemistry, Währingerstrasse 38, A-1090 Vienna, Austria

Received June 8, 2004

Contents

1. Introduction

This is a review of work pertaining to noncovalent binding interactions involving primarily the cationic guanidinium group and the anionic sulf[on]ate and phosph[on]ate groups.[1] To a lesser extent, and where appropriate, interactions by guanidinium with carboxylate groups will also be discussed. The volume of research aimed toward such phenomena and being in progress is evidence of the vast importance that interactions between these complementary groups comprise. A review such as this cannot capture the full extent of available information in every facet of science that utilizes this knowledge. Instead, here we contribute a compilation of those systems where guanidines and phosph[on]ates or sulf[on]ates (and sometimes, for comparison, amines and carboxylates) are explicitly involved. We first discuss the basics behind the systems including structure, environment, mode of interaction, and general techniques for analysis. We next cover specific examples of research and their impact on the development of the field, focusing particularly on advancements in the past few years. The review of existing material is presented in a reciprocal manner. Facts and finds uncovered by studying from each point of view (e.g., sulf[on]ates, phosph[on]ates, and carboxylates interacting with cationic groups or guanidines, amidines, and amines interacting with anionic groups) will be used to elaborate on the collective (acid−base, hydrogen donor−acceptor, negative−positive, etc.) interaction scheme. This review is limited in scope to report mainly on recent discoveries in synthetic and biological systems with a focus on those studies that explicitly address the interactions of interest. Biological systems of interest will be limited to protein and peptide interactions (largely excluding those interactions involved with binding to DNA molecules), whereas synthetic systems will be limited to smaller host−guest receptor systems (mainly "cleft" or "tweezer" systems), excluding larger supramolecular cagelike structures, except where needed for clarification of principles. In the end, we summarize and give some thought on the future of development in this field with the hope of portraying the importance of this topic to readers and encouraging further research.

2. General Features

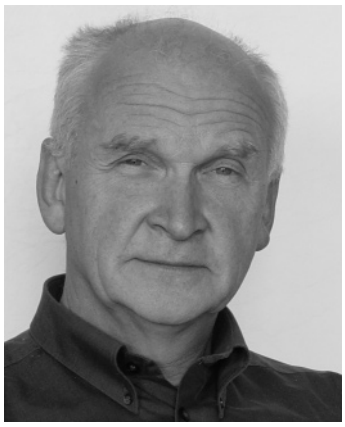
2.1. Structure Implications and Binding

At the heart of the formation of noncovalent complexes or associations between molecules is the complementary geometry and functionality associated with each potentially interacting component. Interactions are optimized for given groups arrayed in a specific geometry. The importance of these arrangements is apparent both in natural processes and for synthetic chemists who seek to mimic these processes. For every interaction that is elucidated and sought to be understood, many "man-made" receptor systems are developed to take advantage of the new information. Focus on these systems is taken from both the anionic-based and cationic-based points

* To whom correspondence should be addressed. Tel: (+43)-1-4277-52300. Fax: (+43)-1-4277-9523. E-mail addresses: schugka@hotmail.com; wolfgang.lindner@univie.ac.at.

10.1021/cr040603j CCC: $53.50
Published on Web 12/09/2004

Kevin A. Schug was born and received his elemental education in Blacksburg, VA. He received his B.S. degree in Chemistry from the College of William and Mary in Williamsburg, VA, in 1998. From there, Kevin returned to Blacksburg to study separation science under Prof. Harold M. McNair at Virginia Tech. In 2002, following dissertation research centered upon the study of formation and detection of noncovalent complexes by electrospray ionization mass spectrometry, he graduated with a Ph.D. in Chemistry. Kevin is now in his second year as a postdoctoral fellow with Prof. Wolfgang Lindner at the University of Vienna (Austria). His work focuses on the study of noncovalent complexation and molecular recognition events through mass spectrometric methodologies. Particular attention is paid to those interactions that are forged by the guanidinium functional unit.



Since 1996, Wolfgang F. Lindner has held the Chair of Analytical Chemistry at the University of Vienna (Austria). Early during his studies, his research interest was influenced by the pharmaceutical (life) sciences and by the separation sciences related to HPLC, GC, CE/CEC, and MS. In this context, particular interest developed in noncovalent interactions and molecular recognition with particular focus on enantiomer discrimination. The development of novel synthetic selectors (receptors), useful for enantioselective separation techniques, lies at the interface of organic, analytical, and biological chemistry, which characterizes best his scientific credo. He was trained in Organic Chemistry at the University of Graz (Austria) receiving his Ph.D. in 1972. He then moved to the Institute of Pharmaceutical Chemistry in Graz, specializing in pharmaceutical analysis. In 1978, he was awarded a Max Kade postdoctoral fellowship, which he held at Northeastern University in Boston, MA, under the supervision of Prof. Barry L. Karger. In 1986, he was visiting scientist at the FDA/NIH in Bethesda, MD, where he was exposed to biological chemistry. Currently, he directs a large and talented multinational research group with a diverse and interdisciplinary focus on the effective combination of analytical, synthetic, and biological chemical methods for molecular recognition and separation purposes.

of view. Each experiment is designed to probe an aspect of these systems with the spotlight placed on a specific functional moiety for the purpose of observing its interaction with other molecules. It is a noble effort for humans to try to mimic the natural systems

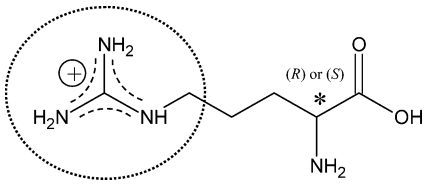


Figure 1. Structure of arginine. The guanidine moiety, highlighted and shown in the singly protonated form, is responsible for the majority of arginine noncovalent interactions.

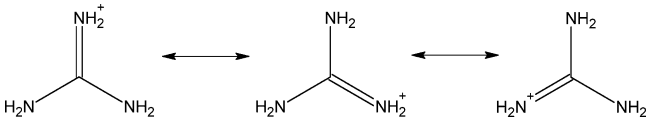


Figure 2. Triple resonance stabilization of a guanidinium molecule.

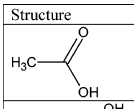
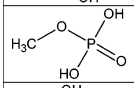
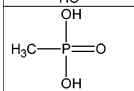
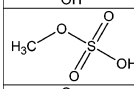
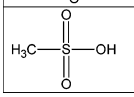
that surround them, however, the shear complexity involved in nature will undoubtedly leave plenty of stones unturned for future generations to ponder.

In the case of systems involving basic amino acids, specifically arginine, characterized by its very basic guanidinium group, much effort has been spent in recent years to unravel the mysteries of its multitudinous interactions in biologically relevant systems. Arginine is a special amino acid that holds the distinction of being the most basic of all natural amino acids. In fact, arginine has the highest proton affinity (PA) of the natural amino acids by more than 14 kcal/mol, lysine being its nearest neighbor (histidine is third on the relative basicity scale).[2] Arginine owes its basicity to its structure, which is based on a proton-loving guanidine moiety. In its free form (shown in Figure 1) or built into peptide, protein, or receptor systems, it is capable of forming both electrostatic and directed hydrogen bond interactions with polar and anionic molecules, mainly through interaction by its planar, forklike guanidinium functionality.

The guanidinium functionality utilizes "Y-delocalization" to establish its great stability as an ion in an aqueous environment.[3,4] Guanidinium, or protonated guanidine, has six potential hydrogen bond donors available, making it highly soluble in aqueous systems. Resonance stabilization of the molecule spreads the positive charge evenly about the three nitrogen atoms as shown in Figure 2.[5] By itself, guanidine has a pK_a of 13.5 in water. In arginine, the pK_a value is attenuated to 12.5. Although it will vary depending on neighboring group effects (as well as due to spatial effects, as in a folded protein), in general, monosubstitutions on the guanidine functionality lower the pK_a by acyl > phenyl > alkyl substitution. Also, substitutions on more than one of the amino groups (e.g., by posttranslational modification) in the guanidine can be expected to change the chargeable properties of the resulting structure. Geometrically, as determined from crystal structures, the C−N single bond length in an alkylguanidine is shorter than normal C−N bonds. In the guanidine, all three bond lengths and bond angles are nearly equal with an average of 1.33 Å and 120°, respectively.

Particularly important interactions formed by the guanidinium moiety of arginine, and consequently

Table 1. Reported and Calculated Values for Acid Dissociation Constants (pK_a) of Methyl Acids with Various Acid Functional Units[a]

Structure	Name	pK_{a1}(calc.)	pK_{a2}(calc.)	pK_{a1}(rep.)	pK_{a2}(rep.)
	acetic acid	4.79	--	4.74	--
	methyl dihydrogen phosphate	1.81	6.22	1.52	6.31
	methylphosphonic acid	2.41	8.07	2.38	7.74
	methyl sulfate	-3.35	--	N/A	--
	methanesulfonic acid	-1.89	--	1.92	--

[a] Reported references are taken from literature sources listed in the ACD/pK_a calculator for solution equilibria at or close to 25 °C. A reported pK_a value for methyl sulfate was not available (N/A).

the subject of this review, are those with anions such as phosph[on]ates, sulf[on]ates, and carboxylates. Most commonly reported are complexes formed between guanidinium and carboxylates; however, owing to their increased acidities relative to the carboxylate moiety, here, phosph[on]ate and sulf[on]ate moieties are given special attention. These functional units have found a niche in molecular recognition schemes, both for purposes of drug development and for biological mimetics. Table 1 lists a comparison of the calculated and reported acidity of some methylated acids based on the differing makeup of their oxoanion structure. Acid dissociation constants (pK_a's) reported here were taken from various sources as well as calculated using ACD/Labs SpecManager software.

From the values in Table 1, it is apparent that the acidic functional unit has a large impact on the ionizability of the molecule. Substitution patterns and neighboring groups will also have marked effects on the pK_a values. The sulf[on]ate group provides a functional unit that will be ionized in aqueous media regardless of the pH of the surrounding medium. Such a group will be expected to interact strongly with cationic and electropositive sites on target molecules. The phosph[on]ate moiety is also substantially more ionizable relative to the carboxylate and adds the ability to form a dianionic species in weakly alkaline media. This functional unit can also be expected to have a high affinity for electropositive groups in neighboring molecules and in some cases might form an anionic salt bridge for complexation between two cationic species. In addition, phosph[on]ates possess the capability to be designed as stereogenic centers for specialized interaction when appropriately esterified. Dihydrogen phosphate and sulfate moiety-containing molecules, such as methyl dihydrogen phosphate and methyl sulfate shown in Table 1 as references, will be included in this review only where applicable in the discussion of biological systems, specifically when information incorporating phosphonates and sulfonates is not explicitly available.

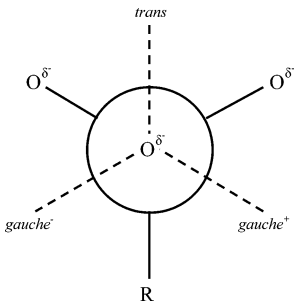


Figure 3. The stereochemistry of Lewis acid interactions with the pyramidal phosphonyl ($\delta = 2e/3$) and sulfonyl ($\delta = e/3$) groups. The trans/gauche designation is preferred to the syn/anti nomenclature used for the planar phosphinyl group.[6]

The preferential stereochemistry of the interaction of the pyramidal anions, phosphonyl dianion ($R-PO_3^{2-}$) and sulfonyl monoanion ($R-SO_3^-$), with Lewis acids can be described by trans/gauche conformational terminology. Christianson and co-workers[6] performed a survey of the Cambridge Structural Database (CSD) for these interacting schemes and found that phosphonyls and hydrogen bond donors display a preference for gauche orientation. Sulfonyls and hydrogen bond donors also show a preference for gauche orientation, but they can be observed in an *eclipsed* orientation in clusters as well. Figure 3 shows the trans/gauche orientation of the Lewis acid interactions with the pyramidal phosphonyl and sulfonyl groups. Studies show that both electrostatic and molecular orbital effects are important in the interaction of these anionic species with Lewis acids. The authors suspect that the broad preference for gauche-oriented interactions allows for a greater number of catalytic conformations in proton transfer schemes. In comparing the binding differences between the phosphonyl and sulfonyl groups, interacting geometries are most likely subject to a difference in preference for hydrogen bonding architecture as well as differences in the Coulombic charges on each pyramidal anion. Clearly, the difference in magnitude, size, and shape of the anion group leads to interaction differences with a given Lewis acid. The use of structures from the CSD allows for statistical characterization of the available intermolecular and intramolecular interactions included in the database and has been used by many researchers for this purpose.

2.2. Systems of Interaction

The range of systems observed and used for studying molecular interaction is widely varied and can be subdivided in multitudinous fashions (by structure, interaction, environment, etc.). In this review, focus will be directed to two basic classifications in hopes of simplifying the reported results. One class will be those interactions that occur in a biological environment (predominantly aqueous physiological conditions), where the investigating scientist seeks to probe systems with measurements of small changes in the local environment and mimetic-based approaches. The second will be those systems designed as molecular recognition schemes and generally

involving synthetic receptors and ligands and specified host−guest schemes. As per different synthetic and analytical schemes, the environment of such interactions can vary considerably. Again, in a reciprocal manner, each class (biological and synthetic) can be addressed from the point of view of those studies that have focused on guanidinium interactions or those that have focused on the interaction of anionic components.

Those systems reviewed and classified as naturally occurring will include protein−protein, protein−peptide, and peptide/protein−ligand (conformational and stereochemical agents and mediators, such as sugar- and pharmaceutical-based molecules) interactions. In these experiments, interactions between arginine-containing and phosph[on]ate- or sulf[on]ate-containing higher order structures will be highlighted. The review will also cover, but will not be limited to, RNA interacting systems. The review will not comprehensively include DNA interacting systems due to limitations in space.

Biological systems in general are complex and difficult to study. Proteins in particular are exceedingly flexible, are very susceptible to influence from the surrounding environment, and interact in a highly specific manner with a variety of complementary molecule types. Intramolecular and intermolecular interactions between the protein and solvent define the native conformations of the proteins.[7] The values associated with these interactions are very large; however, they cancel out in the presence of physiological conditions to amount to overall differences of around tens of kilojoules per mole, adding to the difficulty of elucidating effects of the environment. The specificity of interactions provided by biological systems is based on the flexibility of the molecules allowing for a large degree of "encapsulation" of the targeted substrate.[8] This allows for a large number of interaction points. Binding structures of the host can be connected in a linear, branched, or unbranched fashion and fold about the guest.[9] Rather than cite a simplified lock and key metaphor, this type of specificity can be likened to a large number of teeth in a key (multiple points of contact on the target) and their simultaneous or consecutive interaction with the interior of the lock (the protein) to create an "induced fit" recognition and achieve the proper function. Compared to synthetic hosts, this system has the advantage of the biomolecular host being able to fold about the guest and create a stronger overall interaction through multiple contact points.[5] A difficult question to answer (and consequently, strong reasoning for a reciprocal approach to these systems) is the following: Is the function derived from a specific interaction a result of a host reading the sequence of a guest or vice versa? In this respect, which molecule is really the host or the guest is defined solely by the point-of-view.

In the synthetic realm (and consequently, also in biochemistry and analytical chemistry), the field of molecular recognition has become extremely vast in recent years and is a testament to the originality and creativity of scientists involved in this setting. Much of the early success in this field can be attributed to the work done by Lehn and co-workers in the development of artificial arginine receptors.[10] For the design of synthetic receptors, it is necessary for those creating them to pay attention to a few key points for the designed molecule to be useful and applicable. Small molecules designed to bind selectively to arginine residues could become useful as, for example, antiviral drugs or molecular probes for arginine-rich proteins.[11] To be used in this way, the probe must be reasonably soluble in water and bind with high affinity for an alkylguanidinium moiety. In addition, it must have a significantly lower affinity for other basic residues, such as histidine and lysine—a formidable challenge in many design schemes. Selective receptors for specific peptide sequences in general would have many potential applications, for example, separation of protein mixtures or development of new therapeutics and biosensors.[12] Flexibility (or rigidity), cavity (or spacer) size, and the local arrangement of interacting functional groups in the receptor are the greatest concern to achieve the desired capabilities of a receptor or host molecule.[13] It is obvious that a closer match to the guest topography by the receptor will increase the likelihood of selectivity of the host for that particular guest or class of guests. In a similar manner, complementary character must be exhibited in the design of receptors that are selective for anionic moieties, such as phosph[on]ate and sulf[on]ate. Such a complementarity has been exploited to a larger degree for the selective recognition of carboxylate groups by guanidinium-functionalized receptors. As is shown in the literature presented in this review, strong complementarity is also given by the forked, electropositive units such as guanidines and amidines to phosph[on]ate and sulf[on]ate moieties. The science of amidine-functionalized receptor units, although effective mimics of guanidinium and possessing similar directed binding nature, has been, in the interest of space, largely omitted from this review.

Types of molecular recognition system designs are many. Synthesized receptors are arranged in ways so as to maximize interactions for a particular species or group of molecules to induce *selectivity*. Functional group interactions derived from monitoring biological schemes are implemented in synthetic schemes to invoke molecular recognition abilities. In this review, we will largely address "cleft" or molecular "tweezer" receptor design due to its relative simplicity and relevant current interest. These receptor types are characterized by a headgroup and tweezer-like arms, which provide for stabilization of the complex by secondary interaction groups. Figure 4 shows a general concept of design for such receptors. It is important to note that the dominant interacting groups can be placed at the headgroup or at the tethered ends or both to achieve complementary binding for a particular guest. Also worth note is the ability to design the tweezer arms as structurally rigid or flexible depending on the desired degree of specificity. Some tweezer molecules, despite their inherent flexibility, have shown to be highly selective for specific peptide sequences in both nonpolar and

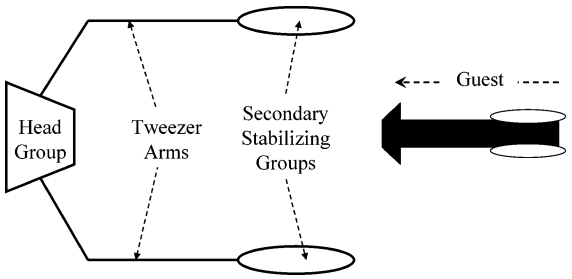


Figure 4. Synthetic design of tweezer receptors indicating the presence of a headgroup and stabilizing groups tethered by tweezer arms.

aqueous media.[12] Information from rigid, cagelike macrocycle receptors will also be addressed where appropriate.

2.3. Interaction Types

At the heart of biological interactions, such as protein and peptide complex formation, and molecular recognition systems, such as receptor−ligand design, is the concept of a noncovalent interaction site. Noncovalent interactions can be simply described as interactions between two species that do not involve covalent bonds or the specific sharing of valence electrons. Associations of this nature cover a variable range of energies and depend explicitly on the structure and functionality of the species involved in the formed complex as well as the local environment of the complex. Table 2 lists the interaction potentials for the major contributions to noncovalent interactions.[14] The equations listed for interaction types represent the free energy associated between two species. All of the interactions are attractive in nature, except for Coulomb−Coulomb and dipole−dipole, which may be attractive or repulsive, depending on the signs of the charges or the relative angular orientation of the dipoles, respectively. Permanent and induced dipoles accommodate the forces acting upon them and do so to cause an attractive interaction. Species with no charge or dipole associated with them can attract each other due to London dispersion forces invoked by the polarizability of the group.

All of the forces are distance-dependent as noted by the presence of the $1/r^n$ term. Also notable is the presence of the dielectric constant ϵ of the medium in the denominator of each equation. The presence of the $1/\epsilon$ term implies that a solvent with a high dielectric constant, such as water, will reduce interaction energies. This term also represents the major difference between forces present in the solution versus in the gas phase, where the dielectric of the solvent is not present. Energies of interactions based on charges, dipoles, or polarizability are expected to increase as the effect of solution (solvation) is removed. An exception to this is a hydrophobic interaction between two nonpolar units, where the presence of water increases the attraction between these moieties. Generally, hydrophobic, hydrophilic, and hydrogen-bonding interactions are not easily described, and a rigorous discussion is beyond the scope of this paper.

Table 2. Noncovalent Interaction Types and Equations Describing Interaction Potentials Involved[14]

interaction type	formula[a]
charge−charge (Coulomb energy)	$\frac{Q_1Q_2}{4\pi\epsilon\epsilon_0 r}$
charge-dipole (fixed dipole)	$-\frac{Qu\cos\theta}{4\pi\epsilon\epsilon_0 r^2}$
charge-dipole (freely rotating dipole)	$-\frac{Q^2u^2}{6(4\pi\epsilon\epsilon_0)^2kTr^4}$
dipole−dipole (fixed dipole)	$-\frac{u_1u_2}{4\pi\epsilon\epsilon_0 r^3}(2\cos\theta_1\cos\theta_2-\sin\theta_1\cos\theta\sin\theta_2)$
dipole−dipole (freely rotating dipole) (Keesom energy)	$-\frac{u_1^2u_2^2}{3(4\pi\epsilon\epsilon_0)kTr^6}$
charge−nonpolar	$-\frac{Q^2\alpha}{2(4\pi\epsilon\epsilon_0)^2r^4}$
dipole−nonpolar (fixed dipole)	$-\frac{u^2\alpha(1+3\cos^2\theta)}{2(4\pi\epsilon\epsilon_0)^2r^6}$
dipole−nonpolar (freely rotating dipole) (Debye energy)	$-\frac{u^2\alpha}{(4\pi\epsilon\epsilon_0)^2r^6}$
nonpolar−nonpolar (London dispersion energy)	$-\frac{3}{2}\frac{\alpha_{01}\alpha_{02}}{(4\pi\epsilon\epsilon_0)^2r^6}\frac{I_1I_2}{I_1+I_2}$
hydrogen bond	special, directed interaction
hydrophilic	special
hydrophobic	special

[a] Q = charge; u = dipole moment; r = distance; α = polarizability; ϵ = dielectric constant; T = temperature (kelvin); ϵ_0 = dielectric permittivity of a vacuum; k = Boltzmann's constant; I = first ionization potential; θ = angle between dipole and vector connecting the interacting particles; φ = polar angle of second dipole relative to the first.

Hydrogen-bonding interactions are primarily electrostatic and directional in nature, meaning the degree of interaction is greatly affected by alignment of potential hydrogen bonding groups and their corresponding geometries. The phenomenon of hydrogen bonding is used in many systems as the main explanation for complex formation. Hydrogen bonds have been traditionally defined by the formalism D−H···A, where donor D and acceptor A are both electronegative atoms, such as N, O, or F.[15] Although, even weaker donor atoms, such as carbon, are currently included in discussions and contribute to the ubiquitous use of hydrogen bonding as a descriptor for binding in systems. Graph-set formalism for classifying hydrogen bonds has become a useful tool for differentiating patterns of hydrogen bonding, especially in crystallographic measurements.[16−18] Though the formation of a hydrogen bond is largely electrostatic in nature, the system need not involve ionization, rather hydrogen-donating and -accepting character of interacting groups are generally discussed. Electrostatic and exchange-repulsion terms are considered to account for more than 80% of the hydrogen-bond interaction, with the other 20% made up of dispersive and polarization energies.[19] The role of solvent in these interactions must be carefully considered as well, since solvent molecules may contain hydrogen-donor and -acceptor character as well. The contribution of the hydrogen bond can be reduced dramatically in solution due to entropic effects, solvation, or dielectric conditions.[20] The importance of the directionality of hydrogen bonding to interactions involving guanidinium has been mentioned previously. The rigid planarity of the functional unit has been shown to induce the axial arrangement of interacting groups, such as phosphate, into this plane to increase stability.[21]

Hydrophobicity, implying the dislike of a molecular unit for water, can be further generalized under the term solvophobicity. Every molecular unit contains some degree of solvophobicity, which is obviously dependent on the particular solvent present. Hydrophobicity is the most commonly cited solvophobic form due to the ubiquitous nature of water. However, due to the diversity of the systems covered in this report and the solution systems employed to study them, it is important to be considerate of the nomenclature in describing particular intermolecular interactions.

Hydrophobic and hydrogen-bonding forces, though sometimes difficult to describe rigorously, are extremely important in the structural arrangement of many molecules and are particularly relevant to complexes formed between biomolecules, as well as synthetic molecular recognition systems in aqueous (polar) media. Hydrophobic forces naturally become stronger in the presence of water, whereas hydrogen-bonding interactions are hampered in aqueous solution by the presence of a large dielectric constant. The induced fit model for biological interactions makes use of both. The synthetic chemist often relies more on selected hydrogen-bonding and electrostatic interactions to associate with high affinity and more easily predictable binding arrangements.

Systems exhibiting cation−π and π−π interactions must also be considered as an important part of the formation of noncovalent complexes, especially in biological systems. Extensive research has been performed by Dougherty and co-workers establishing the methodology and specifics behind cation−π interactions.[22−24] Generally speaking, these interactions are made up from the attractive nature of a cationic species and an adjacent π-bonded system, which might be an aromatic ring moiety, an allylic group, or any other bonding that exhibits π character. The mode of interaction is suspected to be a combination of electrostatic attraction with the polarizability of the π group in a modified ion-induced dipole configuration. Cation−π interactions have been established as one of the fundamental noncovalent forces used by nature in determining protein structures and protein−ligand interactions. Gallivan and Dougherty performed a study modeling interactions found in the Protein Data Bank with ab initio methods.[23] In the structures surveyed, over 70% of the arginine side chains were found in close proximity to an aromatic side chain. Interactions first coined "amino-aromatic" but since generalized as cation−π can be found with the arginine perpendicular or parallel to the aromatic plane. According to calculations, however, parallel interactions provide a more favorable arrangement for binding.

Systems containing π−π interactions generally refer to stacking between aromatic moieties but in concept are not limited to such. In fact, the aromatic ring appears to be an especially important bridge between hydrophobicity and hydrophilicity in an aqueous environment, since it possesses the ability to interact with both cation systems, which are ionic and extremely hydrophilic, and alkyl chains, which comprise the definition of hydrophobicity. Arginine and the arginine residue are known in nature to be involved in stacking interactions with phenyl rings of aromatic amino acid side chains.[25] Both cation−π and π−π interactions can be expected to contribute to more complex interactive systems where an arginine residue or guanidinium unit is involved. As an explicit interaction mode between cationic and anionic residues, these arrangements would be more readily present in secondary interactions (meaning outside of the primary, and here, defining, interaction point) to stabilize the association and help provide the necessary specificity and selectivity.

The concept of a salt-bridge interaction is also commonly encountered in the literature when addressing the interaction between guanidinium and anions. A salt bridge is defined as a central-charge-mediated electrostatic interaction of the form (+)·(−)·(+) or (−)·(+)·(−).[7] This interaction type can be described as a collection of Coulombic interactions but is also commonly referred to in context with hydrogen-bonding contributions. Salt bridges are often attributed to contributing to the helical stabilization (or sometimes, destabilization) of peptides and proteins. Gilson and co-workers have performed an in depth computational study of arginine−phosphate salt-bridge interactions.[26] Through comparison of butylammonium−phosphate^{2-},

Table 3. Binding Free Energies (kJ/mol) of Ion Pairs from Calculated[26] and Experimental[27] Determinations

ion pair	ΔG°_{calcd}	ΔG°_{expt}
guanidinium−phosphate	−4.10	−2.38
butylamine−phosphate	−4.01	−1.76
guanidinium−acetate	−3.26	2.51

guanidinium−phosphate^{2-}, and guanidinium−acetate and the incorporation of similar structures in a simulation of salt-bridge formation in peptides, it was shown that a significant stabilization to helical peptides was offered if arginine and phosphotyrosine were spaced four units away from each other. Table 3 details the comparison of calculated and experimental values[27] for these interactions. Moreover, the guanidinium−phosphonate interaction was found, through calculation and in agreement with experimental values,[28] to be capable of providing a maximal contribution of around −10 kJ/mol to binding between a host and guest. Still, the study stressed the difficulty with assessing the contribution to stability by salt-bridge interactions in proteins. The nature of these salt bridges is highly dependent on neighboring groups, solvent present, and temperature. In accordance with their designation, the thermodynamic stability of a salt bridge is also highly dependent on the ionic strength of the surrounding medium.

2.4. Solvent Systems

The role of a solvent in systems of interaction between cation-based guanidine groups and anion-based molecules, such as phosph[on]ates and sulf[on]ates, can hardly be overstated. The majority of chemical reactions occur in solution. The solvent has several distinct functions in a reacting or interacting system. First and foremost, solvation of analytes or reactants is necessary for dissolution to place molecules in a common medium. Once dissolved, the molecules can collide and interact under restraint of the diffusion of the species through the solvent. The solvent provides a means for moderating temperature of the collisions, either by adding energy to the colliding particles to aid reaction or to absorb the heat generated in an exothermic reaction.

Solvents are generally characterized by their polarity. Mention has already been made previously of the role of the dielectric of the medium in moderating solution-phase interactions. Polar solvents are defined as those with a large dipole moment and a high dielectric constant. Nonpolar solvents have a small dipole moment and low dielectric constant. Table 4 lists the dielectric constant and dipole moment of some common solvents for comparison.[29] There are three categories for classifying solvents by polarity: polar protic, dipolar aprotic, and nonpolar. Protic refers to a hydrogen attached to an electronegative atom, usually oxygen in an ROH configuration. Examples of polar protic solvents are water, methanol, and acetic acid. Aprotic solvents are those that lack an O−H bond and most of which contain a C−O (or C−N) bond. Examples are acetone, acetate, acetonitrile, and dimethyl sulfoxide (DMSO). Nonpolar solvents, such as hexane and benzene, have low dielectric constants and are immiscible with water.

Table 4. Some Dipole Moment and Dielectric Constant Values of Common Solvents[29]

name	dipole moment (D)	dielectric constant
water	1.85	80
methanol	1.70	33
ethanol	1.69	24.3
1-propanol	1.68	20.1
1-butanol	1.66	17.8
formic acid	1.41	58
acetic acid	1.74	6.15
formamide	3.73	109
acetone	2.88	20.7
methyl ethyl ketone	2.78	18.5
ethyl acetate	1.78	6.02
acetonitrile	3.92	36.6
N,*N*-dimethylformamide	3.82	38.3
dimethyl sulfoxide	3.96	47.2
hexane	0	2.02
benzene	0	2.28
diethyl ether	1.15	4.34
tetrahydrofuran	1.63	7.52
methylene chloride	1.60	9.08
carbon tetrachloride	0	2.24

A useful nomenclature to discuss interactions with solvents is to refer to solution media as competitive and noncompetitive solvent systems. In complexes formed between guanidine-based and anionic functional units, where electrostatics and hydrogen bonding are the prominent interaction modes, solvents with a high dielectric constant and that are able to hydrogen bond are classified as competitive solvents. Solvents such as water and dimethyl sulfoxide (DMSO) naturally fall into this category. Hydrophobic solvents such as hexane or benzene would be deemed noncompetitive. There of course exists some gray area when considering the multitude of different solvent systems (and mixtures thereof) that are commonly employed to carry out syntheses and measurements; however, this simple convention can be useful for assessing the overall thermodynamic effect of binding in the systems discussed herein. Often, small amounts of competitive solvents are introduced into host−guest schemes to test the strength and robustness of the interaction.

Since one of the main goals in this field is to develop recognition schemes that can be utilized in a biological- or pharmaceutical-based environment, a move toward the use of more competitive media in systems is apparent. For guanidine−anion systems, binding is often quoted as being hydrogen bond-driven, an enthalpically favorable event. This may very well be the case in noncompetitive solvents. However, once a competitive solvent is introduced, which has the ability to itself hydrogen bond, binding can often become enthalpically unfavorable. In such systems where association is still observed, binding is believed to occur largely due to solvent liberation, an entropically favorable event.[30] In addition, binding can also be the product of desolvation, particularly when hydrophobicity is a significant part of the global binding process. Overall, one must keep in mind that noncovalent binding is always a balance of two issues: favorable attraction between host and guest and the solvation properties of the two.[31]

In light of further discussions of a thermodynamic nature in the following sections, a short introduction

to enthalpy/entropy compensation may here be of use. Enthalpy/entropy compensation addresses the balance between heat transfer and change in order in a chemical reaction. This interplay is discussed in several reviews.[32−34] Briefly, this balance can be described concisely by the Gibbs−Helmholtz equation: $\Delta G = \Delta H - T\Delta S$, where for a given reaction ΔG is the Gibbs free energy, ΔH is enthalpy, T is temperature in kelvin, and ΔS is entropy. For a chemical reaction to proceed spontaneously, ΔG must be less than zero. Therefore, a large negative ΔG for a given complexation reaction indicates a high affinity for the interacting species and, subsequently, the formation of a stable complex. Contributions to this affinity are made through changes in enthalpy (negative ΔH being favorable to interaction) and changes in entropy (positive ΔS being favorable to interaction). Indicated in the equation above, temperature can also have a strong effect of the magnitude of ΔG. It is often difficult to assign a numerical value to each facet of an interaction; however, there are anticipated trends associated with many of such components. For example, in water,[35,36] the partial dehydration (desolvation) of NH_3^+ and SO_3^- gives a positive contribution to ΔH and to $T\Delta S$; loss of conformational degrees of freedom gives a negative (unfavorable) contribution to $T\Delta S$; ion pairing in water is entropically driven; long-range Coulombic bonds or hydrogen bonds are generally enthalpically driven; $\pi-\pi$ interactions give a negative (favorable) contribution to ΔH; hydrophobic interactions give a positive contribution to $T\Delta S$. Given the interplay of these contributions to the overall binding between two partners, it is easy to see that a concerted binding sequence utilizing a sum of interaction types can be a challenge to elucidate thermodynamically.

With respect to the guanidinium−anion interaction of interest, a relevant theoretical study reports density functional theory (DFT) and molecular dynamics (MD) investigations of arginine−phosphate binding interaction in water.[37] Methyl guanidinium and dimethyl phosphate are common prototypes used to investigate arginine−phosphate binding interactions. In this study, two structural arrangements (monodentate and bidentate binding) were considered. A bidentate structure is envisaged to be much stronger than the more commonly encountered monodentate binding structure, but both can be found in natural systems.[38,39] Correlating this hypothesis, the cluster of water molecules connecting the phosphate and the arginine moieties creates a stronger interaction during the dynamics for the bidentate complex. The study shows that water plays a significant role in the formation of noncovalent complexes between the two functionalities. The results of the simulations for the hydrated mono- and bidentate complexes are shown in Figure 5. The hydration is accompanied by significant polarization effects. As we will see with other examples, the incorporation of water actually "tunes" the interaction. A plot of electronic density difference showing redistribution of charge upon hydration illustrates the polarization of the arginine and the phosphate oxygens interacting with it. In other words, arginine hydrogens become much more

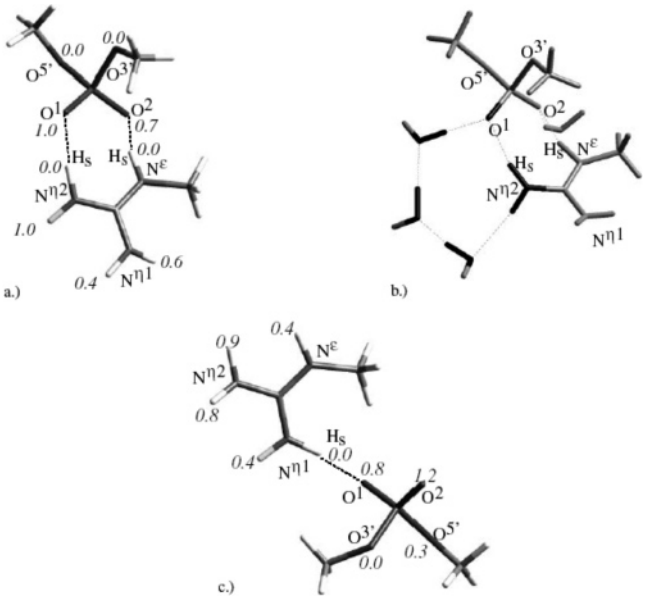


Figure 5. Mono- and bidentate binding between methylguanidinium and dimethyl phosphate is mediated by water molecules. Part a shows water coordination numbers of selected atoms in the bidentate binding arrangement with the water atoms omitted for clarity; part b shows the coordination of the water atoms in the bidentate binding; part c shows monodentate binding with water coordination numbers of selected atoms. Reprinted with permission from ref 37. Copyright 2001 Elsevier.

positively charged in the presence of water in the hydrated complex, causing a stronger binding interaction with the phosphate (i.e., an ion pair). Overall, consideration of solvent effects on these cation−anion interacting systems represents the "cutting edge" for current developmental experiments. To gain access to understanding biological effects and creating effective mimetic systems, focus must be placed on elucidating the molecular-level role of small polar solvent molecules in each respective situation.

Another interesting recent publication provides a different viewpoint from which judgment of the propensity for guanidinium to mediate interactions in a hydrated environment can be made. Recently, Mason and co-workers used neutron scattering experiments to study the hydration of guanidinium cations.[40] What they found is that guanidinium exhibits extremely low hydration with no recognizable hydration shell, which may indicate a reason for the favorable interaction of this functional unit with protein surfaces. It is well-known that the guanidinium ion is one of the strongest denaturant ions, readily inducing the unfolding of proteins with which it interacts. Lack of an ordered hydration structure indicates that the barrier of solvent removal to interact with these proteins would be quite low. These results also demonstrated that the rigid planarity of the guanidinium, which is so often cited from crystal structure determination, is also present in a fully solvated medium.

Finally, to conclude this short discussion on the nature and contribution of solvent to the formation of complexes between guanidinium and oxoanions, the effect of ionic strength in the medium should be briefly discussed. In many cases, the association

between these complementary binding partners is predominantly electrostatically driven. More accurately, what is occurring, especially in an aqueous solution, is an ion exchange between a counterion and the preferred binding partner. This exchange is largely affected by the ionic strength of the bulk medium in which the reaction is taking place. Without going into great detail, an increase in ionic strength can be thought of as increasing the shielding or competition by counterions against the association of the binding partner of interest. In other words, the actual potential (strength of electrostatic binding) between the reactants of interest decreases as the ionic strength of the medium is increased. Because many of the interactions discussed below are taken directly from observations based on or designed to mimic physiological conditions (a high ionic strength medium), consideration of this effect is paramount to assessing the degree of interaction in guanidinium-centered binding schemes.

2.5. Molecular Recognition

Agents possessing stereochemical centers in concert with intra- and intermolecular interactions are Nature's most effective tools for tuning three-dimensional aspects of complex formation. This concept is driven by the chirality (configuration) of chemical entities. Therefore, this is an important consideration in many molecular recognition schemes. Along these lines, an in-depth understanding of stereochemically driven (chiral) interactions, as well as modeling them for novel systems, remains a formidable challenge. A common feature generally accepted of species that interact stereoselectively or enantioselectively, especially in the field of synthetic receptor design, is the phenomenon described by the "three-point contact rule."[41−44] Logically, this rule, driven by geometric considerations, states that for enantioselectivity, a minimum of three simultaneously interacting sites, in concert with a fourth space-discriminating vector, between the chiral analyte of interest and neighboring chiral molecules must be forged to create a sterically defined diastereomeric complex with a binding preference for one particular form over the other. A recent review by Lindner and co-workers addresses the general needs and challenges involved with separation of chiral entities by a wide variety of techniques.[45]

A high degree of conformational homogeneity in the host molecule must exist to preorganize the structure for a higher affinity binding of one enantiomer over another.[46,47] Hydrogen bonding is utilized in many cases as the primary (strongest) interactions to direct the other interactions. Following this, the locking of angles through a rigid structure in side groups achieves stereodiscrimination—the more hydrogen bonding sites (placed judiciously), the greater the interaction strength (but not necessarily the stereoselectivity as well). If there exists a host system that is too flexible, diastereomeric complexes with enantiomers of similar energy will be formed and no enantioselectivity is observed. Also worthy of note, enantioselective binding has been found to be extremely solvent-dependent because small entropic changes in the environment can overwhelm delicate associative arrangements. Prior to synthesis, molecular modeling approaches can be valuable for assessing the possible degree of enantioselectivity.

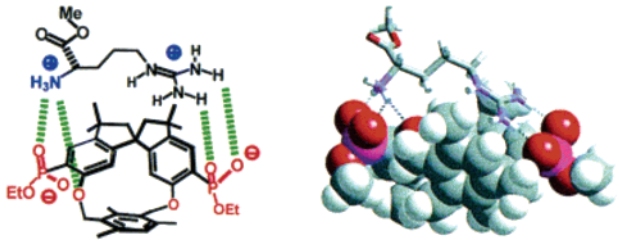


Figure 6. Interaction scheme and simulated structure for the interaction between H-arginine-Me and a novel bisphosphonate receptor. The selector exhibits enantiomeric selectivity as well as selectivity between arginine and lysine residues. Reprinted with permission from ref 50. Copyright 2000 American Chemical Society.

The presence of stereochemical interactions between guanidinium- and complementary oxoanion-containing molecules is not a new discovery. For example, in 1989, simple systems of crystal structures from aggregated amino acids show that the contact between guanidinium and carboxylate groups can mediate stereoselective interaction. Soman et al. reported such structures formed from DL-arginine−DL-glutamate monohydrate and DL-arginine−DL-aspratate systems.[48] Molecular sheets connected in one plane by the α carboxyl and amine groups were connected to similar sheets through guanidinium and carboxylate endgroup interactions. Fundamental differences were observed in the aggregation patterns in the LL−, LD−, and DL−DL complexes.

Schrader et al. also provide an excellent example of chiral selectivity in their work on the enantioselective binding of arginine with one of their novel bisphosphonate-based receptors.[49] The work states that the mechanism of enantioselective recognition relies on the simultaneous interaction of two cation−phosphonate contact points. After docking, part of the arginine residue comes into van der Waals contact with the chiral surface of the chiral bridging unit. By these three interaction sites and due to the creative receptor design, one enantiomer of arginine is bound preferentially over the other. Figure 6 shows the docking of C-terminal methylated arginine onto the novel bisphosphonate receptor.[50]

Other examples of chiral recognition will be discussed in following sections. In general, the importance of this realm is not yet fully understood. However, guanidinium and its complementary anionic groups, being strongly interactive and often mimetic of specific biological interactions, are important contributors and often-used tools for achieving this overall enantioselectivity, especially in the course of chirality-directed drug development.

2.6. Techniques for Analysis

Though complete coverage of the topic of analysis of molecular recognition schemes could easily comprise a complete review in its own right, it is important to lend a small discussion to the types of analyses that are performed to gain useful informa-

tion for assessing and measuring noncovalent binding interactions. This is a short list discussing some of the more prevalent techniques cited in the investigations covered by this review. Several techniques are used, but generally the choice is dictated by analytical fundamentals such as sensitivity, speed, specificity, integration, analyte and solvent compatibilities, preference, accessibility, and the type of information desired.

By far, the most widely reported technique in systems involving complex formation between guanidinium/arginine-based electropositive species and anionic species containing phosphonate, sulfonate, or carboxylate moieties is nuclear magnetic resonance (NMR) spectroscopy titration. In general, NMR experiments are used to follow binding events by observing resonance signals of a host or guest molecule. For example, NMR can follow the formation of a protein−ligand complex in solution through observation of changes in chemical shifts of the protein (host) or the ligand (guest). Such an approach is most useful for monitoring hydrogen-bonding interactions with proton-enhanced NMR (^{1}H NMR); however, this approach (and other NMR-based analytical methodologies designed to probe stereochemical information) can be analytically complicated due to the number of protons present in large biomolecules. Repeating this process for a variety of temperatures allows for the construction of a van't Hoff plot.[51] The van't Hoff plot is then used to deduce enthalpy and entropy of binding involved in complex formation. For studying systems that bind phosphate or phosphonate specifically, ^{31}P NMR has also found use.[5]

When a stoichiometric (e.g., 1:1) complex is formed, this can easily be observed by a chemical shift of qualifying protons in the NMR spectrum. However, complexation of a second equivalent often produces very little change because the ^{1}H NMR chemical shifts observed for dimeric (1:1) and trimeric (1:2) complexes are similar.[51] Therefore, determination of equilibria where multiple binding phenomena are likely is best performed by alternate techniques, such as mass spectrometry or calorimetry. Other drawbacks are that NMR requires a significant sample amount and compatibility with deuterated solvents, which may limit applicability of the technique to some systems. Overall, although titrations performed by NMR are widely reported and appear to be indispensable for collecting a wide array of structural and energetic information on associative solution-phase phenomena, it is believed that thermodynamic determinations of ΔG°_a, ΔH°_a, and ΔS°_a are laborious, insensitive, and error-prone.[52] This however, may be strongly debated, because some authors consider its accuracy to be quite good. Dougherty and co-workers generally attribute free energy change ($-\Delta G^\circ_a$) values determined by NMR to consistently vary by ± 0.2 kcal/mol.[31]

One of the first complete NMR studies to investigate the binding of proteins and peptides involving hydrogen-bond interaction between arginine residues and the SH3 domain protein was performed by Wittekind and co-workers.[53] Results showed the binding to be characterized by transient charge-stabilized hydrogen bonds between the basic guanidine side chain and single or multiple hydrogen-bond acceptor sites in SH3. A comprehensive review of the use of NMR for assessing interactions in biological systems has recently been published by Myers and Peters.[54] In addition to binding information taken from chemical-shift data, this review also covers the techniques for monitoring changes in relaxation times, diffusion constants, NOEs, or exchange of saturation. These approaches are useful for gathering precise knowledge of solution-phase binding in a variety of systems through probing interactions at the atomic level.

Another powerful technique that is often used for determining binding equilibria data is isothermal titration calorimetry (ITC). In general, calorimetry is used to obtain the stability constant (log K) and reaction enthalpy (ΔH°_{rxn}), by simultaneously varying both parameters and determining the minimum of the error square sum. In this manner, both unknown parameters can be fitted to experimental data.[55] ITC measures the heat released or absorbed from the controlled mixing of potential binding components over a series of injections. The resulting binding isotherm is then analyzed based on an appropriate model to determine binding affinity, association enthalpy, and stoichiometry of binding. Such a technique is attractive due to the determination of useful data based on a single ITC experiment.[51] The technique is fast and accurate and gives ΔH°_a directly as a parameter of measurement through integration of heat impulses in each titration step to create a titration curve. ΔG°_a and stoichiometry are estimated based on curve fitting and ΔS°_a can then be calculated from the Gibbs−Helmholtz equation.[56,57]

A comprehensive review that provides information on the use of ITC, as well as differential scanning calorimetry (DSC), for monitoring biological recognition events has been published by Jelesarov and Bosshard.[58] The authors of the review note the popularity of ITC for determination of thermodynamic information but emphasize the more comprehensive use of both techniques in concert for complete coverage of association events. It is important to realize that in biological and biomimetic systems, binding is occurring between large, often rather flexible molecules that generally possess many degrees of freedom. In binding with such systems, a complicated energy profile can be expected, which is a contribution of many large unfavorable as well as favorable interactions. It is the balance of these interactions, with the favorable slightly outweighing the unfavorable, that provides a system where the overall free energy slightly decreases and binding can occur. The medium in which an event takes place obviously has a highly significant contribution to the measured free energy change. In general, calorimetry is used as an approach to gain insight into the thermodynamics of an association reaction, the ultimate goal being the elucidation of its mechanism. To expound the energy profile of a binding event means to study the event under a wide temperature range

with varying conditions, including pH, solvent, cosolutes, etc., and to predict the free energy, enthalpy, and entropy change from these data. Such an approach is important for elucidating the overall mechanism and revealing, for example, the mutual dependence of possible transition states or the intrinsic limits of the overall reaction.[58]

Other popular instrumental techniques for studying noncovalent binding interactions include an array of mass spectrometry (MS)-based methods. Those incorporating fast atom bombardment (FAB),[59−61] electrospray ionization (ESI),[62−64] and matrix-assisted laser desorption/ionization (MALDI)[65−67] techniques are at the forefront of analytical technology for monitoring biological systems, specifically recognition schemes and complex formation.[68] MS detection using these specialized ion sources allows the observation of noncovalent complexes in the gas phase when the association creates a complex with an overall ionic charge. These complexes may be a result of solid-, solution-, or gas-phase binding processes. Once promoted to the gas phase, binding energies can be determined by collision-induced dissociation (CID), which requires instrumentation capable of tandem MS. The first work reported in the literature for the use of ESI-MS to study noncovalent interaction was published by Henion and co-workers.[69] They observed intact receptor−ligand complexes between FK binding protein and macrolides, rapamycin, and FK506. Since then, a multitude of approaches have been reported for measuring binding constants of noncovalent complexes through the combination of MS and CID experiments.[63,70] The methodology integrating the use of MS-based instrumentation is amenable to high-throughput work and is generally regarded as a sensitive, selective, and universal detection technique with excellent analytical capability.

Depending on the ionization technique employed, however, a debate remains as to the reliability of correlating gas-phase binding energies to those present in solution in the native state. Despite the plethora of methods for removal of solvents and formation of fragile gas-phase complex ions for MS analysis, it is critical, especially for biological systems, to use proper solvents and solution conditions (pH, ionic strength, etc.) to preserve native complex conformations in solution to correlate gas-phase data to solution-phase behavior.[71] Still, changes in noncovalent interaction going from solution-mediated to the gas phase occur and must be addressed with each new system.[72]

Proteins, peptides, and other large biomolecules are known to behave quite differently in the gas phase relative to the solution phase.[7] For example, in solution, protonated basic groups on the surface of the protein extend out into the solvent. Upon transfer to the gas phase, the groups may fold back into the protein creating more of a "self-solvating" system. Self-solvation is a result of intramolecular interactions by the basic groups with electronegative groups on the backbone of the protein. The very basic arginine residue appears to favor the neutral nonprotonated form in the gas phase.[73] Still, the protein remains in a folded state upon removal of solvent, and the basic groups may be difficult to access due to the collapsed conformation. Determination of the exact protonation sites is nearly impossible for a moderately sized protein by MS. If a protein has N protonation sites, n of which are protonated, there are $N!/[n!(N - n)!]$ different ways of protonating the sites.[7] Even for a small basic protein such as cytochrome c, the number of permutations is astounding. Those proteins with well-known secondary structures, such as α-helices or β-sheets can be expected to provide more predictable systems.

Adding to the difficulty of MS analysis is the addition of organic modifiers to aid ionization. This is particularly prevalent in ESI-MS where poor ionization efficiency can result from a completely aqueous environment. As such, the analytical environment is not a good mimic of the physiological environment and consequences such as protein denaturing can occur. As a stand-alone amino acid, arginine, with its extremely high proton affinity, has the propensity to form clusters when electrosprayed from solution.[74,75] The appearance of these clusters can also complicate the mass spectrum, especially where quantitative analysis is desired. Dimers and trimers of arginine were modeled by Goddard et al. and shown to correspond to intermolecular interactions between the guanidinium and free carboxyl groups on adjacent residues.[76] Modeled structures are consistent with crystallographic data of interactions between arginine and acetate.[73] Under the same conditions where arginine forms clusters, histidine shows only a small signal for a dimer ion, and lysine shows no cluster formation. On the other side, highly anionic molecules, such as phosphonic and sulfonic acids, are particularly amenable to negative ion formation and analysis by ESI-MS. Still, solvent modification can affect sensitivity through change in pH and promotion (or suppression) of ionization.[77] Recent studies by Schug and Lindner have been performed on guanidinium−oxoanion amino acid-based systems to probe the use of ESI-MS and ESI-tandem MS for routine analyses of interactions mediated by these groups.[78,79] Results show a significant difference in interaction strengths between condensed-phase ionic noncovalent complex formation and those determined from gas-phase CID studies. Phosphonate- and sulfonate-functionalized amino acids show a higher propensity for complex formation compared to carboxylate during the ESI process. However, once these ionic complexes are isolated in the gas phase, the carboxylate-based molecules require higher energy for dissociation relative to the other oxoanionic groups. Acid/base effects in solution versus the gas phase as well as the geometric arrangement of the interacting groups are cited as reasons for explaining the observed differences.

MALDI-MS technology and use has also grown significantly in the past 10 years in the field of biochemical analysis.[67] MALDI allows for soft ionization through the use of matrixes to transfer ionization energy from a laser to the analyte of interest. The matrix has a threefold purpose: (i) cocrystallization with the analyte of interest, (ii) absorption of the energy from the laser pulse, which promotes the

matrix and analyte into the gas phase, and (iii) transfer of charge to the analyte through collisional activation.[80] In this function, the matrix is also commonly used as the pH modifier or buffer of the system to affect denaturing or lack of denaturing of proteins. Basic matrixes have, for example, found use in many biochemical analysis schemes.[80] Another important aspect, the phenomenon of multiple charging as encountered in ESI-MS, is not observed with MALDI-MS experiments. As such, detectors such as time-of-flight (TOF) mass spectrometers are widely applied for larger biomolecules. Karas and Krüger have published an interesting article dealing with the adduct formation between arginine-containing peptides and various anions.[81] Results indicated that adduct ions could easily survive into the gas phase, provided that the gas-phase basicities[82] of the anions were relatively low. Another interesting example has been reported recently by Shiea and co-workers.[83] They used a novel hexasulfonated fullerene as (1) a selective precipitating reagent for arginine in a mixture of amino acids and, subsequently, (2) a matrix for direct MALDI-MS of the precipitated analyte. This specialized fullerene was also found to be amenable to selecting peptides with varying degrees of arginine present in their structure.

MS-based approaches, specifically collision-induced dissociation (CID), have also found wide use in the field of proteomics, where sequencing of proteins has become a primary issue recently. Fragmentation of proteins that contain arginine can suffer from a "charge-remote" process, where the localization of the charge on the protein is about the arginine residue due to its much higher proton affinity compared to the other amino acids.[2] The consequence is nonrandom fragmentation where proteins containing arginine, which are subjected to dissociation processes, will always fragment about the arginine site, most often leaving arginine on the C-terminus of the fragment. This limits the amount of sequential information that can be gained by CID, since random fragmentation of a protein would be more useful. Still, strategies such as arginine modification and covalent tagging using arginine-specific reagents have found use in surmounting this problem.[84] Fragmentation studies have also been performed on phosphonate-based molecules, such as aminophosphonic acids, and show some interesting fragmentation techniques that could be beneficial in identifying phosphonylated molecules and phosphonyl-containing molecules in future situations.[85]

Often, one of the most useful pieces of information is a visual representation of the interacting species, where binding distances and angles can be analyzed for information. Experimentally, this is achieved through X-ray diffraction, where a diffraction pattern can be used to localize specific atoms in a solid crystal. The largest drawback of this technique is most certainly the need for having the complex form and be stable in the solid phase. Currently, the most sought after component (to borrow a phrase used previously, "the holy grail")[15] of crystallography is prediction of crystal structure based on molecular structure.[86] Such a level of application has not yet been achieved and is a testament to the amount of development still needed with this technique. However, information gained through the use of crystallography in viable systems makes it an important tool for assessing specific binding information at the atomic level. Chakrabarti used structures derived from crystal diffraction to investigate the binding of arginine and lysine side chains to anions in general.[87] The report emphasizes the conformational flexibility of the groups in binding anions, although both basic residues show well-defined patterns in the systems studied. Nearly half of the known proteins incorporating arginine and lysine bind compounds possessing phosphoryl groups. Also, it was important to note from this study that the two end nitrogens (ω-NHs) on guanidinium are not structurally equivalent due to the stereochemical nature of arginine. Numerous other examples of the use of X-ray crystallography are present in the literature and will be cited where applicable in the following sections.

Deserving mention in this compilation of analysis techniques related to guanidinium−anionic interactions is the realm of separation science. Though diverse in its own right, specialized chromatographic techniques have been targeted at these specific interaction systems. In the general sense, exploiting specific interaction, specificity, and selectivity by employing a functionalized ligand is known as affinity chromatography. Affinity chromatography has found a very lucrative niche as a purification scheme for pharmaceutical and biopharmaceutical products. A review of some of the more recent developments in affinity chromatography is provided by Lowe and co-workers.[88] The work emphasizes the need in biopharmaceutical development to reduce the cost associated with processing and purifying a biologically active drug candidate. Affinity chromatography in this case offers the use of biomolecules, such as heparin, gelatin, and other binder or receptor proteins to induce molecular recognition processes in the purification of drug targets. Despite showing high selectivities, the major drawbacks of this approach are currently low binding capacities, limited life cycles, and low scale-up potential. Still, there exists promise in the technique because many researchers are focused on improving various aspects. For example, Fassina and co-workers are developing chimeric assemblies based on C-terminal arginine peptides for multiple (and not necessarily related) recognition and affinity purposes.[89] It should be noted also that many of the investigations cited below involve affinity purifications of some form or the other to obtain crucial starting materials for further experiments.

More widely applicable, however often less reliable and more time-consuming, are theoretical calculations of optimized geometries of associated species. Theoretical and simulation-based approaches can offer useful visual information as well as insight into reaction specifics, such as elucidation of potential energy surfaces present. A recent review has been published that covers the use of computational simulations of biomolecular systems, specifically noncovalent interaction in proteins.[90] Older reviews are

also available.[91,92] In molecular modeling techniques, potentials between interacting molecules are generated mathematically using specified algorithms and simulated by computer. Though the accuracy of this method is dependent on the choice of appropriate interaction potential and ability to find a global minimum in potential energy, molecular modeling, in a wide variety of forms, continues to be a useful addition to instrumental-based methods.[55] More complex algorithms are available and used widely as well. These include semiempirical, ab initio, and density functional model[93] calculations. One of the main drawbacks is the compromise between the cost of highly accurate simulations (time) and the level of information desired (interaction geometries, binding energies, etc.). For the case of large biomolecular-type systems, the cost generally rises quickly with the size of the system, and a lower level of calculation must be substituted. Here, where large biomolecules contain a multitude of interaction points and degrees of freedom, you get what you pay for, but the cost can easily be impractical.

With the advent of the ability to solve more and more complex systems and with more powerful computers and algorithms continually being developed, structure-based modeling has become a powerful tool in aiding the understanding of important biological processes such as signal transduction, enzyme cooperativity, and metabolic reactions—all systems where the functional units of interest being discussed here are present.[90] Molecular dynamics approaches have also become popular for analysis of solvation, particularly hydration, effects on proteins, as well as in other interacting systems.[37] Compared to continuum solvation models where the solvent is treated as a bulk dielectric, MD allows the interaction of a few specific solvent molecules in the presence of a particular complex system to be monitored.

Simulations are not always completely accurate and should generally be utilized as supplemental material. In the recognition of a phosphate anion by a macrocyclic bicyclic guanidinium receptor studied by de Mendoza and co-workers, a discrepancy in the point of interaction was found between NMR and MD experiments.[94] Theoretical MD predicted the binding of phosphate within the cavity of the receptor, whereas experimental NMR data showed the phosphate anion to bind with the exterior of the macrocyclic cage and shift quickly from side to side. Figure 7 shows (A) the MD simulation of the encapsulation of the diphenyl phosphate molecule into the guanidinium-functionalized macrocycle and (B) the idealized representation of the macrocycle for structural clarification. Results such as these advocate the need for cautious interpretation of correlations between theoretical and experimental results. Though such discrepancies are sometimes uncovered, molecular modeling and theoretical computer simulations have been used for many years[95,96] to study biological interactions and still make important contributions to the field.

Overall, the chemist's toolbox for the assessment of interactions between these systems of interest is widely varied and very powerful. In the future, there will be a move to more efficient, specialized, and miniaturized analytical techniques. That is not to say that the current techniques have lost their flavor. A solid foundation has been laid to build upon and much advancement will undoubtedly be small improvements on those tools already being used. The diversity available should allow for the further exploitation and observation of interactions between guanidinium groups and anionic-based structures of interest and increase the number of significant developments in this field.

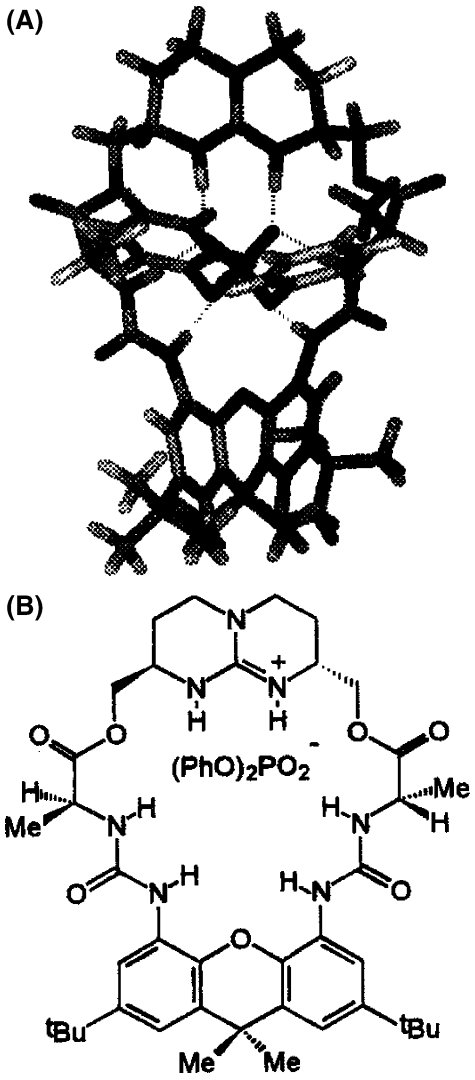


Figure 7. Part A shows a molecular dynamic simulation of the encapsulating of a diphenyl phosphate salt by a guanidinium-functionalized macrocycle. NMR studies showed this picture to be inconsistent with actual events. Panel B shows an idealized picture of the receptor structure and the encapsulation of the diphenylphosphine molecule. Reprinted with permission from ref 94. Copyright 1998 Elsevier.

3. Specific Receptor Systems

For complete coverage of the systems of interaction and the current interpretation of the specifics of interaction between guanidinium-based and phosph[on]ate- or sulf[on]ate-based molecules, we will follow a chemical reciprocity scheme. In this manner, we will first address guanidinium-based

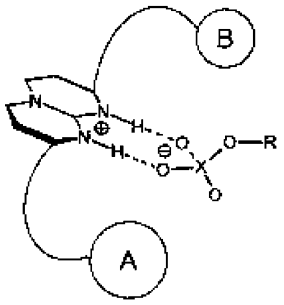
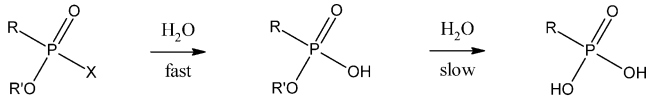


Figure 8. Example of the geometrical alignment of a bicyclic guanidinium group with an arbitrary oxoanion in the context of a molecular tweezer design. Reprinted with permission from ref 52. Copyright 1999 American Chemical Society.

receptors in biological and then in nonbiological-based systems. Next, we will address research that has been focused about phosph[on]ate groups and their specific interaction with guanidinium-based moieties in, similarly, both natural and synthetic environments. The same approach will be taken for addressing the studies of similar interactions by sulf[on]ate moieties. By following the results presented in the literature, we may then provide a summary of guidelines established for interaction between these complementary groups. Thus, in a reciprocal manner, we can combine informative material to gain greater insight in further work on these associative systems.

To begin collecting information on the guanidinium group−anion group interaction systems, we first focus on investigations centered on elucidating the behavior of the guanidinium group as an integral component in arginine residues in biological systems and as a central molecular recognition functional unit. As we have eluded to earlier, the guanidinium group is able to bind to a wide variety of electronegative groups, specifically anionic species, such as carboxylates, phosph[on]ates, and sulf[on]ates. Figure 8 shows a bicyclic guanidinium depiction in the context of a molecular tweezer receptor and the complimentary arrangement it provides an arbitrary oxoanion.[52] In biological and synthetic systems, a multitude of these interaction sites may be adjacent (linearly or topologically) to increase the overall binding effect to complimentary substrates or guests.[30] Synthetically, arginine may be most useful in developing promising leads for membrane transporters, which can deliver biomolecules and drugs to various cellular targets.[97]

Based on ionizability, phosph[on]ate-based receptors show a distinct advantage over those based on carboxylate units. As an acidic unit, capable of Coulombic interactions and hydrogen bonding, phosphonates are fully dissociated in all but the most extremely acidic aqueous medium due to their much lower pK_a values (see Table 1). This is an important consideration for molecular recognition under a wide variety of solution-phase conditions but ensures a completely ionized group under physiological conditions. The possibility for forming a doubly charged anion under slightly basic conditions is also present. In addition, the phosphonate group can be esterified to incorporate lateral functionality that may aid in



X = F, $SCH_2CH_2NR"_2$

R = C_1-C_3, R' = C_1-C_{10}

Figure 9. Hydrolysis products of nerve agents to phosphonic acid.

the formation of favorable interactions with various functional arrangements.[98,99] The phosphonate group and the related phosphate group are important components in biological processes. Phosphorylation is a prominent molecular switch used by Nature for signal transduction and enzyme catalysis.[100−102] Consequently, these groups make excellent molecular recognition functionalities for forming directed hydrogen bonds with guests containing electropositive groups. Also, with the recent focus on chemical warfare agents, many of which contain oxygenated phosphorus functionality, the study of phosphonic acids, prominent hydrolysis products of these agents, has met with increased interest in forensic-based science. Figure 9 details the conversion of phosphorus compounds, such as nerve agents, to phosphonic acids.[103] Aminophosphonic acids have also found a fair amount of interest in recent publications due to their ability to mimic natural carboxylate-based amino acids. Similarly, phosphorylated amino acids have been used to build phosphorylated peptides.[85]

Sulfonate-based recognition systems for arginine are not as widely reported relative to phosphonate-based systems. However, their diverse biological activity makes them potential leads as active pharmaceutical agents. The effectiveness of the new sulfonate drugs depends on their ability to bind to specific sites on proteins.[104] Since the sulfonate group is a strong acid that is deprotonated under the entire pH range in aqueous conditions, interactions with electropositive groups is a likely scheme. Consequently, the arginine residue and other basic sites on proteins will be prominent targets. In addition to their use in drug development, sulfonated compounds have a history as cocrystallization or coprecipitation agents for biomolecules. This indicates solubility of formed complexes to be of great importance for solution-based systems that can be used by other analytical techniques, such as ESI-MS. The use of sulfonate dyes is the most commonly reported agent used for recognition of guanidine-based systems by sulfonates in the literature. Currently, the selectivity of simple sulfonate compounds for arginine residues is not completely understood. Interactions with proteins also include the need for determining which residues are accessible for a given probe and protein in the system studied. Protonated amino acid residues on the surface of biomolecules form noncovalent complexes with deprotonated sulfonate moieties.[105] These systems will be discussed along with other sulfonate-based receptors for probing arginine residues.

In the following text, the strong interest in manipulation of these complementary interacting groups in recent research will become apparent. All realms of chemistry, from synthetic organic to analytical to

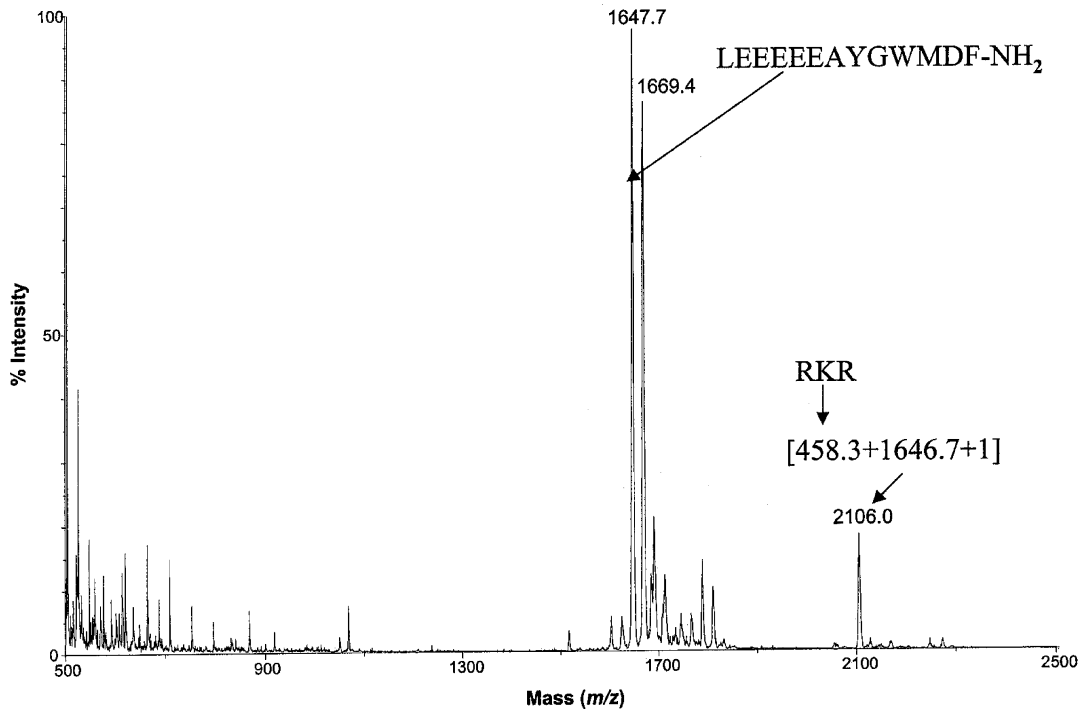


Figure 10. MALDI analysis showing complex formation between the RKR (Arg-Lys-Arg) motif and minigastrin. Reprinted with permission from ref 112. Copyright 2001 Elsevier.

applied industrial chemistry, can identify with the information provided here. By gathering and presenting the topic in a way where the reciprocal complementary of these groups can be better understood, we hope to provide not only a firm understanding of the extensive fundamentals already elucidated but also a better insight and an increased interest in this topic for further work to build upon.

3.1. Arginine-Derived Groups Interacting in Biological Systems

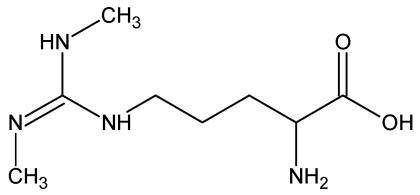
Arginine residues and their function as anion binding sites are ubiquitous in nature. They are found in the binding region of a large number of enzymes and signaling proteins. These proteins employ arginine residues, more specifically, guanidine moieties, to interact with negatively charged anionic[106] or π-electron-rich aromatic[22,24] moieties of substrates or cofactors. The interaction or interface between two proteins in particular have been found to include a disproportionately large number of arginine residues. Such sites have been commonly referred to as "hot spots" in recent literature discussions.[107] Here we focus specifically on guanidine−anion interactions in biological systems. Although a large majority of anion-types in these systems are carboxylate-based, we will focus mainly on the interactions of guanidine (specifically in the form of arginine as an amino acid residue in a sequence or as a base in a free amino acid) with phosph[on]ate- and sulf[on]ate-based oxoanions. In biological systems, most of the material will focus on interactions with the phosphate, rather than the phosphonate group. Though these groups may differ slightly in their ionizability (see Table 1), geometrically, phosphate and phosphonate are quite similar and are widely reported as having high affinity for guanidinium groups. Similarly, sulfate groups are more commonly encountered in natural systems than sulfonate groups, and we will therefore approach this discrepancy in a manner similar to the phosph[on]ate group-based interactions. Proteins that bind phosphate and sulfate in biological systems in particular are important receptors for active transport of these ions in cells.[108] Other interactions cited are hypotensive and adrenergic neuron blocking effects.[109,110] A comprehensive review was published previously that covers many of the relevant biological interaction systems, such as those in staphalococcal nuclease, deoxyribonuclease, and alkaline phosphatase.[5] These systems will therefore be largely ignored in this section and addressed in following sections with the discussion of anion-based systems.

One peptide and protein sequence found and recently studied due to its specific biological activity is the arginine−arginine (RR in one letter nomenclature) motif in peptides. This is often referred to in the biochemical literature in conjunction with the twin arginine translocase (Tat) binding domain.[111] Dynorphin, an opioid peptide, and five of its fragments contain this motif and consequently show consistent interaction with peptides that contain two to five adjacent acidic residues, namely, aspartic acid and glutamic acid during MALDI-MS analysis.[112] Other peptides not related to dynorphin but still containing the RR motif or even the arginine−lysine−arginine (RKR) motif exhibited similar affinity for complexation. Figure 10 is an example of this complex formation, showing the MALDI mass spectrum generated by Woods and Huestis of a complex formed between the RKR peptide and minigastrin (a principle form of gastrin that in biological systems stimulates gastric acid production). Peptides containing only adjacent lysine or histidine residues did not

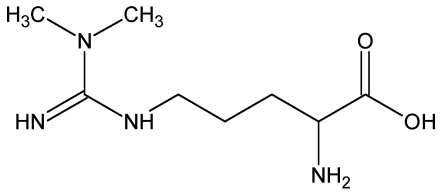
complex acidic residues on other peptides. Similarly, when an arginine in the binding motif was switched to another residue, such as phenylalanine, complex formation was disrupted. In this study, strong ionic bonds were formed by dynorphin and the RR-motif-containing peptides with the carboxylate units of the acidic peptides. These bonds remained intact during enzymatic digests and returned a method for selective "peptide footprinting" where the point of interaction could be determined. The effect of the MALDI matrix was also investigated, showing matrixes with a less acidic pH to be less disruptive to complex formation. This study is a prime example of the formation of complementary peptide−peptide complexes by electrostatic attraction (salt bridges and hydrogen bonding) between a specific arginine (guanidine) motif on one peptide and carboxylate groups on another.

Another good example of the role of the Tat motif in binding biological phosphates is given by Loo and co-workers.[111] Using ESI in both the positive and negative mode, they tested the specificity of Tat interactions with TAR RNA using TAR mutants. They showed that TAR mutants and their binding with Tat peptide could be differentiated through competitive binding experiments. Since Tat protein, from HIV, is a viral transactivator that is essential to the replication of the virus, compounds that are developed that interfere with this complexation to TAR RNA may exhibit beneficial antiviral activity. This is only one such example of the wealth of attention paid in recent years to binding by the Tat region peptide sequence motif.

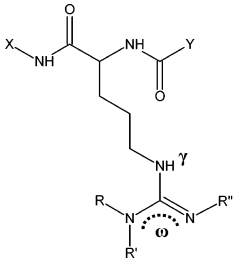
Modified arginine residues are commonly encountered in natural systems, and it has been postulated that the methylation of arginine residues is an important molecular switch used by nature to modify selectivity. Compared to phosphorylation and sulfation, methylation is another example of a posttranslational modification.[113] Methylation of arginine was discovered over 30 years ago,[114] but it has not been until more recent work was completed that the full function of the methylation of arginine as a pathway for signal transduction, transcription activation, and protein sorting has come to light.[115] It is for this



1



2



R	R′	R″	Name	$pK_a(\gamma)$	$pK_a(\omega)$
-H	-H	-H	Arg	-1.51	13.31
$-CH_3$	-H	$-CH_3$	sDMA	1.64	13.66
$-CH_3$	$-CH_3$	-H	aDMA	-1.40	14.27

- X, Y denote general peptide linkages.
- Values calculated using ACD/Labs pKa calculator.
- Charge placed on γ indicates an overall +2 charge on the group.

Figure 11. Calculated pK_a values for the modified and unmodified arginine residue. Results show overall a very small change in calculated values of the groups and therefore little change in chargeability with change in pH upon modification.

reason that the methylation of arginine, a common posttranslational modification in eukaryotes,[116] has received considerable attention in the past few years. Methylation or, more commonly, dimethylation of arginine can occur to create a symmetrically or asymmetrically dimethylated arginine molecule (sDMA **1** and aDMA **2**, respectively). Dimethylation of arginine has several physicochemical effects:[117] (a) it makes the guanidinium moiety, and hence the arginine residue, slightly more basic; (b) it increases arginine hydrophobicity and molecular volume; (c) it decreases the hydrogen bonding ability of the functional unit, due to removal of potential donor sites; (d) it subtly changes the pK_a of arginine.[118] Still, through dimethylation, the total charges on the nitrogens remain unchanged and only a slight change in partial charges on each atom is found. Figure 11 illustrates the small change experienced in the chargeability of the arginine residue upon dimethylation. These values were calculated using ACD/Labs pK_a calculator program. Though imperfect with respect to resonance structuring of the modified guanidinium group, the calculation shows that there is little change in the calculated pK_a of the groups and that, over a reasonable pH range, the overall charging of the group will remain unchanged.

An interesting study was performed by Pongor and co-workers on the assessment of the effect that methylation has in hydrogen bonding between arginine in peptide sequences and phosphate in nucleic acids.[117] Such a scenario accounts for the majority of these interaction types. Quantum chemical calculations using HBPLUS were used to find and analyze H-bond interactions between arginine and phosphate in 95 nonhomologous protein−DNA complexes.[119] In this system, 702 contacts were found between arginine and DNA, 658 of which were hydrogen bonds formed by guanidinium. Of these 658 instances, 395 (~60%) were nonspecific interactions with the sugar−phosphate backbone. These bonds were grouped into seven different hydrogen-bonding patterns as stipulated previously by Shimoni and Glusker.[18] The occurrence of dimethylarginine (DMA) hydrogen bonding was significantly less abundant; however, aDMA could still be found in a few cases and would generally only interact through a few of the hydrogen-bonding patterns available to unmodified arginine residues. It was conjectured that dimethylation could be used by natural systems to pick certain conformational arrangements (increase specificity) over

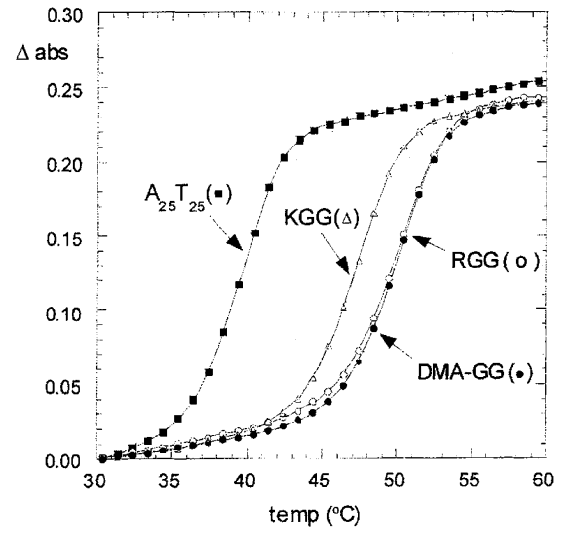


Figure 12. Melting curve at 260 nm of $A_{25}T_{25}$ dsDNA oligonucleotide alone and in the presence of a 50-fold molar excess of KGG, RGG, and DMA-GG peptides. This shows the similarity in effect that arginine- and DMA-containing peptides can have on structural interactions with other proteins and peptides. Reprinted with permission from ref 117. Copyright 2001 Oxford University Press.

those nonspecific interactions encountered with unmodified arginine. However, this was not found to be the case, and apparently the dimethylated systems studied here bind in a similar manner as unmethylated arginines to nucleic acids. Figure 12 shows the similar effect of RGG and DMA-GG on the melting of $A_{25}T_{25}$ dsDNA oligonucleotide, relative to that without any peptides added or in the presence of KGG.[117] More likely, the difference in the effect that arginine- and DMA-containing peptides have on binding to various substrates is the exclusion of some hydrogen-bonding patterns in the methylated systems allowing for a particular arrangement of secondary structure outside of the primary interaction point.

Concerning the analytical aspects of dimethylated arginine, as might be expected, aDMA and sDMA, being isomeric molecules, have been difficult to identify between and separate. Recently, a procedure for fragmentation of peptides, based on tandem MS, has been reported to elucidate the nature of post-translational modifications on arginine, specifically DMA.[113] The use of parallel monitoring of ions in precursor ion scanning MS showed five characteristic fragments that could be used to differentiate sDMA and aDMA. This development is a significant contribution to the field of proteomics, specifically through the sequencing of proteins of important function. Also interesting is the future application of precursor ion scanning for differentiation of other isobaric protein modifications. Very recently, two other reports have been published on tandem MS separation of dimethylated arginines.[120,121] These approaches also rely on identification through the formation of distinct fragments ions depending on the presence of symmetric or asymmetric modification.

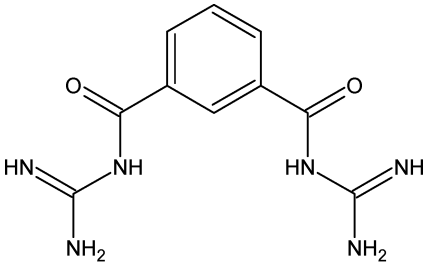
Other arginine variants have also been shown to be biologically active. A recent study by Balz et al. investigated the in vitro effects of arginine, homoarginine, *N*-acetylated arginine, and argininic acid on nucleoside triphosphate diphosphohydrolase (NTPDase) and 5′-nucleotidase activities.[122] NTPDase hydrolyzes the extracellular nucleosides tri- and diphosphates and has been well characterized in the central nervous system. The arginine derivatives listed above have been shown to have epileptogenic properties when accumulated in hyperargininemic patients.[123,124] All of the tested arginine variants were shown to promote an increase in nucleotide hydrolysis at concentrations commonly found in plasma of cerebral fluid of afflicted patients. The results provide insight into the understanding of the fundamental biological processes that promote and lead to hyperargininemia in humans.

Though we have only mentioned a couple of basic motifs encountered in biological investigations where the guanidinium group plays a prominent role, specifically here the RR (and other similar cationic arrangements) and the DMA arrangements, such interactions in these systems are of course not limited to these. For example, many permutations of cationic residue sequences are important to the forging of noncovalent interactions that drive recognition and transport in biological systems. As stated previously, a more complete coverage of the current work into developing an understanding of the interaction of guanidinium groups with phosph[on]ate and sulf[on]ate groups will be given in the sections focused on those anionic moieties, which follow.

3.2. Synthetic Guanidinium Groups/Selectors/Ligands/Receptors

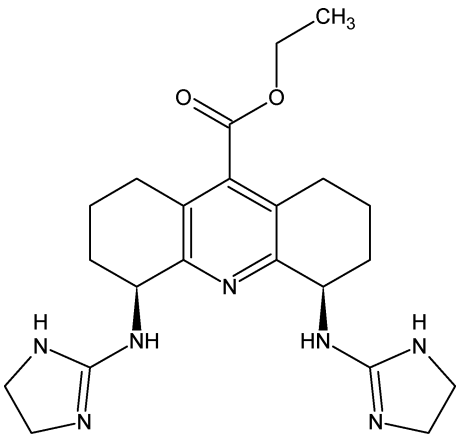
The field of arginine- and guanidine-based synthetic receptors has been reviewed several times in the not-to-distant past. A thorough recent review by Anslyn and co-workers focuses on the field of guanidinium-based receptors for anionic ligands.[30] An older review by Hannon and Anslyn has also blanketed the material related to the earlier development of the guanidine receptor field, as well as the role of the guanidinium group in biological functions.[5] Because so many excellent reviews have been published on this topic, we seek here to only present highlights and some more recent discoveries. In short, the factors associated with binding to guanidinium receptors that are currently being addressed (and with good success) are utility in aqueous media, increased enantioselectivity, new systems types including catalytic ability, and a better understanding of underlying chemical principles that drive recognition ability. Since not all of the guanidinium-based receptors are used exclusively for binding phosph[on]ates and sulf[on]ates, it is necessary here to also consider those receptors designed specifically for binding carboxylates as well. The receptor types considered here will largely be cleft or tweezer-type receptors due to their relative simplicity compared to encapsulating macrocyclic constructions.

In 1992, Hamilton and co-workers synthesized a bisguanidinium-based receptor, **3**, to accelerate phosphodiester cleavage in RNA by mimicking staphylococcal nuclease (SNase). The work of Anslyn and co-



3

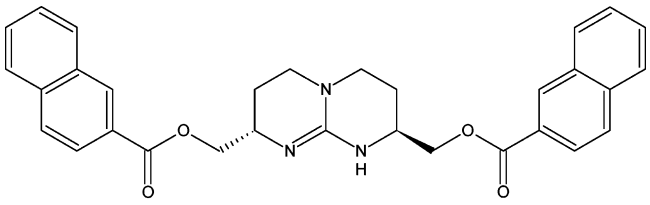
workers in 1993 showed the synthesis of another bisguanidinium receptor, **4**, able to recognize phos-



4

phodiester groups of RNA to induce selective cleavage.[125] These works showed the role of the arginine residues in staphylococcal nuclease for RNA cleavage and its ability to mimic this interaction with synthetic receptors. Orientation about the phosphate was designed to mimic the "arginine fork" motif.[126] The specifics of the "arginine fork" motif are discussed in a later section. Following Hamilton's work, Muche and Göbel constructed bisguanidinium alcohol receptors to mimic the structural and functional features of the phosphodiesterase SNase.[127] Reaction rates and binding strengths were shown to increase with the addition of a second binding site in the receptor architecture. The kinetics of the reaction rate for phosphorylation (a covalent binding event) was studied by ^{31}P NMR.

Lehn, De Mendoza, and co-workers are accredited with development of the first bicyclic guanidinium receptor **5** to achieve enantioselective recognition of



5

anions.[128,129] Figure 13 shows the chemical shift data representative of the complex formation between the first chiral bicyclic guanidinium receptor and an aromatic carboxylate. Bicyclic guanidiniums are con-

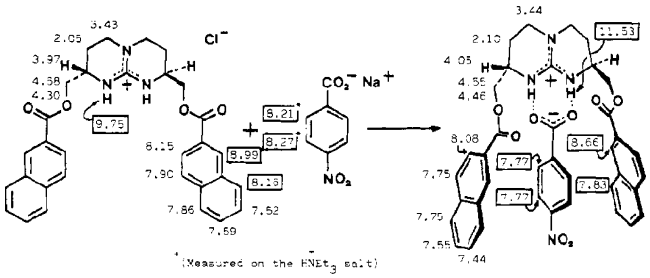
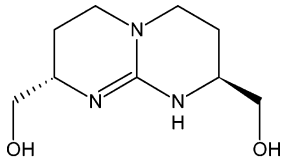


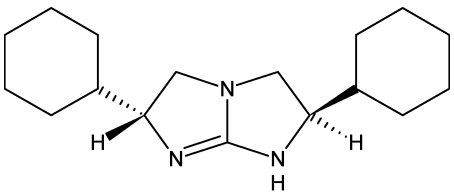
Figure 13. Chemical shift data relevant to the first recognition of carboxylate by a chiral bicyclic guanidinium receptor developed by De Mendoza and co-workers. Reprinted with permission from ref 128. Copyright 1989 American Chemical Society.

ceptually simple as they orient the guest of interest into a well-defined region of the receptor without the encapsulation necessary for larger macrocycle receptors.[130] Aromatic carboxylic guests interacted via ion pairing with the guanidinium group and $\pi-\pi$ interactions with aromatic side chains. The development of the receptor is a natural extension on the synthesis of previously developed chiral bicyclic guanidine molecules by Echvarren et al. (**6**, (*S*,*S*)-isomer).[131] Use



6

of asparagine allowed for production of (*R*,*R*)-, (*S*,*S*)-, and *meso*-isomers of this receptor scaffold. Corey and Ohtani contributed as well to this development with the synthesis of chiral bicyclic guanidine molecules, of which **7** is an example, from enantiomerically pure

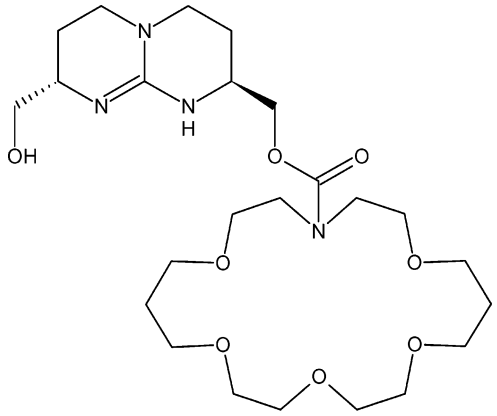


7

amino acids in eight steps.[132] This strategy could also be utilized for the preparation of unsymmetrically substituted chiral receptors by starting from two different amino acid molecules. In general, for diguanidinium receptors, Lehn and co-workers assessed that the dominant factor controlling the stability and selectivity of binding in water was the charge density on the host and guest.[133] Greater charge created greater binding. Thus for binding of phosphates to di- and triguanidinium receptors, $P_2O_7^{4-} > HPO_7^{3-} > H_2P_2O_7^{2-}$, and with respect to the guanidiniums, trifunctionalized hosts bound the guests more strongly than difunctional ones.

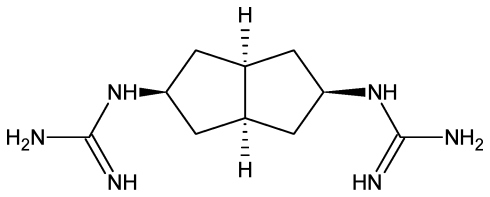
Recently, de Mendoza and co-workers have demonstrated the first example of chiral recognition through bicyclic guanidinium-based selectors of underivatized amino acids as carriers under neutral conditions.[134] The impetus for this work was the

preparation of artificial carriers for zwitterionic aromatic amino acids across bulk model membranes. To achieve chiral discrimination, a simultaneous binding of the positive ammonium group and the anionic carboxylate of the amino acid of interest should be performed. In this configuration, the side chain of one enantiomer should bind into a more favorable orientation than that of the other enantiomer, thus creating a marked energy difference in the binding of the two enantiomers. The carrier, a bicyclic guanidinium developed 10 years previously (substituted analogues of **5** and **6**) and functionalized with a crown ether unit, achieves up to 80% ee for some amino acids and surmounts the formidable barrier of transporting a zwitterionic species out of aqueous media. Structure **8** shows one of these crown ether analogues.



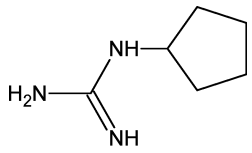
8

Hamilton and co-workers were among some of the earlier researchers who recognized the potential of guanidinium groups for sequence-selective binding in proteins. One example of their work is the use of a synthetic, well-characterized bisguanidinium receptor (**9**, with an intramolecular guanidine distance



9

of 4−5 Å) for recognition of aspartate groups.[135] Short helical peptides were used as models of a protein surface where aspartate residues were present with increased spacing in the peptide series. The spacing of aspartic acid groups in the peptides ($i + 3$, $i + 4$, and $i + 11$ units between groups) was used to model the differences in peptide structure that would cause the formation of various secondary structures (α-helix, β-sheet, β-turn, etc.). The remainder of the peptide was composed mainly of hydrophobic residues. The monoguanidinium analogue **10** of the receptor was used as a control. Though the monoguanidinium receptor was found to bind well in a 1:1 complex, the bisguanidinium showed poor binding



10

and was most likely a structural mismatch for the peptide conformation. Results indicated that increased selectivity in protein recognition was most likely going to be achieved through targeting of polar residues on the protein surface with complimentary interactions, such as hydrophobicity, providing additional binding energy. Another example of a receptor designed to interact with carboxylate functionalities on a peptide or protein and based on a guanidinium functional unit was reported in collaboration between Hamilton, de Mendoza, and coworkers.[136] The tetraguanidinium linear chainlike receptor was shown to bind aspartate residues and was able to stabilize α-helical orientation in the peptide. Figure 14 is a MD simulation of the stabilization offered by the tetraguanidinium receptor. It shows the simulated conformation of the peptide in the presence of the receptor (Figure 14A) and alone (Figure 14B). These studies were performed in a 90/10 methanol/water solvent mixture.

In the hunt for suitable receptors for binding protein carboxylate structures under physiological conditions, again Hamilton and co-workers have published an excellent article detailing the thermodynamics associated with binding of different hydrogen bond donor groups (ureas, thioureas, and guanidinium groups) in increasingly competitive solvent

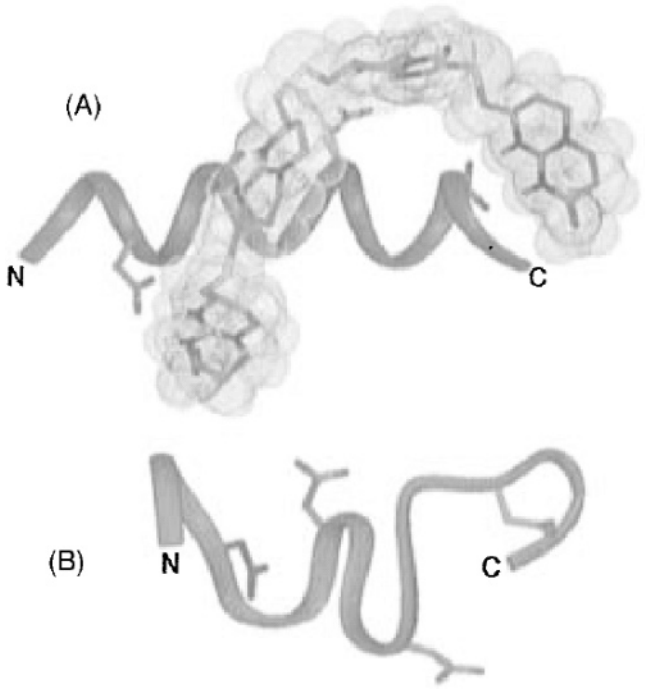


Figure 14. MD simulation of the stabilization of a α-helix conformer by a tetraguanidinium receptor. The panels show structure of the peptide (A) in the presence of the synthetic receptor and (B) alone. Reprinted with permission from ref 136. Copyright 1997 American Chemical Society.

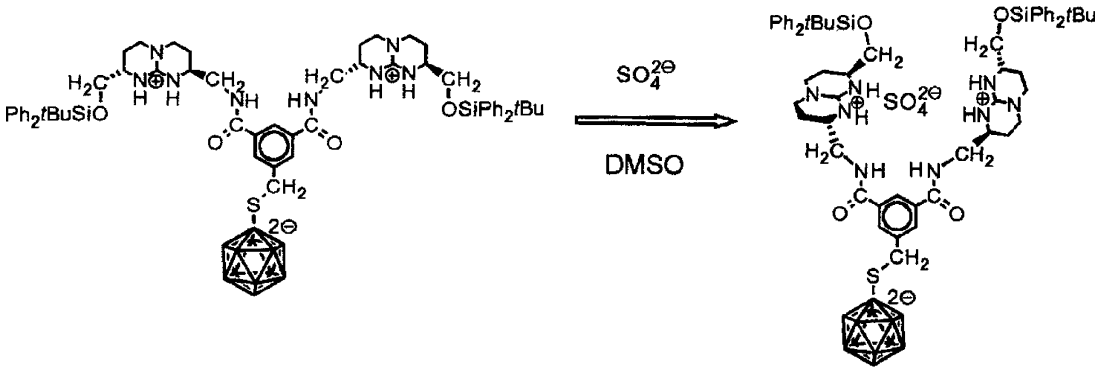
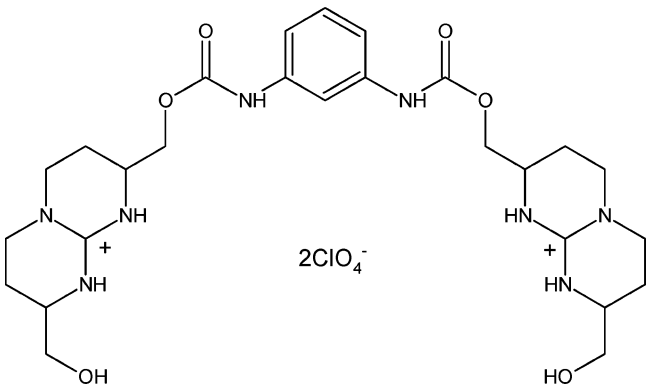


Figure 15. Extraction of sulfate from solution by a ditopic guanidinium host in DMSO. This receptor also shows the capability to operate under dilute aqueous conditions. Reprinted with permission from ref 56. Copyright 1998 Wiley-VCH.

media (from DMSO to water).[51] Association with carboxylate groups increased with hydrogen bond donor strength and, overall, decreased with increasing amounts of water. The decline in association with increase in water amount reinforces the knowledge that a hydrogen bond donating solvent will reduce binding affinity through increased competition by the solvent. Recognition was still observed, though diminished, in high percentages of water. Enthalpic binding between guanidinium and carboxylates in DMSO, a less polar medium, was shown to be enthalpically driven, whereas, in more polar media, such as methanol and water, the enthalpic contribution to binding was positive, indicating entropical contributions due to solvent liberation as the main driving force during the binding and induced fitting events. Berger and Schmidtchen have also reported the entropically driven formation of a complex between ditopic guanidinium based hosts and sulfate ion.[56] One variation was able to extract sulfate ions from very dilute (10^{-4} M) aqueous solution into chloroform with a 99.8% efficiency. Figure 15 shows the extraction of sulfate and the arrangement of the ditopic guanidinium host in DMSO. Another ditopic guanidinium receptor using supporting ligands to provide supplementary interaction outside of the primary electrostatic interaction showed selective complex formation for phosphate over sulfate.[52] ITC and NMR studies were performed to assess the binding in these systems. This information and the further development of such systems become exceedingly important as there is a greater push toward operation of recognition schemes in physiological conditions.

An important point to consider when examining the scheme shown in Figure 15 is the actual dynamic nature of the system under consideration. What is shown is a change in overall conformation through alteration of torsion neighboring the guanidinium moieities upon association between the receptor and the sulfate anion. In reality, and in polar solvents, this arrangement may not be the most dominant associative structure. Many 3D structures are in principle possible that incorporate the receptor, the sulfate, and the solvent. Although the depicted structure may contribute a significant amount to the favorable free energy for extraction of sulfate by this receptor, there are likely numerous contributions from alternative conformational arrangements and stoichiometries (i.e., one or more sulfates attached to one or the other or both guanidiniums in unfixed or unlocked configurations). The negative free energy measured in this case may be due to the ensemble of host−guest associates (the collection of different arrangements) delivering more binding modes than the separated host and guest would have in solution. In other words, in the case of this host−guest system, the complex formation is entropically driven and enthalpically opposed (i.e., enthalpy/entropy compensation occurs); therefore, the design of the receptor merely allows for a greater number of degrees of freedom when sulfate is complexed versus when all components are solvated. It is important not to take such a representation at its face value but rather to remember that there are competing processes involved, such as association by counterions and polar solvent molecules. These effects change from medium to medium due to the different solvation, dielectric, and ionic strength effects, but they still must be considered when assessing the solution structure of a predominantly electrostatically driven ion-paired complex.

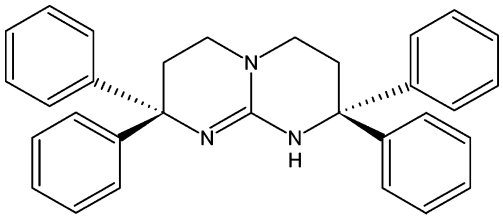
Under physiological conditions, one of the main difficulties for molecular recognition of amino acids is their charged nature. Receptors have to be designed with the ability to recognize a zwitterionic structure without collapsing in on themselves through intramolecular binding. Work by Schmidtchen and co-workers has become increasingly important to the development of such receptors. Their approach relies largely on a guanidinium scaffold (cation-based with affinity for negative component of the zwitterion) supplemented by other structural features to gain affinity for the cationic portion of a zwitterion. Schmidtchen provided early examples of ditopic hosts **11** able to bind tetrahedral anions such as phosphate species in chloroform and in water.[137] This work provided the first example of specific complexation of mononucleotides in water. Stereospecificity was also built into the host. Variations on the guanidinium receptors were made using crown ethers as complementary arrangements for the formation of an array of polytopic hosts.[55] These hosts showed good selectivity for zwitterionic amino acids and laid a foundation for important aspects in further development of the field. These aspects include a strong electrostatic binding functionality, hydrophobic groups for stabilization of the electrostatic interaction, and



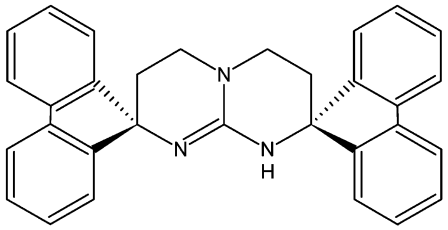
11

selected spacing/spacers to create the proper span necessary for ditopic binding.

More recent work by the group of Schmidtchen has focused on the design of guanidinium-based receptors capable of overcoming the enthalpic penalty for association in competitive media.[57] They have developed a series of receptors, **12** and **13**, wherein the



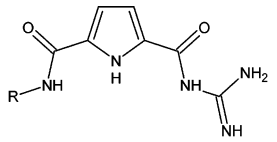
12



13

receptor site is lined with aromatic residues to reduce the solvation in the vicinity of the binding site. The design of the synthetic receptors started with development of allyl, alcohol, and Tos-protected functional side chain groups on the guanidinium receptors in a one-pot synthesis scheme that reduced the complexity of synthetic methods to date.[138,139] With the addition of aromatic side chains in a subsequent synthesis, the enthalpic penalty for interrupting the solvation shell was shown to be reduced, which consequently reduced the hindrance of the solvent to the overall exothermicity of the binding. It was conjectured that this decrease in solvation was accompanied by a lower local dielectric constant near the binding site, which also aided in Coulombic attraction of an anionic guest. A strong emphasis in this investigation was the inadequacy of the geometry- and enthalpy-based Fischer "lock and key" metaphor[140] commonly used to describe recognition processes, and the consequent focus on the role of solvation and entropy in many binding systems.

The laboratory of Schmuck and co-workers has been very successful in creating systems of guanidinium-based receptors, capable of strongly complexing carboxylate groups in competitive media, such as water. Much of their methodology focuses on the addition of multiple hydrogen-bonding groups, as well as secondary scaffolding (side chain interactions), in the receptors to overcome solvent competition during association in aqueous systems.[141] Variations in the side chains also offer the ability to create chiral selectivity in recognition systems. Structure **14** shows one

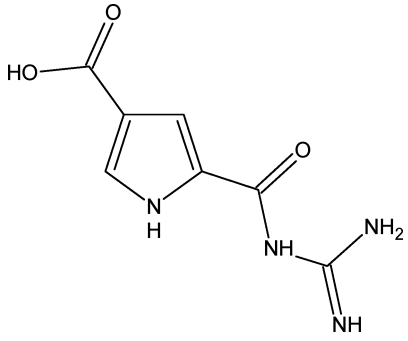


R = ethyl, butyl, amidomethyl, 1-amido-2-methylpropyl

14

of the scaffolds, 2-(guanidiniocarbonyl)-1*H*-pyrroles, investigated for selective binding. MD, simulated in chloroform and water, was used to evaluate the lowest energy binding structure with carboxylates. In chloroform, there was a large discrepancy, 13 kcal/mol, in the difference between the lowest (most-stable) and highest (least-stable) energy structures calculated. In water, this difference was calculated to only be 2 kcal/mol, indicating the barrier for arranging the groups in a favorable orientation for binding is significantly lowered in water. This difference is attributed to a decrease in the repulsion between groups on the receptor due to the effect of a more polar solvent. Binding differences among carboxylates are attributed to the basicity of the carboxylate molecule. A more basic molecule means a higher propensity for forming hydrogen bonds and thus the formation of a stronger complex. In studies performed by NMR titration in 40/60 H_2O/DMSO, binding constants of $K = 360-1700$ mol^{-1} were achieved for the various carboxylates.[142] In 100% DMSO, the binding is so strong that NMR titration just showed a linear increase of the shift changes until a 1:1 molar ratio (clearly showing 1:1 binding stoichiometry) was reached. An additional feature of these receptor/ligand systems is the ability to turn association on and off with pH modification and its effect on deprotonation and protonation of carboxylate groups.[143] It is apparent that this on/off strategy would be less successful in the binding of more acidic sulfonate and phosphonate groups, which remain largely ionized throughout the practical range of aqueous pH modification.

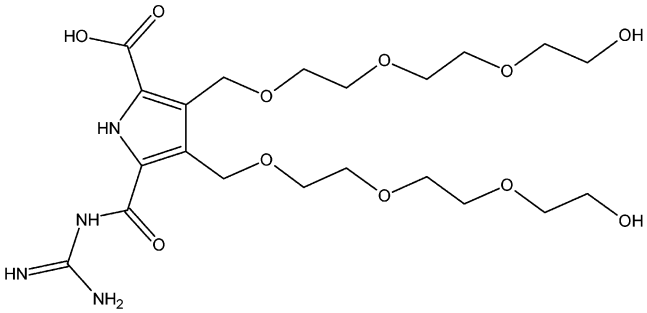
Much work on the 2-(guanidiniocarbonyl)-pyrrole-based receptors has dealt with the functionalization of the receptor by addition of carboxylate groups to assess self-assembly. Molecules such as **15** form linear chainlike oligomers in solution.[144] The structures, monitored by NMR chemical shift data, showed a significant decrease in binding as the polarity of the solvent was increased. In highly polar solvents, such as DMSO, the energy necessary for desolvation can be greater than that gained from the specific binding. However, in this system, it is found that



15

liberation of the solvent creates an entropically favorable binding event. It is pointed out that NMR is a poor tool for studying oligomerization, since signals become increasingly broad with increasing concentration of species. Other examples of controlling self-assembly through dimerization versus oligomerization have been discussed previously through the synthesis of flexible alkyl-linked receptors.[145] By incorporation of a longer alkyl group (butylenes), two similar end groups could fold into each other and self-assemble into a dimer. Use of a shorter linker (ethylene) did not allow enough flexibility for dimer formation, and hence oligomerization was observed instead.

In contrast to other synthesized receptors, **15** will not form discreet dimers. An example of an extremely hydrophilic variation on Schmuck's guanidine-based receptor design is shown in **16**.[146] This molecule binds



16

as a 1:1 dimer in DMSO with a binding constant too high to measure by NMR. In water, the dimer shows a high association constant of 170 M^{-1}, one of the most efficient self-assembling systems in water reported so far. The binding is attributed to a very high contribution of mutual ion-pair formation between the complementary guanidinium and carboxylate groups. In structures based on functionalizing **15**, poor interaction orientation is involved for the formation of dimers, and hence oligomerization results. Termae and co-workers have reported the self-assembly of a pyrene-functionalized monoguanidinium receptor capable of recognizing phosphate ions.[147] The guanidinium receptor forms 1:1 and 2:1 complexes with a biologically relevant pyrophosphate ($P_2O_7^{4-}$, PPi). Figure 16 depicts the binding of the aromatic guanidine receptor to PPi in a 2:1 arrangement. Aromatic $\pi-\pi$ interactions align two receptors in a stacked configuration where the guanidinium groups in each receptor are bridged by PPi.

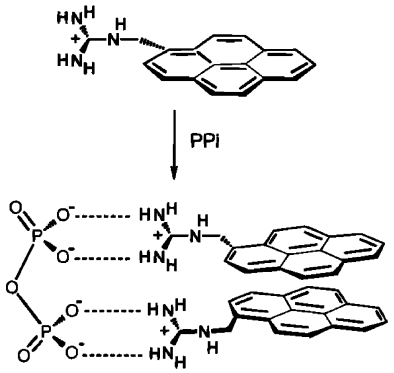


Figure 16. Binding of aromatic guanidinium to pyrophosphate (2:1). Aromatic $\pi-\pi$ stacking is evident in the dimeric arrangement of the two guanidinium receptors. Reprinted with permission from ref 146. Copyright 1999 American Chemical Society.

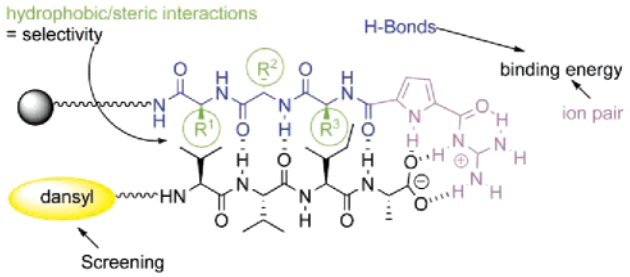


Figure 17. Tripeptide library binding scheme for guanidiniocarbonyl pyrrole receptors with a tetrapeptide representing the C-terminus of Aβ. The guanidine-based receptors are anchored onto beads, which can be monitored by fluorescence for binding events. Reprinted with permission from ref 149. Copyright 2003 The Royal Society of Chemistry.

In more recent work in Schmuck's laboratory, there has been a focus on automated parallel synthesis and evaluation of the binding of different members in a library. In general, there has been a strong move toward high-throughput combinatorial methods. All of the members of the library incorporate a guanidine moiety and are functionalized to maximize the binding of carboxylates in an aqueous environment. The impetus for this research is to contribute to the design of biosensors, targeting of specific cellular processes (e.g., cancer, Alzheimer's disease, and bacterial infections), and the design of new therapeutics.[148] The general features of this receptor class are a rigid, planar binding motif, additional hydrogen bonding due to the presence of the pyrrole NH group, and added selectivity and binding from a variation in side chain groups. Results are generally assessed by NMR titration and temperature-dependent studies,[144] as well as MD and molecular modeling simulations. More recently, fluorescent and UV binding assay determination on functionalized beads have been used for evaluation of the generated receptor libraries.[149,150] Figure 17 generalizes the approach for the library screening by using fluorescence detection to monitor binding events between the guanidinocarbonyl pyrrole (linked covalently to beads) and the tetrapeptide library representative of the target protein terminus (dansylated to provide a chromophore).

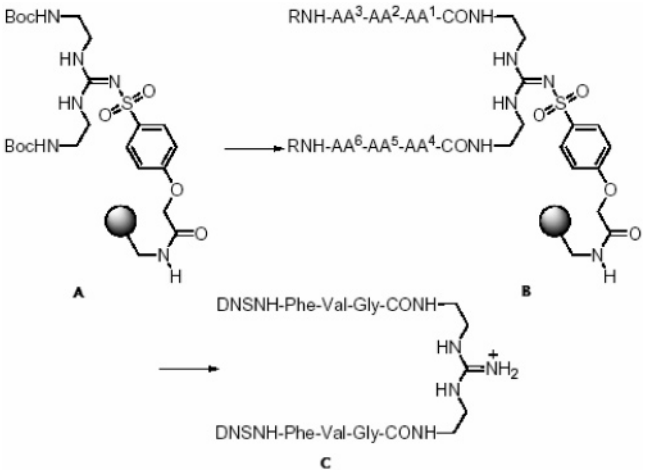


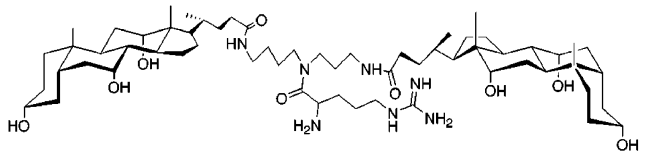
Figure 18. Solid-phase synthetic scheme of flexible tweezer receptor for the carboxylate terminus of peptides in aqueous media. Reprinted with permission from ref 152. Copyright 2002 Wiley-VCH.

Other work, also originating from solid-phase peptide synthesis, has come from Fernández-Carneado, Giralt, and co-workers.[151] Using a novel fluorescence tagging procedure, they synthesized a set of amphipathic peptides of the form $(VXLPP)_n$, where X = histidine, arginine, and lysine, and evaluated their ability to cross cell membranes. Interaction in this setting was between the basic residues in the peptides and the free oxygens of the phosphate diester in the outer part of the cell membrane. The results showed that the arginine-containing peptides exhibited a much higher transport due to their greater interaction with the phosphate diester. Plans are to extend this knowledge to the development of proline-rich peptides for use in their transport of plasmid DNA.

The use of guanidine-based receptors, following their study as self-assembling systems, has been shown in the case of binding peptide sequences to be important in the study of a variety of neurodegenerative diseases. Conditions such as scrapie, bovine spongiform encephalopathy (BSE), Creutzfield-Jakob disease, and Alzheimer's disease are believed to be a result of the self-aggregation of various peptides in the brains of animals and humans.[149] As a model for the protein that is responsible for Alzheimer's disease, the lipophilic peptide sequence Val-Val-isoleucine-Ala was focused upon for development of tripeptide sequences incorporating guanidine functionality that would bind to the peptide of interest, preventing self-aggregation. Receptor libraries of 125 and 512 members were synthesized and evaluated for binding in aqueous media.[149,150] Major conclusions to the study included the role of the side chains and the ability for different variations on the receptor to tune relative binding selectivity over a range of up to 2 orders of magnitude. Since all library members investigated included a guanidinium end-group moiety, this emphasized the importance of secondary interactions for efficient binding, especially under physiological conditions. Binding of a fluorescent- or UV-tagged receptor to the beads functionalized with the tripeptide could also be used to assess affinity for different library members. This approach proved to be an efficient way to measure binding affinity for a modest-sized library on solid supports. The work of Schmuck and co-workers is an excellent application of combinatorial guanidine technology and should continue to add to the knowledge base for a wider array of use of this structure in biochemical and biomimetic applications.

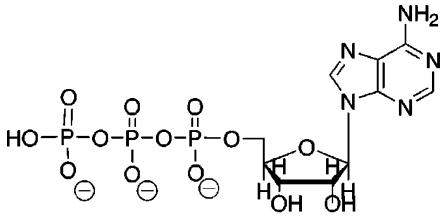
Kilburn and co-workers have reported their version of molecular tweezers based on recognition by a guanidinium headgroup and stabilizing by amino acid sidearm groups.[152] Despite the inherent flexibility of the tweezer sidearms, this receptor scheme has proven highly selective for certain peptide sequences in both nonpolar and aqueous systems. The guanidinium group binds to a free carboxylate group in peptides, and the sidearms provide the bulk recognition media to influence selectivity. Figure 18 shows the steps involved in the generalized solid-phase synthetic scheme for the peptide-based flexible tweezer receptor. Use of the guanidinium headgroup makes the receptor particularly amenable to complex formation in aqueous systems. Peptide receptors in general have been recently reviewed by Peczuh and Hamilton.[153] Original work on tweezer receptors was reported by Whitlock and Chen.[154]

Another variation on the tweezer or cleft motif has been named a molecular "umbrella" receptor by Regen and co-workers.[155] Based on the combination of three biogenic precursors (cholic acid, spermidine, and arginine), this receptor, **17**, has been shown to



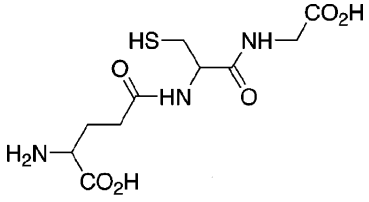
17

transport ATP, **18**, a phosphorylated biomolecule,



18

across phospholipids bilayers at the same time showing selectivity against glutathione, **19**, transport,



19

which exhibits carboxyl functionality. The results of this work suggest that umbrella-type receptors may be applicable as efficient and selective drug delivery devices.

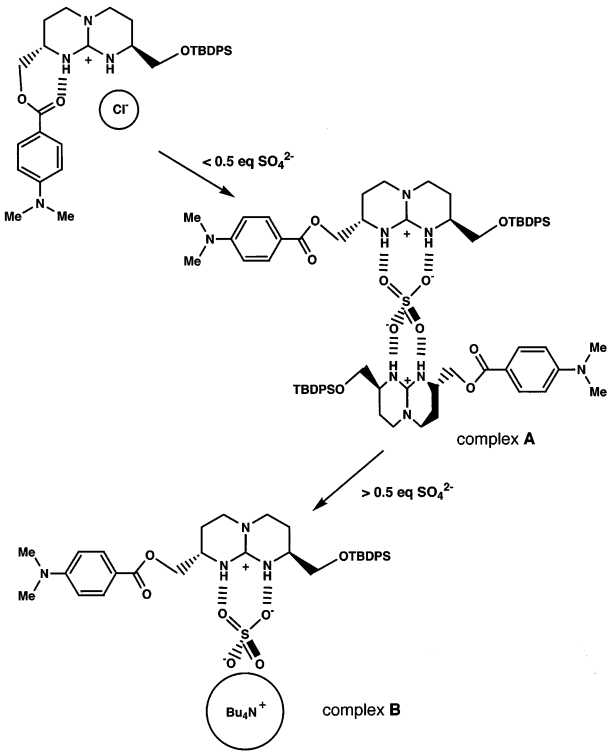
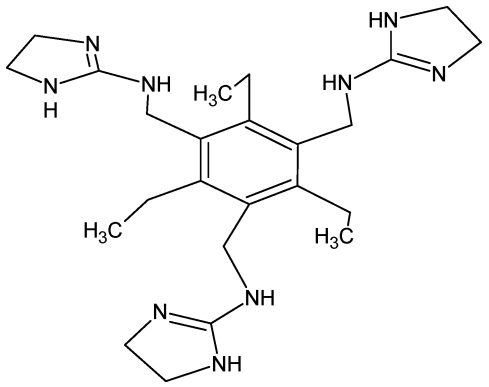


Figure 19. Chiral guanidinium receptor by Kobiro and Inoue binds sulfate. Reprinted with permission from ref 157. Copyright 2003 American Chemical Society.

Another novel recognition system reported by Liu and co-workers is based on a bisguanidinium receptor with a calixarene headgroup that is shown to form monolayers at an air−water interface in the presence of 5′-AMP^{-} and 5′-GMP^{2-}.[156] This biphasic receptor, 5,11,17,23-tetra-*tert*-butyl-25,27-bis-(2-guanidinoethoxy)-26,28-dihydroxy calix[4]arene hydrochloride (see reference for structure), was studied with film balance and relaxation experiments using Langmuir−Blodgett films. Binding constants estimated for 5′-AMP^{-} and 5′-GMP^{2-} using a barrage of spectroscopic techniques (circular dichroism (CD), FT-IR, UV, and X-ray photoelectron spectroscopy (XPS)) were 1×10^6 (in a 1:1 molar ratio) and 6×10^5 (in a 2:1 molar ratio), respectively.

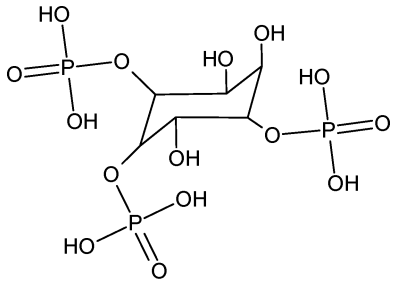
A recent report from Kobiro and Inoue introduced a new chiral probe for sulfate anion based on a chiral guanidinium−*p*-dimethylaminobenzoate conjugate.[157] The receptor incorporates a UV chromophore for spectroscopic analysis, which, studied in a variety of hydrogen bonding solvents, showed a nearly constant λ_{max} value. Chiral information was elucidated from concentration studies related to the observed stoichiometries of the sulfate complexes. UV, CD, fluorescence, and NMR spectroscopy were all used to characterize the system. The binding of sulfate anion by this receptor is shown in Figure 19. Prior to this work, Inoue and Anslyn collaborated to study the binding of citrate to a synthetic trisguanidinium receptor, **20**.[158] This system exhibited different thermodynamic behavior depending on the number of guest molecules associated with it. In a 1:1 complex, association was due to electrostatic attraction (enthalpically favorable); however, as higher aggregates (2:1) were formed, the authors deduced that solvent



20

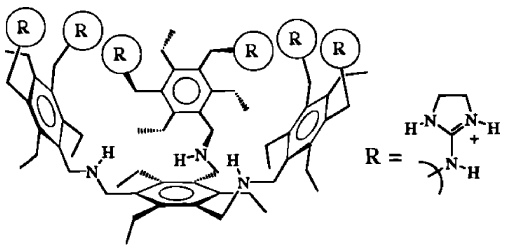
liberation, an entropically driven force, was responsible for the observed binding. Formation of aggregates increased as the concentration of citrate molecules decreased. Binding through hydrophobic interactions were ruled out by addition of small amounts (10%) of methanol, which had no effect on the binding. This system serves as an example for building supramolecular systems at low concentration with a self-controlled affinity toward a particular substrate.

In more recent work by Anslyn and co-workers, a synthetic chemosensor array has been reported based on an array of guanidinium recognition sites for phosphates, specifically for inositol triphosphate (IP3), **21**.[159] Though the structures of natural IP3



21

receptors have not been fully elucidated, it is known that modifying arginine with suitable reagents (such as *p*-hydroxyphenylglyoxal) can block IP3 binding.[160] This suggests that interaction with guanidinium groups is essential to the natural activity of IP3. The synthetic receptor **22** was thus designed using steric

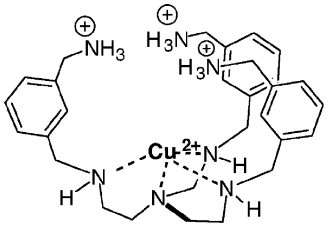


22

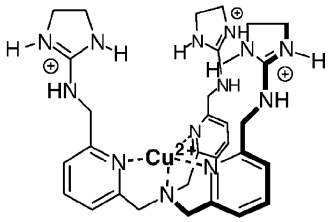
gearing to orient six guanidinium groups toward the interior cavity of the host molecule. Among the

phosphate-based compounds tested (IP3, adenosine triphosphate (ATP), inositol, and phytic acid), only phytic acid (inositol hexaphosphate, IP6), a compound containing six phosphate groups, bound more strongly to the synthetic receptor than IP3. When guanidinium sites on the receptor were substituted with ammonium functional units, all guests showed equal or lower binding than prior to substitution. Also, the specificity of binding to the ammonium-based receptor was much diminished compared to that of the guanidinium-based receptor. By addition of NaCl, a salt effect was used to show that the dominant binding pattern to the ammonium receptor was largely electrostatic in nature, whereas the guanidinium-based receptor showed a higher robustness with increasing salt concentration (an indication of geometrically driven direction-oriented binding). These results indicate that guanidinium-based receptors show less nonspecific binding interactions compared with ammonium-based receptors, thereby allowing the shape of the receptor and the directed interaction groups to guide the overall recognition process.

Other current work by Anslyn and co-workers has focused on the development of metalloreceptors with a high affinity for phosphate. A good example of this work is shown with their comparison of Cu^{2+} metalloreceptors functionalized with ammonium, **23**, and guanidinium groups, **24**, operating in water.[161] Both



23



24

receptors show a high affinity for phosphate. The majority of the binding affinity is attributed to the metal ion, whereas the side chain basic residues serve to tune the interactions for stability and selectivity. Binding affinities were determined by UV/vis and shown to be remarkably similar between the two receptors for phosphate. The major difference between their associations with phosphate was elucidated by evaluating the enthalpy/entropy contributions using ITC. Due to the rigidity, poor solvation, and higher preorganization of the guanidinium moieties in these scaffolds, binding was shown to be mainly enthalpy-driven. In contrast, the more highly solvated and flexible ammonium side chains created a binding event that was determined to be entropy-driven. Other anionic guests were studied as well.

Table 5. Binding Affinities for Different Anions by 23 and 24 Determined by UV/vis Titriations[161]

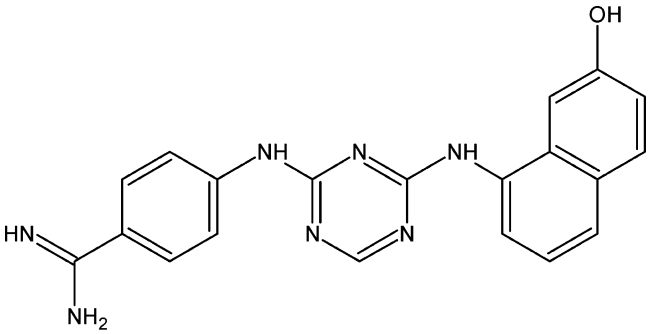
anion	binding constant **23** (M^{-1})	binding constant **24** (M^{-1})
HPO_4^{2-}	2.5×10^4	1.5×10^4
$HAsO_4^{2-}$	2.5×10^4	1.7×10^4
ReO_4^-	2.0×10^3	<100
AcO^-	<900	<100
NO_3^-	<20	<100
HCO_3^-	*a*	<100
Cl^-	*a*	<100

[a] Not determined.

The results from these experiments are shown in Table 5. This work serves as a nice comparison between the different effects of the basic residues in a supported binding role, rather than as the main site for interaction.

Affinity chromatography is currently a popular niche for development, due to its incorporation in many purification strategies. Investigation of systems incorporating arginine or guanidinium moieties are an important part of this development. In an attempt to keep up with the high-throughput world, affinity chromatography systems have also recently improved efficiency using combinatorial techniques. Fassina and co-workers have provided a review of ligand design and synthesis for this purpose.[162] Outside of their more current work on chimeric affinity assemblies, they reference their own work on a combinatorial approach to design a ligand able to compete with the interaction between protein A and biotinylated immunoglobulins.[163] The library, composed of 5832 randomized tripeptide tetramers, was produced by solid-phase synthesis. The first screening identified a sublibrary of the most active compounds with an arginine residue at the N-terminus (60% inhibition). Subsequent cycles led to the most active multimer, (Arg-Thr-Tyr)$_4$−(Lys)$_2$−Lys-Gly or TG19318, an effective Protein A mimetic. Though an effective inhibitor, the authors go on to explain in their more recent review how affinity ligands based on peptides can be unstable and that cyclized analogues would provide more robust designs. Their work however provides one such example of the explicit incorporation of arginine residues into affinity ligands, due to their high activity in biomimetic systems.

Another interesting example of the exploitation of biomimetic affinity ligands based on the specificity of arginine in biological systems is offered by Lowe and co-workers.[164,165] Kallikrein, a trypsin-like protease, acts on kininogen as its natural substrate. In doing so, kallikrein shows preference for binding to a dipeptide phenylalanine−arginine residue structure in the substrate. A ligand based on the phenylalanine−arginine dipeptide template was designed to affinity purify kallikrein. By orienting phenethylamine and *p*-aminobenzamidine on a triazine scaffold, affinity chromatography with the ligand **25** resulted in 110-fold purified kallikrein in good yield. This is due to the similar shape, size, and polar/apolar characteristics as the parent dipeptide. The difference is that the purified ligand is ultrastable and sterilizable relative to the phenylalanine−arginine moiety, since it contains no fissile bonds.



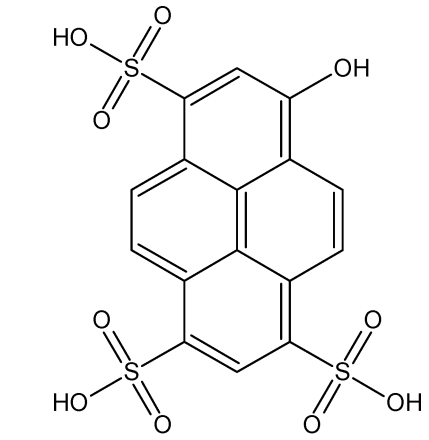
25

Though this ligand design is based on a para-substituted benzamidine unit, the ability to mimic a natural guanidinium group interaction through creative synthetic design (in this case, a slight variation incorporated into a rigid scaffold) should serve as an excellent example of how modern technology can be used to probe biological interactions.

In a different approach to achieving selective recognition, arginine residues have recently been arranged in arrays to effect stochastic sensing. Work with this new concept has been performed chiefly by Bayley and co-workers.[166,167] In stochastic sensing, single-molecule qualitative and quantitative detection can be achieved by functionalizing a nanopore and measuring current signatures as analytes of interest pass through the pore and reversibly bind to the receptor placed in the pore. Qualitatively, each analyte can be differentiated based on its electrical current signature when bound. Quantitatively, the number of these binding events (strength of signal) can be related to the concentration of the analyte present. The use of a pore functionalized with 14 arginine residues was shown to be a sensitive sensor element for the detection of nanomolar concentrations of IP3 in the presence of magnesium−adenosine triphosphate.[168] Stochastic sensing is likely to find wide application in systems varying from the detection of warfare agents to high-throughput screening of protein structures soon.

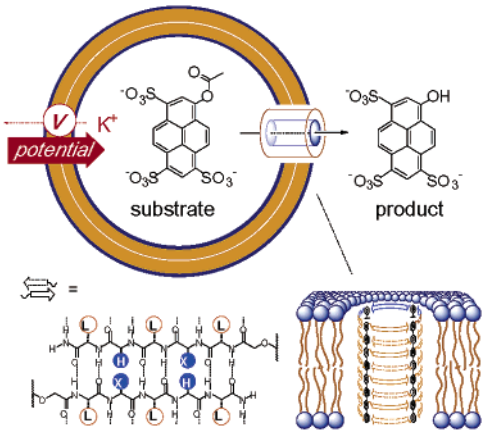
In a similar approach, reports have been made by Matile and co-workers of the use of arginine-functionalized synthetic ion channels. A study preliminary to incorporation of arginine in the ion channels was performed in 1997.[169] This report laid the groundwork for the possibility of creating artificial biomimetic ion channels and used egg yolk phosphatidylcholine (EYPC) small unilamellar vesicles containing entrapped 8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS) **26** to study pH-dependent binding with channels formed by substituted oligo(*p*-phenylene)s. The results showed that it was indeed possible that synthetic rigid-rod molecules carrying hydrophilic substitutions could facilitate ion transport across hydrophobic lipid bilayers.

The use of these rigid-rod ionophores as mimics for cell membrane recognition by antibiotics was an impetus for driving the research forward. Synthetic characteristics were defined to incorporate (a) a rigid scaffold to minimize available conformations, (b) a unimolecular pathway for ion conduction for simplicity of monitoring the function of the channel, (c) an extended fluorophore for monitoring the structure, (d) a variable axial rod dipole, and (e) variable terminal charges for the purpose of creating axial symmetry.[170] This study also established some important differences between the nature of two structurally similar polypeptides: melittin, a known toxin; and magainin 2, an antibiotic. Though both are α-helical peptides that can mediate ion transport, they differ in axial electrostatic asymmetry and their specific cationic motif (magainin relies on a KK motif, whereas melittin contains a KRKR motif). HPTS was again used as an internal pH-sensitive fluorescent probe. Other interesting data was also collected through ESI-MS in the form of nonspecific adduct formation. The mass spectra for these polypeptides showed their ability to bind multiple cations, thus hinting at their ability to mediate ion transfer by the same "cation-hopping" or "wire" mechanism.



26

Perhaps most interesting, in coherence with this review, is the work by Matile's group in the incorporation of arginine−histidine dyads in their synthetic rigid-rod β-barrel ion channels.[171] The internal arginines were introduced to recognize organic ions



synthetic catalytic pore **SCP$_1$**: X = R
synthetic multifunctional pore **SMP$_1$**: X = H

Figure 20. SCP created with an arginine−histidine internal dyad structure for the conversion of AcPTS to HPTS. Reprinted with permission from ref 172. Copyright 2003 American Chemical Society.

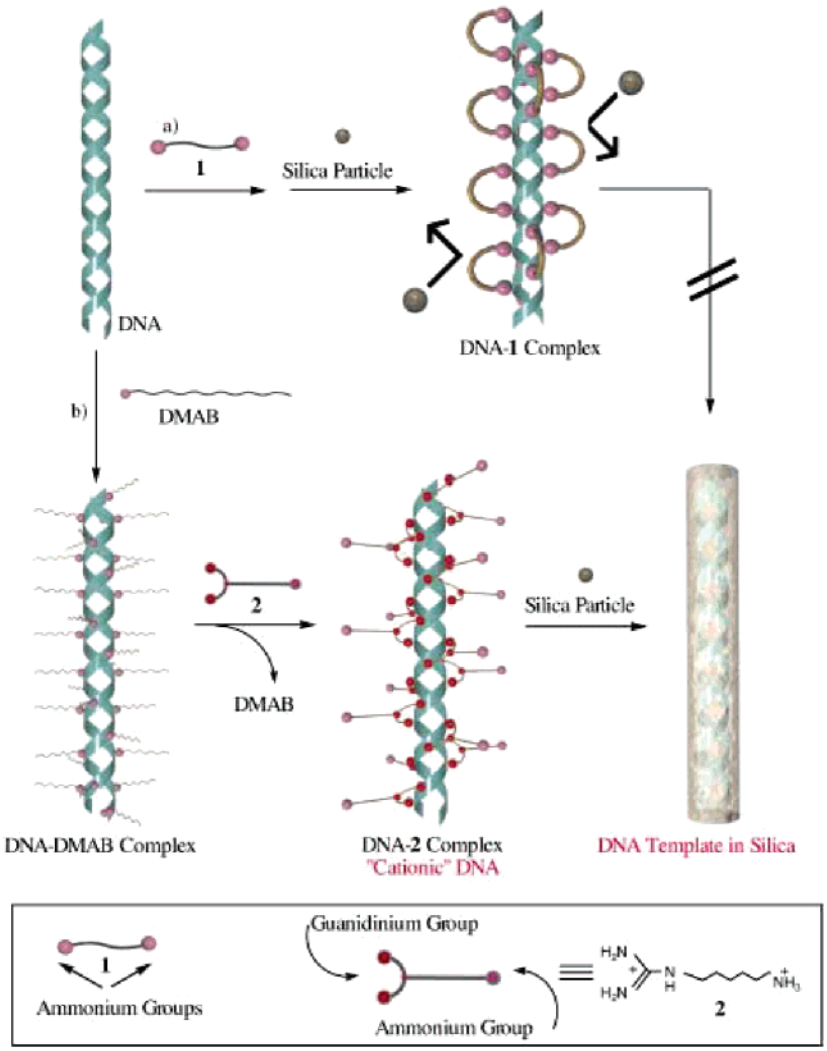


Figure 21. Sol-gel templating of plasmid DNA using a guanidinium surfactant. The figure demonstrates the inability of ammonium surfactant to provide a free positive charge for reaction. Reprinted with permission from ref 178. Copyright 2004 Wiley-VCH.

such as α-helical poly(L-glutamic acid) (PLGA) passing through the polarized bilayers membrane pores. PLGA acts as a good α-helix model. The helicity of the model compound changes as a function of pH, which can be used as a tool to distinguish the ion channel recognition of PLGA as a function of α-helical versus anionic content. These rigid rods were shown to be easily programmable with the incorporation of arginine; however, the channel appeared to recognize PLGA based on its anionic nature rather than α-helical content. Even more recently, the arginine−histidine synthetic catalytic pore (SCP) was studied with a sulfonate as a model compound.[172] AcPTS (8-acetoxy-1,3,6-trisulfonate) was loaded into EYPC to study its interaction with the SCP. SCPs are interesting because of their ability to generate products where the rate of substrate binding and substrate or product release or both can be controlled and manipulated by, for example, membrane potentials and concentration gradients. Figure 20 shows the conversion of AcPTS to HPTS by the arginine−histidine SCP. These results, and other similar studies,[173,174] provide encouraging results in the application of synthetic arginine−anion recognition for biomimetic purposes.

In 2003 and 2004, Matile and co-workers have continued their development of the internal arginine−histidine dyad pores and have shown a number of impressive advancements.[175,176] With the use of phosphate to block the strong cationic activity of the arginine component, pH-dependent selectivity for cations and anions through complex formation with the histidine residue afforded unique function of the pores. This result suggested that "molecular recognition within multifunctional pores can be regulated by the selectivity of molecular translocation across the same pore." In this case, the effect is provided by the complementary arginine−phosphate interaction. It is most interesting perhaps that this polycationic pore can be used for cation-selective transport, as well as being a small anion channel (~3.3 Å) that can host large anionic guests (HPTS). Also, in comparison to other internal functional arrangements, these pores show an extended single-channel lifetime of up to 20 s. In a somewhat related study, using thermal CD experiments and *p*-octiphenyl rods carrying complementary tripeptide strands functionalized with arginine residues, it was shown that annealing of the rods with temperature lead to an inversion of supramolecular chirality.[177] This took

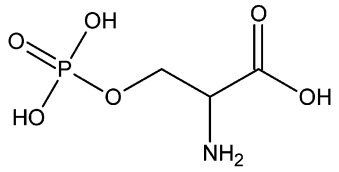
place because the high temperature allowed exchange of phosphate from the guanidinium residue by glutamate residues in a programmed assembly environment. The authors describe the phenomenon as "Arginine Magic", a fitting designation. This demonstration of irreversible "noncovalent" chirality switching is interesting because it initiates thoughts on the development of reversible systems that might be used for chiral sensors. Supramolecular chirality can be designed to be responsive to lower activation energies than those required to break bonds. In other words, chiral sensors could conceivably be developed that are responsive to pH, ionic strength, light, heat, electron transfer, or molecular recognition events.

It is not difficult to see the potential for the development of novel materials through exploiting the uniqueness of "Arginine Magic." From affinity chromatography to the development of synthetic ion channels, the guanidinium motif is an ideal functional unit to employ for creating excellent binding and selectivity properties. One last example of the use of synthetic guanidinium systems here will suffice to indicate one of the many promising directions of materials development. Recent work by Shinkai and co-workers has shown the ability to template DNA helices with silica using guanidinium surfactants.[178] The surfactants, which interact with the negatively charged phosphate moieties on DNA, solubilize the DNA and act as active template for a sol−gel reaction. In this reaction, guanidinium surfactants replace ammonium surfactants leaving a positive charge available to initiate reaction. The general scheme of this reaction is shown in Figure 21. Here, a tubule material is shown; however, the authors point out that it is possible to create a large variety of different shapes and sizes by affecting the arrangement of the plasmid DNA prior to templating.

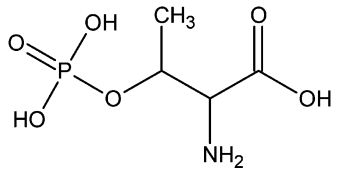
To summarize the literature presented in this section, we must emphasize the incompleteness of this information. A complete review would be much too lengthy. Much work has been performed in the past to lay the groundwork for these more recent applications that we have highlighted. With the number of publications addressing this subject, it is indeed an important subject that has not been fully exploited. Arginine-residue- and guanidinium-based functional units will continue to be incorporated in unique and innovative ways to take advantage of the powerful recognition ability these groups and arrangements afford. The current work by Matile and co-workers on the development of functionalized biomimetic ion channels is only one such, though exciting, example. The strong hydrogen-bonding nature will allow these receptors a strong niche in aqueous-based environments and no doubt open the door to some new and exciting biochemical- and pharmaceutical-based agents.

3.3. Biologically-Relevant Phosph[on]ate-Based Interactions with Guanidinium

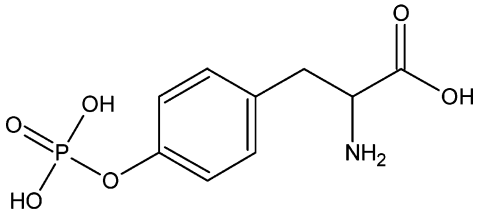
Phosphorylation is one of the most common and physiologically important posttranslational modifications in proteins and peptides. The process plays a crucial role in a number of biochemical interactions that control normal cell operation.[179] As an example, in biological catalysis it has been hypothesized that the participation of the phosphate group in an intramolecular proton abstraction step is important to the mechanisms of aldolase,[180] aspartate transcarbolymase,[181] and dehydroquinate synthase.[182,183] In many schemes, individual phosphate anions are transported into cells and organelles by specialized carrier proteins. The binding of the anion to the carrier molecules and differences in selectivity for cognate binding are believed to stem mainly from differences in hydrogen-bonding architectures.[6] In the case of binding between peptides and proteins, often it is the interaction between a phosphorylated residue and a neighboring basic amino acid residue that creates a favorable interaction for complex formation. Although this review focuses mainly on the role of the phosphonate group in interacting systems with guanidine, in biological systems it is the phosphate group (found largely as a product of phosphorylated amino acid residues, such as phosphoserine (pSer) **27**, phosphothreonene (pThr) **28**, and phosphotyrosine (pTyr) **29**, and other important phosphorylated



27



28



29

species, such as IP3) that is encountered frequently in the literature and will largely be discussed below.

Much of the biological implications of interactions between phosphorus-containing groups, such as pTyr, and arginine centers in neighboring molecules have been elucidated in the past 10−15 years. A system that has been shown to be favorable for studying this type of interaction is the binding of SH2 domain proteins with phosphopeptides. SH2 domains are pTyr-binding modules found in a variety of important signal-transducing molecules and, with protein tyrosine kinases and their cellular substrates, mediate protein interactions.[184] The SH2 domain, a series of approximately 100 amino acids in length, is lined by

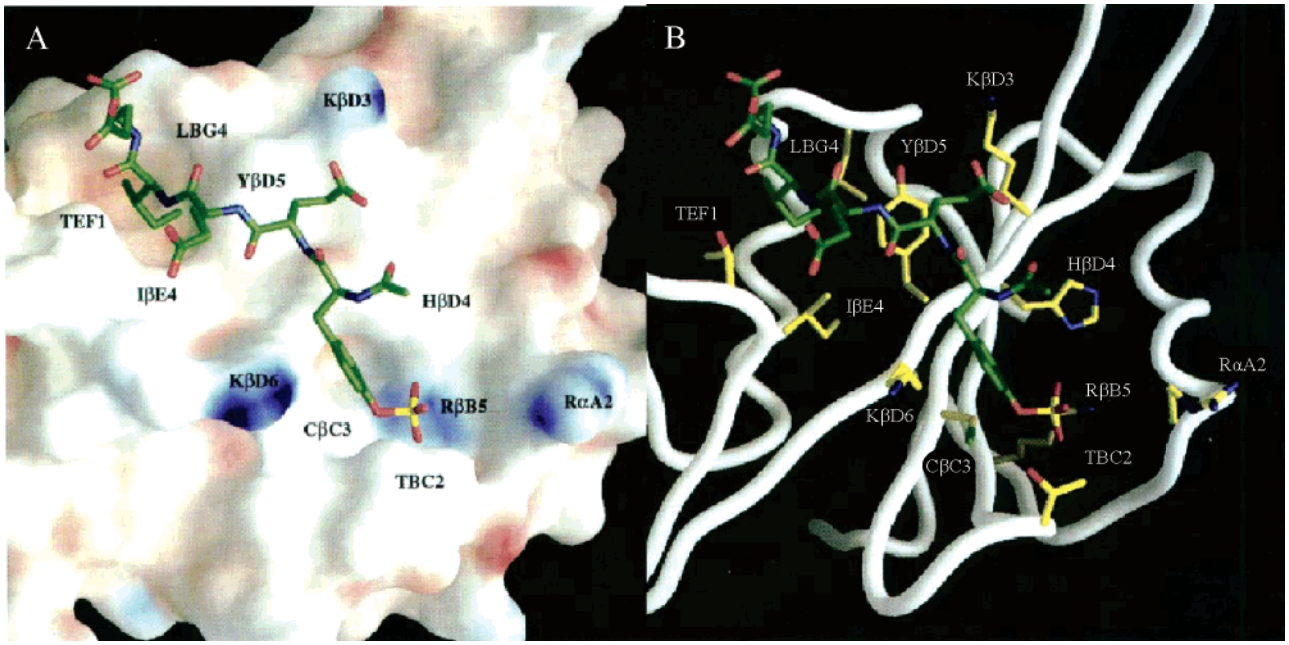


Figure 22. Panel A shows the electric surface representation of the binding of Ac-pYEEIE-OH, a phosphorylated pentapeptide, with the src SH2 binding domain. Red indicates regions of negative charge, whereas blue indicates regions of positive charge. Panel B is the same view as in that in panel A except with the surface removed to display the orientations of the protein residue side chains. Reprinted with permission from ref 187. Copyright 1995 American Chemical Society.

basic and hydrogen bond-donating amino acid residues that effect binding to pTyr (specifically, three histidine, eight lysine, eight arginine, and the N-terminal amine).[179] Phosphorylation at particular tyrosine residues in proteins adjacent to SH2 serves as an on/off switch for binding. These interactions are correlated with a plethora of experimental data, including X-ray data, NMR, and MS-based studies. X-ray structures clearly show ionic interactions between the phosphate group on pTyr and arginine, as well as between arginine and the π cloud on tyrosine.[185]

The first group to publish X-ray structures of this binding phenomenon was Kuriyan and co-workers in 1992.[186] The detailed and complex structural alignment of the sequences of the SH2 domains can be found in the cited work. The three-dimensional structures demonstrated that the phosphate group of pTyr is tightly bound in a pocket by multiple cationic interactions with arginine and lysine residues, as well as additional hydrogen-bonding interactions through residues such as glutamine, serine, and threonine. The structure of the SH2 domain is highly conserved and thought to be important for the specific recognition of pTyr, controlling its activity. Close contacts (3.1 Å) in the structures indicate the prevalence of amino−aromatic interactions ("amino" generalized here to mean predominantly interaction through a guanidinium cationic group, that is, a cation−π interaction) for stabilizing the interaction. The total structure was likened to a right-hand grasping the pTyr residue with the fingers corresponding to the sheet structure of the SH2 domain. An obvious distinction between the binding of pTyr and other phosphorylated residues, such as phosphoserine (pSer) and phosphothreonine (pThr), could also be made. In comparison to pTyr, pSer and pThr lack the side chain length necessary for reaching into the binding pocket and interacting with the buried arginine residue. In a study by Xu and co-workers, a comparison between the use of NMR and X-ray techniques provides a comparison between the solution and crystal state of the SH2 domain complexed with a phosphorylated pentapeptide.[187] In solution, the arginine residue does not show an appropriate signal for the interaction observed in the solid state. This was attributed to the rapid proton exchange of arginine, specifically the guanidinyl group, with the solvent. In comparison of NMR and X-ray data, secondary and tertiary structures in the solution state appear to be conserved in the solid state. Figure 22 shows an electrostatic surface representation of the phosphorylated pentapeptide interacting with the SH2 domain. The pTyr unit on the pentapeptide is shown to be in close proximity with an arginine residue on the protein backbone. The interaction site is likened to a two-pronged plug (the pTyr and the isoleucine residue three units from it on the ligand) engaging a two-holed socket (the conserved binding secondary structure of the SH2 domain incorporating the buried arginine residue). Very recently, Heck and co-workers have used ESI with tandem MS and CID to characterize the binding of synthetic inhibitors of spleen tyrosine kinase, targeted toward SH2 domains.[188] Molecular recognition between the pTyr motifs and catalytic and noncatalytic proteins was shown to be fully characterizable by fragmentation in the positive ionization mode.

A computational evaluation (ab initio and semiempirical) of the interaction of phosphorylated amino acids with arginine and lysine was used to assess binding differences between these two basic residues.[189] Methyl, ethyl, and phenyl phosphate were used as prototypes to approximate pSer, pThr, and pTyr, respectively. In general, greater interaction energy was found for interactions with lysine than

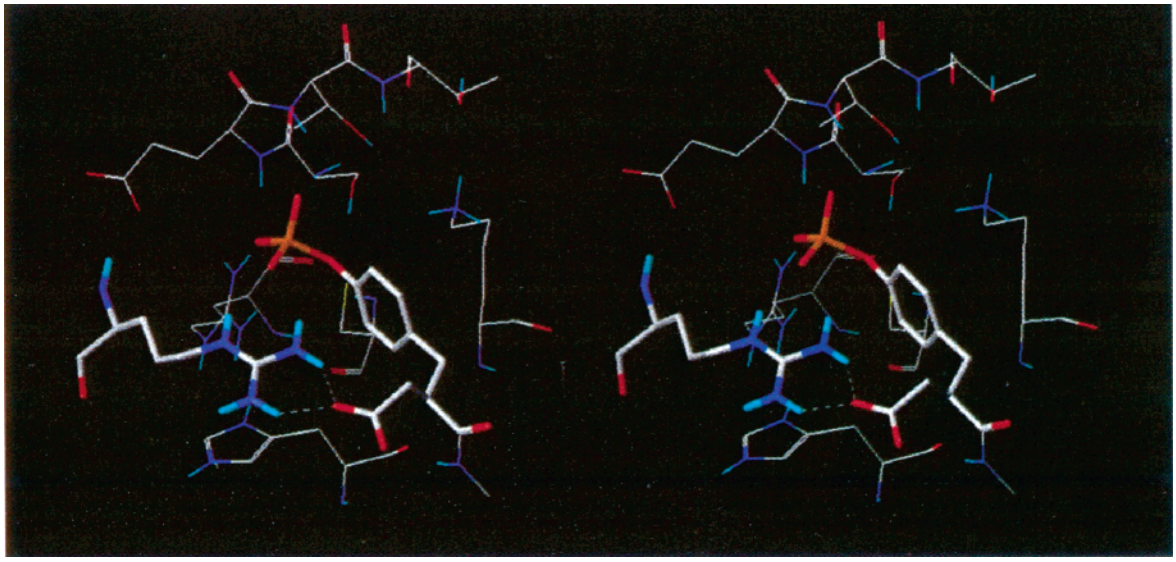


Figure 23. Modeled interactions between Src SH2 domain and one in a series of Ac-pTyr (α-carboxymethyl pTyr) peptide mimics. Hydrogen binding between the carboxyl unit and the Arg-155 residue is shown in white. The phosphoryl group is also shown in the vicinity of the guanidinium group. Reprinted with permission from ref 192. Copyright 1996 Elsevier Science.

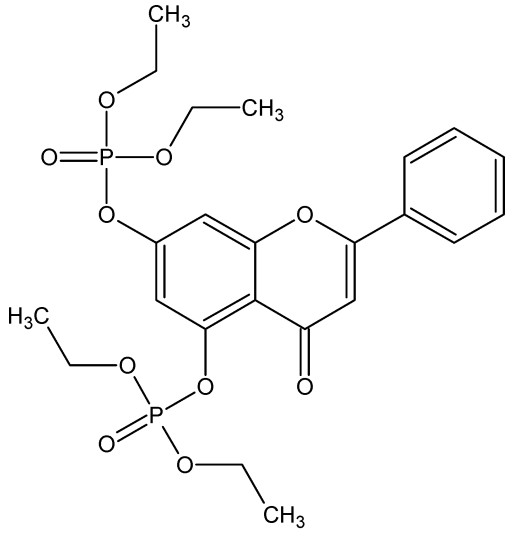
with arginine, although arginine is known to be more strongly basic. These interactions were assessed with and without the addition of solvation models. Reasonable results were recorded with the use of bulk solvent system modeling. In the end, the differences in binding, specifically the increased interaction found with lysine, were attributed to the shortcoming of the basis sets used for accurately assessing hydrogen-bonding interactions. Lysine was simulated as a more compact point charge and, therefore, with its higher charge density, showed a stronger electrostatic component than did guanidinium with the phosphorylated amino acids. This is a common misconception arrived at when calculating binding energies using algorithms that focus only on the electrostatic component of interaction. This work emphasizes the need to consider the directional and specific nature of hydrogen-bonding contributions, a shortcoming of many current theoretical treatments that is now beginning to be addressed. Comparing the relative binding of the phosphorylated species shows a stronger interaction for the alkyl versus the aryl phosphates. A similar study comparing the binding of formate and methyl phosphate showed formate to bind more strongly relative to the phosphate with methyl guanidinium.[190] Overall, these theoretical investigations provide a basis for comparison with experimental data when the formation of salt bridge structures between guanidinium and anions are suspected. Still, such comparisons should be made with care.

More recently, work on studying the binding by the SH2 domain protein was performed through mass spectrometric means. In a study by Loo and coworkers, the relative abundance of multiply charged ions generated by ESI-MS for the Src SH2 domain protein with several phosphorylated and nonphosphorylated peptides was compared.[179] ESI-MS has found a unique niche in biochemical analysis with its ability to form multiply charged ions of large biomolecules.[62] This achievement earned John Fenn a share of the Nobel Prize in 2002. Correlations can often be drawn between the number of charges observed in a mass spectrum and the number of basic amino acid residues that are unbound and available for charging. Loo calculated dissociation constants and compared them to derived solution-phase values, showing there was a decent correlation. This indicated the ability of ESI-MS to measure solution-phase binding through gas-phase processes in this system (an ongoing debate concerning the capabilities of ESI-MS).[191] Noncovalent complexes were observed for particularly strong interacting molecule pairs and even at 7:1 peptide/Src SH2 concentration mixtures, the relative abundance for noncovalent complexes between nonphosphorylated peptides and the protein were very low. Competitive binding experiments with complex mixtures containing the phosphorylated peptides were also reported to be able to distinguish between relative affinities of stereoisomeric peptides (though this was not the focus of the research and not tested rigorously). Chiral information was conjectured to be available because the binding between the two molecules incorporates multiple interaction types, including hydrophobic, electrostatic, and hydrogen-bonding modes.

To specifically mimic the interaction between pTyr and the pp60src SH2 domain, a series of phosphorylated penta- and tripeptides were synthesized. Generation of three-dimensional structures showed that the phosphate group is tightly bound within a binding pocket by multiple cationic and hydrogen-bonding interactions. Specifically, the Arg-155 unit is shown to simultaneously recognize both the phosphate group and the aromatic ring of the pTyr residue.[179] This work had followed a similar investigation by Shahripour et al.[165] Incorporation of various pTyr mimetics into peptides included the study of the effect of enantiomeric phosphorylated tri- and pentapeptide isomers to binding in the SH2 domain. In some cases, inversion of chirality (exchange of D for L amino acids) in the mimetic species led to a

substantial decrease in binding. The structure−activity relationships generated with these mimetics approaches were in agreement with the previously acquired three-dimensional X-ray structures. Figure 23 shows a modeled complex between an Ac-pTyr mimic and Src SH2. The hydrogen bond with Arg-155 is shown in white. An important point impressed in this work was the apparent contribution by the secondary structure remote from the chiral center to the overall binding affinity observed. Also during this time period, Berman and co-workers probed the activity of Src SH2 binding to phosphopeptides. By replacing pTyr moieties with α-dicarbonyl analogues, they recorded a decrease in effective binding of the systems.[106] It appears to be clear that the preference of binding between the SH2 domain and ligands is aided in a significant way by the phosphate−arginine binding motif. The binding also appears to incorporate a large degree of selectivity with respect to the primary electrostatic or hydrogen bonding interaction site but also with the steric nature of the secondary structure.

Also recently reporting information on the specificity for the interaction of phosphorylated molecules with basic peptides and proteins were Zhao and co-workers.[193] They studied the interaction of a flavanoid, chrysin, which was phosphorylated through a modified Atherton−Todd reaction. The structure of the molecule **30** was determined by NMR, ESI-



30

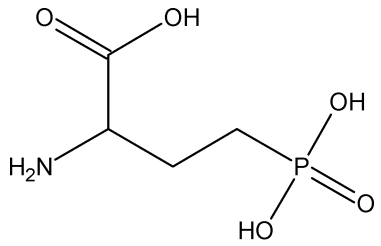
MS/MS, and X-ray data. Using ESI, it was shown that the effect of phosphorylation was a greatly increased interaction with basic proteins, such as lysozyme, myoglobin, bovine insulin, and cytochrome *c*. To correlate the ESI results to solution-phase behavior, fluorescence spectroscopy was also employed. Results were shown to be consistent with interaction in the solution phase.

It is not uncommon for interactions between phosphorylated and basic side chains of amino acid residues to be involved in the formation of α-helical structures in proteins and peptides. A study by Chenault and Liehr asserts that the charge−charge

Table 6. Phosphorylated Peptide Structures Used to Study the α-Helix-Forming Ability by pSer versus pAbu and the Spacing with Arginine in the Peptide[194]

no.	structure	phos/Arg distance
1	Ac-YEAAAK(pSer)AARAEAAAKA-NH_2	$i/(i+3)$
2	Ac-YEAAAK(pSer)AAAREAAAKA-NH_2	$i/(i+4)$
3	Ac-YEAAAK(pAbu)AARAEAAAKA-NH_2	$i/(i+3)$
4	Ac-YEAAAK(pAbu)AAAREAAAKA-NH_2	$i/(i+4)$

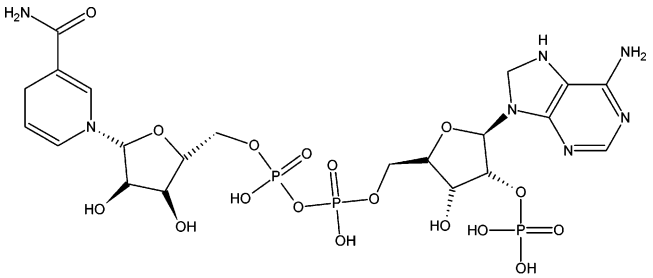
and hydrogen-bonding interactions between pSer or α-amino-γ-phosphonobutyric acid (pAbu), **31**, groups



31

and arginine residues helps to stabilize the formation of a helical structure in peptides.[194] The pSer and pAbu are considered to be largely similar in promoting this interaction, provided the arginine residues are properly spaced to align the interacting groups. Table 6 shows the peptidic structures compared in this experiment. Comparison of the α-helix-forming propensity of these phosphopeptides shows pAbu to be a good mimic of pSer, even increasing the α-helical content of the peptide relative to pSer in the case where an arginine residue was spaced four units from the phosphorylation site ($i/(i+4)$). The α-helical content of the peptides was determined by circular dichroism.

To illustrate the formation of multiple interactions with arginine in the binding of a specific structure, McKeever et al. showed that the incorporation of arginine in 17β-hydroxysteroid dehydrogenase allowed for multipoint interactions in the binding of NADPH, **32**.[195] In this system, the guanidinium



32

moiety is able to establish one or two hydrogen bonds with the 2′-phosphate of NADPH. This interaction is stabilized by a π−π stacking interaction between the guanidinium π system and the adenyl heterocycle of NADPH. This study purports that "a single residue within a given constellation of side chains can have a significant impact on the selectivity and avidity of the enzyme for a given cofactor."[195] Comparisons of

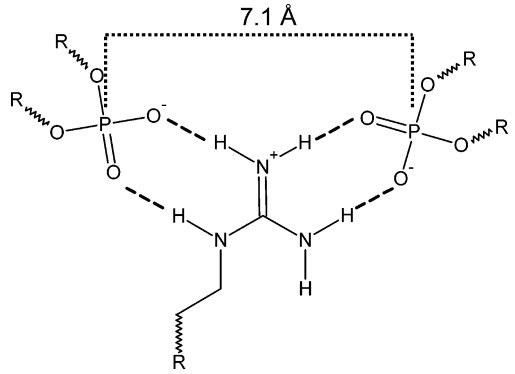


Figure 24. Arrangement of phosphate groups about the guanidine group to form an "arginine fork".[126]

interactions between arginine and lysine in the same placement show the ability for arginine to form a bidentate association, as well as a hydrophobic pocket for the purine ring of adenosine, whereas lysine cannot. Lysine is structurally unable to create a bidentate complex to accommodate binding in the configuration.

Frankel and co-workers performed some early work on the elucidation of contact points between the HIV-1 Tat protein and a bulged region in TAR RNA.[126] The RNA binding region of Tat is a heavily cationic amino acid residue-based unit from residues 49−57 (RKKRRQRRR) in the protein and is well-known to be a natural drug transporter unit through the cell membrane. A mutant protein was used to discover the basic residue in the Tat protein responsible for binding RNA. By systematic replacement of each residue in a string of nine lysine residues with arginine, a single arginine was identified to be responsible for binding and transactivation. By replacement of all nine lysine with arginine residues, a wild-type transactivation was obtained where the activity was 100-fold greater than when the nine lysines were present. Molecular modeling was used to show the binding between the arginine in the Tat protein and two phosphates in the TAR RNA bulge. The authors likened the arrangement to an "arginine fork", where the distance between the phosphate residues was tuned for specific interaction with the guanidine group. Figure 24 details the arrangement and distance between phosphate groups generated by the energy-minimized geometrical configuration calculation. This study parallels the information reported by Sannes-Lowery, Loo, and co-workers, which was presented previously.[111]

Before moving the focus to other interactive systems, a brief mention should be made on the current interest in analysis of the phosphopeptides and phosphorylated structures used, for example, in the previous works mentioned. A heavy focus has been placed on fast and efficient techniques for the characterization of these components in a move to better comply with the need for high-throughput analytical techniques. It has been shown quite recently by Cotter and Woods that the presence of arginine or lysine in a phosphopeptide induces preferential loss of the phosphate group during collision-induced dissociation (CID).[196] The goal was to provide a means for differentiating various phosphorylated species through their unique pathways of dissociation following mass spectrometry. The data support two routes for the dephosphorylation of peptides in the presence of these basic residues. The electron-withdrawing nature of the pTyr aromatic phenyl ring may allow for a partial positive charge on the phosphate that can be easily attacked by a nucleophile. Alternatively, the loss of phosphate may be a result of cation$-\pi$ interactions between the phenyl group of the pTyr and arginine or lysine, which may leave the phosphate susceptible to gas-phase loss. It is clear, however, that the structural arrangement is responsible for this anomaly and that such a pathway does not occur when neither arginine nor lysine are present in the phosphoprotein.

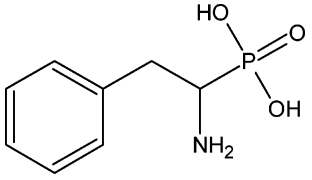
Work in the area of fragmentation of phosphopeptides has also pointed out the fact that phosphate group-specific fragmentation can be a useful tool for the detection of phosphopeptides.[197] It was demonstrated by Qin and co-workers that fragmentation by tandem MS of phosphopeptides can be used to determine the actual site of phosphorylation as well.[198] In comparison of the ion trap fragmentation pathways of pSer-, pTyr-, and pThr-containing peptides, for the latter two, the fragmentation was complicated and dependent on the charge state of the ion, whereas for the former, loss was through a simple fragmentation pattern and found to be independent of charge state. Thus, by this method, discriminating results could be obtained to direct the identification of the site of phosphorylation in a specific phosphorylated peptide by studying the mass spectra obtained from CID. Knapp and co-workers analyzed mono- and diphosphorylated peptides by ESI-tandem MS to elucidate phosphorylation sites in a similar manner.[199]

As has been shown, the majority of biological interactions focused on phosphorus-based oxoanions deal with phosphate rather than phosphonate interactions. However, concepts such as the "arginine fork" motif are expected to correlate well between these geometrically related units. As more focus is placed on phosphonate-based species, as can be seen below with the preponderance of new and interesting arrangements of synthetic molecular recognition phenomena, more and more likenesses will be drawn with biological information involving these related moieties. Particularly interesting could be the rise in interest of affinity chromatography ligands based on phosphonate moieties because of the stronger ionizability and more stable and predictable arrangements.

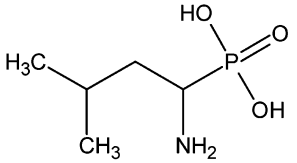
3.4. Design and Use of Phosphonate-Based Receptors for Guanidinium Binding

Overall, synthetic approaches to the design of phosphonate-based receptors for arginine and guanidine moieties are limited in number. It appears the reciprocal recognition system based on guanidine receptors for anions has gotten far more attention in recent years than the reverse. Still, the work that has been performed, whether designed specifically for guanidine recognition or other hydrogen donor groups, is of high quality. A good review addressing the

general synthesis and interaction modes of α-amino- and α-hydroxyphosphonates has been published.[200] Especially interesting is the information on α-aminophosphonates, which can self-assemble through intermolecular interactions to form cyclic complexes. As potential competitors for structurally related carboxylic acids, α-aminophosphonic acids are of high biological and pharmacological significance and interest. Concurrently, they can be assembled to form phosphonylated peptides and, as such, can be considered key common precursors to specialized proteins and nucleotides. Some of the molecules are believed to exhibit neuroactive properties and anticancer, herbicidal, and antibacterial activity. Structures **33** and **34** show representative phosphonic acid



33

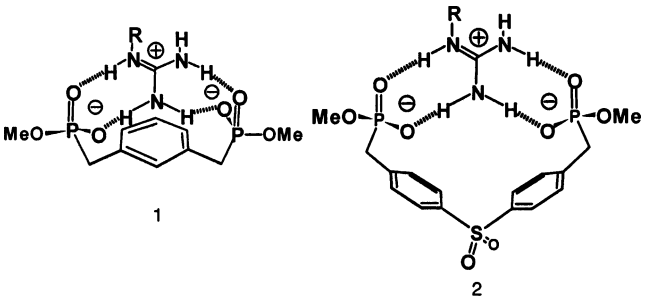


34

analogues of the amino acids phenylalanine (pPhe) and leucine (pLeu), respectively. Other examples are given in publications such as that by Hammerschmidt and co-workers who present a new route to synthesis of α- and β-aminophosphonic acids based on inversion of configuration through a Mitsunobu reaction[201] to prepare a series of phosphonic acid molecules for separation on a chiral anion exchanger stationary phase.[202] Additional synthetic schemes have also come to light as the recent interest in these phosphorylated amino acid mimics has increased.[203,204]

More pioneering work in the field of synthetically designed molecular recognition hosts based on bisphosphonate-based receptors and their affinity for arginine and other basic functional moieties has been performed by Schrader and co-workers.[99] Synthetic receptors containing multiple P=O bonded groups are rare in the literature due to the isomeric complexity of incorporating a stereogenic center in new host molecules. This problem can be bypassed by using achiral phosphonate groups. The work by Schrader et al. demonstrated the usefulness of this synthetic approach and, consequently, created an impressive series of effective receptor molecules that (a) increase specificity and binding affinity through the incorporation of more and varied structural and functional constituents and (b) are able to perform well in competitive media, making the receptors attractive for selective extraction of molecules in a diverse range of environments.

Incorporation of two phosphonate moieties into a tweezer-based receptor allows for appreciable selectivity and strength of binding for guanidine functionalities. Figure 25 shows the geometric complementarity in binding of two of Schrader's bisphosphonate-based receptors with a generic guanidinium residue.[205] The two phosphonate functionalities arrange in a concerted forklike fashion with guanidinium, similar to that shown previously in the work done by Frankel and co-workers.[126] For the sake of comparison, interactions between a series of basic ammonium-based functional units and phosphonate moieties were studied by NMR. Simple monophosphonate/ammonium interactions were shown to be relatively weak with binding constants typically in the 10^2 M^{-1} range. By contrast, a bisphosphonate binds to a benzylammonium ion with a 14-fold increase, indicating a chelating effect that aids the interaction strength. Interactions between alkylammoniums and bisphosphonates were simulated with molecular modeling to demonstrate the nature of linear hydrogen-bonding involved.[98]

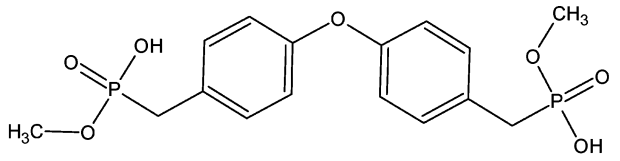


1 2

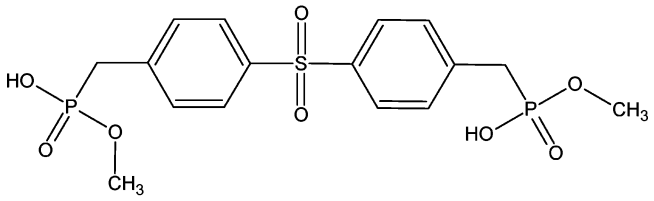
Figure 25. Binding of a guanidinium unit by synthetic bisphosphonate receptors. The explicit structures of these bisphosphonate receptors are shown as **37** (1) and **38** (2). Reprinted with permission from ref 205. Copyright 2001 American Chemical Society.

Extending this methodology to guanidine-based systems was straightforward, since secondary amines appear to bind more strongly than primary amines. Bisphosphonate-based molecules were shown to strongly bind to the guanidine group in arginine-containing peptides, simultaneously forming a cooperative hydrogen bond with the amide nitrogen, five atoms away from the guanidine. The additional hydrogen bond leads to an order of magnitude increase in binding and demonstrates that proper orientation through organization can achieve complementary interactions to strengthen the primary directed hydrogen-bonding interactions, which account for the majority of the interaction strength. This arrangement can be considered, in other words, to provide "conformational locking".[206]

Extension of these "molecular tweezer" systems into biomimetic applications seemed a logical next step. An excellent study was published in 1997 based on alkylguanidinium recognition.[206] Force-field calculations showed that the phosphonate ligands arranged themselves to form a nearly completely planar network of hydrogen bonds with guanidinium groups, much in the same way RNA−protein interactions are perceived to act in the recognition of the AIDS virus. By employment of different hinge groups in the tweezer structure (O and SO_2 in **35** and **36**, respectively), variation in the angles of interaction of the tweezer arms could be exploited. The goal of choosing



35

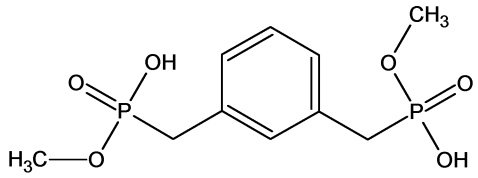


36

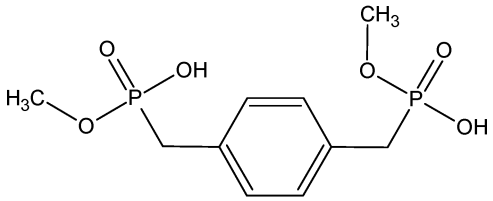
such groups was to gain selectivity for guanidinium groups over closely related structures such as ammonium ions. NMR titration showed the bisphosphonate receptor with the central SO_2 hinge **36** to have a strong affinity and high selectivity for arginine. In addition, an added advantage could, in the future, be realized by linking the phosphonate-based receptor with an appropriate nucleophile to mimic enzymes, such as trypsin and thrombin, and create a molecule capable of providing selective cleavage of proteins and peptides.

Attempting to further increase the selectivity and biomimetic nature of these receptor systems, additional modifications to the structure were sought. To optimize recognition properties, a new host should be capable of forming interactions through at least three binding points, whether the interactions are electrostatic, hydrogen-bonding, or polarization-based in nature.[98] For this purpose, benzylic bisphosphonates, based on an aromatic tweezer headgroup, were designed and shown to increase binding with aromatic guanidines via $\pi-\pi$ interactions.[207] Binding studies were performed by NMR titration in DMSO; however, the author notes that $\pi-\pi$ interactions would be strengthened in the presence of water. Such a design was used to affect recognition of amino sugars using a *meta*-xylylene bisphosphonate host.[208] Cooperative hydrogen bonds between the bisphosphonates and the hydroxyl groups on the amino sugar were formed and studied by solution phase NMR in DMSO, water, and methanol. This study demonstrates the ability of the bisphosphonates to operate effectively and selectively in highly polar media.[99] Comparisons between meta- and para-arranged xylylene bisphosphonates (**37** and **38**, respectively) were also made. The meta-arranged bisphosphonate was shown by force-field calculations to bind to a higher degree than the para-arranged receptor due to a decreased $PO^{-}\cdots NH^{+}$ interaction distance, indicating both stronger electrostatic and directed hydrogen bonds.[98]

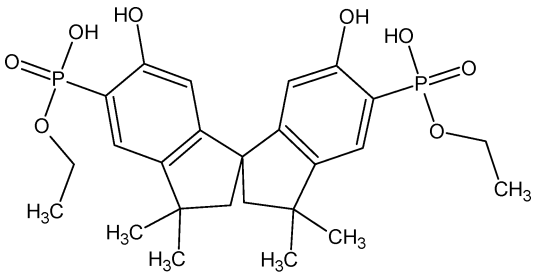
To build further on the elaborated concepts, receptor design by Schrader and co-workers was extended to continue to study selectivity for arginine versus other basic amino acids. Spirobisindane bisphosphonates **39** and **40** were synthesized to create host structures that exhibited rigidity, thus leading to a more highly preorganized structure.[50] Both receptors



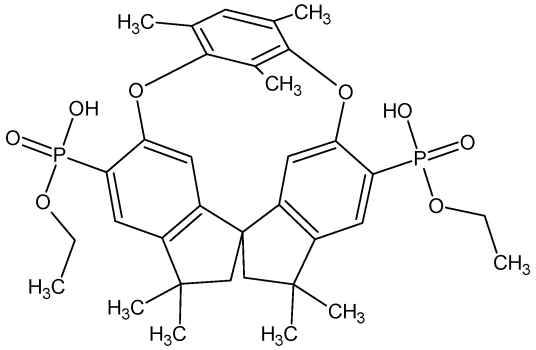
37



38



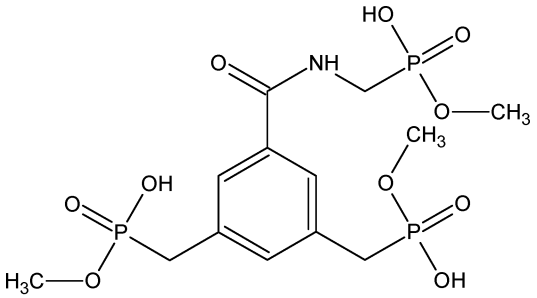
39



40

are chiral and were investigated for their enantiomer recognition properties. Structure **39** showed no enantioselectivity toward arginine residues, whereas **40** was enantioselective, provided that proper distance relationships were present.

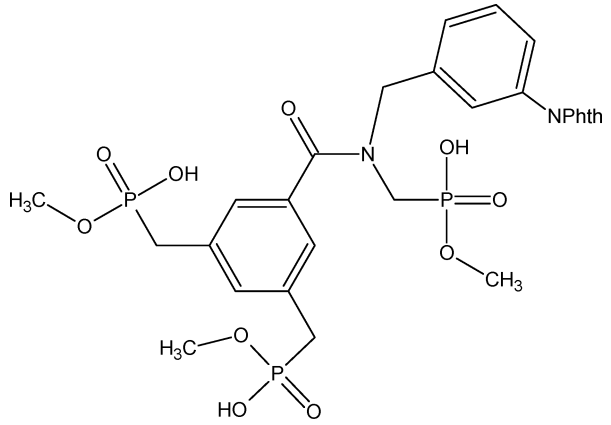
Efficient binding to arginine and increased selectivity against lysine were also achieved by increasing the number and strength of cation$-\pi$ and electrostatic interactions. By modification of the π face of the aryl bisphosphonates **37** and **38**, the existence and importance of cation$-\pi$ interactions could be determined experimentally.[205] The resulting molecule **41** forms five hydrogen bonds with guanidinium ion upon addition of a third phosphonate functional unit. The contribution to stabilizing energy of the cation$-\pi$ interaction in this substituted aryl triphosphonate molecule was determined to be around 0.5−0.6 kcal/mol. The only drawback cited for this



41

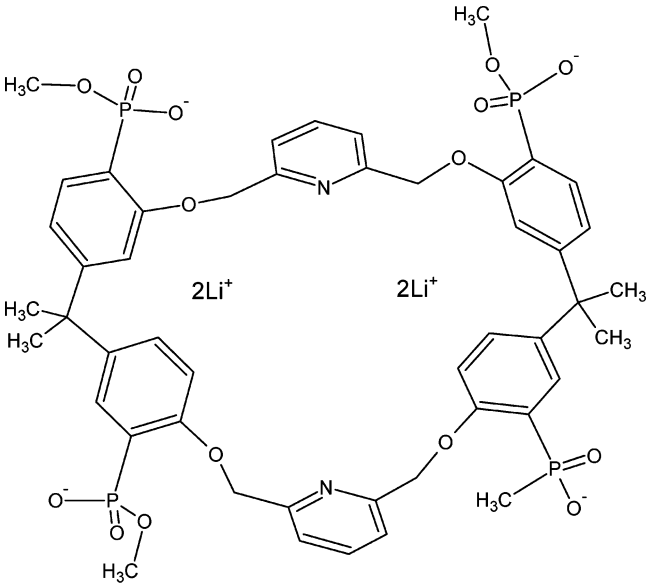
receptor, with respect to selectivity, was its inability to distinguish between methylguanidine and amide-protected arginine ester (H-arginine-NH_2) when methanol was the solvent. Data showed that in this case, one of the cooperative hydrogen bonds that previously allowed this receptor to distinguish between these two similar guest molecules was absent in methanol.

The functionalized aryl trisphosphonate **41** shows remarkable selectivity for arginine (specifically, the guanidinium subunit), making it applicable to screening for specific arginine sites in peptides and proteins. To demonstrate this, Schrader and co-workers chose to focus on the recognition of the RGD sequence, which is known to be essential for many cell surface recognition processes.[205] The sequence is, for example, an important part of the structure of integrin. The trisphosphonate showed remarkable recognition of RGD peptide in water.[209] The addition of the third phosphonate group greatly facilitated its recognition performance in aqueous media. To gain selectivity, the receptor was further functionalized to create an additional recognition point for the aspartic acid (D) group. This was accomplished by incorporating another aryl group for the purpose of secondary $\pi-\pi$ interactions to stabilize binding in **42**.



42

As shown above, increasing the phosphonate functionality of the receptor has afforded an increase in selectivity for guanidinium units and the ability for the receptor to operate in aqueous media. Continuing with this trend, the performance of phosphonate-based receptors in aqueous environments was further enhanced by designing a macrocycle that contained two bisphosphonate moieties, or four phosphonate units.[49] The macrocycle **43** is characterized by a rigid



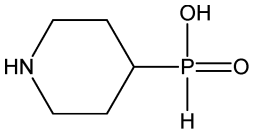
43

distance between interacting groups. The results show 2:1 stoichiometric binding for arginine and histidine and a 1:1 binding for lysine. Due to the arrangement of the receptor functional groups, only lysine can span the distance between the anionic groups. The arrangement, however, allows two arginine to bind simultaneously. Incidentally, two histidines, also in a 2:1 configuration, are bound remarkably well due to the special hydrogen bond donor arrangement of the imidazolium ion in their structures. The two histidine molecules are involved in a double chelate with three of the four phosphonate groups. All of these results were correlated with simulated results from geometric optimizations using molecular modeling. Differences were also observed depending on the type of solvent used. The highest binding constant in methanol was exhibited by histidine, whereas, in water, lysine was bound the strongest. In both solvent mediums, the binding of arginine was inferior to lysine and histidine, indicating a structural preorganization of the groups that favored these two amino acids over arginine in this receptor. In water, the overall binding for all of the tested amino acids was 5−7 times weaker relative to methanol-based experiments.

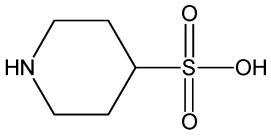
To summarize the contributions by Schrader and his colleagues, there have been significant recent advances in the development of synthetic receptors based on multiple phosphonate functionalities that are selective for the recognition of guanidine-based molecules and are able to operate in competitive media. A variety of bisphosphonate tweezer molecules have been shown to bind strongly and, in some cases, enantioselectively to arginine. Selection of appropriate headgroups was necessary to create the proper geometric complementarity. Additional functionality has been used to increase interactions outside of the directed hydrogen bonds for creating multiple point interactions that exhibit strength and

selectivity in binding as well as providing for use of the receptor in more polar media, such as water. Receptors based on these motifs could also be tuned biomimetically to select for biologically relevant peptide fragments, such as RGD. Perhaps one of the most attractive features of these systems is their relative ease of synthesis compared to other host molecules with different functional arrangements. A good example of this is the relatively simple modular two-step synthetic procedure used to create the tetraphosphonate macrocycle **43**.[49]

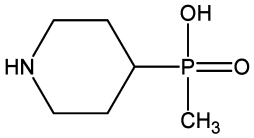
Other good examples of synthetic phosphonate-based molecules interacting with arginine residues exist as well. γ-Aminobutyric acid (GABA)-A is believed to interact through arginine residues at GABA recognition sites by binding between the guanidinium unit on the arginine residue and the anionic portion of GABA-A.[210] Krogsgaard-Larsen and co-workers studied the efficacy and potency of a series of bioisosterically modified GABA analogues by electrophysiological determination.[211] Among the structure variations were three phosphinic acids **44**, **45**, and **48**, a phosphonic acid **46**, and a sulfonic acid **47** analogue. Though there were small differences



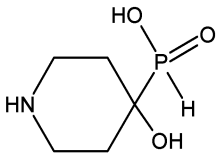
44



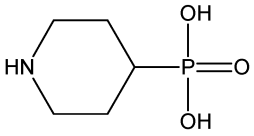
47



45



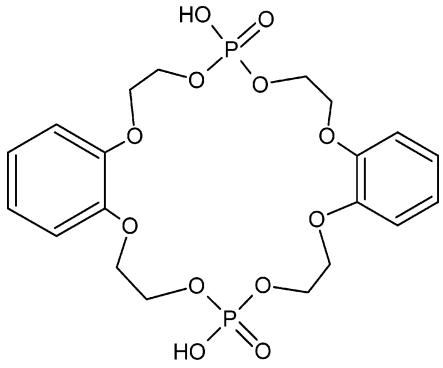
48



46

exhibited in efficacy among the analogues, the specific interaction with arginine and the guanidinium group remains to be studied in detail.

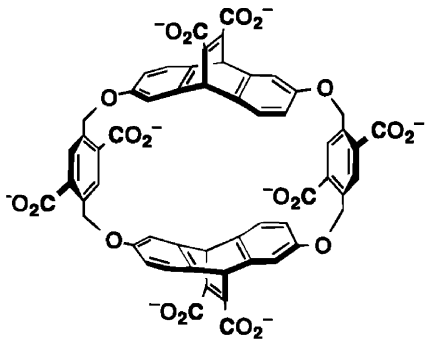
Another system has recently been reported that is based upon the selective complexation of arginine by a phosphate-functionalized crown ether in the gas phase. Studies designed to probe interactions in the gas phase, generally performed with MS- and tandem MS-based instrumentation, can provide useful information about the interacting partners without the interference of solvent. Julian and Beauchamp showed previously that the protonated alkyl-guanidinium side chain of arginine forms a stable noncovalent complex with dibenzo-30-crown-10 by ESI-MS.[212] Following this, they modified the crown ether structure by incorporating two endocyclic dialkylhydrogenphosphate esters (DP) **49** and studied its interaction with a series of arginine-containing peptides.[213]



49

Dissociation of the DP/peptide adduct ions, correlated with DFT calculations, revealed the proportional relationship of peptide chain length with acidity of the peptide. Comparisons of sequence selectivity for binding also showed greater selectivity and specificity for binding DP to arginine-containing peptides versus other peptides.

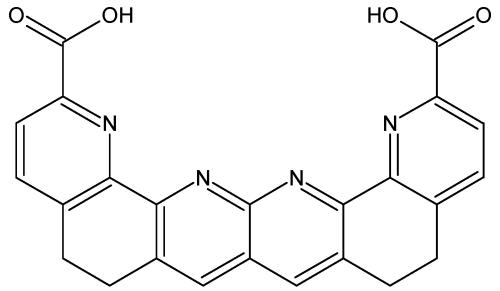
Not based on phosphonate or phosphate functionality, but worth mention due to their effective recognition of arginine residues in polar media, such as water, are a set of carboxylate-functionalized macrocycle receptors, represented here by the highest affinity binding structure **50** designed by Dougherty



50

and co-workers.[31] Central to the mechanism of association provided by these molecules are cation$-\pi$ interactions, where positive charges on the guanidinium group in arginine-containing dipeptides are stabilized by their proximity to electron-rich faces of aromatic rings. The studies emphasize that there exists a more complex binding relationship than just simple electrostatic attraction provided by the carboxylate$-$guanidinium interaction. Rather, results from studies where dielectric of the medium is changed through addition of a cosolvent, such as acetonitrile, and ionic strength is changed through addition of salt support a strong component of the association due to a novel induced dipole mechanism rather than just ion pairing. Variation in binding observed between arginine and lysine with the receptor was attributed to the better structural match of the guanidinium moiety with the aromatic π-systems incorporated in the macrocycle. This complementary match has been discussed previously in a review of cation$-\pi$ interactions by Dougherty.[22]

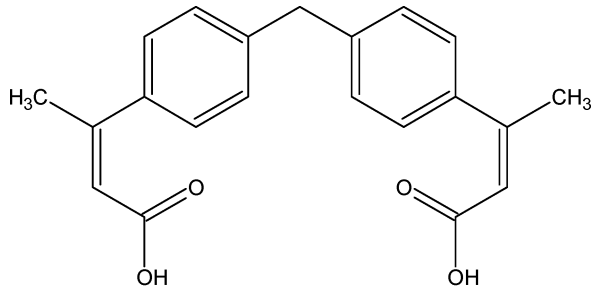
Another specialized receptor for the guanidinium group, specifically, for arginine residues and the dipeptide sequence RR, was developed by Bell and co-workers.[11] This hexagonal lattice receptor **51** uses



51

carboxyl end groups to orient the guanidinium receptor of arginine into the cleft of the receptor where it is bound by hydrogen bonding to a set of nitrogen heterocycle structures. The end-group carboxyl groups also provide a contact site for a second guanidinium group when the RR peptide sequence is present. The receptor shows a 3-fold selectivity for guanidinium over ammonium in methanol and boasts a 200-fold improvement in binding for methylguanidinium compared to one of Schrader's first molecular tweezer receptors based on phosphonate funationality.[206] Arginine by itself binds in a 2:1 stoichiometry, whereas the dipeptide RR binds 1:1 with the cleft receptor.

An additional interesting and creative approach to the selection of specialized receptors is a study published by Eliseev and Nelen on the use of Darwinian mutation and survival.[214] In this study, immobilized arginine is used to select the preferred isomer of a carboxylate receptor and concentrate it on the solid support. Although it is the arginine residue that is immobilized in this experiment, the focus is on enriching an isomeric carboxylate receptor for the guanidinium unit. By passing a mixture of cis/cis, cis/trans, and trans/trans isomers of a carboxylate-based cleft receptor over the immobilized arginine substrate, the preferred configuration or most effective binder (cis/cis) **52** becomes enriched on the solid



52

support. The isomeric forms are interconverted through irradiation with UV light. The method is designed to provide an automatic amplification of the component with the highest affinity through a dynamic equilibrium process. Results indicated an overwhelming success, where the HPLC of the mixture of receptors taken off of the substrate showed an 85:13:2 (cis/cis−cis/trans−trans/trans) enrichment (starting from an initial 3:28:69 distribution prior to cycling).

The results assembled from the use of phosphonate-based molecules and receptors for selectivity of guanidinium units are encouraging. Schrader and coworkers have shown perhaps the most in-depth research into developing these systems, as they applied their creative multiphosphonate unit receptors directly to biologically oriented systems. As can be seen from the geometries of interaction and the ability to bind guanidinium units in a wide variety of solution media, the possibility of applying phosphonate-functionalized species to a diverse range of systems, including affinity and biomimetic systems, is very feasible and attractive. With no uncertainty, these phosphonate receptor schemes will broaden to a large degree soon, much like the guanidinium-based receptors have in the past 10−15 years.

3.5. Biologically-Relevant Sulf[on]ate-Based Interactions with Guanidinium

Sulf[on]ation, related to the biological environment, tends to center around elucidation of binding specificity for large polysulfonated biomolecules. Related to binding between sulf[on]ate and guanidinium, a large amount of the literature is focused on the interactions of heparin and its analogues. Heparin is, due to its biological function, a widely used pharmacological agent for anticoagulation during surgery. Structurally, heparin is homogeneous in nature and is composed of long and highly negatively charged, unbranched polysaccharide chains.[215] Figure 26 shows the structure of the major repeating unit of heparin and structurally related heparin sulfate.[216] The anticoagulant activity of heparin in biological systems is mainly due to its interaction with antithrombin III (ATIII), a strongly basic protein. This activity will only be broken if heparin is given a more basic protein or peptide with which to bind. In work performed by Yang and co-workers, heparin bound to ATIII was subjected to a protamine containing 67% arginine in its structure.[217] Due to the stronger affinity of the protamine for heparin, its introduction

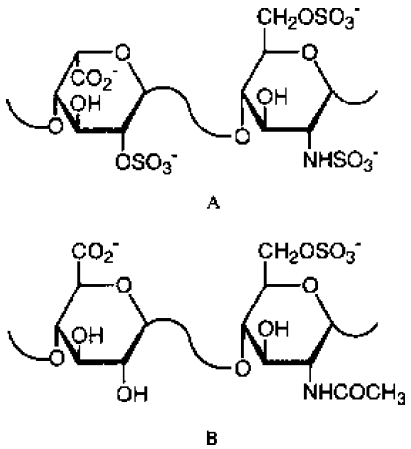


Figure 26. Structure of the major repeat unit of heparin (A) and heparin sulfate (B). Reprinted with permission from ref 216. Copyright 1996 Elsevier Science.

into the system was able to dissociate the heparin−ATIII complex, thereby reversing the anticoagulant function of heparin. Although the exact sequence of the protamine was not established, it was conjectured that a string of six adjacent arginine residues were responsible for the heightened heparin activity. This is in contrast to earlier studies where it was believed to be a bank of lysine residues in ATIII that are responsible for binding with heparin.[218,219] Binding of specific amino acid sequence patterns, such as XBBXBX and XBBBXXBX have been suggested, where B is a basic residue such as arginine or lysine, and X is a nonbasic residue.[220,221] This arrangement may be either in a linear row or topologically as part of the orientation of a three-dimensional structure.[222−225] Currently, a row of arginine active sites are believed to be the main interaction site between these two biomolecules. Beliefs about the role of other interactions between secondary structures in the interacting proteins have not changed and are believed to stem from an organized three-dimensional structure that is instrumental to stabilizing the conformation of the complex.[218] Also important to the interactions is the degree of sulfation and the net negative charge on the mucopolysaccharide.

Early work on the binding of lactoferrin, an anti-inflammatory response protein, to glycogaminoglycans (GAGs), such as heparin, dubbed the arginine-rich site of complex formation the "cationic cradle."[226] GAGs are classified as anionic polysaccharides and are present at the surface of most cells. Molecular modeling experiments of the system showed the arrangement of the cradle structure by the arginine residues to accommodate guest docking. Lactoferrin contains a GRRRRS sequence at its amino terminus, which acts synergistically with a RKVR sequence further down the chain (positions 28−31) to bind with GAGs. In contrast, transferrin, which possesses a similar structure to lactoferrin but lacks the clusters of basic residues, failed to bind GAGs under physiological pH and ionic strength conditions. Work by Ingham and co-workers also correlated this finding through analysis of fibronectin binding to heparin but included the stipulation that the tertiary structure of the basic proteins plays an important role in the stability of the cationic cradle.[227] Comparison of binding to synthetic peptide analogues that lacked the tertiary structure of complete fibronectin showed a decline in binding when only the specific basic residues were present. Elucidation of these concerted binding structures emphasizes the greater prevalence of an induced fit-based mechanism over a simplified lock and key scheme. Molecular modeling was used to examine the three-dimensional structure of the heparin binding domain.

More recently, further work on elucidating the binding of heparin to fibronectin was performed by Miyamoto and co-workers.[228] The binding between extradomain A containing fibronectin (EDA(+)FN) and heparin, along with fibrinogen, creates an aggregate gel called a cryogel, which has been detected in the blood of rheumatoid arthritis patients. Work was aimed at determining the precise interaction points between EDA(+)FN and heparin to devise a plausible inhibitory scheme for the process. Using oligo- and desulfonated heparin, as well as synthetic peptides to mimic the binding domain of FN, inhibitory activity was found using a Gly-Arg-Lys-Lys-Thr (GRKKT) synthetic peptide arrangement. Concurrently the binding affinity of heparin could be destroyed by desulfonation. These results demonstrated that the use of synthetic basic peptides and their interaction with the sulfonated biomolecule heparin could be a useful inhibition scheme for cryogelation in plasma, as well as a useful treatment for rheumatoid arthritis.

The binding of heparin to vitronectin, a plasma glycoprotein that circulates in human blood and is involved in a number of physiological processes including blood coagulation, fibrinolysis, and cell adhesion,[229] is one of the most widely studied interaction systems in vitronectin literature. A short review of this research and a multifaceted contribution to this field was published by Peterson and co-workers.[230] Experiments involving solution biochemistry, spectroscopy, and recombinant approaches were performed to evaluate ionic and nonionic interactions and identify potential residues contributing to complex formation between heparin and vitronectin. Compared to binding between ATIII and heparin, the complex formed between vitronectin and heparin is approximately 2 orders of magnitude weaker.[231] The reason for this was partially elucidated in this study. While ATIII binds heparin through a bank of five or six adjacent arginine residues, the residues for binding by vitronectin are confined to a single C-terminal cationic cluster. The effect of ionic strength was previously used to study this hypothesis, showing the interaction to be largely ionic in nature.[232] Binding between heparin and vitronectin was performed by labeling heparin with a fluorescent coumarin probe and titrating with vitronectin. Also elucidated through NMR studies with synthetic peptides were the nonionic contributions of glutamine and asparagine units in the binding complex. Another protein, human activated protein C (APC), has also been investigated and shown to interact with heparin mainly through ionic means and been accredited to a bank of lysine, rather than arginine, residues.[215] Again, the ionic nature of the interaction was tested by systematically increasing the salt concentration in the system.

In 1996, Weiler and co-workers prepared an affinity separation system to assess the affinity of heparin and heparan sulfate for randomly generated seven-mer peptide molecules.[216] Heparin is known to be the strongest acid in the human body and the most acidic polysaccharide in nature. It has, consequently, found wide use as an effective affinity ligand. Proteoglycan heparan sulfate is found in cellular membranes and is active in cell to cell communication through the binding of growth factors. Along with establishing a general difference between the binding of specific sequences of amino acids for heparin and heparan sulfate, this study further elucidated the nature of selectivity of these highly sulfonated molecules for arginine residues relative to other amino acid residues. The frequency of various amino acids in randomly sequenced peptides was examined to define the

primary structural requirements for binding to these anionic biomolecules. Amino acid residues that were found to be depleted, and therefore not active for promoting heparin binding, were histidine, isoleucine, methionine, and phenylalanine. Arginine and lysine were found to be, by far, the most commonly encountered amino acids; however, it was confirmed that arginine forms tighter interactions than lysine.[233] In comparison to histidine residues, arginine and lysine possess longer chains and can access binding pockets in heparin more readily. In comparison of arginine with lysine residues, the main difference must be due to the forklike nature of the guanidinium unit and its highly basic and directed nature. Overall, there appeared to be a little more disparity in assessing binding to heparan sulfate compared to heparin, suggesting interactions between heparin and arginine exhibit a slightly higher frequency than those between heparan sulfate and arginine. Along with binding information with basic amino acids, there appeared to be a slight increase in the occurrence of neutral amino acid residues, such as serine and glycine, which are known to impart flexibility in peptide sequences. The peptides containing these residues may allow for a better access by cationic residues for anionic binding sites as well as "inducing fit" as an integral part of secondary structure interactions.

Another example of heparin binding to a band of cationic residues has been published by Gabay and co-workers with their work on Azuricidin/HBP. Azuricidine/CAP37/HBP is an antimicrobial and chemotactic protein and aids in defense against human neutrophils.[234] The three-dimensional structure of azuricidin, a member of the serprocidin family, had been previously elucidated.[235,236] The structure shows a highly cationic region of 16 basic amino acid residues located at one pole of the molecule. At least two clusters within this section resemble the heparin binding motif. The clusters appear to form two loops, adopting a conformation that can also be described by a cationic cradle arrangement.[226, 227]

A common theme that is evident in the binding between anionic polysulfated molecules such as heparin and a variety of cationic-based biomolecules in biological systems is the relationship between the degree of basicity of the cationic molecules and their strength of binding with polysulfates. An increase in the number and proximity of basic residues available for binding can be correlated with an increased affinity for anionic biomacromolecules. In a reciprocal nature, the converse relationship was also demonstrated for the degree of sulfation on the anionic substrate, showing analytically a high degree of sulfation could be used to augment complex formation with basic proteins.

Approximately 10 years ago, Juhasz and Biemann used MALDI-MS to analyze the interactions by large biomolecules such as suramin (an aromatic polysulfonic acid) and heparin (a large, highly sulfated polysaccharide) with a variety of basic peptides including cytochrome *c*, melittin, and bovine insulin.[237] Highly acidic compounds are difficult to analyze by MALDI-MS due to their high affinity for cations, which gives rise to broad unresolved peaks in spectra. More favorable results were achieved when the anion-based biomolecules were mixed with basic peptides or proteins. It was found that the extent of complex formation correlated with the number of phosphate and sulfate groups in the acidic component and the number of arginine residues in the basic component, although sulfates were found to bind more strongly than phosphates. The significance of the number of complexed arginine units was also related to their distribution throughout the protein. For larger proteins and peptides, their tertiary structure appeared to play a role in the availability of binding sites. Through this methodology, it became possible to efficiently ionize and detect polysulfated molecules by using basic peptides and proteins as "carriers" into the gas phase. Though the specifics of binding were not explicitly elucidated in this study, a significant foundation for analysis of strongly anionic species and, in a reciprocal manner, strongly cationic species by analytical techniques (MALDI-MS in this case) was cemented. This information found use very quickly in subsequent investigations, as we will see in the next section.

In biological systems, the vast majority of interactions documented include polysulfated molecules such as heparin and a variety of cationic-based components. Though not covered explicitly here, the sulfation of residues as a biological modification scheme is an important process in living organisms. We have shown that a diverse array of interactions at various degrees of strength can be expected between sulfated species and basic amino acid residues, specifically with a cation cradle-type motif. The degree of basicity (and similarly, the degree of acidity) has been correlated with the strength of binding under physiological conditions. In addition, the effects on binding observed with small structural changes, sometimes unrelated to the hydrogen-bonding or ionic interaction site, in one or the other interacting partners, further emphasizes the more modern notion of an induced fit recognition mechanism, rather than the lock and key metaphor previously used. These themes, increasing binding sites for increased specificity and the importance of secondary and tertiary structure, are important concepts realized and incorporated by scientists involved in synthetically engineered recognition systems for biomimetic purposes.

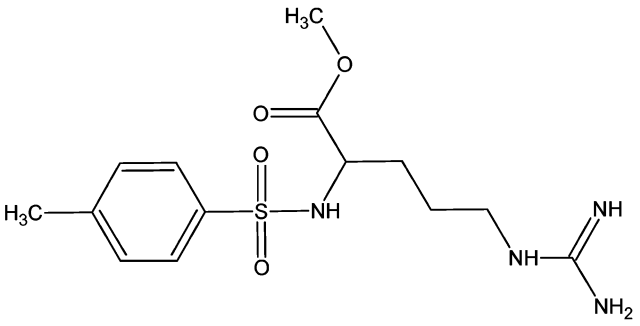
3.6. Design and Use of Sulfonate-Based Molecules for Guanidinium Binding

The main constituency of interactions of synthetic sulfonated molecules reported in the literature comes with the use of sulfonated dye molecules to bind with proteins and peptides for the purpose of elucidating conformation and selective protein capturing. In the majority of the examples cited, probing techniques, such as mass spectrometry, are used to study and shed light on biological implications. Some polysulfonated azo dyes (containing aryl sulfonyl groups) were analyzed by Sullivan and Gaskell by MALDI-MS and ESI-MS and shown to behave nicely in MS-

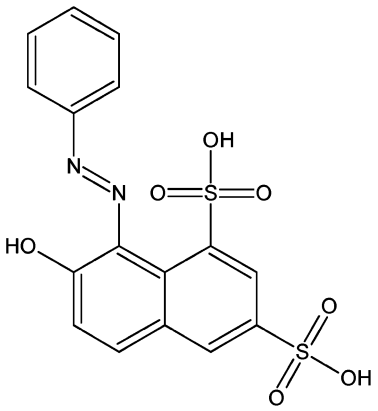
based experiments; however, they did not exploit the usefulness of the dyes as noncovalent probes as will be discussed below.[238] Other application notes for the analysis of azo dyes by ESI-MS are given by Slater and co-workers.[77] Notes on the analysis of sulfonated dyes and intermediates are given by Holčapek and co-workers.[239] Dyes are favorable for complexation studies because they are usually nontoxic, form a stable complex, are themselves stable, and provide a means for spectroscopic analysis through addition of a chromophore.[240] Much of this information is provided in a short review by Zenobi and co-workers in one of their recent publications.[241] The review addresses the use of noncovalent probes for determination of arginine residues in peptides and proteins by MS. Another review of dye−ligand affinity systems has been published recently by Pişkin and Denizli.[242]

Gleason and co-workers recognized fairly early the significance of azo dyes as potential models for biomolecular recognition.[104] A molecular level understanding of the ways in which sulfonated compounds interact with proteins would be of great utility for the design of new drugs, particularly in their use for purification of proteins, antibodies, and other related molecules by affinity chromatography. They used graph-set formalism to study the interactions observed in crystallographic data for binding between GAGs and sulfonate groups. Graph-set formalism has been used in a number of cases to study hydrogen bonding in crystallographic structures.[16−18] These studies were supplemented by the study of interactions between tosyl arginine methyl ester (TAME) (**53**) and sulfonated dyes such as the Orange G dianion (**54**), Little Rock Orange (**55**), and HABS anion (**56**) using X-ray diffraction. TAME is an ideal model for guanidinium-based peptidic interaction sites, since it lacks charged amino and carboxyl termini to interfere with binding to the guanidinium moiety. It was found that bridging interactions are the favored mode of interaction and that these bridges are likely to form between both small molecules and large sulfated biomolecules with proteins. Also revealed from the data were numerous water interactions incorporated into the bridging structures.

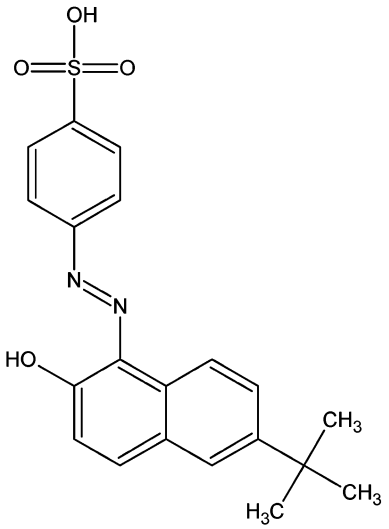
A similar study performed by Ward and co-workers on crystallography of self-assembled guanidinium alkane- and arenesulfonates showed the propensity of the system to form hydrogen-bonded sheets, bilayers, or both, the size of the substituted R of the sulfonate group incorporated into the structure being the major factor in association to form ordered assemblies.[243] Crystals of guanidinium sulfate salts can be formed simply by slow evaporation of guanidinium chloride and the appropriate sulfonic acid or sulfonate salt. Directed hydrogen-bonding interactions are responsible for the binding between the sulfonate and guanidinium moieties and packing patterns result from the partitioning of hydrophobic regions contained in the sulfonates. Adjacent substituents were seen to interpenetrate to maximize van der Waals contacts. Overall, these ordered motifs are characterized by a large number of hydrogen



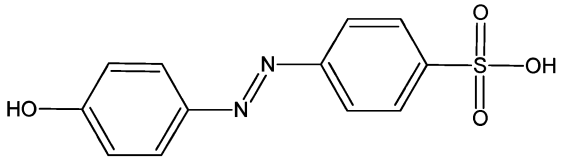
53



54



55



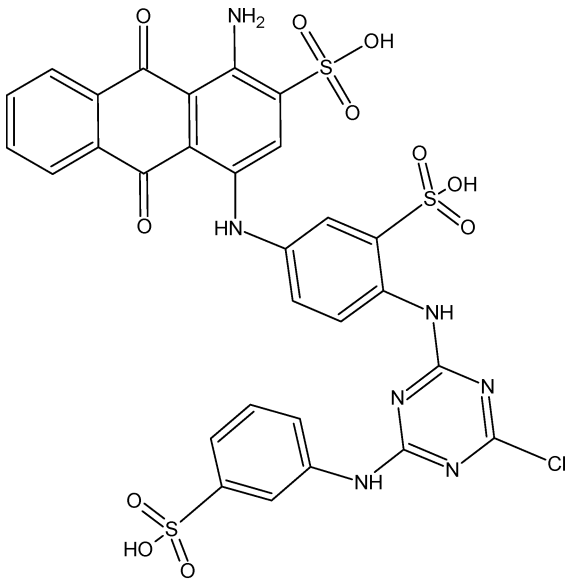
56

bonds and matched numbers of donors and acceptors. The addition of added donor or acceptor sites disrupts the ordering of the system.

Much of the work dealing with the use of sulfonic acid probes in a MALDI-MS setting has been carried

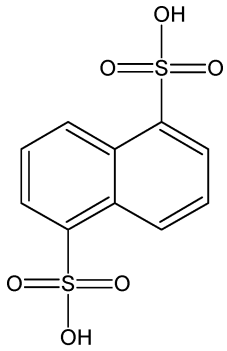
out in Zenobi's laboratory. In addition to the use of noncovalent probes, the concept of "surface topology probing"[244,245] was also discussed, where accessible sites on a protein could be covalently modified prior to protein digestion and sequence mapping. Such reference to the conversion of residues such as lysine to homoarginine were discussed;[246] however, rigorous coverage of this topic is beyond the scope of this review. It is obvious that the development of sequence-specific covalent or noncovalent probes is an invaluable tool for high-throughput screening and sequencing of peptides and proteins, particularly with the current interest in proteomics. Sulfonates provide good complementary binding character for electropositive residues and, as will be shown below, can possess structural character to enhance specificity for select residues, namely, arginine.

Cibacron Blue F3G-A (CCB) **57**, a trisulfonated dye common in affinity chromatography, was shown to



57

bind to accessible basic sites of a folded protein or peptide.[240] The biomolecules investigated were melittin (helix with all six positive sites available for binding at neutral pH), ubiquitin (12 arginine, lysine, or histidine groups plus the N-terminus), bombesin, insulin (large distance between basic sites), and cytochrome *c* (large number of basic sites with a highly variable charge state with changing pH). In this approach, the number of CCB adducts observed in the mass spectrum of the protein or peptide of interest could be likened to the number of free arginine, histidine, or lysine residues plus the amino terminus. CCB was not, however, found to be selective among the different basic sites, and steric effects with this large, bulky dye molecule had to be considered. In an attempt to study smaller sulfonic acid derivatives as probes, naphthalene-disulfonic acid (NDS), **58**, was employed in MALDI-MS experiments.[105] Results showed NDS to be selective only for arginine residues and the amino terminus in the peptide molecules investigated. Also, as the dye complexity and molecular volume were reduced, the

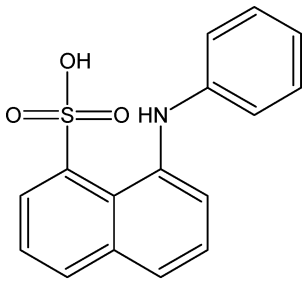


58

number of dye adducts to the peptide could be increased, eventually up to the number of positively charged groups on the molecule.[240] Work performed by Zenobi and co-workers has also led to the common use of basic or neutral matrixes, such a *p*-nitroaniline (pH = 6.5) when examining biomolecules to avoid protein denaturing effects observed with acidic matrixes.

Supplemental work on the topic of using sulfonic acid-functionalized molecules for noncovalent probing of peptides and proteins has elucidated a few basic rules governing their use.[241] These are as follows: (1) multifunctional dyes, such as CCB, act as a chelating ligand and probe for all accessible basic sites; (2) smaller, more specific probes, such as NDS, will recognize arginine, as well as the amino terminus, as long as an adjacent amide bond is present and can be involved in a cooperative hydrogen bond (NDS does not probe lysine or histidine and will only bind to surface-accessible arginine residues); (3) the binding of NDS to the arginine residue consists of ionic and hydrogen-bonding interactions that are largely preserved upon transfer to the gas phase. These observations were made based on the use of MALDI-MS. Comparison with results from ESI-MS showed, in contrast to MALDI spectra, the occurrence of multiple charging. The abundance of a +7 charge state on cytochrome *c* in the presence of NDS during ESI-MS indicates that the protein remains largely in its natural solution-phase state upon transfer to the gas phase for MS analysis. In addition, the presence of a +7 charge indicates that at least seven basic units are free and not bound with the sulfonate probe. This simple comparison shows one example of the very different results that can be obtained using MALDI- and ESI-MS ion sources for analysis of these systems in the gas phase.

A similar study was performed by Friess and Zenobi to investigate the use of 1-anilino-naphthalene-8-sulfonic acid (ANS), **59**, as a selective probe for arginine, as similarly found for NDS.[105] ANS, a fluorescent hydrophobic dye, was used previously and is still used for the detection of globular protein states.[247] Binding in this context was previously thought to occur between hydrophobic portions of the ANS probe and the folded or partially folded protein.[248] Recent investigations, however, show that ANS binds to cationic residues through its sulfonate group in such systems.[249] ESI-MS based studies have confirmed this.[250,251] ANS, similarly to NDS, binds



59

specifically to arginine with exclusion of lysine and histidine in a MALDI-MS analysis system. Reasoning for specificity of these interactions goes beyond simple ion-pairing schemes. Between guanidinium and sulfonate, a salt bridge is formed, which is further enhanced by ion−dipole and hydrogen-bonding interactions. Hydrogen bonds are highly favorable due to the complementary shapes of the interacting groups. The complementarity exhibited by the sulf[on]ate moiety is very similar to that shown previously for phosph[on]ate (see Figure 24) and is a manifestation provided by mutual forklike structures. Histidine and lysine lack the functional geometry to bind through this complementary interaction scheme.

As another example, Table 7 shows the selectivity of ANS determined by Matulis and Lovrien for differences in binding between ANS and polylysine, polyarginine, and polyhistidine peptides.[249] The result clearly shows the preference of the sulfonated dye for the arginine residue relative to lysine and histidine in addition to the predominantly 1:1 binding stoichiometry. In binding to arginine residues, crystallographic measurements clearly show the interaction between the sulfonate and guanidinium groups. ANS was also shown to bind to the N-terminus of the peptides. General interactions and specificity of sulfonate groups for arginine residues in this manner are complementary to the results presented earlier on the site-specific binding of large biomolecules such as suramin, heparin, and heparin sulfate.[237] However, a recent communication by Dif and co-workers on the use of ANS for site-specific probing by ESI-MS showed a lack of specificity for the sulfonate in binding with the different positive sites on proteins.[252] In fact, ANS was observed to bind to all basic residues and the amino terminus on a series of peptide and protein molecules. This is in contrast to the explicit specificity of ANS for the arginine residue reported by MALDI analysis[105] and should be investigated further.

Used as a probe for cytochrome *c*, ANS is shown to bind electrostatically in acidic media.[253] Proteins are well-known to denature and unfold in acidic media due to the presence of destabilizing repulsive interactions.[254] In the presence of ANS, cytochrome *c* was observed to refold in acidic media while being monitored with fluorescence spectroscopy and ESI-MS. Cytochrome *c* carries a large portion of basic residues (24 out of 105 are lysine, arginine, or histidine). The negatively charged groups in the protein are neutralized as pH is lowered, resulting in denaturing due to repulsion of the positively charged basic residues. Anions, such as ANS, reduce the repulsion by Debye−Huckel screening and through interaction with the positive residues. As this happens, the intrinsic hydrophobic interactions induce refolding.

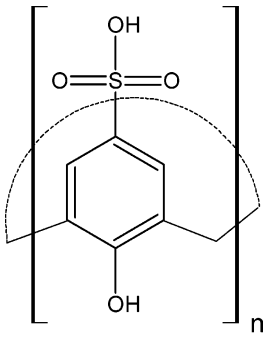
Table 7. Stoichiometries of Binding (*n*), Association Constants (K_b), and Enthalpies of Association (ΔH) of ANS^- to Poly(amino acids) at 25 °C and pH 2.0[a]

poly(amino acid)	n (±0.1)	K_b (±10%) (M^{-1})	ΔH (±10%) (kcal/mol)
poly-Lys	0.97	19 000	−2.7
poly-His	0.95	150 000	−4.5
poly-Arg	0.88	260 000	−6.0

[a] ANS^- binds 1:1 with every positively charged amino acid residue.[249]

As an affinity ligand, CCB, **57**, can be considered as a good model for ligand design. Lowe and coworkers have shown through X-ray crystallography and affinity labeling studies that CCB binds to enzymes with the anthraquinone, diaminobenzene sulfonate, and triazine rings adopting similar positions as the adenine, adenosine ribose, and pyrophosphate groups of NAD^+.[165] More specifically, the terminal aminobenzene sulfonate ring of the dye was bound to the side of the main NAD^+-binding site, adjacent to the side chains of basic arginine and histidine residues. Though much of this work was performed nearly 10 years ago, it is obvious that the specificity of the binding of the dye has provided an impetus to the more recent investigations cited above. Other similar dye ligands have also shown affinities and abilities to mimic biological systems. CCB, a specific form of Cibacron, is a product of Ciba-Geigy. Differing only in the position of the sulfonate group on the aniline ring, Procion H is a similar dye series marketed by ICI. The sulfonate group is present ortho in the Cibacron dyes and either meta or para in the Procion H dye series.[242] Pişkin and Denizli point out that all proteins, under some solution condition, can be adsorbed to a dye−ligand affinity sorbent, meaning that the class of molecules is very diverse in their interactive nature. Besides CCB, other anthraquinone-containing aromatic sulfonated dyes, such as Cibacron Blue 3GA, Procion Blue H-B and MX-R, and Vilmafix A-R, bind preferentially to the nucleotide binding site of many proteins, mimicking the naturally occurring binding of anionic coenzymes, such as NADH and FAD. Between the work on dye-affinity systems and the use of dyes for noncovalent labels, their diverse binding specificities allow for a multitude of favorable interaction with guanidine units while probing biomolecular systems.

An interesting biomimetic system has been recently investigated by Morel-Desrosiers and coworkers.[35,255] Using sulfonated calix[n]arenes **60** of various sizes (n = 4, 6, and 8) as mimetics for GAGs, they studied the binding to arginine and lysine amino acids and peptides. Using NMR and ITC, they studied binding and entropy/enthalpy relationships in aqueous solutions. For the free amino acids, binding of arginine was found to be about twice as strong with the hosts compared to that of lysine, even though both were enthalpically driven. Interactions



60

between the π face of the sulfonatocalix[4]arene and guanidinium ($\pi-\pi$) were conjectured to be responsible for the increased binding. In the case of the peptides, association constants of RK and KR were found to be similar to that of KK, but only half as strong as that for RR with calix[4]arene. Again, additional $\pi-\pi$ interactions were expected to be the cause. Lysine, however, was shown to have the ability to insert into the cavity of the host, having a marked effect on the measured enthalpy/entropy relationships. Binding of the peptides with the sulfonatocalix[6]arene was similar to that for sulfonatocalix[4]arene, except for an additional signal indicating an additional 2:1 stoichiometric association. Very little meaningful data could be extracted from binding with sulfonatocalix[8]arene.

Studies utilizing molecular modeling and simulation techniques have been used to guide studies of the interaction between the guanidinium and sulfonate groups. In an attempt to mimic the side chains of lysine and arginine, ab initio calculations (with sufficiently diffuse basis sets, HF/6-31G** and B3LYP/6-31++G**) were performed to assess the binding of methylamine and guanidine groups, respectively, with methylsulfonate.[256] This methodology lies parallel to similar studies presented earlier for phosphonates. Lowest energy structures obtained for these arrangements were a three-point hydrogen-bonding contact between all three amine hydrogens on methylamine and the three oxygens of the sulfonate anion (methyl sulfonate−amine interaction), and a two-point salt-bridge interaction between two of the planar guanidine nitrogens and two of the sulfonate oxygens (methyl sulfonate−guanidine interaction). Figure 27 shows the geometrical representation of the two lowest energy conformations calculated for the interaction between methyl sulfonate and guanidinium. Also notable were the poor results for a simulation of the guanidine−sulfonate interaction where the starting geometry included a symmetrical three-point contact of the sulfonate oxygens above all three of the planar nitrogens. This indicates the strength of the guanidine planarity, where the hydrogens prefer not to bond axially, as well as the effect of the π electron density and its poor interaction with an anionic species.

More recently, improvements were made in modeling solvated interactions between carboxylate, sulfonate, guanidinium, and ammonium functional units. The problem previously was the use of continuum solvation models in ab initio calculated interaction

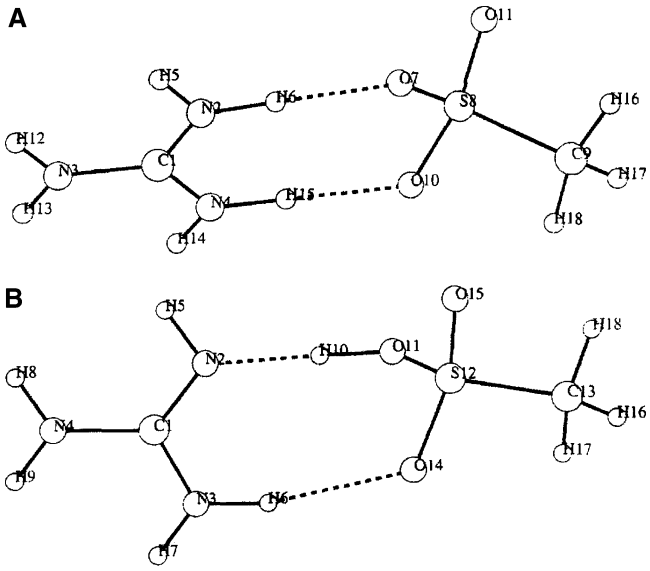
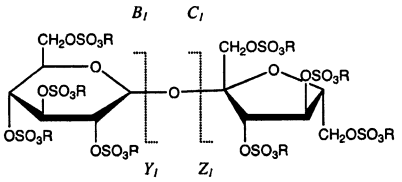


Figure 27. Lowest energy geometric conformations calculated for the interaction of methyl sulfonate with guanidine. Reprinted with permission from ref 256. Copyright 1998 Elsevier.

energies between these cationic and anionic moieties. Commonly used continuum models are unable to detect differences due to specific chemical changes, such as the stronger retention of basic proteins on sulfate- (strong) versus carboxylate-based (weak) cation exchangers. As cited earlier, similar problems were encountered for differentiation of ammonium and guanidinium interactions with carboxylate functionalities. Lenhoff and co-workers, using methylated anionic and cationic functional units to model the active residues present in larger systems, used a revised "heuristic" approach to study these specific interactions.[257] The heuristic approach relies more specifically on the bulk solvation energies of each group and seeks to incorporate specific solvent molecules into calculations of interaction energy between the moieties of interest. The ion-pairing interaction with solvent molecules is not accounted for in continuum models. By incorporation of several water molecules into the system local to the binding partners, as well as the conventional continuum model outside of this (bulk), binding energies for the cationic systems (and for methylguanidinium, in particular) were shown to be $1-3kT$ (where k is Boltzmann's constant and T is absolute temperature) larger with sulfated systems relative to those with carboxylated systems. The locally placed solvent molecules actually tune the interaction and show, besides an improved calculation method for solvated systems, the importance of the explicit interaction of polarized solvents in the interaction. This technique could certainly be applied to improve theoretical treatments in other similar ion-pairing systems.

As a model for the analysis of highly sulfonated biomolecules, such as heparin and heparin sulfate, Linhardt and co-workers have recently reported the positive mode ESI-MS analysis of sucrose octasulfate (SOS) in the positive ionization mode.[258] Though other recent studies have successfully shown the analysis of isomeric heparin biomolecules by ESI-MS,[259] the study of SOS is the first report of the



R	R	MW	M+Na	M+R	M-R	$B_1(Z_1)$+R	$C_1(Y_1)$+R
H	1.01 1	982.80 982	1005.79 1005	983.81 983	981.79 981	484.41 484	500.41 500
Na	22.99 23	1158.66 1158	1181.65 1181	1181.65 1181	1135.67 1135	594.32 594	610.32 610
$(CH_3)_4N$	74.15 74	1567.91 1566	1590.90 1589	1642.02 1640	1493.76 1492	850.10 849	866.10 865
$(CH_3CH_2)_4N$	130.25 130	2016.77 2014	2039.76 2037	2147.03 2144	1886.52 1884	1130.64 1129	1146.64 1145
$(CH_3CH_2CH_2)_4N$	186.36 186	2465.64 2462	2488.63 2485	2652.00 2648	2279.27 2276	1411.18 1409	1427.18 1425
$(CH_3)_4P$	91.11 91	1703.65 1702	1726.64 1725	1794.76 1793	1612.53 1611	934.94 934	950.94 950

Upper: average mass
Lower: nominal mass

Figure 28. Sucrose octasulfate (SOS) and its various structural fragments (*m/z*) and salts encountered by ESI-MS analysis. Reprinted with permission from ref 258. Copyright 2003 American Chemical Society.

successful use of positive mode ionization for analysis of the highly anionic molecules. Figure 28 depicts the structure and major fragment ions of the various SOS salts encountered by ESI-MS. Analysis of sulfonated molecules by ESI-MS is consistently hampered by the loss of sulfo groups and sensitivity problems. The use of ammonium cations allowed for the formation of stable positive ions in the analysis of SOS and has opened a new mode for the analysis of sulf[on]ated biomolecules. Particularly interesting may be the selective complexation of cationic species to provide discrimination between isomeric or chiral molecules. The use of polyamines and basic peptide counterions in SOS analysis in this study met with little success.

The introduction of small sulfated molecules as probes for cationic structure, as well as immobilized for affinity separation purposes, has opened a very useful avenue for mimicking biological and pharmacological interactions in conjunction with high-throughput analytical schemes. The character of these small dyes (nontoxic, inexpensive, and highly soluble) will continue to drive their use in conjunction with molecular recognition type systems. In addition, movement to mimics of larger sulfated biomolecules also presents favorable avenues. Detailed observations in recent reports have shed light on many of the underlying principles that govern the molecular-scale interaction between sulfonates and guanidinium groups. Building upon these fundamentals to gain useful information in more complex systems through common and advantageous analytical techniques, such as mass spectrometry, is expected to help this interest grow soon.

4. Summary and Outlook

In determining the nature of interacting molecules, whether biologically or synthetically derived or both, it is absolutely necessary to describe the complementary interactive nature of both species (or in some cases, more) involved. For peptides and proteins in biological media, much of this interaction is provided by ion-pair binding events coupled with secondary stabilizing interactions to provide specificity and selectivity. In synthetic schemes, this ion-pair complementarity again becomes a dominant goal as chemists seek to mimic both biological functional arrangements and their ability to bind in more polar (aqueous) or competitive media. In both cases, the role of the solvent is a key consideration for understanding the thermodynamics associated with the noncovalent complex formation.

In this review, we have focused both on guanidinium interactions (synthetic types and arginine side chain residues) with anionic systems and vice versa. Biological as well as synthetic cases have been used to illustrate the importance and utility of these interactions. Though biological systems are generally flexible and operate via a complex induced fit recognition pattern, the use of synthetically engineered receptors, based mostly upon rigid structures with predictable interactions, offers an exciting way to mimic biological interactions. Through reversible noncovalent binding and creative recognition schemes, scientists are able to reproduce (biomimetic) and even fix (biopharmaceutical) biological activation, inhibition, and transfer processes. To supplement and study these systems, an increasingly impressive tool kit for analysis continues to develop. Future aspects of these techniques move to high-throughput combinatorial schemes, miniaturization, and higher compatibility with physiological media. Theoretical investigations are beginning to become reliable and oft-used checks on the experimental evidence.

From the point-of-view of guanidinium-based receptors, a formidable amount of research has been performed. These systems are able to recognize with high selectivity (even, enantioselectivity in some cases), in a variety of media (solid, liquid, and gas phases), and with robustness and reproducibility. The extensive activity of arginine residues and cationic domains in biological systems continue to increase the awareness and creativity in systems designed to exploit the specificity of those interactions. Doubtless, interesting applications along the lines of synthetic ion channels, affinity screening, and materials development will continue to evolve to the benefit of everyone.

When we move next to the work involving anionic-based receptors for guanidinium- and cationic-functionalized units in general, more work is needed. The groundwork has been laid for an increase in the use of anionic-functionalized receptors to mimic the interactions of closely related biological systems. From the studies covered in this review, success in development of anionic-functionalized receptors that are selective for guanidinium residues over other basic moieties seems more likely than the development of guanidinium-functionalized receptors that can select between the different anionic units discussed in this review. The use of and selectivity provided by dye-derived sulfonate molecules in particular is a testament to the diversity of application available to such interaction schemes. As more of the biological processes involving switches such as phos-

phorylation and sulfation are better understood, a concurrent increase in the number of synthetic recognition schemes will be realized. These synthetic systems will continue to be implemented in novel ways, strengthening important and useful application protocols, such as in screening devices and sensor arrays.

With the methodological toolbox in hand (MS, NMR, calorimetry, molecular modeling, etc.) and continuing to develop, greater insight into synthetic and biological interactions between these complementary units can be envisioned. Spectroscopic, biomimetic, molecular interaction modeling, and bioinformatics tools, among others, are greatly varied and used in a wide variety of fields. The information gained by the biochemist should enhance the productivity of the organic chemist and so on into all fields of science. The sharing of each person's tool of expertise is an important part of the development of any subject but especially in those such as what has been discussed above, because of its inherent applicability and utility across so vast a domain.

5. Acknowledgments

The authors acknowledge the Austrian Science Fund (FWF Project No. P15482) for support of this work. All structures included that have not been redrawn are taken from American Chemical Society journals and, as such, are copyrighted by the ACS in the respective year of the journal from which they are taken.

6. References

(1) Interactions where sulfate, sulfonate, phosphate, and phosphonate groups are employed or observed are all believed to be relevant to this discussion; hence, the coalescent nomenclature "sulf[on]ate" and "phosph[on]ate". Specific discussions will be clarified.
(2) Cox, K. A.; Gaskell, S. J.; Morris, M.; Whiting, A. *J. Am. Soc. Mass Spectrom*. **1996**, *7*, 522.
(3) Gund, P. *J. Chem. Educ*. **1972**, *49*, 100.
(4) Klein, J.; Medlik, A. *J. Chem. Soc., Chem. Commun*. **1973**, 275.
(5) Hannon, C. L.; Anslyn, E. V. *Bioorganic Chemistry Frontiers*; Dugas, H. and Schmidtchen, F. P. Eds.; Springer-Verlag: Berlin and Heidelberg, Germany, 1993; Vol. 3, pp 194−255.
(6) Kanyo, Z. F.; Christianson, D. W. *J. Biol. Chem*. **1991**, *266*, 4264.
(7) Jarrold, M. E. *Annu. Rev. Phys. Chem*. **2000**, *51*, 179.
(8) Pieters, R. J.; Cuntze, J.; Bonnet, M.; Diederich, F. *J. Chem. Soc., Perkin Trans. 2* **1997**, 1891.
(9) Schmidtchen, F. P.; Berger, M. *Chem. Rev*. **1997**, *97*, 1609.
(10) Lehn, J.-M.; Vierling, P.; Hayward, R. C. *J. Chem. Soc., Chem. Commun*. **1979**, 296.
(11) Bell, T. W.; Khasanov, A. B.; Drew, M. G. B.; Filikov, A.; James, T. L. *Angew. Chem., Int. Ed.* **1999**, *38* (27), 2543.
(12) Fessman, T.; Kilburn, J. *Angew. Chem., Int. Ed*. **1999**, *38*, 1993.
(13) Whitlock, B. J.; Whitlock, H. W. *J. Am. Chem. Soc*. **1990**, *112*, 3910.
(14) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom*. **2002**, *216*, 1.
(15) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210.
(16) Etter, M. C.; MacDonald, J. M.; Bernstein, J. *Acta Crystallogr.* **1990**, *B46*, 256.
(17) Bernstein, J. *Acta Crystallogr*. **1991**, *B47*, 1004.
(18) Shimoni, L.; Glusker, J. P. *Protein Sci*. **1995**, *4*, 65.
(19) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: New York, 1997.
(20) Williams, D. H.; Westwall, M. S. *Chem. Soc. Rev*. **1998**, *27*, 57.
(21) García-Pérez, M.; Pinto, M.; Subirana, J. A. *Biopolymers* **2003**, *69*, 432.
(22) Ma, J. C.; Dougherty, D. A. *Chem. Rev*. **1997**, *97*, 1303.
(23) Gallivan, J. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A*. **1999**, *96*, 9459.
(24) Zacharias, N.; Dougherty, D. A. *Trends Pharmacol. Sci*. **2002**, *23*, 281.
(25) Burley, S. K.; Petsko, G. A. *FEBS Lett*. **1986**, *203*, 139.
(26) Luo, R.; David, L.; Hung, H.; Devaney, J.; Gilson, M. K. *J. Phys. Chem. B* **1999**, *103*, 727.
(27) Springs, B.; Haake, P. *Bioorg. Chem*. **1977**, *6*, 181.
(28) Kato, Y.; Conn, M. M.; Rebek, J., Jr. *J. Am. Chem. Soc*. **1994**, *116*, 3279.
(29) O=Chem Solvents, http://www.usm.maine.edu/~newton/Chy251_253/Lectures/Solvents/Solvents.html (accessed September 2003).
(30) Best, M. D.; Tobey, S. L.; Anslyn, E. V. *Coord. Chem. Rev*. **2003**, *240*, 3.
(31) Ngola, S. M.; Kearney, P. C.; Mecozzi, S.; Russell, K.; Dougherty, D. A. *J. Am. Chem. Soc*. **1999**, *121*, 1192.
(32) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2749.
(33) Dunitz, J. D. *Chem. Biol*. **1995**, *2*, 709.
(34) Williams, D. H.; Weetwell, M. S. *Chem. Soc. Rev*. **1998**, *27*, 57.
(35) Douteau-Guével, N.; Coleman, A. W.; Morel, J.-P.; Morel-Desrosiers, N. *J. Chem. Soc., Perkin Trans. 2* **1999**, 629.
(36) Schneider, H.-J. In *Methods and Principles in Medicinal Chemistry*; Böhm, H.-J., Schneider, G., Eds.; Wiley-VCH: Weinheim, Germany, 2003; Vol. 19, p 21.
(37) Frigyes, D.; Alber, F.; Pongor, S.; Carloni, P. *J. Mol. Struct. (THEOCHEM)* **2001**, *574*, 39.
(38) Swaminathan, K.; Flynn, P.; Reece, R. J.; Marmorstein. R. *Nat. Struct. Biol*. **1997**, *4*, 751.
(39) Parraga, A.; Bellsolell, L.; Ferre-D'Amare, A. R.; Burley, S. K. *Structure* **1998**, *6*, 661.
(40) Mason, P. E.; Neilson, G. W.; Dempsey, C. E.; Barnes, A. C.; Cruickshank, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4557.
(41) Webb, T. H.; Wilcox, C. S. *Chem. Soc. Rev*. **1993**, *22*, 383.
(42) Rebek, J., Jr. *Angew. Chem., Int. Ed. Engl*. **1990**, *29*, 245.
(43) Pirkle, W. H.; Pochapsky, T. C. *Chem. Rev*. **1989**, *89*, 347.
(44) Sogah, G. D. Y.; Cram, D. J. *J. Am. Chem. Soc*. **1979**, *101*, 3035.
(45) Maier, N. M.; Franco, P.; Lindner, W. *J. Chromatogr., A* **2001**, *906*, 3.
(46) Lustenberger, P.; Martinborough, E.; Denti, T. M.; Diederich, F. *J. Chem. Soc., Perkin Trans. 2* **1998**, 747.
(47) Cuntze, J.; Owens, L.; Alcazar, V.; Seiler, P.; Diederich, F. *Helv. Chim. Acta* **1995**, *78*, 367.
(48) Soman, J.; Vijayan, M.; Ramakrishnan, B.; Guru Row, T. N. *Biopolymers* **1989**, *29*, 533.
(49) Grawe, T.; Schrader, T.; Finocchiaro, P.; Failla, S.; Consiglio, G. 4th International Electronic Conference on Synthetic Organic Chemistry (ECSOC-4) **2000**, www.unibas.ch/mdpi/ecsoc-4/c0026/c0026.htm (accessed February 2003).
(50) Wehner, M.; Schrader, T.; Finocchiaro, P.; Failla, S.; Consiglio, G. *Org. Lett*. **2000**, *2*, 605.
(51) Linton, B. R.; Goodman, M. S.; Fan, E.; van Arman, S. A.; Hamilton, A. D. *J. Org. Chem.* **2001**, *66*, 7313.
(52) Berger, M.; Schmidtchen, F. P. *J. Am. Chem. Soc*. **1999**, *121*, 9986.
(53) Wittekind, M.; Mapelli, C.; Lee, V.; Goldfarb, V.; Friedrichs, M. S.; Meyers, C. A.; Mueller, L. *J. Mol. Biol*. **1997**, *267*, 933.
(54) Meyer, B.; Peters, T. *Angew. Chem., Int. Ed.* **2003**, *42*, 864.
(55) Czekalla, M.; Stephan, H.; Habermann, B.; Trepte, J.; Gloe, K.; Schmidtchen, F. P. *Thermochim. Acta* **1998**, *313*, 137.
(56) Berger, M.; Schmidtchen, F. P. *Angew. Chem., Int. Ed.* **1998**, *37*, 2694.
(57) Haj-Zaroubi, M.; Mitzel, N. W.; Schmidtchen, F. P. *Angew. Chem., Int. Ed.* **2002**, *41*, 104.
(58) Jelesarov, I.; Bosshard, H. R. *J. Mol. Recognit*. **1999**, *12*, 3.
(59) Dobberstein, P.; Korte, E.; Meyerhoff, G.; Pesch, R. *Int. J. Mass Spectrom. Ion Phys*. **1983**, *46*, 185.
(60) Sawada, M. *Mass Spectrom. Rev*. **1997**, *16*, 73.
(61) De Brouwer, A. P. M.; Bouma, B., van Tiel, C. M.; Heerma, W.; Brouwers, J. H. F. M.; Bevers, L. E.; Westerman, J.; Roelofsen, B.; Wirtz, K. W. A. *Chem. Phys. Lipids* **2001**, *112*, 109.
(62) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. F. *Science* **1989**, *246*, 64.
(63) Kempen, E. C.; Brodbelt, J. S. *Anal. Chem*. **2000**, *72*, 5411.
(64) Kebarle, P. *J. Mass Spectrom*. **2000**, *35*, 804.
(65) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Phys*. **1987**, *78*, 53.
(66) Karas, M.; Hillenkamp, F. *Anal. Chem*. **1988**, *60*, 2299.
(67) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. *Rapid Commun. Mass Spectrom*. **1988**, *2*, 151.
(68) Schalley, C. A. *Int. J. Mass Spectrom*. **2000**, *194*, 11.
(69) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc*. **1991**, *113*, 6294.
(70) Shields, S. J.; Oyeyemi, O.; Lightstone, F. C.; Balhorn, R. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 460.
(71) Loo, J. A. *Int. J. Mass Spectrom*. **2000**, *200*, 175.
(72) Wu, Q.; Gao, J.; Joseph-McCarthy, D.; Sigal, G. B.; Bruce, J. E.; Whitesides, G. M.; Smith, R. D. *J. Am. Chem. Soc*. **1997**, *119*, 1157.

(73) Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc*. **1998**, *120*, 12956.
(74) Zhang, D. X.; Wu, L. M.; Koch, K. J.; Cooks, R. G. *Eur. J. Mass Spectrom*. **1999**, *5*, 353.
(75) Julian, R. R.; Hodyss, R.; Beauchamp, J. L. *J. Am. Chem. Soc*. **2001**, *123*, 3577.
(76) Julian, R. R.; Beauchamp, J. L.; Goddard, W. A., III. *J. Phys. Chem. A* **2002**, *106*, 32.
(77) Ballantine, J. A.; Games, D. E.; Slater, P. S. *Rapid Commun. Mass Spectrom*. **1995**, *9*, 1403.
(78) Schug, K.; Lindner, W. *Int. J. Mass Spectrom*. **2004**, *235*, 213.
(79) Schug, K.; Lindner, W. *Int. J. Mass Spectrom*., submitted for publication.
(80) Fitzgerald, M. C.; Parr, G. R.; Smith, L. M. *Anal. Chem*. **1993**, *65*, 3204.
(81) Krüger, R.; Karas, M. *J. Am. Soc. Mass Spectrom*. **2002**, *13*, 1218.
(82) Hunter, E. P.; Lias, S. G. In *NIST Chemistry WebBook*; Mallard, W. G., Linstrom P. J., Eds.; NIST Standard Reference Database Number 69; National Institute of Standards and Technology: Gaithersburg MD, 2000; http://webbook.nist.gov.
(83) Shiea, J.; Huang, J.-P.; Teng, C.-F.; Jeng, J.; Wang, L. Y.; Chiang, L. Y. *Anal. Chem*. **2003**, *75*, 3587−3595.
(84) Leitner, A.; Lindner, W. *J. Mass Spectrom*. **2003**, *38*, 891−899.
(85) Cao, S.; Zhang, J.; Xu, J.; Liao, X.; Zhao, Y. *Rapid Commun. Mass Spectrom*. **2003**, *17*, 2237−2240.
(86) Gavezzotti, A. *Synlett* **2002**, 201.
(87) Chakrabarti, P. *Int. J. Protein Pept. Res*. **1994**, *43*, 284.
(88) Lowe, C. R.; Lowe, A. R.; Gupta, G. *J. Biochem. Biophys. Methods* **2001**, *49*, 561−574.
(89) Rossi, M.; Manfredi, V.; Ruvo, M.; Fassina, G.; Verdoliva, A. *Mol. Immunol.* **2002**, *39*, 443−451.
(90) Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. *Annu. Rev. Biophys. Biomol. Struct*. **2001**, *30*, 211.
(91) Jorgensen, W. L. *Acc. Chem. Res*. **1989**, *22*, 184.
(92) McCammon, J. A. *Curr. Opin. Struct. Biol*. **1991**, *1*, 196.
(93) Ziegler, T. *Chem. Rev*. **1991**, *91*, 651−667.
(94) Alcázar, V.; Segura, M.; Prados, P.; de Mendoza, *J. Tetrahedron Lett*. **1998**, *39*, 1033.
(95) Kumar, N. V.; Govil, G. *Biopolymers* **1984**, *23*, 1995.
(96) Gresh, N.; Pullman, B. *Theor. Chim. Acta (Berlin)* **1979**, *52*, 67.
(97) Merdan, T.; Kopeček, J.; Kissel, T. *Adv. Drug Delivery Rev*. **2002**, *54*, 715.
(98) Schrader, T. *J. Org. Chem*. **1998**, *63*, 264.
(99) Schrader, T. *J. Inclusion Phenom. Macrocyclic Chem*. **1999**, *34*, 117.
(100) Nurse, P. *Nature* **1990**, *344*, 503.
(101) Pot, D. A.; Dixon, J. E. *Biochim. Biophys. Acta* **1991**, *1136*, 35.
(102) Fisher, E. H.; Charbonneau, H.; Tonks, N. K. *Science* **1991**, *235*, 401.
(103) Black, R. M.; Read, R. W. *J. Chromatogr., A* **1998**, *794*, 233.
(104) Ojala, W.; Sudbeck, E.; Lu, L.; Richardson, T.; Lovrien, R.; Gleason, W. *J. Am. Chem. Soc*. **1996**, *118*, 2131.
(105) Friess, S. D.; Zenobi, R. *J. Am. Soc. Mass Spectrom*. **2001**, *12*, 810.
(106) Mehrotra, M. M.; Sternbach, D. D.; Rodriguez, M.; Charifson, P.; Berman, J. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 1941.
(107) Bogan, A. S.; Thorn, K. S. *J. Mol. Biol*. **1998**, *280*, 1.
(108) Xie, H.; Yi, S.; Wu, S. *J. Chem. Soc., Perkin Trans. 2* **1999**, 2571.
(109) Gilman, A. G., Goodman, L. S., Gilman, A., Eds. *The Pharmacological Basis of Therapeutics*, 6th ed.; Macmillan Co.: New York, 1980; p 198.
(110) Cotton, F. A.; Day, V. W.; Hazen, E. E., Jr.; Larsen, S. *J. Am. Chem. Soc*. **1973**, *95*, 4834.
(111) Sannes-Lowery, K. A.; Hu, P.; Mack, D. P.; Mei, H.-Y.; Loo, J. A. *Anal. Chem*. **1997**, *69*, 5130.
(112) Woods, A. S.; Huestis, M. A. *J. Am. Soc. Mass Spectrom*. **2001**, *12*, 88.
(113) Rappsilber, J.; Friesen, W. J.; Paushkin, S.; Dreyfuss, G.; Mann, M. *Anal. Chem*. **2003**, *75*, 3107.
(114) Paik, W. K.; Kim, S. J. *J. Biol. Chem*. **1968**, *243*, 2108.
(115) McBride, A. E.; Silver, P. A. *Cell* **2001**, *106*, 5.
(116) Aletta, J. M.; Cimato, T. R.; Ettinger, M. J. *Trends Biochem. Sci*. **1998**, *23*, 89.
(117) Raman, B.; Guarnaccia, C.; Nadassy, K.; Zakhariev, S.; Pintar, A.; Zanuttin, F.; Frigyes, D.; Acatrinei, C.; Vindigni, A.; Pongor, G.; Pongor, S. *Nucleic Acids Res*. **2001**, *29*, 3377.
(118) Liu, Q.; Dreyfuss, G. *Mol. Cell. Biol*. **1995**, *15*, 2800.
(119) McDonald, I. K.; Thornton, J. M. *J. Mol. Biol*. **1994**, *238*, 777.
(120) Gehrig, P. M.; Hunziker, P. E.; Zahariev, S.; Pongor, S. *J. Am. Soc. Mass Spectrom*. **2004**, *15*, 142.
(121) Brame, C. J.; Moran, M. F.; McBroom-Cerajewski, L. D. B. *Rapid Commun. Mass Spectrom*. **2004**, *18*, 877.
(122) Balz, D.; de Souza Wyse, A. T.; Morsch, V. M.; da Silvia, A. C.; Vieira, V. L.; Morsch, A. L. B.; Schetinger, M. R. C. *Int. J. Dev. Neurosci.* **2003**, *21*, 75.
(123) Mori, A. *Pavlov J. Biol. Sci.* **1987**, *22*, 85.
(124) Wierchert, P.; Marescau, B.; De Deyn, P. P.; Lowenthal, A. *Neuroscience* **1987**, *12*, 35.
(125) Smith, J.; Ariga, K.; Anslyn, E. V. *J. Am. Chem. Soc*. **1993**, *115*, 362.
(126) Calnan, B. J.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. D. *Science* **1991**, *252*, 1167.
(127) Muche, M.-S.; Göbel, M. W. *Angew. Chem., Int. Ed. Engl*. **1996**, *35*, 2126.
(128) Echvarren, A.; Galán, A.; Lehn, J. M.; De Mendoza, J. *J. Am. Chem. Soc*. **1989**, *111*, 4994.
(129) Rebek, J. J. *Chemtracts: Org. Chem*. **1990**, *3*, 240.
(130) Müller, G.; Riede, J.; Schmidtchen, F. P. *Angew. Chem., Int. Ed. Engl*. **1988**, *27*, 1516.
(131) Echvarren, A.; Galán, A.; De Mendoza, J.; Salmerón, A.; Lehn, J.-M. *Helv. Chim. Acta* **1988**, *71*, 685.
(132) Corey, E. J.; Ohtani, M. *Tetrahedron Lett.* **1989**, *30*, 5227.
(133) Dietrich, B.; Fyles, D. L.; Fyles, T. M.; Lehn, J.-M. *Helv. Chim. Acta* **1979**, *62*, 2763.
(134) Breccia, P.; Van Gool, M.; Pérez-Fernández, R.; Martín-Santamaría, S.; Gago, F.; Prados, P.; de Mendoza, J. *J. Am. Chem. Soc*. **2003**, *125*, 8270.
(135) Albert, J. S.; Goodman, M. S.; Hamilton, A. D. *J. Am. Chem. Soc*. **1995**, *117*, 1143.
(136) Peczuh, M. W.; Hamilton, A. D.; Sánchez-Quesada, J.; de Mendoza, J.; Haack, T.; Giralt, E. *J. Am. Chem. Soc*. **1997**, *119*, 9327.
(137) Schmidtchen, F. P. *Tetrahedron Lett*. **1989**, *30*, 4493.
(138) Schmidtchen, F. P. *Tetrahedron Lett*. **1990**, *31*, 2269.
(139) Kurzmeier, H.; Schmidtchen, F. P. *J. Org. Chem*. **1990**, *55*, 3749.
(140) Fischer, E. *Chem. Ber*. **1894**, *27*, 2985.
(141) Schmuck, C. *Chem.—Eur. J*. **2000**, *6*, 709.
(142) Schmuck, C. *Chem. Commun*. **1999**, 843.
(143) Schmuck, C. *Eur. J. Org. Chem*. **1999**, 2397.
(144) Schmuck, C. *J. Org. Chem*. **2000**, *65*, 2432.
(145) Schmuck, C.; Wienand, W. *J. Am. Chem. Soc*. **2003**, *125*, 452.
(146) Nishizawa, S.; Kato, Y.; Teramae, N. *J. Am. Chem. Soc*. **1999**, *121*, 9463.
(147) Schmuck, C.; Heil, M.; Dechantstreiter, M.; Hackler, U. E. *J. Assoc. Lab. Autom*. **2001**, *6*, 51.
(148) Schmuck, C. *Tetrahedron* **2001**, *57*, 3063.
(149) Schmuck, C.; Heil, M. *Org. Biomol. Chem.* **2003**, *1*, 633.
(150) Schmuck, C.; Heil, M. *ChemBioChem* **2003**, *4*, 1232−1238.
(151) Fernández-Carneado, J.; Kogan, M. J.; Castel, S.; Giralt, E. *Angew. Chem., Int. Ed.* **2004**, *43*, 1811.
(152) Jensen, K. B.; Braxmeier, T. M.; Demarcus, M.; Frey, J. G.; Kilburn, J. D. *Chem.—Eur. J.* **2002**, *8*, 1300.
(153) Peczuh, M. W.; Hamilton, A. D. *Chem. Rev*. **2000**, *100*, 2479.
(154) Chen, C.-W.; Whitlock, H. W., Jr. *J. Am. Chem. Soc*. **1978**, *100*, 4921.
(155) Janout, V.; Jing, B.; Staina, I. V.; Regen, S. L. *J. Am. Chem. Soc*. **2003**, *125*, 4436.
(156) Liu, F.; Lu, G.-Y.; He, W.-J.; Liu, M.-H.; Zhu, L.-G.; Wu, H.-M. *New J. Chem*. **2002**, *5*, 601.
(157) Kobiro, K.; Inoue, Y. *J. Am. Chem. Soc*. **2003**, *125*, 421.
(158) Rekharsky, M.; Inoue, Y.; Tobey, S.; Metzger, A.; Anslyn, E. *J. Am. Chem. Soc*. **2002**, *124*, 14959.
(159) Niikura, K.; Metzger, A.; Anslyn, E. V. *J. Am. Chem. Soc*. **1998**, *120*, 8533.
(160) O'Rourke, F.; Feinstein, M. B. *Biochem. J*. **1990**, *267*, 297.
(161) Tobey, S. L.; Anslyn, E. V. *J. Am. Chem. Soc*. **2003**, *125*, 14807.
(162) Fassina, G.; Ruvo, M.; Polombo, G.; Verdoliva, A.; Marino, M. *J. Biochem. Biophys. Methods* **2001**, *49*, 481−490.
(163) Ruvo, M.; Scardino, P.; Cassani, G.; Fassina, G. *Protein Pept. Lett*. **1994**, *1*, 187−192.
(164) Gupta, G.; Lowe, C. R. In *Encyclopedia of Separation Science*; Wilson, I. D., Poole, C. F., Adlard, T. R., Cooks, M., Eds.; Academic Press: New York, 2000; Vol. 2, p 297.
(165) Lowe, C. R.; Burton, S. J.; Burton, N. P.; Alderton, W. K.; Pitts, J. M.; Thomas, J. A. *Trends Biotechnol*. **1992**, *10*, 442−448.
(166) Bayley, H.; Martin, C. R. *Chem. Rev*. **2000**, *100*, 2575.
(167) Bayley, H.; Cremer, P. S. *Nature* **2001**, *413*, 226.
(168) Cheley, S.; Gu, L.-Q.; Bayley, H. *Chem. Biol*. **2002**, *9*, 829.
(169) Sakai, N.; Brennan, K. C.; Weiss, L. A.; Matile, S. *J. Am. Chem. Soc*. **1997**, *119*, 8726.
(170) Sakai, N.; Gerard, D.; Matile, S. *J. Am. Chem. Soc*. **2001**, *123*, 2517.
(171) Sorde, N.; Matile, S. *J. Supramol. Chem*. **2002**, *2*, 191.
(172) Sakai, N.; Sorde, N.; Matile, S. *J. Am. Chem. Soc*. **2003**, *125*, 7776.
(173) Sakai, N.; Houdebert, D.; Matile, S. *Chem.—Eur. J*. **2003**, *9*, 223.
(174) Som, A.; Sakai, N.; Matile, S. *Bioorg. Med. Chem*. **2003**, *11*, 1363.
(175) Sakai, N.; Sordé, N.; Das, G.; Perrottet, P.; Gerard, D.; Matile, S. *Org. Biomol. Chem*. **2003**, *1*, 1226.
(176) Talukdar, P.; Sakai, N.; Sordé, N.; Gerard, D.; Cardona, V. M. F.; Matile, S. *Bioorg. Med. Chem*. **2004**, *12*, 1325.
(177) Sakai, N.; Matile, S. *Chirality* **2004**, *16*, S28−S35.
(178) Numata, M.; Sugiyasu, K.; Hasegawa, T.; Shinkai, S. *Angew. Chem., Int. Ed*. **2004**, *43*, 3279.
(179) Loo, J. A.; Hu, P.; McConnell, P.; Mueller, W. T.; Sawyer, T. K.; Thanabal, V. *J. Am. Soc. Mass Spectrom*. **1997**, *8*, 234.

(180) Periana, R. A.; Motiu-DeGrood, R.; Chiang, Y.; Hupe, D. J. *J. Am. Chem. Soc*. **1980**, *102*, 3923.
(181) Gouaux, J. E.; Krause, K. L.; Lipscomb, W. N. *Biochem. Biophys. Res. Commun*. **1986**, *142*, 893.
(182) Bender, S. L.; Widlanski, T.; Knowles, J. R. *Biochemistry* **1989**, *28*, 7560.
(183) Widlanski, T.; Bender, S. L.; Knowles, J. R. *Biochemistry* **1989**, *28*, 7572.
(184) Burke, T. R., Jr.; Smyth, M. S.; Otaka, A.; Nomizu, M.; Roller, P. R.; Wolf, G.; Case, R.; Shoelson, S. E. *Biochemistry* **1994**, *33*, 6490.
(185) Gilmer, T.; Rodriguez, M.; Jordan, S.; Crosby, R.; Alligood, K.; Green, M.; Kimery, M.; Wagner, C.; Kinder, D.; Charifson, P.; Hassell, A. M.; Willard, D.; Luther, M.; Rusnak, D.; Sternbach, D. D.; Mehrotra, M.; Peel, M.; Shampine, L.; Davis, R.; Robbins, J.; Patel, I. R.; Kassel, D.; Burkhart, W.; Moyer, M.; Bradshaw, T.; Berman, J. *J. Biol. Chem*. **1994**, *269*, 31711.
(186) Waksman, G.; Kominos, D.; Robertson, S. C.; Pant, N.; Baltimore, D.; Birge, R. B.; Cowburn, D.; Hanafusa, H.; Mayer, B. J.; Overduin, M.; Resh, M. D.; Rios, C. B.; Silverman, L.; Kuriyan, J. *Nature* **1992**, *358*, 646.
(187) Xu, R. X.; Word, J. M.; Davis, D. G.; Rink, M. J.; Willard, D. H., Jr.; Gampe, R. T., Jr. *Biochemistry* **1995**, *34*, 2107.
(188) Catalina, M. I.; Dekker: F. J.; Liskamp, R. M.; Versluis, C.; Maier, C. S.; Heck, A. J. R. *Int. J. Mass Spectrom*. **2003**, *228*, 879.
(189) Mavri, J.; Vogel, H. J. *Proteins: Struct., Funct., Genet*. **1996**, *24*, 495.
(190) Deerfield, D. W., II; Nicholas, H. B., Jr.; Hiskey, R. G.; Pedersen, L. G. *Proteins: Struct., Funct., Genet*. **1989**, *6*, 168.
(191) Last, A. M.; Robinson, C. V. *Curr. Opin. Chem. Biol*. **1999**, *3*, 564.
(192) Shahripour, A.; Plummer, M. S.; Lunney, E. A.; Para, K. S.; Stankovic, C. J.; Rubin, J. R.; Humblet, C.; Fergus, J. H.; Marks, J. S.; Herrera, R.; Hubbell, S. E.; Saltiel, A. R.; Sawyer, T. K. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 1209.
(193) Chen, X.-L.; Qu, L.-B.; Zhang, T.; Liu, H.-X.; Yu, F.; Yu, Y.; Liao, X.; Zhao, Y.-F. *Anal. Chem*. **2004**, *76*, 211.
(194) Liehr, S.; Chenault, H. K. *Bioorg. Med. Chem. Lett*. **1999**, *9*, 2759.
(195) McKeever, B. M.; Hawkins, B. H.; Geissler, W. M.; Wu, L.; Sheridan, R. P.; Mosley, R. T.; Andersson, S. *Biochim. Biophys. Acta*. **2002**, *1601*, 29.
(196) Moyer, S. C.; Cotter, R. J.; Woods, A. S. *J. Am. Soc. Mass Spectrom*. **2002**, *13*, 274.
(197) Tholey, A.; Reed, J.; Lehman, W. D. *J. Mass Spectrom.* **1999**, *34*, 117.
(198) DeGnore, J. P.; Qin, J. *J. Am. Soc. Mass Spectom*. **1998**, *9*, 1175.
(199) Busman, M.; Schey, K. L.; Oatis, J. E.; Knapp, D. R. *J. Am. Soc. Mass Spectrom*. **1996**, *7*, 243.
(200) Failla, S.; Finocchiaro, P.; Consiglio, G. A. *Heteroat. Chem*. **2000**, *11*, 493.
(201) Hammerschmidt, F.; Wunnenig, F. *Tetrahedron: Asymmetry* **1999**, *10*, 1709.
(202) Zarbl, E.; Lämmerhofer, M.; Hammerschmidt, F.; Wuggenig, F.; Hanbauer, M.; Maier, N. M.; Sajovic, L.; Lindner, W. *Anal. Chim. Acta* **2000**, *404*, 169.
(203) Palacios, F.; Aparicio, D.; De Retana, A. M. O.; De los Santos, J. M.; Gil, J. I.; De Munain, R. L. *Tetrahedron: Asymmetry* **2003**, *14*, 689.
(204) Kaboudin, B. *Tetrahedron Lett*. **2003**, *44*, 1051.
(205) Rensing, S.; Arendt, M.; Springer, A.; Grawe, T., Schrader, T. *J. Org. Chem*. **2001**, *66*, 5814.
(206) Schrader, T. *Chem.—Eur. J*. **1997**, *3*, 1537.
(207) Schrader, T. H. *Tetrahedron Lett*. **1998**, *39*, 517.
(208) Schrader, T. *J. Am. Chem. Soc*. **1998**, *120*, 11816.
(209) Rensing, S.; Schrader, T. *Org. Lett*. **2002**, *4*, 2161.
(210) Frølund, B.; Tagmose, L.; Liljefors, T.; Stensbøl, T. B.; Engblom, C.; Kristiansen, U.; Krogsgaard-Larsen, P. J. *J. Med. Chem*. **2000**, *43*, 4930.
(211) Ebert, B.; Mortensen, M.; Thompson, S. A.; Kehler, J.; Wafford, K. A.; Krogsgaard-Larsen, P. *Bioorg. Med. Chem. Lett*. **2001**, *11*, 1573.
(212) Julian, R. R.; Akin, M.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L. *Int. J. Mass Spectrom*. **2002**, *220*, 87.
(213) Julian, R. R.; Beauchamp, J. L. *J. Am. Soc. Mass Spectrom*. **2004**, *15*, 616.
(214) Eliseev, A. V.; Nelen, M. I. *J. Am. Chem. Soc*. **1997**, *119*, 1147.
(215) Friedrich, U.; Blom, A. M.; Dahlbäck, B.; Villoutreix, B. O. *J. Biol. Chem*. **2001**, *276*, 24122.
(216) Caldwell, E. E. O.; Nadkarni, V. D.; Fromm, J. R.; Linhardt, R. J.; Weiler, J. M. *Int. J. Biochem. Cell Biol*. **1996**, *28*, 203.
(217) Byun, Y.; Singh, V. K.; Yang, V. C. *Thromb. Res*. **1999**, *94*, 53.
(218) Pecon, J. M.; Blackburn, M. N. *J. Biol. Chem*. **1984**, *259*, 935.
(219) Smith, J. W.; Knauer, D. J. *J. Biol. Chem.* **1987**, *262*, 11964.
(220) Cardin, A. T.; Weintraub, H. J. *Arteriosclerosis* **1989**, *9*, 21.
(221) Montserret, R.; Aubert-Foucher, E.; McLeish, M. J.; Hill, J. M.; Ficheux, D.; Jaquinod, M.; van der Rest, M.; Deleage, G.; Penin, F. *Biochemistry* **1999**, *38*, 6479.
(222) Sali, A.; Matsumoto, R.; McNeil, H. P.; Karplus, M.; Stevens, R. L. *J. Biol. Chem*. **1993**, *268*, 9023.
(223) Matsumoto, R.; Sali, A.; Ghildyal, N.; Karplus, M.; Stevens, R. L. *J. Biol. Chem*. **1995**, *270*, 19524.
(224) Murakami, M.; Matsumoto, R.; Urade, Y.; Austen, K. F.; Arm, J. P. *J. Biol. Chem*. **1995**, *270*, 3239.
(225) Ghildyal, N.; Friend, D. S.; Stevens, R. L.; Austen, K. F.; Huang, C.; Penrose, J. F.; Sali, A.; Gurish, M. F. *J. Exp. Med*. **1996**, *184*, 1061.
(226) Mann, D. M.; Romm, E.; Migliorini, M. *J. Biol. Chem*. **1994**, *269*, 23661.
(227) Busby, T. F.; Argraves, W. S.; Brew, S. A.; Pechik, I.; Gillilang, G. L.; Ingham, K. C. *J. Biol. Chem*. **1995**, *270*, 18558.
(228) Miyamoto, K.; Kobayashi, D.; Maeda, R.; Ito, T.; Komai, T. *Int. J. Biol. Macromol.* **2003**, *31*, 207.
(229) http://www.haemtech.com/plat_08_vitro.html (accessed November 2004).
(230) Gibson, A. D.; Lamerdin, J. A.; Zhuang, P.; Baburaj, K.; Serpersu, E. H.; Peterson, C. B. *J. Biol. Chem*. **1999**, *274*, 6432.
(231) Peterson, C. B.; Blackburn, M. N. *J. Biol. Chem*. **1985**, 610.
(232) Zhuang, P.; Chen, A. I.; Peterson, C. B. *J. Biol. Chem*. **1997**, *272*, 6858.
(233) Fromm, J. R.; Hileman, R. E.; Caldwell, E. E. O.; Weiler, J. M.; Linhardt, R. J. *Arch. Biochem. Biophys*. **1995**, *323*, 279.
(234) McCabe, D.; Cukierman, T.; Gabay, J. E. *J. Biol. Chem*. **2002**, *277*, 27477.
(235) Iversen, L. F.; Kastrup, J. S.; Bjorn, S. E.; Rasmussen, P. B.; Wiberg, F. C.; Flodgaard, H. J.; Larsen, I. K. *Nat. Struct. Biol*. **1997**, *4*, 265.
(236) Karlsen, S.; Iversen, L. F.; Larsen, I. K.; Flodgaard, H. J.; Kastrup, J. S. *Acta Crystallogr., Sect. D: Biol. Crystallogr*. **1998**, *54*, 598.
(237) Juhasz, P.; Biemann, K. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4333.
(238) Sullivan, A. G.; Gaskell, S. J. *Rapid Commun. Mass Spectrom*. **1997**, *11*, 803.
(239) Holčapek, M.; Jandera, P.; Přikryl, J. *Dyes Pigm*. **1999**, *43*, 127.
(240) Salih, B.; Zenobi, R. *Anal. Chem*. **1998**, *70*, 1536.
(241) Friess, S. D.; Daniel, J. M.; Hartmann, R.; Zenobi, R. *Int. J. Mass Spectrom*. **2002**, *219*, 269.
(242) Denizli, A.; Pişkin, E. *J. Biochem. Biophys. Methods* **2001**, *49*, 391.
(243) Russell, V. A.; Etter, M. C.; Ward, M. D. *J. Am. Chem. Soc*. **1994**, *116*, 1941.
(244) Suckau, D.; Mak, M.; Przybylski, M. *Proc. Natl. Acad. Sci. U.S.A*. **1992**, *89*, 5630.
(245) Glocker, M. O.; Borchers, C.; Fiedler, W.; Suckau, D.; Przybylski, M. *Bioconjugate Chem*. **1994**, *5*, 583.
(246) Kimmel, J. R. *Methods Enzymol*. **1967**, *11*, 584.
(247) Ohgushi, M.; Wada, A. *FEBS Lett*. **1983**, *164*, 21.
(248) Spector, S.; Young, P.; Raleigh, D. P. *Biochemistry* **1999**, *38*, 4128.
(249) Matulis, D.; Lovrein, R. *Biophys. J*. **1998**, *74*, 422.
(250) Hamdan, M.; Curcuruto, O.; Molinari, H.; Zetta, L.; Ragona, L. *J. Mass Spectrom*. **1996**, *31*, 1261.
(251) Ray, S. S.; Balaram, P. J. *J. Phys. Chem. B* **1999**, *103*, 7068.
(252) Dif, K.; Pepe, C.; Lavedrine, B. *Rapid Commun. Mass Spectrom*. **2003**, *17*, 1002.
(253) Ali, V.; Prakash, K.; Kulkarni, S.; Ahmad, A.; Madhusudan, K. P.; Bhakuni, V. *Biochemistry* **1999**, *38*, 13635.
(254) Tan, Y.-J.; Oliverberg, M.; Davis, B.; Fersht, A. R. *J. Mol. Biol*. **1995**, *254*, 980.
(255) Douteau-Guével, N.; Perret, F.; Coleman, A. W.; Morel, J.-P.; Morel-Desrosiers, N. *J. Chem. Soc., Perkin Trans. 2* **2002**, 524.
(256) Fang, D.-C.; Fabian, P.; Szekely, Z.; Fu, X.-Y.; Tang, T.-H.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1998**, *430*, 161.
(257) Asthagiri, D.; Schure, M. R.; Lenhoff, A. M. *J. Phys. Chem. B* **2000**, *104*, 8753.
(258) Gunay, N. S.; Tadano-Aritomi, K.; Toida, T.; Ishizuka, I.; Linhardt, R. J. *Anal. Chem*. **2003**, *75*, 3226.
(259) Saad, O. M.; Leary, J. A. *Anal. Chem*. **2003**, *75*, 2985.

CR040603J