Aromatic Amino Acids and Related Substances: Chemistry, Biology, Medicine, and Application

Cation-π Interactions Involving Aromatic Amino Acids1–4

Dennis A. Dougherty *

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

Abstract

The cation-π interaction is a general, strong, noncovalent binding force that is used throughout nature. The side chains of the aromatic amino acids [phenylalanine (Phe), tyrosine (Tyr), and tryptophan (Trp)] provide a surface of negative electrostatic potential that can bind to a wide range of cations through a predominantly electrostatic interaction. In this brief overview, the fundamental nature of the cation-π interaction will be described, relying on fundamental, gas phase studies of the effect. Then, several examples of cation-π interactions involving aromatic amino acids will be described. These include contributions to protein secondary structure, in which Phe/Tyr/Trp-lysine (Lys) interactions are common. We will also describe several examples of protein-ligand interactions that make use of cation-π interactions. We will place special emphasis on the binding of quaternary ammonium ions, such as trimethylated Lys and the neurotransmitter acetylcholine. J. Nutr. 137: 1504S–1508S, 2007.

Aromatic amino acids have unique and important properties. Phenylalanine (Phe)1, tyrosine (Tyr), and tryptophan (Trp) are generally hydrophobic, but compared with simpler hydrophobic residues, such as leucine or valine, the aromatic amino acids have additional capabilities. Both Tyr and Trp can contribute hydrogen bonds, an important feature. However, there is another reason that aromatic amino acids are substantially overrepresented at protein binding sites and that Trp, in particular, is the most conserved of all amino acids. That is the cation-π interaction, a strong, noncovalent binding interaction that contributes to protein secondary structure and to diverse receptor interactions.

Here we present a brief overview of the cation-π interaction. More detailed reviews are available elsewhere. (1–4) We begin with a description of the fundamental nature of the interaction, emphasizing gas phase studies. Then we describe the contribution of cation-π interactions to protein secondary structure and recent examples of cation-π interactions in the binding of ligands to diverse proteins. It is clear from such studies that aromatic amino acids can make special contributions in a number of ways, including the cation-π interaction.

Fundamentals of the cation-π interaction

Figure 1 summarizes several studies of the cation-π interactions in the gas phase (5–8). Two features of these results are important. First, these are very large binding energies for a clearly noncovalent binding interaction. Importantly, pioneering work by Kebarle in 1981 (6) measured not only the benzene···K⁺ interaction but also the water···K⁺ interaction. Everyone would agree that water is a potent ligand for ions, and it is, binding K⁺ with a ΔH of 18 kcal/mol. Remarkably, benzene binds the K⁺ ion more tightly than water, and this is the first indication that the cation-π interaction is a potentially important binding force.

The second feature of Figure 1 is that the results follow a classical electrostatic trend, much as one would see in aqueous solvation energies or crystal lattice energies (5). That is, smaller ions with more focused charges have the larger affinity. These results, and many more, have led us to advocate a primarily electrostatic model for the cation-π interaction (9). While it is certainly true that van der Waals and polarization effects contribute to the cation-π interaction, the defining feature is electrostatic. In fact, most observations concerning the cation-π interaction (and other noncovalent interactions involving simple aromatics) (10) can be rationalized by the following observation: sp² carbon is more electronegative than hydrogen. As shown in Figure 2, this creates 6 local C⁶–H⁶⁺ bond dipoles around the benzene ring. When summed, these 6 bond dipoles create an overall charge distribution that is a build-up of negative charge in the center of the ring and a belt of positive charge around the edge.

---

1 Published in a supplement to The Journal of Nutrition. Presented at the “Conference on Aromatic Amino Acids and Related Substances: Chemistry, Biology, Medicine, and Application” held July 20–21, 2006 in Vancouver, Canada. The conference was sponsored by Ajinomoto Company, Inc. The organizing committee for the symposium and Guest Editors for the supplement were: Katsuji Takai, Dennis M. Bier, Luc Cynober, Sidney M. Morris, Jr., and Yoshitaru Shimomura. Guest Editor disclosure: Expenses to travel to the meeting were paid by Ajinomoto Company, Inc. for K. Takai, D. M. Bier, L. Cynober, S. M. Morris, Jr., and Y. Shimomura; D. M. Bier has consulted for Ajinomoto Company, Inc. on scientific issues.

2 Supported by the NIH (NS 34407).

3 Author disclosures: D. A. Dougherty, The Ajinomoto Company, Inc. paid the author’s travel expenses to the TICAAA meeting.

4 Color versions of Figures 4 and 5 are available with the online posting of this paper at jn.nutrition.org.

5 Abbreviations used: ACh, acetylcholine; AChBP, acetylcholine-binding protein; AChE, acetylcholine esterase; Arg, arginine; Cys, cysteine; GABA, γ-amino butyric acid; His, histidine; Lys, lysine; nAChR, nicotinic acetylcholine receptor; Phe, phenylalanine; Trp, tryptophan; Tyr, tyrosine.

* To whom correspondence should be addressed. E-mail: dadougherty@caltech.edu.
The colorful electrostatic potential surface of Figure 2 results from high level quantum mechanical calculations and shows the build-up of negative electrostatic potential (red) in the center of the ring. This is the region to which cations bind and it is the origin of the cation-π interaction. Quite simply, cations, with their positive charge, are attracted to the negative electrostatic potential over the face to the benzene ring. Importantly, the cation must interact with the face, not the edge, of the ring.

A few other general features of the cation-π interaction deserve mentioning. First, whereas polarizability of the ring may contribute to the magnitude of the binding, polarizability is not the defining feature of the cation-π interaction. For example, cyclohexane is substantially more polarizable than benzene, but it is a decidedly weaker binder of cations than benzene. Second, the aromaticity of benzene is also not a factor. Simple π systems such as ethylene and acetylene also bind cations and the origin of the effect is the same; sp² (and sp) carbon is more electronegative than hydrogen. It is a cation-π interaction, not a cation-aromatic interaction.

Finally, when considering aromatic amino acids, electrostatics are again useful. Electrostatic potential maps of the sort shown in Figure 2 have been compared for benzene (corresponding to Phe), phenol (Tyr), and indole (Trp) (11). Interestingly, benzene and phenol are very similar to each other in this comparison, indicating that the fundamental cation-π-binding abilities of Phe and Tyr are similar. However, if the OH of Tyr is involved in a hydrogen bond as a donor, the cation-π-binding ability of Tyr increases (12). Indole shows much greater negative electrostatic potential over its ring and its cation-π-binding ability is correspondingly greater. As we will see, this carries over to the aromatic amino acids, because Trp is the optimal partner for a cation-π interaction.

**Contributions to protein secondary structure**

In proteins, cation-π interactions can arise between Phe/Tyr/Trp as the π component and lysine (Lys)/arginine (Arg) as the cation (Fig. 3). Because of the diverse structures of the 6 possible pairwise interactions, evaluating the frequency of cation-π interactions in protein structures can be challenging. Early work by Burley and Petsko (13) provided important insights but may have produced unreliable statistics. Therefore, we devised a novel approach to enumerate possible cation-π interactions in proteins (12). Using a combination of ab initio quantum mechanical calculations and modern force field methods, we developed an energy-based method to evaluate potential cation-π interactions. We feel this is a much better way to establish whether such interactions are or are not significant.

The results of this study were that cation-π interactions, that is Phe/Tyr/Trp···Lys/Arg, are common in proteins. In fact, on average there is 1 cation-π interaction for every 77 amino acids in the protein data bank. As a result, essentially all proteins of significant size have at least 1 cation-π interaction. Of the cations, Arg is more frequently used than Lys. Most impressively, over 25% of all Trp residues are involved in a cation-π interaction to a Lys or Arg. For a collection of images of representative cation-π interactions in proteins, see: http://www.ccco.caltech.edu/~dadgrp/research/cation-pi.pdf.

**Cation-π interactions in protein-ligand interactions**

Over the past decade, innumerable examples of protein-ligand interactions that use cation-π interactions have been described. Most rely on crystallography, which reveals a clear contact between a cationic center and the face of a Phe, Tyr, or Trp. Some caution should be exercised in such cases, because the observation of a molecular contact in a crystal structure provides no information on the energetic consequences of the interaction. In favorable cases, structure-function studies establish a clearly important interaction. Still, in some cases, the X-ray evidence alone is quite compelling. We note that a large number of drug-receptor interactions involve cation-π interactions and workers
interested in designing/optimizing pharmaceutical structures should especially be aware of the cation-π interaction. Here we cite just a few examples of aromatic amino acids participating in cation-π interactions.

All types of cations are known to participate in cation-π interactions. However, because aromatics are innately hydrophobic, we might expect that interactions with more nearly hydrophobic cations would be more common. For example, the preference for Arg over Lys, noted above, can in part be rationalized by the fact that the face of the guanidinium ion of Arg, being a delocalized π system, is to some extent hydrophobic. Another type of cation that is relatively hydrophobic is a quaternary ammonium ion, such as RNMe₃⁺. Indeed, we find that there are many compelling examples of strong cation-π interactions involving such groups.

Chromatin is the ~1:1 mixture of protein:DNA that defines the highly compacted form of DNA in chromosomes. The major proteins of chromatin are histones, named H1, H2, H3, etc. Surface-exposed Lys residues of histones undergo a post-translational modification that involves methylation of the terminal amine, producing alkylated ammonium ions of the type RNMe₃⁺ (also RNHMe₂⁺ and RNH₂Me⁺). Trimethylated (but not dimethylated) Lys marks the transcriptional start site of eukaryotic genes. Trimethylated Lys recruits proteins to the chromosome and several crystal structures have clearly established that cation-π interactions play an essential role in the binding a trimethylated Lys by these transcriptional activators (14,15). A representative image is shown in Figure 4. It shows Lys 4 of histone H3, trimethylated and binding to a nucleosome remodeling factor called BPTF. The ammonium ion is completely surrounded by aromatics, leaving no doubt that cation-π interactions are important in the binding. The binding arrangement is quite reminiscent of interactions we documented almost 20 y ago between RNMe₃⁺ groups and artificial cyclophane receptors that we developed and were critical in the early establishment of the cation-π interaction. (16,17)

The most pervasive RNMe₃⁺ group in nature is that of acetylcholine (ACh), the longest known, most-studied neurotransmitter. Based on our studies of cyclophane receptors, which bind ACh with strengths comparable to those of natural receptors, we predicted in 1990 that ACh-binding sites would be rich in aromatic amino acids and would make use of cation-π interactions (18). Since then, that prediction has been confirmed repeatedly.

The first natural ACh-binding site to be revealed was that of ACh esterase (AChE), the enzyme that terminates synaptic transmission by hydrolyzing the ester group of ACh. In 1991, Sussman and coworkers reported the crystal structure of AChE (19). The active site is that of a typical hydrolase with the archetypical triad of serine, histidine (His), and glutamic acid residues. However, the binding region did not contain the long-anticipated anionic site for binding the quaternary ammonium of ACh. Instead, the cation sits directly on the face of Trp-84 of the esterase. Clearly, a cation-π interaction is important to binding ACh to this essential enzyme (Fig. 5). The active site lies at the bottom of a deep tunnel. This region is lined with aromatic amino acids, so it was termed the aromatic gorge. AChE displays an extremely rapid rate of turnover, as is essential for its role in synaptic transmission, and it seems likely that cation-π interactions also play a role in guiding ACh to the enzyme active site.

The predominant focus of our current research program is the vast array of neuroreceptors and ion channels that establish synaptic transmission. These are large, integral membrane proteins that are not amenable to X-ray crystallography or NMR spectroscopy. As such, chemical approaches are best suited to unraveling the mechanisms of action of these complex but essential proteins. As described in detail elsewhere, we have developed the ability to site-specifically incorporate unnatural amino acids into such proteins (20,21). The receptors are expressed in vertebrate cells and probed with the unsurpassed precision and sensitivity of electrophysiology.

A major focus of our efforts has been the cysteine (Cys)-loop superfamly of receptors (22,23). The prototype is the nicotinic ACh receptor (nAChR). At the neuromuscular junction and throughout the central nervous system, when ACh (Fig. 6) is released from a nerve terminal and traverses the synapse, it binds to the nAChR, a ligand-gated ion channel. The binding of ACh (the agonist) induces a structural change in the receptor that opens a self-contained ion channel, repropagating the electrical signal that caused the original release of ACh. Biochemical studies identified a number of aromatic amino acids that appeared to contribute to the ACh-binding site, but determining their precise role was problematic.

The unnatural amino acid methodology is especially well suited to this problem (24). The electrostatic model described above not only rationalizes the basic interaction, it also nicely explains substituent effects. In particular, fluorine is deactivating in the cation-π interaction (9). The high electronegativity of

---

**FIGURE 4** Binding of trimethylated Lys from histone H3 to BPTF. The trimethylated Lys is shown as a stick figure; the protein residues in space filling.

**FIGURE 5** Binding of ACh to the active site of AChE. Residues of the catalytic triad are shown in stick form; ACh and the Trp-84 are shown in space filling.
the histone-binding protein of Figure 4 and it has been suggested above in the nAChR. This cluster closely resembles that seen in site, including 1 that directly corresponds to Trp-149 identified 1 but 5 aromatic amino acids that contribute to the ACh-binding the nAChR, the region that binds ACh. The AChBP contains not tantly, it is somewhat homologous to the extracellular domain of (26,27). As the name implies, AChBP binds ACh and, impor- small, soluble protein secreted from the glial cells of snails membrane receptor.

fluorine pulls electron density away from the aromatic ring, diminishing the negative electrostatic potential that is the source of attraction for cations. Importantly, the deactivateing affect of fluorine substitution is additive; a series of monofluoro, difluoro, trifluoro, etc. derivatives shows a progressive diminution in cation-π-binding ability. In addition, fluorine is accepted to be quite a small substituent, so any steric perturbation introduced by fluorine substitution is minimal.

We have exploited the fluorine effect to identify cation-π interactions in neuroreceptors. Beginning with the nAChR of the neuromuscular junction, we evaluated a number of aromatic amino acids that were implicated in agonist binding. At 1, and only 1, site, Trp-149 of the α subunit was a compelling result (25). As we moved from Trp to F-Trp to F₂-Trp to F₃-Trp to F₄-Trp, we saw a consistent drop in agonist affinity, over almost 2 full orders of magnitude in response. In fact, a quantitative correlation between agonist affinity and anticipated cation-π-binding ability of the aromatic ring could be obtained. This established unambiguously that when ACh binds to the neuro- muscular nAChR, the quaternary ammonium of ACh makes van der Waals contact with the side chain of Trp-149 of the α subunit. This is very high precision structural information obtained, not through X-ray crystallography or NMR but through chemical scale structure-function studies of a complex, integral membrane receptor.

A related structure is the ACh-binding protein (AChBP), a small, soluble protein secreted from the glial cells of snails (26,27). As the name implies, AChBP binds ACh and, import- tantly, it is somewhat homologous to the extracellular domain of the nAChR, the region that binds ACh. The AChBP contains not 1 but 5 aromatic amino acids that contribute to the ACh-binding site, including 1 that directly corresponds to Trp-149 identified above in the nAChR. This cluster closely resembles that seen in the histine-binding protein of Figure 4 and it has been suggested that such multiple cation-π-binding sites might be common (28).

We noted above that there is a superfamily of Cys-loop receptors. Along with an array of nAChR (the neuromuscular form just discussed and a number of neuronal receptors associated with the central nervous system), there are Cys-loop receptors that respond to glycine, γ-aminobutyric acid (GABA), and serotonin as neurotransmitter agonists. Along with the nAChR, we have studied a GABA receptor and 2 different serotonin receptors (29–31). In each case, we were able to use a fluorination approach to identify a particular aromatic amino acid that makes a cation-π interaction with the neurotransmitter. Note that GABA and serotonin are not quaternary ammonium ions like ACh, but rather are cations of the form RNH₃⁺. These are certainly much less lipophilic (much more water soluble) cations, but still the cation-π interaction is an important contributor to their binding. Clearly, the cation-π interaction is exploited for neurotransmitter recognition throughout the nervous system.

The aromatic amino acids play many special roles in biology. Although Phe, Tyr, and Trp comprise <9% of our amino acids, they are substantially over-represented at binding sites. One of the reasons for this is the cation-π interaction. We would argue that along with the hydrophobic effect, hydrogen bonding, and ion pairing (salt bridges), the cation-π interaction constitutes the 4th key force that contributes to macromolecular structure and molecular recognition in biology. The cation-π interaction has a large electrostatic component that is most useful in rationalizing trends seen in various data. Cation-π interactions between amino acids, Phe/Tyr/Trp→Lys/Arg, contribute significantly to stabilizing protein secondary structure. In addition, drug-receptor interactions across a wide array of systems use cation-π interac- tions. One especially well-documented case is the binding of neurotransmitters to their cognate neuroreceptors, where many cation-π interactions have been established. We can anticipate still more examples of cation-π interactions through- out biology as this special feature of aromatic amino acids becomes more widely appreciated.

**Literature Cited**


