THE ROLE OF AROMATIC INTERACTIONS IN
BIOMOLECULAR RECOGNITION:
CONTRIBUTIONS TO AFFINITY AND SPECIFICITY

MARCEY L. WATERS

Department of Chemistry, CB 3290, University of North Carolina, Chapel Hill, NC
27599, U.S.A.

E-Mail: mlwaters@email.unc.edu

Received: 7th March 2007 / Published: 5th November 2007

ABSTRACT

Using beta-hairpin model systems, we have investigated the role of aromatic interactions in providing selectivity to protein folding and protein-protein interactions. In particular, we have explored the role of cation-π and amide-π interactions and their role in controlling protein-protein interactions with post-translationally modified proteins. Specific modifications that have been studied are Lysine methylation and acylation and Arginine methylation. The molecular recognition of these modified amino acids has particular relevance to understanding the “histone code”, which has been proposed to control chromatin structure and gene expression.

INTRODUCTION

There are two factors which are critical to biomolecular recognition, whether involving protein-protein, protein-DNA, or protein-RNA interactions, or intramolecular interactions which give rise to a single low energy protein structure: affinity and selectivity. Often these two factors are coupled, but not always; for example, high affinity can be achieved with a large number of non-specific interactions, such as the binding of polycations to DNA through electrostatic interactions. Thus, one can argue that specificity is the more difficult of the two to achieve. Early work in de novo protein design provides an example. The de novo design of α-helical coiled coils using binary coding of hydrophobic and hydrophilic residues lead to tertiary interactions between the two helices. However, the hydrophobic interface was described as being “liquid-like”; there was not one single low energy struc-
ture as in a native protein. This points to a general problem in that the hydrophobic effect, which is considered a major contributor to protein folding and protein-protein interactions, is a *nonspecific* interaction. A common solution in the design of α-helical coiled coils has been to bury a hydrogen bond at the hydrophobic interface [1, 2]. This provides specificity, but at the expense of affinity due to the desolvation costs associated with burying such an interaction. Another approach has been to use a “knobs-in-holes” approach, in which the hydrophobic surfaces “fit together” in a specific way [3]. We were interested as to whether there may be more subtle approaches that Nature uses to address this problem: one such approach is the use of aromatic interactions to provide specificity. The term aromatic interaction includes a number of interactions that are specific to the π-system of an aromatic ring, including π-π, cation-π, amide-π interactions, among others [4, 5]. The key feature of these interactions which we expected would provide specificity is that an aromatic ring is not non-polar: it has a quadrupole moment which gives rise to a fixed non-uniform charge distribution on its surface (Fig. 1). Nonetheless, an aromatic ring is still a hydrocarbon, and hence is still relatively hydrophobic (but not as hydrophobic as aliphatic groups). This gives rise to a number of unique properties in its non-covalent interactions. Aromatic groups have a preference for interacting with other aromatic groups due to quadrupole-quadrupole interactions, which results in preferential orientations of the two rings in an offset-stacked or edge-face orientation (Fig. 1b and c). This is not what would be expected from a simple hydrophobic interaction, in which maximal surface area would be buried. Because these interactions occur in a specific geometry, they provide directionality like hydrogen bonds, yet there is no significant desolvation penalty for incorporating them in a hydrophobic environment.

![Figure 1](image)

Figure 1. (a) Electrostatic potential map of benzene generated with MacSpartan; HF/6-31 g*; isodensity value = 0.02; range = -20 (red, electron rich) to 20 kcal/mol (blue, electron poor).

**CATION-π INTERACTIONS IN PROTEINS**

Cation-π interactions between an aromatic sidechain and the ammonium group on lysine (Lys) or the guanidinium group on arginine (Arg) are common in protein structures [6]. In this case, the full positive charge on the cation is attracted to the partial negative charge on the face of the aromatic ring. Once again, these interactions occur in a specific geometry, providing directionality. Interestingly, however, the most common geometry for the interaction of Lys with an aromatic sidechain is the packing of the ε-CH2 of the Lys sidechain into the face of the aromatic ring (Fig. 2a) [6, 7]. This leads to the question as to whether this can still be considered to be a cation-π interaction or whether this is an example in
which a hydrophobic interaction is preferred. This question becomes even more relevant with respect to post-translationally modified Lys and Arg sidechains. Site-specific methylation of Lys and Arg in histone proteins induces protein-protein interactions that are critical for mediating gene expression (Fig. 2b) [8, 9]. There is evidence that both methylated Lys and Arg are recognized via key aromatic interactions. For example, when Lys 9 of histone 3A is trimethylated, it is bound in an aromatic pocket made up of one Trp and two Tyr residues in the HP1 chromodomain (Fig. 2b) [10, 11].

**Figure 2.** (a) Lys-Trp pair in typical cation-π geometry (PDB code: 1ONR); (b) Trimethyllysine of histone 3A bound to the aromatic pocket of the HP1 chromodomain (PDB code: 1KNE).
**AMIDE-\(\pi\) INTERACTIONS**

Amide-\(\pi\) interactions are also common in protein structures, in which the amide stacks against the surface of the aromatic ring. With regard to protein-protein interactions, there is evidence from protein crystal structures that acylated Lys (KAc) forms amide-\(\pi\) interactions which may be important in its recognition (Fig. 3) [12, 13]. KAc is another common post-translational modification in histone proteins and mediates chromatin condensation [14].

![Figure 3. KAc-Tyr pair in typical amide-\(\pi\) geometry between the acetylated histone H4 peptide and GCN5 bromodomain (PDB code: 1E6I).](image)

We were interested in gaining a molecular level understanding of the magnitude, driving force, and specificity of such interactions and their role in protein folding and protein-protein interactions. We have done so by investigating these interactions in a minimalist peptide model system in which we can isolate and investigate a single sidechain-sidechain interaction. The results of such studies are discussed below.

**RESULTS**

**System design**

\(\beta\)-Hairpins have been shown to make excellent model systems for studying non-covalent interactions in water within a biologically relevant context [15–22]. Moreover, \(\beta\)-hairpins have a number of distinct advantages over other peptide-based model systems. Unlike \(\alpha\)-helices, \(\beta\)-hairpins are highly amenable to characterization by NMR, allowing for information at each residue of the peptide to be obtained [23, 24]. NMR allows for characterization
of the folded state and also of sidechain-sidechain interactions. Importantly, sidechain-sidechain interactions have been shown to directly contribute to β-hairpin stability. Because folding of β-hairpins has been shown to be two-state, energetic values for sidechain-sidechain interactions can be obtained easily from measurements of the hairpin stability.

**Cation-π interactions**

Our earliest work on cation-π interactions was aimed at investigating cation-π interactions between unmodified sidechains and their potential contribution to specificity in protein folding [19, 20]. To this end, we were intrigued by the finding that Lys interacts with Trp most often via its ε-CH$_2$ group, and not directly via the ammonium group. Does this CH$_2$-π interaction have cation-π character or is it simply a (nonspecific) hydrophobic interaction? We found that in the context of a β-hairpin peptide, placing an aromatic residue (Phe or Trp) in the i position and a basic residue (Lys or Arg) in the j-2 position allowed for a diagonal sidechain-sidechain interaction (Fig. 4) [19, 20] NOE data as well as upfield shifting of the Lys sidechain due to ring current effects indicated that the Lys interacts with the face of the aromatic sidechain primarily through its alkyl chain (see below). The ε-CH$_2$, which is adjacent to the aromatic ring, provides the site for greatest interaction (Fig. 5). We hypothesized that if this interaction is strictly a hydrophobic interaction, then an alkyl sidechain missing the ammonium group should interact in the same way as Lys. However, if there is a cation-π component to the Trp-Lys interaction, then a sidechain missing the ammonium group would not interact in the same way. Hence, we compared the peptides containing Lys or norleucine (Nle) using a β-hairpin model system (Fig. 4) [20].

![Figure 4](image-url)

**Figure 4.** (a) Structure of the model β-hairpin peptide used to study cation-π interactions. (b) Structure of Lys and Nle.
Characterization of the Lys and Nle-containing peptides (Fig. 4) indicated that both residues resulted in similarly well-folded β-hairpins. However, their interactions with Trp were distinctly different. Upfield shifting of the Lys and Nle sidechains provided insight into their interactions with the π-cloud of Trp (Fig. 6). The Lys sidechain is significantly upfield shifted due to its interaction with the π-cloud of Trp, with a maximum interaction at the ε-CH₂ position, as discussed above. In contrast, Nle exhibits little upfield shifting, indicating that it is not as well packed against the Trp sidechain. NOEs are consistent with this: Lys exhibits a number of strong NOEs to Trp, whereas Nle displays weak NOEs to Trp as well as to the other sidechains on the same face of the hairpin. This indicates less preference for Trp to interact with Nle than with Lys.

Thermal denaturation of the β-hairpins also provides insight into the difference between the Lys and Nle on sidechain-sidechain interactions (Fig. 7). The curves themselves are quite different, despite the fact that the only difference between the two peptides is the presence or absence of an ammonium group. Fitting of these data allows for abstraction of ΔH°, ΔS°, and ΔCp°, which provides insight into the effect that the Lys-Trp interaction has on the driving force for folding (Table 1). In comparison to WNle, the driving force for folding of WK is more enthalpically favourable and less entropically favourable. The more favourable
enthalpy is consistent with the addition of the cation-π interaction, which is enthalpic in nature. The less favourable entropy is consistent with a reduced hydrophobic driving force for folding, but likely also reflects that fact that the preferential interaction between Trp and Lys results in less conformational freedom for Lys than for Nle.

Figure 7. Thermal denaturation profiles of \( WK \) and \( WNle \) peptides as determined by NMR. The fraction folded was determined from the Gly splitting. Error is ±0.5 K in temperature and ±1% in fraction folded. Conditions: 50 mM NaOAc-d4 buffer, pD 4.0 (uncorrected).

Table 1. Thermodynamic parameters\(^a\) for hairpin folding at 298 K for peptides \( WK \) and \( WNle \).\(^a\)

<table>
<thead>
<tr>
<th>Entry</th>
<th>Peptide</th>
<th>( \Delta H^o )</th>
<th>( \Delta S^o )</th>
<th>( \Delta C_p^o )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( WK )</td>
<td>-2.8 (0.03)</td>
<td>-6.8 (0.1)</td>
<td>-163 (3)</td>
</tr>
<tr>
<td>2</td>
<td>( WNle )</td>
<td>-1.41 (0.07)</td>
<td>-1.2 (0.2)</td>
<td>-170 (9)</td>
</tr>
</tbody>
</table>

\(^a\) Determined from the temperature dependence of the Gly chemical shift from 0 to 60 °C. Units are: \( \Delta H^o \): kcal/mol; \( \Delta S^o \): cal/mol K; \( \Delta C_p^o \): cal/mol K. Errors (in parentheses) are determined from the fit.

Taken together, these data indicate that replacement of Lys with Nle causes significant changes in the sidechain’s preference for interaction with Trp, and on the overall driving force for folding. Furthermore, the data indicate that the interaction between Trp and Lys is not strictly hydrophobic in nature. We propose that the Lys interacts via the ε-CH₂ of its sidechain, rather than the NH₃⁺, because the ε-CH₂ has a partial positive charge due to inductive effects of the NH₃⁺, resulting in a polar yet hydrophobic group. This matches perfectly with the polar hydrophobic face of the aromatic ring, resulting in a specific interaction at that position in the chain. In contrast, the NH₃⁺ group does not interact directly with Trp because it is highly solvated by water.
Next, we asked whether methylation of Lys influenced its interaction with an aromatic residue, in this case Trp, to address the role of cation-π interactions in the recognition of trimethyllysine by the aromatic pocket of chromodomain proteins (Fig. 8) [10, 11]. Using the same methods described above, we found that the interaction with Trp was enhanced significantly when Lys is methylated. This is most clearly seen from the upfield shifting of the KMe3 sidechain, particularly at the ε-CH₂ and the terminal methyl positions (Fig. 8c).

As can be seen from Fig. 9, methylation of Lys has a significant stabilizing effect on the β-hairpin. It also results in some cold denaturation, as was seen in WNle. Fitting of the data indicates that Lys methylation has an impact on ΔH°, ΔS°, and ΔCp° as well (Table 2). The driving force for folding of WKMe3 is enthalpically less favourable than WK, suggesting that the contribution from the cation-π interaction is decreasing. This is not surprising based on inspection of the electrostatic potential maps of Lys and KMe3 (Fig. 10). Methy-
lation effectively spreads the positive charge over a larger surface area, such that the magnitude of the $\delta^+$ at any one site is reduced. The more favourable entropy of folding in WKMe3 can be explained by the fact that there are now multiple sites for favourable interaction with Trp (three degenerate methyl groups) and also that the sidechain has become more hydrophobic. In summary, alkylation of the ammonium group increases the net driving force for interaction with Trp via changes in both entropy and enthalpy.

**Figure 9.** Thermal denaturation profiles of WK and WKMe3 peptides as determined by NMR. The fraction folded was determined from the Gly splitting. Error is $\pm 0.5$ K in temperature and $\pm 1\%$ in fraction folded. Conditions: 50 mM NaOAc-$d_4$ buffer, pD 4.0 (uncorrected).

**Table 2.** Thermodynamic parameters$^a$ for hairpin folding at 298 K for peptides WK and WKMe3.$^a$

<table>
<thead>
<tr>
<th>Peptide</th>
<th>$\Delta H^\circ$</th>
<th>$\Delta S^\circ$</th>
<th>$\Delta C_p^\circ$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WK</td>
<td>-2.8 (0.03)</td>
<td>-6.8 (0.1)</td>
<td>-163 (3)</td>
</tr>
<tr>
<td>WKMe3</td>
<td>-0.1 (0.1)</td>
<td>+4.5 (0.3)</td>
<td>-243 (4)</td>
</tr>
</tbody>
</table>

(a) Determined from the temperature dependence of the Gly chemical shift from 0 to 80 °C. Units are: $\Delta H^\circ$: kcal/mol; $\Delta S^\circ$: cal/mol K; $\Delta C_p^\circ$: cal/mol K. Errors (in parentheses) are determined from the fit.
Arg-π Interactions.

Methylation of Arg is another common post-translational modification found in histone proteins. Because Arg and dimethylarginine (DMA) have been observed to stack with Trp in protein crystal structures (Fig. 11), we chose to investigate the role of Arg methylation on cation-π interactions using the same β-hairpin model system (Fig. 12) [25]. We had previously shown that an Arg-Trp interaction occurs via stacking of the guanidinium group with the Trp, and that the interaction is more favourable than the Lys-Trp interaction discussed above [19]. As with Lys [21], we found that Arg methylation enhances the cation-π interaction, as indicated by significant upfield shifting of the guanidinium group and the methyl groups (Fig. 12c) [25].

Figure 10. Electrostatic potential maps of (a) Lys and (b) KMe3 generated with MacSpartan; HF/6–31 g*; isodensity value = 0.02; range = 20 to 120 kcal/mol (red, electron rich; blue, electron poor).

Figure 11. (a) Arg-Trp pair in typical cation-π geometry (human growth hormone and receptor complex: 3HHR) [26]. (b) DMAα-Trp pair from histone H3-chromodomain complex (2B2U) [27].
Thermal denaturations of WR, WDMAa, and WDMAs indicate that methylation of Arg has a similar stabilizing effect on β-hairpin stability as was seen with Lys, although the magnitude is not as great (Fig. 13, Table 3). Methylation of Arg results in a less favourable enthalpy of folding and a more favourable entropy of folding, as was observed in WKMe3 as compared to WK (Table 3). Thus, it appears that methylation-induced enhancement of cation-π interactions may be a common biological method for enhancing biomolecular recognition between two proteins or a protein and another biomolecule.
Figure 13. Thermal denaturation profiles of WR, and WDMAa, and WDMAs peptides as determined by NMR. The fraction folded was determined from the Gly splitting. Error is ±0.5 K in temperature and ±1% in fraction folded. Conditions: 50 mM NaOAc-4 buffer, pD 4.0 (uncorrected).

Table 3. Thermodynamic parameters for hairpin folding at 298 K Peptides WR, WDMAs, and WDMAa

<table>
<thead>
<tr>
<th>Peptide</th>
<th>$\Delta H^{b}$</th>
<th>$\Delta S^{b}$</th>
<th>$\Delta C_p^{c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WR</td>
<td>-3.7 (0.1)</td>
<td>-9.1 (0.2)</td>
<td>-193 (29)</td>
</tr>
<tr>
<td>WDMAs</td>
<td>-2.4 (0.2)</td>
<td>-2.6 (0.8)</td>
<td>-409 (61)</td>
</tr>
<tr>
<td>WDMAa</td>
<td>-2.3 (0.1)</td>
<td>-2.1 (0.4)</td>
<td>-355 (53)</td>
</tr>
</tbody>
</table>

(a) Determined from the temperature dependence of the Gly chemical shift from 0 to 80 °C. Units are: $\Delta H^{b}$: kcal/mol; $\Delta S^{b}$: cal/mol K; $\Delta C_p^{c}$: cal/mol K. (b) Errors (in parentheses) are determined from the fit. (c) Error for $\Delta C_p^{c}$ values are estimated to be ±15%.

Amide-π Interactions

Lastly, we investigated the role of Lys acylation on its interaction with Trp (Fig. 14) [28]. In this case, the positive charge is lost, so the comparison is between a cation-π interaction in WK and an amide-π interaction in WKAc. One might expect that the cation-π interaction would be substantially stronger than the amide-π interaction, but that was not what was observed in the context of the β-hairpin model system. Upfield shifting of the AcLys sidechain is similar in magnitude to that of Lys, and unlike that observed for Nle (Fig. 14). However, the position of maximal upfield shifting differs for Lys and AcLys: AcLys exhibits maximal shifting at the NH rather than at the ε-CH2 group (Fig. 14c). Although not definitive, this suggests stacking of the amide group with the Trp ring, such that the NH is still free to hydrogen bond with water, similar to the stacking of Arg with Trp.
Despite the difference in geometry of the Trp-Lys and Trp-AcLys interactions, thermal denaturation studies (Fig. 15) indicate that the driving force for folding is virtually identical, with a favourable enthalpy of folding and a concomitant entropic cost (Table 4). These findings suggest that the Trp-AcLys interaction may compensate for the loss of the charge-quadrupole component of the cation-π interaction through the additional π-π component as well as the dipole-quadrupole interaction between the amide bond and aromatic ring.
Figure 15. Thermal denaturation profiles of WK and WKAe peptides as determined by NMR. The fraction folded was determined from the Gly splitting. Error is ± 0.5 K in temperature and ± 1% in fraction folded. Conditions: 50 mM NaOAc-d4 buffer, pD 4.0 (uncorrected).

Table 4. Thermodynamic parameters\(^{a}\) for folding of WK, WKAe, and WNle at 298 K [10].

<table>
<thead>
<tr>
<th>Peptide</th>
<th>(\Delta H^o) (kcal/mol)</th>
<th>(\Delta S^o) (cal/mol K)</th>
<th>(\Delta C^o_p) (cal/mol K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WK</td>
<td>-2.8 (0.03)</td>
<td>-6.8 (0.1)</td>
<td>-163 (3)</td>
</tr>
<tr>
<td>WKAe</td>
<td>-2.91 (0.07)</td>
<td>-6.1 (0.2)</td>
<td>-250 (40)</td>
</tr>
<tr>
<td>WNle</td>
<td>-0.85 (0.07)</td>
<td>-0.7 (0.2)</td>
<td>-170 (25)</td>
</tr>
</tbody>
</table>

\(^{a}\) Determined from the temperature dependence of the Gly chemical shift from 0 to 60 °C. Errors (in parentheses) are determined from the fit. Error is ± 0.5 K in temperature and ± 1% in fraction folded. Conditions: 50 mM NaOAc-d4 buffer, pD 4.0 (uncorrected).

CONCLUSION

Using a β-hairpin model system, we have been able to study a wide range of subtle interactions which are important to protein folding and biomolecular recognition. These findings provide insight into how Nature gains selectivity in molecular recognition, and suggests that simplifying assumptions of hydrophobic and hydrophilic interactions are incomplete for understanding biomolecular recognition.
The Role of Aromatic Interactions in Biomolecular Recognition

REFERENCES


