Tetrahedron Letters 50 (2009) 7035–7037

Contents lists available at ScienceDirect

Tetrahedron Letters

journal homepage: www.elsevier.com/locate/tetlet

A tryptophan-analog host whose interactions with ammonium ions in water are dominated by the hydrophobic effect

Amanda L. Whiting, Nicole M. Neufeld, Fraser Hof *

Department of Chemistry, University of Victoria, PO Box 3065, Victoria, BC, Canada V8W 3V6

ARTICLE INFO

Article history:
Received 23 June 2009
Accepted 28 September 2009
Available online 1 October 2009

Keywords:
Cation–π interaction
Hydrophobic effect
Quaternary ammonium ions
NMR titration
Aromatic cage

ABSTRACT

The binding of quaternary ammonium guests to a flexible, indole-based host has been studied in both aqueous and organic solvents. Binding was shown to depend strongly on the hydrophobic effect and less on the cation–π interaction.

The combination of an electron-rich π surface with a significant hydrophobic character makes indoles key players in a variety of biologically important recognition events. Surveys of protein structures have shown that over 25% of all tryptophan side chains are in close contact with cationic neighbors,[1] including a subset that are particularly important for protein–protein interactions.[2] Quaternary ammonium cations, including a variety of neurotransmitters,[3] phospholipid head groups,[4] and post-translationally methylated lysine side chains,[5] constitute a specific class of cations that are routinely engaged by a family of protein-binding pockets called 'aromatic cages'. The known examples of these pockets, expressed in a variety of evolutionarily distinct domains, typically contain at least one tryptophan indole ring alongside other aromatic residues. Figure 1 shows two examples of such binding motifs, with one example each of a cationic inhibitor binding to an enzyme (1a) and a cation-mediated protein–protein interaction (1b).

Taking inspiration from the aromatic cages of biology, tris(indole) host **1** was designed while considering the general goals of rapid synthesis, the presentation of multiple electron-rich indole rings, and water solubility (Fig. 2). We report herein, the synthesis of host 1 and studies on its binding with various quaternary ammonium cations in pure water. Solvent effects were examined by using the analogous chloroform-soluble host **2**. Differing trends were found in the two solvent systems, allowing us to comment on the roles of cation–π interactions, the hydrophobic effect, and preorganization.

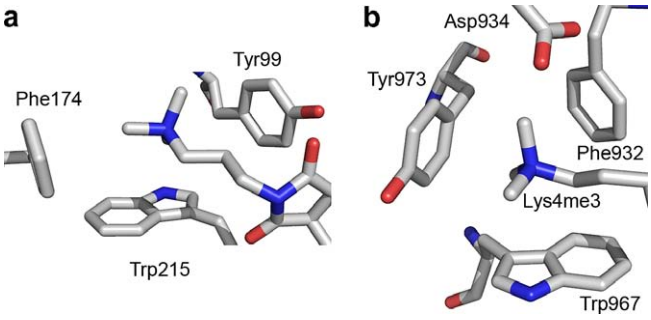


Figure 1. Examples of quaternary ammonium binding in biological systems. (a) $RNMe_3^+$ group of an inhibitor bound in the S4-pocket of Factor Xa (PDB code 2BOK).[16] (b) Histone H3 trimethylated Lys4 (Lys4me3) bound to the double tudor domain of JMJD2A (PDB code 2GFA).[19] Grey, carbon; red, oxygen; blue, nitrogen.

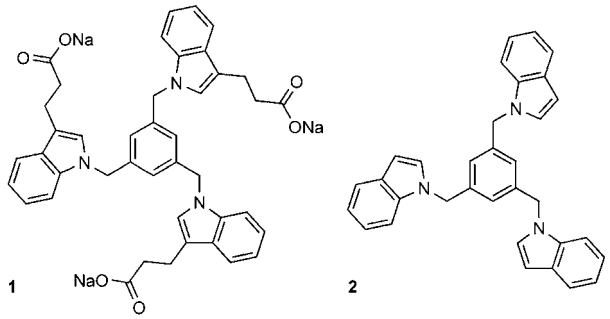


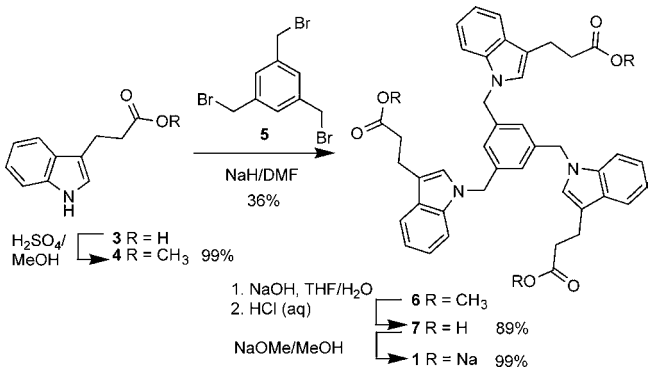
Figure 2. Hosts studied in this work.

* Corresponding author. Tel.: +1 250 721 7193.
E-mail address: fhof@uvic.ca (F. Hof).

0040-4039/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2009.09.161

The desired water-soluble indole host **1** is synthesized in three steps starting from commercially available indole-3-propionic acid **3** (Scheme 1). After protection as the methyl ester,[6] three equivalents of indole **4** are coupled to 1,3,5-tribromomethylbenzene (**5**) upon deprotonation with NaH in DMF in 36% yield. Basic hydrolysis of methyl ester **6** is followed by acidification and isolation of the triacid product. Formation of the tri-sodium salt by treatment with stoichiometric NaOMe gives **1** in an overall yield of 31%. Chloroform-soluble indole host **2** is similarly synthesized from **5** and indole in 39% yield.[7]

Host **1** is soluble in water to a concentration greater than 15 mM, and NMR dilution studies revealed self-association in phosphate-buffered D_2O with a $K_{homodimer}$ of 400 M^{-1}. The degree of homodimerization, though small at the concentrations of the following experiments (1–2 mM), was taken into account when calculating binding constants in water (see ESI for details). We investigated the binding of various cationic guests by **1** by NMR titration of solutions of ammonium salts (2–150 mM) into host solutions (1–2 mM) in D_2O buffered with sodium phosphate (50 mM) at pD 7.6/pH 8.0 (Fig. 3).[8] Host concentration was matched in both titrant and receiving solution, and chemical shift data were fit to a 1:1 binding isotherm using a standard protocol.[9] The average K_{assoc} values resulting from three replicate titrations of each host-guest pair are shown in Table 1. In all cases, the guest signals moved upfield while the host signals moved downfield, as expected for interactions between the cationic guests and the aromatic surfaces of the host.



Scheme 1. Synthesis of host **1**.

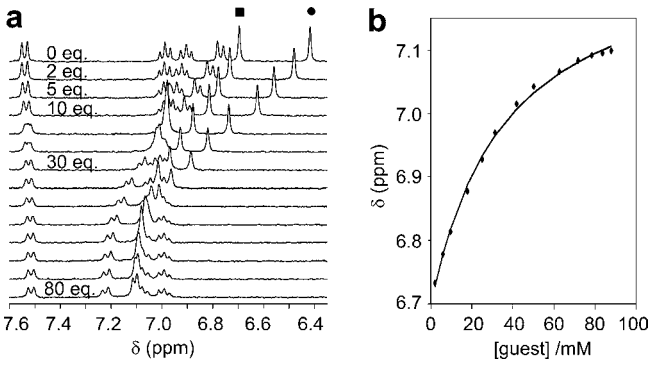


Figure 3. (a) Representative ^1H NMR titration data for the complexation of acetylcholine chloride by host **1**. Titrations carried out at 295 K in phosphate-buffered D_2O (50 mM Na_2HPO_4/NaH_2PO_4 at pH 8.0). Titrant containing guest (125 mM) and host (1 mM) was added to a solution containing a matched concentration of host. Tracked signals are ■ (6.72 ppm, *PhH*) and ● (6.44 ppm, indole 2-*H*). (b) Titration curve resulting from (a) using shifts from ■. The line represents fitted 1:1 binding isotherm for this equilibrium (K_{assoc} = 38 ± 3 M^{-1}).

Table 1
Cation affinities of hosts **1** and **2**

Entry	Host	Guest	Solvent[a]	K_{assoc}[b] /M^{-1}
1	**1**	Me_4NI	D_2O	35 ± 6
2	**1**	Me_4NCl	D_2O	32 ± 2
3	**1**	Me_4NOAc	D_2O	26 ± 8
4	**1**	$LysMe_3Cl$	D_2O	63 ± 13
5	**1**	AChCl	D_2O	42 ± 5
6	**1**	$MeNH_3Cl$	D_2O	7 ± 5
7	**1**	Me_2NH_2Cl	D_2O	20 ± 18
8	**1**	Me_3NHCl	D_2O	34 ± 16
9	**1**	Et_4NCl	D_2O	66 ± 7
10	**1**	Pr_4NCl	D_2O	450 ± 250
11	**1**	Bu_4NCl	D_2O	(2.7 ± 3.5) × 10^4
12	**2**	AChCl	$CDCl_3$	5 ± 5
13	**2**	AChCl[c]	$CDCl_3$	4 ± 5
14	**2**	Et_4NCl	$CDCl_3$	2 ± 1
15	**2**	Pr_4NCl	$CDCl_3$	0[d]
16	**2**	Bu_4NCl	$CDCl_3$	0[d]

[a] D_2O = phosphate-buffered D_2O (50 mM Na_2HPO_4/NaH_2PO_4 at pH 8.0).
[b] Values determined at 295 K by fitting of ^1H NMR titration data to a 1:1 binding isotherm. All values are the average of 3–4 titrations. The errors reported are the standard deviations of the averaged K_{assoc} values. In D_2O, [host] = 1–2 mM, [guest] = 2–150 mM. In $CDCl_3$, [host] = 1–2 mM, [guest] = 45–400 mM.
[c] Value determined by inverse titrations in $CDCl_3$, where [guest] = 1–2 mM, [host] = 20–300 mM.
[d] Small chemical shifts and nonexistent curvature indicate that the K_{assoc} is below the detectable limit for NMR.

In phosphate-buffered D_2O, host **1** showed similar affinity for Me_4N^+ ions regardless of counter anion (Table 1: entries 1–3), which stands in contrast to the strong counterion effects typically observed in $CDCl_3$.[10] Nevertheless, all subsequent titrations were carried out with Cl^- salts to avoid the possibility of even small counterion effects. Host **1** also displayed similar association constants for the biologically relevant quaternary cations trimethyllysine and acetylcholine ($LysMe_3$ and ACh, entries 4 and 5), suggesting a similar mode of binding involving the engagement of the $RNMe_3^+$ portion of each compound. A second series of guests was investigated to examine the effect of increasing alkylation on binding to **1** (entries 6–11). The stepwise increasing methylation from $MeNH_3^+$ (K_{assoc} = 7 M^{-1}) to Me_4N^+ (K_{assoc} = 32 M^{-1}) resulted in a trend of increasing affinity with methylation. A small additional increase in affinity was found for Et_4N^+, while Pr_4N^+ and Bu_4N^+ gave rise to much more dramatic increases in affinity, with Bu_4N^+ (K_{assoc} = 2.7 × 10^4 M^{-1}) binding 1500-fold more strongly than Me_4N^+.

NMR titrations on host **2** were performed in $CDCl_3$ using protocols similar to those reported for host **1**. Acetylcholine chloride was used as a replacement for Me_4N^+ Cl^- due to the insolubility of the latter in $CDCl_3$. Weak but consistently measurable binding constants were observed (Table 1, entries 12–16). The strongest binding was found between **2** and acetylcholine chloride, with progressively weaker binding observed upon increasing size of quaternary ammonium ion.

The interactions of indoles (in the form of tryptophan side chains) and quaternary ammonium ions (in a variety of neurotransmitters and in trimethyllysine side chains) are increasingly recognized as critical components of a variety of biochemical signaling pathways,[11,12] and the cation–π interaction is a primary driving force for these binding events.[13,14] In general, larger cations should produce weaker cation–π interactions because (a) the larger cations have a larger radius of interaction between π surface and positive charge and (b) larger cations bear a more diffuse positive charge, creating a weaker electrostatic interaction with electron-rich π surfaces. Yet it is often observed in protein binding events that more highly methylated ammonium ions bind more strongly to their aromatic partners (indeed that is the basis for

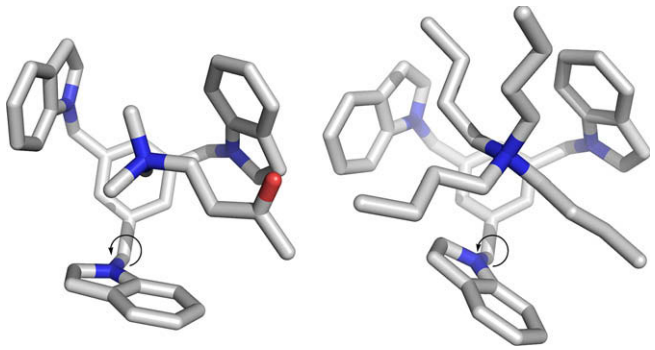


Figure 4. Optimized structures of host **2** bound to acetylcholine (left) and Bu_4N^+ (right) (DFT B3LYP/6-31G*). Key rotational degrees of freedom are indicated.

selective signaling by trimethylated lysine residues[15]), and that trend is reproduced in the current model system. The subtle increase observed upon moving from $MeNH_3^+$ to Me_4N^+ and the more dramatic increase in binding for the greasy Bu_4N^+ ion both run counter to the trends expected for cation–π interactions and are better explained by the surface-area-dependent hydrophobic effect. The complete lack of binding of Pr_4N^+ and Bu_4N^+ in $CDCl_3$—where no hydrophobic effect can participate—confirms the importance of the hydrophobic effect for the binding of cations by **1** in water.

Mutational studies of aromatic cage sites have demonstrated that tryptophan side chains participate in cation–π interactions.[3] Comparisons of *t*-butyl groups [$RC(Me_3)$] and nearly isosteric trimethylammonium groups [$RN(Me_3)^+$] have been elegantly used to demonstrate the importance of the cation–π interaction in protein–protein and drug-protein contacts.[13,16] The authors of these studies conclude that the hydrophobic effect plays no role in the observed binding events.[13] Yet our studies of tris(indole) hosts **1** and **2** suggest that the hydrophobic effect can operate in this artificial tryptophan analog given a greasy enough guest, and that the cation–π interaction, if operative, is scarcely measurable. What differentiates the indoles of host **1** and those of proteins? Crystal structures of cation-binding proteins in bound and free states reveal almost no movement of the aromatic cage side chains, suggestive of a highly rigid and preorganized binding pocket.[17] It has been recently suggested that this rigidity is a hallmark of aromatic cage binding sites.[18] This rigidity is certainly lacking in the flexible hosts **1** and **2**—DFT calculations of **2** bound to both acetylcholine and Bu_4N^+ reveal that the host can adapt its shape to complement the size of the guest cation (Fig. 4). The result is a system that complements the larger hydrophobic surface area of Bu_4N^+ and produces binding that is remarkably strong given the complete flexibility of both host and guest. It is possible that this strong binding motif, very easily prepared, may be exploited in other unnatural aqueous-phase recognition systems. We are continuing to explore the links between preorganization, hydrophobicity, and high-affinity binding in water in this and other Trp-derived systems.

Acknowledgments

A.W. thanks NSERC for PGS M and PGS D fellowships. F.H. is a Career Scholar of the Michael Smith Foundation for Health Research. This research was supported by NSERC and the University of Victoria.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tetlet.2009.09.161.

References and notes

1. Gallivan, J. P.; Dougherty, D. A. *Proc. Nat. Acad. Sci. U.S.A.* **1999**, *96*, 9459–9464.
2. Crowley, P. B.; Golovin, A. *Proteins: Struct. Funct. Bioinf.* **2005**, *59*, 231–239.
3. Zhong, W.; Gallivan, J. P.; Zhang, Y.; Li, L.; Lester, H. A.; Dougherty, D. A. *Proc. Nat. Acad. Sci. U.S.A.* **1998**, *95*, 12088–12093.
4. Martin, S. F.; Follows, B. C.; Hergenrother, P. J.; Trotter, B. K. *Biochemistry* **2000**, *39*, 3410–3415.
5. Jacobs, S. A.; Taverna, S. D.; Zhang, Y.; Briggs, S. D.; Li, J.; Eissenberg, J. C.; Allis, C. D.; Khorasanizadeh, S. *EMBO J.* **2001**, *20*, 5232–5241.
6. Pedras, M. S. C.; Jha, M. *Bioorg. Med. Chem.* **2006**, *14*, 4958–4979.
7. Rajakumar, P.; Swaroop, M. G. *Tetrahedron Lett.* **2006**, *47*, 3019–3022.
8. Glasoe, P. K.; Long, F. A. *J. Phys. Chem.* **1960**, *64*, 188–190.
9. Sanderson, J. M. http://www.dur.ac.uk/j.m.sanderson/science/downloads.html, The Centre for Bioactive Chemistry, Department of Chemistry, University Science Laboratories: Durham, UK.
10. Bartoli, S.; Roelens, S. *JACS* **2002**, *124*, 8307–8315.
11. Xiu, X.; Puskar, N. L.; Shanata, J. A. P.; Lester, H. A.; Dougherty, D. A. *Nature* **2009**, *458*, 534–537.
12. Lin, Z.; Wang, W.; Uhl, G. R. *Mol. Pharmacol.* **2000**, *58*, 1581–1592.
13. Hughes, R. M.; Wiggins, K. R.; Khorasanizadeh, S.; Waters, M. L. *PNAS* **2007**, *104*, 11184–11188.
14. Zacharias, N.; Dougherty, D. A. *Trends Pharmacol. Sci.* **2002**, *23*, 281–287.
15. Jacobs, S. A.; Khorasanizadeh, S. *Science* **2002**, *295*, 2080–2083.
16. Schärer, K.; Morgenthaler, M.; Paulini, R.; Obst-Sander, U.; Banner, D. W.; Schlatter, D.; Benz, J.; Stihle, M.; Diederich, F. *Angew. Chem., Int. Ed.* **2005**, *44*, 4400–4404.
17. Li, H.; Ilin, S.; Wang, W.; Duncan, E. M.; Wysocka, J.; Allis, C. D.; Patel, D. J. *Nature* **2006**, *442*, 91–95.
18. Taverna, S. D.; Li, H.; Ruthenburg, A. J.; Allis, C. D.; Patel, D. J. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025–1040.
19. Huang, Y.; Fang, J.; Bedford, M. T.; Zhang, Y.; Xu, R.-M. *Science* **2006**, *312*, 748–751.