J Incl Phenom Macrocycl Chem (2008) 62:251–254
DOI 10.1007/s10847-008-9464-y

ORIGINAL ARTICLE

Scope of amino acid recognition by cucurbit[8]uril

P. Rajgariah · A. R. Urbach

Received: 22 February 2008 / Accepted: 15 May 2008 / Published online: 13 June 2008

Abstract This paper describes the molecular recognition of amino acids by cucurbit[8]uril (Q8) and by the 1:1 complex between Q8 and methyl viologen (MV) in purely aqueous solution. These hosts are known to bind aromatic peptides with high affinity and sequence specificity, but prior work has focused on only a small subset of amino acids. In an effort to elucidate the scope and limitations of amino acid recognition by Q8 and Q8•MV, a comprehensive examination of the 20 genetically encoded amino acids was carried out by ^{1}H NMR spectroscopy and isothermal titration calorimetry. We find that both Q8 and Q8•MV bind measurably to only tryptophan, phenylalanine, and tyrosine. These results demonstrate that Q8 and Q8•MV are highly selective in the context of all genetically encoded amino acids and are therefore promising for the development of recognition-intensive applications involving peptides, proteins, and proteomes.

Keywords Cucurbit[n]uril · Cucurbit[8]uril · Molecular recognition · Amino acid · Amino acid recognition · Aqueous solution · Tryptophan · Phenylalanine

Electronic supplementary material The online version of this article (doi:10.1007/s10847-008-9464-y) contains supplementary material, which is available to authorized users.

P. Rajgariah · A. R. Urbach (✉)
Department of Chemistry, Trinity University,
San Antonio, TX 78212, USA
e-mail: adam.urbach@trinity.edu

Introduction

The cucurbit[n]uril (Qn) family of synthetic macrocycles has shown great promise in the areas of molecular recognition and molecular devices due largely to the capacity to bind organic amines over a large range of affinities (10^3 M^{-1} to 10^{12} M^{-1}) in aqueous solution [1–4]. Cucurbit[8]uril (Q8, Fig. 1) comprises eight methylene-bridged glycoluril units and has sufficient inner capacity to accommodate the simultaneous inclusion of two aromatic guests [5–8]. For example, Q8 binds tightly to methyl viologen (MV), and the resulting Q8•MV complex binds selectively to electron-rich aromatic guests, including dihydroxynaphthalene, dopamine, tyrosine (Tyr), and tryptophan (Trp), offering in some cases a substantial optical response upon binding [6, 9].

We extended this work to include Trp-containing peptides and showed that Q8•MV is selective for N-terminal Trp on the basis of electrostatic charge [10]. In the absence of MV, Q8 can recognize and noncovalently dimerize peptides containing N-terminal Trp or phenylalanine (Phe) with extraordinary sequence selectivity [11]. In these studies, we focused solely on the aromatic amino acids due to their similarity to guests studied in prior work by Kim and coworkers [6, 12]. Tao and coworkers looked at the binding of Q8 to a subset of nine amino acids including Trp, Phe, and Tyr and found binding only to Trp, Phe, and Tyr. In order to define the scope of molecular recognition in the context of peptides and proteins, however, it is necessary to consider all 20 genetically encoded amino acids. Moreover, Q8•MV as a host for amino acids deserves further exploration due to its success in the context of peptides. To address these concerns, we present here a comparative binding study of Q8 and Q8•MV with all 20 amino acids.

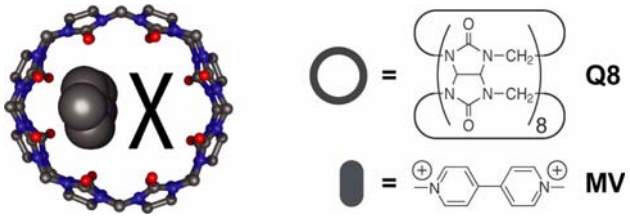

Fig. 1 Chemical structures of cucurbit[8]uril (Q8) and methyl viologen (MV). X denotes an amino acid

Methods

Materials

All reagents were of analytical purity grade and were used without further purification. Deuterium oxide was from Cambridge Isotope Laboratories. Methyl viologen, sodium phosphate (mono and dibasic) and all amino acids were from Sigma-Aldrich. Water was from US Biochemicals (nuclease-free, ultrapure, MB-grade, 0.2 μm filtered). Cucurbit[8]uril (Q8) was synthesized by the group of Dr. Anthony Day (University of New South Wales, Australia) and purchased from Unisearch.

A stock solution of 1.0 M sodium phosphate buffer was adjusted to pH 7.0 and sterile filtered. The pH was checked periodically. ITC experiments were carried out in 10 mM phosphate buffer, which was made as needed by diluting the 1.0 M stock. Fresh analyte solutions were prepared every couple of days and were thoroughly dissolved by heating at 60 °C. All analytes were massed to ±0.02 mg with an accuracy of at least three significant digits. The purities of Q8 and MV were determined by ^{1}H NMR using anhydrous *tert*-butyl alcohol as standard.

Isothermal titration calorimetry (ITC)

Titration experiments were carried out in 10 mM sodium phosphate buffer (pH 7.0) at 27 °C on a VP–ITC calorimeter from Microcal, Inc. (http://www.microcalorimetry.com). In a typical experiment, Q8 or Q8•MV was in the sample cell at a concentration of 0.05–0.1 mM, and the amino acid was in the injection syringe at a concentration of 0.5–4 mM. The titration schedule consisted of 29 or 57 consecutive injections of 2–10 μL with at least a 200 s interval between injections. Heats of dilution, measured by titrating beyond saturation, were subtracted from each data set. All solutions were degassed prior to titration. The data were analyzed using Origin 7.0 software.

^{1}H NMR Spectroscopy

One-dimensional spectra were collected in deuterium oxide at 25 °C on a Varian Inova 400 MHz spectrophotometer using a presaturation pulse to suppress the signal from residual protiated solvent. Amino acids were 5 mM in concentration, and 1:1 (amino acid:Q8•MV) mixtures contained all components at 1 mM.

Table 1 Thermodynamic binding data

Amino acid (AA)	Q8•MV + AA K_a (M^{-1})[a]	Q8 + AA K_{ter} (M^{-2})[b]
Trp	4.3 (±0.3) × 10^4	6.9 (±1.3) × 10^7
Phe	5.3 (±0.7) × 10^3	1.1 (±0.2) × 10^8
Tyr	2.2 (±0.1) × 10^3	<10^3 M^{-1}[c]
All 17 others	No binding obs	No binding obs

[a] Values are from Reference [10]

[b] Mean values obtained from at least three ITC experiments at 27 °C in 10 mM sodium phosphate, pH 7.0

[c] The units given here assume a 1:1 host:guest stoichiometry

Results and discussion

Binding to Q8•MV

Previously [10], we measured the binding of Q8•MV to Trp, Phe, Tyr, and histidine (His) using isothermal titration calorimetry (ITC, 27 °C, 10 mM sodium phosphate, pH 7.0) and found the following trend in the values of equilibrium association constant (K_a): Trp > Phe > Tyr >> His (Table 1). In these studies, the stoichiometry of binding was determined by fitting the enthalpy data to a one-set-of-sites model in Origin software; a 1:1:1 (Q8:MV:AA) stoichiometry was observed consistently [10]. Species of this stoichiometry were also observed by electrospray mass spectrometry.

A more efficient method of screening for the presence or absence of binding is to look for perturbation in the NMR spectrum of the amino acid upon adding Q8•MV. This is made possible by the millimolar solubility of Q8•MV. To this end, ^{1}H NMR spectra in deuterium oxide were obtained for each of the 20 amino acids[1] in the absence and presence of an equimolar quantity of Q8•MV. Only Trp, Phe, and Tyr showed spectral perturbation upon treatment with Q8•MV (see Supporting Information), thus demonstrating that the other 17 amino acids do not bind appreciably.

[1] We used the L-enantiomer of arginine, cysteine, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, threonine, tryptophan, tyrosine, and valine; and the DL-racemate of alanine, asparagine, aspartic acid, glutamine, glutamic acid. The 50% reduction in concentration for the L-enantiomer in these five racemic mixtures was not considered to be significant given that no binding was observed in all cases.

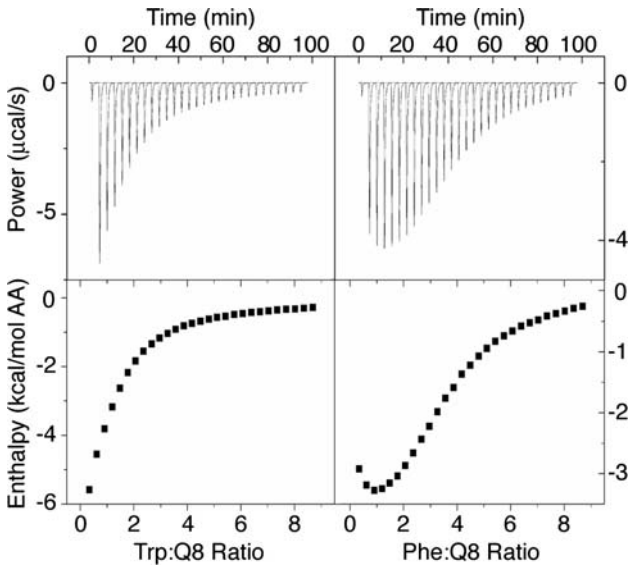

Fig. 2 ITC data for the complexation of Q8 with tryptophan (left) and phenylalanine (right) at 27 °C in 10 mM sodium phosphate, pH 7.0. Amino acids were titrated at 4 mM into a 0.1 mM solution of Q8. The top plot shows the raw data for power versus time. The integrated enthalpy values are plotted at bottom as a function of the mole ratio of amino acid:Q8

Binding to Q8

Due to the poor aqueous solubility (≤0.1 mM) of Q8 [2, 3, 10], NMR spectroscopy was inefficient for screening a large set of compounds. Therefore, ITC (27 °C, 10 mM sodium phosphate, pH 7.0) was used to look for possible binding of Q8 to the 20 amino acids (see Supporting Information). Only Trp, Phe, and Tyr showed measurable binding, which corroborates the findings of Tao [13] and additionally rules out any binding to the remaining eleven amino acids. Trp and Phe bind in a 2:1 (amino acid:Q8) stoichiometry (Fig. 2), as expected from prior work with the analogous Trp-Gly-Gly and Phe-Gly-Gly peptides [11]. A binding transition was observed for Tyr, but the affinity was too low ($<10^3$ M^{-1}) to allow for accurate determination of affinity or stoichiometry.[2] Both 1:1 and 2:1 (amino acid:Q8) complexes were observed by mass spectrometry for each of these three amino acids (see Supporting Information).

We were unable to determine the nature of cooperativity or to separate the stepwise binding equilibria for these complexes. Therefore, Table 1 lists thermodynamic data for the overall formation of the ternary complexes, such that K_{ter} has units of M^{-2}.[3] These values of K_{ter} are, surprisingly, 50–1000-fold smaller than those for the analogous N-terminal tripeptides, Trp-Gly-Gly ($K_{ter} = 3.6 \times 10^9$ M^{-2}) and Phe-Gly-Gly ($K_{ter} = 1.5 \times 10^{11}$ M^{-2}) [11]. The additional stability of the tripeptides is likely due to the hydrogen bonds that form between the amide NH groups and the Q8 carbonyls [11], as well as to the lack of a negatively charged carboxylate group proximal to the aromatic ring.

Q8 vs. Q8•MV

These results reinforce our prior observation that peptide:host stoichiometry can be controlled by the presence (1:1) or absence (2:1) of MV. Interestingly, Q8 prefers Trp over Phe in the presence of MV, whereas there is a small preference for Phe in the absence of MV. This change in specificity is likely due to two factors: (1) a favorable charge–transfer interaction between indole and MV; and (2) the ability to pack two Phe side chains completely within the cavity of Q8 in a fashion which maintains attractive interactions between the two phenyl rings and between the N-terminal ammonium groups and proximal carbonyl groups on Q8. This structural argument is supported by prior work [11].[4]

Factors governing binding

In considering how each amino acid might bind to Q8 or Q8•MV, we would expect: (1) all polar groups to remain outside the cavity of Q8 due to favorable aqueous solvation and potential attraction to the carbonyl groups of Q8; and (2) all nonpolar groups to enter the cavity of Q8 due to hydrophobic inclusion and to van der Waals attraction with the inner surface of Q8. Although the N-terminal ammonium group has been shown to be important in the binding of Trp to Q8•MV [10], this group is common to all 20 amino acids tested here and cannot, therefore, explain differences in binding.

Figure 3 plots the solvent-exposed surface area [14] of each of the 20 amino acid sidechains against its free energy of transfer from cyclohexane to water, as reported by Radzicka and Wolfenden [15]. The upper right corner of this plot, indicated in grey, is occupied by Trp, Phe, and Tyr, suggesting that binding may be governed by a combination of size and hydrophobicity. For example, leucine and isoleucine are more hydrophobic but much smaller than Trp, and

[2] Cong et al. report a 1:1 (Q8:amino acid) stoichiometry for Q8 with tryptophan and tyrosine and a 2:1 stoichiometry with phenylalanine on the basis of continuous variation experiments by UV-visible spectroscopy. This method, however, does not provide reliable stoichiometric constants if higher order dissociation constants are significantly greater than the working concentration.

[3] The methodology used for determining the thermodynamic constants for the formation of ternary complexes with Q8 is described in detail in reference 11.

[4] We have an unpublished crystal structure of Q8 bound to two copies of Trp-Gly-Gly, which shows that both indoles cannot fit completely within the cavity of Q8 while maintaining contacts between N-terminal ammonium groups and proximal carbonyl groups on Q8.

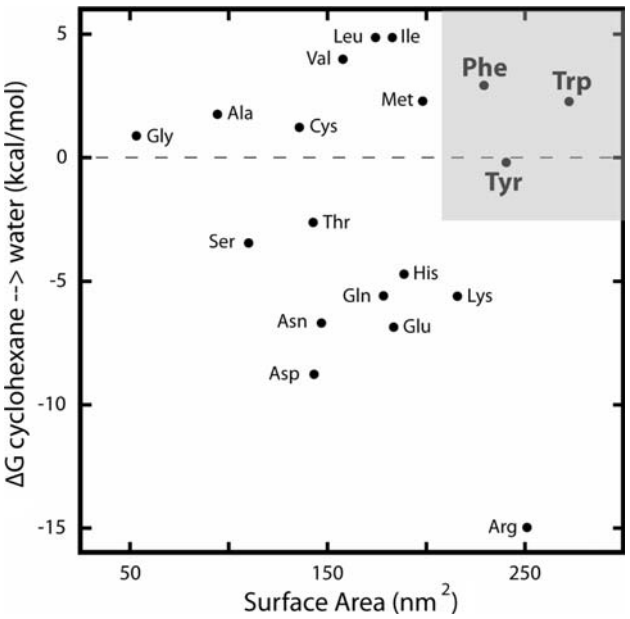

Fig. 3 Plot of the free energy of transfer from dilute cyclohexane solution to dilute aqueous solution, pH 7, versus the solvent-exposed surface area of the sidechains of the 20 genetically encoded amino acids. The transfer energies were reported by Radzicka and Wolfenden for side chain analogues, e.g., methane for alanine and 3-methylindole for tryptophan [15]. The surface areas were calculated by Chothia for tripeptides Gly-X-Gly [14]

thus they do not bind. Conversely, arginine and lysine are large and positively charged, but they are too hydrophilic. The water-to-cyclohexane transfer energy represents a model for the transfer of a molecule from aqueous solution to the hydrophobic interior of a host. Beyond considerations of size and hydrophobicity, shape may also play an important role in binding. The inner surface of Q8 has low curvature, and molecular modeling and considerations of the nonclassical hydrophobic effect [16] suggest that a flat, aromatic group should provide the best fit.

Conclusions

This article provides the first quantitative and comprehensive evaluation of the binding of all 20 amino acids by the two synthetic hosts, Q8 and Q8•MV. We show that both hosts prefer to bind the relatively large and hydrophobic Trp, Phe, and Tyr, in great preference to the other 17 genetically encoded amino acids. We also show a reversal of preference for Trp vs. Phe for Q8•MV vs. Q8, respectively. These results demonstrate that Q8 and Q8•MV offer greater promise than expected for the development of recognition-intensive applications in the context of peptides, proteins, and proteomes.

Acknowledgements This research was funded in part by grants from the Welch Foundation (W-1640) and Research Corporation (CC6517). We also thank the Welch Foundation for a departmental grant that supported summer research for PR. We thank Dr. Joseph J. Reczek and Prof. Jerry Yang for helpful discussions and Lisa M. Ryno and Gretchen A. Vincil for technical support.

References

1. Mock, W.L., Shih, N.-Y.: Organic ligand–receptor interactions between cucurbituril and alkylammonium ions. J. Am. Chem. Soc. **110**, 4706–4710 (1988)
2. Lee, J.W., Samal, S., Selvapalam, N., Kim, H.-J., Kim, K.: Cucurbituril homologues and derivatives: new opportunities in supramolecular chemistry. Acc. Chem. Res. **36**, 621–630 (2003)
3. Lagona, J., Mukhopadhyay, P., Chakrabarti, S., Isaacs, L.: The cucurbit[n]uril family. Angew. Chem. Int. Ed. **44**, 4844–4870 (2005)
4. Liu, S.M., Ruspic, C., Mukhopadhyay, P., Chakrabarti, S., Zavalij, P.Y., Isaacs, L.: The cucurbit[n]uril family: prime components for self-sorting systems. J. Am. Chem. Soc. **127**, 15959–15967 (2005)
5. Kim, J., Jung, I.-S., Kim, S.-Y., Lee, E., Kang, J.-K., Sakamoto, S., Yamaguchi, K., Kim, K.: New cucurbituril homologues: synthesis, isolation, characterization, and X-ray crystal structures of cucurbit[n]uril (n = 5, 7, and 8). J. Am. Chem. Soc. **122**, 540–541 (2000)
6. Kim, H.-J., Heo, J., Jeon, W.S., Lee, E., Kim, J., Sakamoto, S., Yamaguchi, K., Kim, K.: Selective inclusion of a hetero-guest pair in a molecular host: formation of stable charge-transfer complexes in cucurbit[8]uril. Angew. Chem. Int. Ed. **40**, 1526–1529 (2001)
7. Jeon, W. S., Kim, H., Lee, C., Kim, K.: Control of the stoichiometry in host–guest complexation by redox chemistry of guests: inclusion of methylviologen in cucurbit[8]uril. Chem. Commun. 1828–1829 (2002)
8. Sindelar, V., Cejas, M.A., Raymo, F.M., Chen, W., Parker, S.E., Kaifer, A.E.: Supramolecular assembly of 2,7-dimethyldiazapyrenium and cucurbit[8]uril: a new fluorescent host for detection of catechol and dopamine. Chem. Eur. J. **11**, 7054–7059 (2005)
9. Kim, K., Kim, J., Jung, I.-S., Kim, S.-Y., Lee, E., Kang, J.-K.: Cucurbituril derivatives, their preparation methods and uses United States Patent, 6,365,734, Pohang University of Science and Technology (2000)
10. Bush, M.E., Bouley, N.D., Urbach, A.R.: Charge-mediated recognition of N-terminal tryptophan in aqueous solution by a synthetic host. J Am. Chem. Soc. **127**, 14511–14517 (2005)
11. Heitmann, L.M., Taylor, A.B., Hart, P.J., Urbach, A.R.: Sequence specific recognition and cooperative dimerization of n-terminal aromatic peptides in aqueous solution by a synthetic host. J. Am. Chem. Soc. **128**, 12574–12781 (2006)
12. Kim, K., Kim, J., Jung, I.-S., Kim, S.-Y., Lee, E., Kang, J.-K., Method of preparation and uses of cucurbituril devices. United States Patent, 6,639,069. Pohang University of Science and Technology (2000)
13. Cong, H., Tao, L.-L., Yu, Y.-H., Yang, F., Du, Y., Xue, S.-F., Tao, Z.: Molecular recognition of aminoacid by cucurbiturils. Acta Chim. Sin. **64**, 989–996 (2006)
14. Chothia, C.: The nature of the accessible and buried surfaces in proteins. J. Mol. Biol. **105**, 1–12 (1976)
15. Radzicka, A., Wolfenden, R.: Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. Biochemistry **27**, 1664–1670 (1988)
16. Meyer, E.A., Castellano, R.K., Diederich, F.: Interactions with aromatic rings in chemical and biological recognition. Angew. Chem. Intl. Ed. **42**, 1210–1250 (2003)