Molecular Recognition of Amino Acids, Peptides, and Proteins by Cucurbit[n]uril Receptors

Adam R. Urbach
Trinity University, aurbach@trinity.edu

Vijayakumar Ramalingam
Trinity University, vramalin@trinity.edu

Molecular Recognition of Amino Acids, Peptides and Proteins by Cucurbit[n]uril Receptors

Prof. Adam R. Urbach[*] [a] and Dr. Vijayakumar Ramalingam[a]

Department of Chemistry, Trinity University, 1 Trinity Place, San Antonio, TX, 78212, USA

Keywords: sequence-specific, site-specific, host, water, aqueous, binding

Abstract: At the forefront of the endeavor to understand and manipulate living systems is the design and study of receptors that bind with high affinity and selectivity to specific amino acids, peptides, and proteins. Cucurbit[n]urils are among the most promising class of synthetic receptors for these targets due to their high affinities and selectivites in aqueous media and to the unique combination of electrostatic and hydrophobic interactions that govern binding. The fundamental supramolecular chemistry in this area has been explored in depth, and novel, useful applications are beginning to emerge.

[*] to whom correspondence should be addressed: aurbach@trinity.edu, phone +1 210 999 7660, fax +1 210 999 7569

[a] Department of Chemistry, 1 Trinity Place, San Antonio, TX, 78212, USA

1. Introduction

Molecular recognition is the process by which one molecule associates with another molecule via *specific* noncovalent interactions. The specificity of these interactions allows molecules to assemble in manner that is predetermined by their structural attributes, including size, shape, and polarity. In the study of living systems, this topic represents the next level of structural hierarchy in building from molecules to cells— that is, chemists have a relatively well developed understanding of how the arrangements of atoms in molecules influences their physical properties and the covalent chemical reactivity *within* them, but in order to understand the details of their biochemical function, one must also study the associations *between* them. Indeed, the selectivity of molecular recognition exhibited in living systems is exquisite and will fascinate scientists for generations to come.

The study of molecular recognition of biological molecules by synthetic receptors is a burgeoning field that merges the principles and applications of supramolecular chemistry with structurally complex targets in aqueous solution.[1] The cucurbit[n]urils (Qn's) are a family of synthetic, macrocyclic receptors that have been shown to bind organic guests with equilibrium association constant (K_a) values over an enormous range of affinities (up to 10^{15} M^{-1}) in aqueous solution.[2] Therefore, the Qn family is among the most promising class of receptors for targeting biological molecules with affinities and selectivities that are necessary for applications *in vivo*.[3]

The cucurbit[n]urils are cyclic oligomers (of length n) of bis(methylene)-bridged glycoluril (Figure 1). The most commonly studied homologues (Q5, Q6, Q7, Q8, and

Q10)[2f, 4] are pumpkin-shaped containers with similar depth (9.1 Å) but varying

cavity volume (82 Å^3 to greater than 500 Å^3).[2h] The nonpolar cavity may be reached

via entry through either of two negatively charged, constricted portals lined with

ureido-carbonyl oxygens. These features drive the binding of Qn receptors to guests

that contain cationic and nonpolar groups via ion-dipole interactions with the portal

oxygens and hydrophobic interactions within the cavity (Figure 1). High-affinity

guests, many of which are discussed in this review, have both types of groups

arranged for simultaneous contact with the cavity *and* portals.

This paper provides a comprehensive review of the literature pertaining to

cucurbit[n]uril interactions with amino acids, peptides, and proteins. The much

broader topic of peptide and protein recognition by synthetic receptors is beyond the

scope of this review. The reader is directed to other reviews on the subject,[1b, 1g, 5] as

well as a number of interesting recent developments in the area with respect to the

design of receptors for sequence-specific peptide recognition, for pattern recognition,

and for the disruption of protein-protein interactions by targeting sites on protein

surfaces.[6]

2. Amino Acids

The genetic code dictates protein composition from twenty common amino acid

building blocks. Although the structural diversity among the amino acids is sufficient

to produce myriad globular and structural protein products, much of the diversity

arises from the variation in sequence combinations. At the level of a single amino

acid, relatively minimal diversity in the structures of the twenty sidechains presents a

major challenge for molecular recognition. For example, for a receptor that binds to hydrophobic guests (i.e., cyclodextrins), the sidechains of alanine, valine, leucine, isoleucine, proline, phenylalanine, and tryptophan are all potential guests. Subtleties will therefore define selective interaction with a single amino acid or a small subset of amino acids.

2.1. Amino Acid binding by Cucurbit[6]uril

The binding of amino acids to cucurbit[n]uril receptors was first studied by Buschmann and coworkers using isothermal titration calorimetry (ITC) (Table 1).[7] K_a values for Q6 binding to Gly, Ala, Val, and Phe were in the range 1×10^3 - 5×10^3 M^{-1}, showing little variance due to size or hydrophobicity. It is therefore believed that in these studies the amino acids bound as *exclusion* complexes to the portals of Q6 and were not stabilized by appreciable interactions with the Q6 cavity. In all cases, binding was both enthalpically and entropically favorable. In this study, the binding of amino acids to Q6 was compared to that of α-cyclodextrin, which has approximately the same size cavity, and they observed similarly weak affinities among the amino acids. In numerous other cases, Qn's have been been shown to be superior to cyclodextrins for targeting certain guests.[2h]

Table 1. Binding of Q6 to Amino Acids.

Amino Acid	K_a (M^{-1})	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)
Gly	4.7×10^3	-3.1	-1.9
Ala	1.0×10^3	-1.7	-2.4
Val	1.4×10^3	-1.0	-3.2
Phe	1.4×10^3	-1.6	-2.7

ITC experiments in 50% (v/v) aqueous formic acid at 25 °C.[7]

2.2. Amino Acid Binding Mediated by an Auxiliary Guest.

In 2001, Kim and coworkers reported a seminal study in which Q8 was shown to bind simultaneously to two different guests (Figure 2).[8] Q8 binds to only one equivalent of methyl viologen (MV), and the resulting Q8•MV complex binds to 2,6-dihydroxynaphthalene (HN). Binding of HN produces a new visible charge-transfer absorbance and the quenching of HN fluorescence. A crystal structure shows that the two aromatic groups are stacked face-to-face in the cavity of Q8. In a patent on this work, they describe that Q8•MV also binds to the amino acids Trp and Tyr.[9] In these studies, MV is already bound to Q8 when the second guest binds, and thus MV can be thought of as an auxiliary guest that mediates the binding of the second guest.

Our group followed up on this work and quantified the binding of Q8•MV to the aromatic amino acids.[10] The K_a value for the binding of Q8 to MV was determined by ITC to be 8.5×10^5 M^{-1}. The Q8•MV complex was then found to bind selectively to Trp ($K_a = 4.3 \times 10^4$ M^{-1}) with 8- and 20-fold selectivity over Phe and Tyr, respectively (Table 2). No binding was observed for His. NMR studies showed that the aromatic chemical shifts of the amino acid and MV were perturbed upfield upon

binding, which indicates that MV and the amino acid sidechains bound

simultaneously within the cavity of Q8.

Table 2. Binding affinities of Q8 mediated by an auxiliary guest.

Auxiliary	Amino Acid	K (M^{-1})
MV	Trp	4.3×10^{4} [a]
MV	Phe	5.3×10^{3} [a]
MV	Tyr	2.2×10^{3} [a]
DPT	Trp	4.2×10^{5} [b]
MBBI	Trp	3.4×10^{4} [c]

[a] ITC experiments in 10 mM sodium phosphate, pH 7.0, at 27 °C.[10] [b] UV titration experiments; solvent and temperature not reported.[11] [c] ITC experiments in 10 mM sodium phosphate, pH 7.0, at 25 °C.[12]

A broader study found no binding to the remaining 16 amino acids under these

conditions.[13] It is believed that the selectivity of Q8•MV for only three of the twenty

amino acids is based on a combination of hydrophobicity of the sidechain and van der

Waals contacts within the receptor cavity. A plot (Figure 3) of the free energy for

transferring sidechain analogues (e.g., 3-methylindole for Trp, toluene for Phe) from

cyclohexane solution to aqueous solution (a measure of hydrophobicity)[14] vs. the

surface area of the sidechains[15] shows that Trp, Phe, and Tyr are effectively

separated from the other seventeen amino acids on the basis of these properties.

Detailed studies were carried out on tryptophan to explore the effects of the

electrostatic charges of zwitterionic Trp on its binding to Q8•MV.[10] A series of

singly charged Trp derivatives that vary in the number, type, and location of

electrostatic charges (Figure 4) was tested for binding to Q8•MV by ITC, and it was

found that derivatives containing a positive charge bound to Q8•MV at least an order

of magnitude in K_a more tightly than those without a positive charge (Table 3).

Remarkably, there was little difference between groups that varied significantly in the amount of steric bulk, for example tryptamine vs. tryptophan methyl ester. Therefore, the enhancement in binding was likely mediated by the positively charged ammonium group. Binding of all derivatives was enthalpically favorable and entropically unfavorable.

Table 3. Thermodynamic data for Q8•MV binding to derivatives of Trp.

Derivative	K (M^{-1})	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)
Trp	4.3×10^4	-10.6	4.2
1	6.3×10^4	-10.7	4.1
2	5.4×10^4	-12.2	5.7
3	3.1×10^3	-11.1	6.2
4	2.3×10^3	-12.7	8.0

ITC experiments in 10 mM sodium phosphate, pH 7.0, at 27 °C.[10]

That study also confirmed that indole binding to Q8•MV results in a visible charge-transfer (CT) absorbance and the quenching of indole fluorescence, as first indicated by Kim and coworkers.[8] Remarkably, the molar absorptivity of the CT band and the extent of quenching as a function of the extent of binding was consistent among the five indole derivatives, even though the binding constants varied significantly. This result indicated that the mode of binding is likely similar among the series, with the indole and MV groups stacked face-to-face inside the Q8 cavity (Figure 5), and thus similar to the crystal structure reported by Kim and coworkers for the Q8•MV•HN complex. A crystal structure of Q8•MV•indole has yet to be reported.

Kaifer and coworkers developed an excellent replacement for MV as an auxiliary guest for assisting Q8 in the binding of amino acids and other guests.[11] They showed that Q8 binds to 2,7-dimethyldiazaphenanthrenium (DPT, Figure 6) with an affinity of 10^5 M^{-1}, and that the resulting Q8•DPT complex binds 10-fold more tightly to tryptophan than the corresponding Q8•MV complex (Table 2), although it is not clear if the experimental conditions were identical for this comparison. Importantly, they showed that the intrinsic fluorescence of the Q8•DPT complex is quenched upon tryptophan binding. The optical properties of this system are highly advantageous because detection did not depend on the fluorescence of the indole, and thus the system is likely amenable to the sensing of nonfluorescent analytes. More recently, our group in collaboration with the Bielawski and Scherman groups examined tetramethylbenzobis(imidazolium) (MBBI, Figure 6) as an alternative auxiliary guest and found that the Q8•MBBI complex bound to Trp with very similar to that of Q8•MV (Table 2).[12] MBBI is discussed in more detail in the context of peptide binding in Section 3.4.

2.3. Amino Acid Binding without an Auxiliary Guest.

Following the studies described above on Q8•MV binding to amino acids, several reports on amino acid binding by Qn analogues and homologues have appeared. Isaacs and coworkers synthesized an expanded and fluorescent analogue of Q6 and studied its interaction with a range of guests including Trp, Phe, Tyr, and His by fluorescence titration experiments (Table 4).[16] Remarkably, the elongated, aromatic

Q6 analogue bound to Trp with high affinity (K_a = 3.2 x 10^6 M^{-1}) and 1-2 orders of magnitude selectivity versus Phe and Tyr, likely due to stronger binding of the indole sidechain. This order of selectivity had been observed previously for binding to Q8•MV,[10] but not to the same extent of selectivity exhibited with the elongated Q6 analogue. No fluorescence response was observed for His, suggesting that the sidechain is protonated and does not pi-stack with the walls of the host inside the cavity. In a subsequent study another Q6 analogue, (±)-bis-nor-seco-cucurbit[6]uril, was found to bind stereoselectively to Phe.[17]

Table 4. Binding data for binary complexes with amino acids.

Host	Amino Acid	K_a (M^{-1})	Ref
Q6 analogue	Phe	4.2 x 10^4	16[a]
Q6 analogue	Tyr	5.7 x 10^4	16[a]
Q6 analogue	Trp	3.2 x 10^6	16[a]
Q7	Phe	8.2 x 10^5	18[b]
Q7	Tyr	2.3 x 10^5	18[b]
Q7	Trp	3.7 x 10^5	18[b]
Q7	His	8.0 x 10	18[b]
Q7	Glu	1.0 X 10^2	18[b]
Q7	Met	2.7 x 10^2	18[b]
Q7	Val	4.4 x 10^2	18[b]
Q7	Leu	1.5 x 10^2	18[b]
Q7	Ala	3.6 x 10	18[b]
Q7	Phe	1.5 x 10^5	2i[c]
Q7	Phe	1.8 x 10^6	2j[d]

Q7	Trp	1.9×10^3	19[e]
Q7	Tyr	2.2×10^4	19[e]
Q7	Lys	8.0×10^2	19[e]
Q7	Arg	3.3×10^2	19[e]
Q7	Trp	1.6×10^3	19[f]
Q7	Tyr	2.4×10^4	19[f]
Q7	Lys	8.7×10^2	19[f]
Q7	Arg	3.1×10^2	19[f]
Q7	His	4.0×10^2	19[f]

[a] Fluorescence titration in 50 mM NaOAc, pH 4.74 at 25 °C.[16] [b] UV titration.[18] [c] Competitive ^{1}H NMR titration in water at 25 °C.[2i] [d] ITC in water at 25 °C.[2j] [e] ITC in 10 mM NH$_4$OAc, pH 6.0 at 30 °C.[19] [f] Competitive fluorescence titration in 10 mM NH$_4$OAc, pH 6.0 at 25 °C.[19]

Tao and coworkers studied the binding of Q7 with a series of amino acids by UV-visible spectroscopy.[18] They reported K_a values for the binding of Q7 to Phe, Tyr, and Trp on the order of 10^5 M^{-1} (Table 4), as well as 1:1 ratios of binding of Q7:amino acids. They report much lower affinities for His, Glu, Met, Val, Leu, and Ala ($K_a < 500$ M^{-1}). It is interesting that the difference in binding of Q7 to aromatic vs. non-aromatic amino acids was so similar to the pattern observed for Q8•MV, suggesting a similar mechanism of selectivity. The affinity of Q7 for Phe has also been reported as 1.5×10^5 M^{-1} by competitive NMR titration[2i] and 1.8×10^6 M^{-1} by isothermal titration calorimetry, both in water,[2j] which highlights the importance of the salt content in measuring binding affinities.

In the absence of MV, Q8 was found to bind to two equivalents of Trp and Phe with overall equilibrium constants of 6.9×10^7 and 1.1×10^8 M^{-2}, respectively.[13] The binding of Q8 to two equivalents of Trp- and Phe-containing peptides is described in

detail in Section 3.5. Tao and coworkers reported three crystal structures of Q8 in complex with two equivalents of Tyr, His, or Leu,[20] showing that the sidechains bind within the cavities, and the ammonium groups interact with the carbonyl oxygens on Q8, as had been observed previously for Q8•peptide complexes (discussed in Section 3.5).

2.4. Amino Acid-Related Applications of Cucurbit[n]urils.

Nau and coworkers developed a breakthrough application of cucurbit[n]urils for the measurement of enzyme activity.[21] Their "supramolecular tandem enzyme assay" is based on the selective and competitive displacement of a reporter dye from the cavity of Q7 by the product of an enzyme-catalyzed reaction (Figure 7). They demonstrated this assay on amino acid decarboxylases, in which the substrate amino acids (Lys, Arg, His, Tyr, and Trp) bind significantly more weakly to Q7 than their decarboxylated products (cadaverine, agmatine, histamine, tyramine, and tryptamine, respectively). The reporter dye was dapoxyl, which binds tighter than substrate but weaker than product, and thus the dye is displaced as the enzyme-catalyzed reaction proceeds. The kinetics of competitive binding are faster than that of the enzyme, and thus the reaction is reported in real time by fluorescence spectroscopy. Using this assay and the intrinsic enantioselectivities of the decarboxylases for their L-amino acid substrates, Nau and coworkers subsequently reported a multiparameter sensor assay that signals only in the presence of a reactive pair of an L-amino acid and its corresponding decarboxylase.[19] This paper also reports binding affinities of Q7 for

Trp, Tyr, Lys, Arg, and His using ITC and a competitive fluorescent indicator displacement assay (they observed close correlation of the fluorescence and ITC experiments). The reported K_a values for binding of Q7 to Trp and Tyr (Table 4) are lower than previously reported, which may be due to the 10 mM ammonium acetate pH 6.0 buffer.

The action of lysine decarboxylase on lysine was subsequently used by Du and coworkers to mediate the release of protein from magnetic mesoporous silica particles.[22] Fe_3O_4-embedded magnetic mesoporous silica decorated with silane-tethered alkylammonium groups was complexed with Q7 at the cationic sites. Calcein dye was loaded in the porous particles before Q7 complexation, and lysine was added to the mixture. Addition of lysine decarboxylase produces cadaverine as the product, which competitively binds to Q7 and releases it from the surface of the porous silica, thereby releasing the calcein dye. The extent of dye release could be controlled by the lysine concentration.

3. Peptide Binding by Cucurbit[n]urils

Peptides differ structurally from amino acids in several ways. In addition to their increased size, peptides have an oligoamide backbone with only one ammonium group at the N-terminus and one carboxylate group at the C-terminus, they have multiple (sometimes numerous) sidechains, and their structure depends on the sequence of amino acids. These features create both challenges and opportunities for

molecular recognition. This section reviews fundamental and applied studies in the supramolecular chemistry of cucurbit[n]urils with peptides.

3.1. Peptide Binding by Cucurbit[6]uril

Early work on peptide binding by the cucurbit[n]uril family was reported by Buschmann and coworkers, who studied the interaction of Q6 with dipeptides and tripeptides by ITC (Table 5).[7b, 23] All peptides had affinities in the range 3.7×10^2 M^{-1} - 1.5×10^3 M^{-1}. Given the relatively small binding constants, the minimal variance with respect to the size and sequence of the peptide, and the small size of Q6, it is likely that these peptides form exclusion complexes with the portals of Q6 as with the amino acids described in section 2.1. An interesting study of lysine and oligo(Lys) binding in the gas phase was carried out by Dearden and coworkers,[24] who deduced from electrospray mass spectrometry experiments and Monte Carlo calculations that Q6 bound particularly well to the N-terminal Lys but that all Lys sidechains in the Lys$_5$ peptide are potential binding sites.

Table 5. Binding of Q6 to Peptides.

Peptide	K_a (M^{-1}) [a]
Gly-Phe	1.1×10^3
Gly-Gly	7.9×10^2
Gly-Leu	3.7×10^2
Gly-Val	1.5×10^3
Gly-Ala	6.3×10^2
Leu-Val	6.2×10^2
Gly-Asn	6.6×10^2
Leu-Phe	6.0×10^2
Leu-Trp	8.3×10^2
Gly-His	6.2×10^2

Leu-Gly-Phe	5.5×10^2
GSH [b]	6.3×10^2

[a] ITC experiments in 50% (v/v) aqueous formic acid at 25 °C.[7b, 23] [b] Reduced glutathione.

3.2. Molecular Recognition of N-Terminal Tryptophan

In the studies described in Section 2.2 on the binding of Q8•MV to tryptophan

derivatives, our group observed that tryptophan derivatives containing a positively

charged ammonium group bound more tightly than those without this group (Figure 4

and Table 3) likely due to stabilizing interactions between the positively charged

ammonium group and the negatively charged carbonyl oxygens of Q8. On the basis

of these results, a critical connection between amino acids and peptides was first

made (Figure 8).[10] It was observed that tryptophan located at the N-terminal position

of a peptide chain mimics the chemical structure of the positively charged tryptophan

derivatives, and that a tryptophan located at the C-terminal position mimics the

chemical structure of the negatively charged tryptophan derivatives. Therefore, it was

hypothesized that Q8•MV should bind selectively to peptides containing an N-

terminal Trp residue versus those with a C-terminal Trp, thus providing a mechanism

for sequence-specific peptide recognition.[10]

To test this hypothesis, a series of peptides was designed to place one tryptophan

at N-terminal (**WGG**), C-terminal (**GGW**), or non-terminal positions (**GWG** and

GGWGG). In this design **WGG** is analogous to Trp-OMe and TrpA, whereas **GGW**

is analogous to N-AcTrp and IPA (Figure 8). In addition, **GWG** and **GGWGG** were

designed to examine the effects of moving the terminal charges one and two residues,

respectively, from the indole sidechain. ITC experiments on these peptides (Table 6)

showed that the Q8•MV complex binds with highest affinity to the N-terminal **WGG** with 6-fold selectivity over **GWG** and **GGWGG** (which had virtually identical binding thermodynamics), and 40-fold selectivity over **GGW**. As with the monomeric tryptophan derivatives, binding was enthalpically driven and entropically unfavorable, and increased binding was driven by increasing exothermicity with some compensation from entropy. NMR spectra in D_2O of mixtures of Q8, MV and peptides showed upfield perturbations and broadening of the indole peaks, indicating binding inside the cavity of the Q8. These data suggested that recognition of the N-terminal Trp residue is mediated by a combination of ion-dipole interactions between the N-terminal ammonium group and proximal carbonyl oxygens on Q8 in addition to hydrophobic interactions of the indole sidechain with the cavity of Q8 (Figure 9).

Table 6. Thermodynamic data for Q8•MV binding to peptides.

Peptide	K (M^{-1})	ΔH (kcal/mol)	-TΔS (kcal/mol)
WGG	1.3×10^5	-14.8	7.8
GWG	2.1×10^4	-11.4	5.5
GGWGG	2.5×10^4	-12.1	6.1
GGW	3.1×10^3	-8.8	4.0

ITC experiments in 10 mM sodium phosphate, pH 7.0, at 27 °C.[10]

This result was the first demonstration of sequence-specific peptide recognition by the cucurbit[n]uril family of synthetic receptors.[10] In the molecular recognition of biopolymers (e.g., DNA and proteins) by synthetic receptors, predictive binding on basis of the sequence of building blocks (e.g., nucleotides or amino acids) is highly advantageous because it reduces the need for three-dimensional structural information, which is more difficult to obtain than the corresponding sequence. This

concept has been most elegantly demonstrated on DNA,[1a] where small molecules are shape-matched with the curvature of the major or minor grooves, and sequence discrimination is accomplished with complementary patterns of hydrogen bonds between ligand and DNA. In the context of peptides, there are several excellent examples of sequence-specific recognition by synthetic receptors in aqueous solution.[1b, 6c, 6d, 25] Surprisingly few of these reports include binding affinities in excess of 10^4 M^{-1} or significant selectivity for a target sequence. Therefore, the Q8•MV system, with its high affinity and sequence selectivity, was a significant contribution to the field.

As with the binding of Q8•MV to Trp derivatives described in Section 2.2, it was observed that the binding of Q8•MV to peptides containing Trp results in a new visible charge-transfer absorbance and the quenching of indole fluorescence. This "built-in" capacity to detect peptides by commonly available optical techniques is useful for measuring peptide binding and could be useful for the development of peptide-specific sensing devices.

Scherman and coworkers recently demonstrated an interesting application of this system to the reversible capture and release of peptides.[26] Viologen-terminated alkanethiols were assembled on a Au substrate and used to trap Q8 noncovalently. The substrate was then treated with fluorophore-conjugated peptides containing an N-terminal Trp residue, and they observed selective capture of these peptides on the Q8•viologen surface. A negative potential was applied to convert the viologen dication to the cation radical, and this electrical stimulus induced the release of the peptide. The active capture surface could then be reactivated by electrochemical

oxidation, and this process was shown to be repeatable numerous times without degradation of the substrate.

3.3. Multivalent Binding of Peptides by Modular Self-Assembled Receptors

Optical sensing in the Q8•MV•Trp system was made possible by the properties conferred to the receptor by the viologen guest, not just the Q8 host. Another type of application involving an auxiliary guest was demonstrated in the study of multivalent interactions, which involve the association of molecules via the simultaneous interaction of multiple host sites with multiple guest sites.[27] In theory, the energy of these interactions are approximately additive, and thus multivalent binding has the potential to dramatically stabilize complex formation.[27c] In practice, however, this has rarely been observed,[28] and much remains to be understood about how to control the energetics of multivalent systems by design. Two significant challenges slow progress toward this end: 1) the difficult chemical synthesis of water-soluble multivalent receptors;[29] and 2) the measurement of the number of simultaneous contacts that stabilize multivalent complexes. These challenges were addressed by our group by making use of the auxiliary viologen guest in the Q8•MV•Trp system.[10, 30]

Instead of linking macrocycles covalently, a scaffold was used to assemble the macrocycles noncovalently (Figure 10).[30] Peptide-based scaffolds presenting multiple viologen groups (**5** - **7**) were synthesized on solid support (Figure 11)[31] and shown to recruit an equivalent number of Q8 molecules in a non-cooperative manner (K_a = 2 x 10^6 M^{-1} per binding site regardless of the number of sites per scaffold). The

resulting multivalent receptors bound peptides containing an equivalent number of tryptophan residues in a discrete, multivalent fashion with 20-200-fold increase in affinity (Table 7). UV-visible spectroscopy was used to confirm the simultaneous binding of all tryptophans with all Q8•viologen groups based on the precisely additive charge-transfer absorptivity in comparing 1+1, 2+2, and 3+3 complex formation. Therefore, this system has a built-in observable for measuring the valency of a multivalent interaction.[30]

Table 7. Thermodynamic binding data for monovalent and multivalent complexes.

Receptor	Peptide	valency	K (M^{-1})	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)
5•Q8	GGWGG	1+1	2.2×10^4	-10.8	4.9
6•Q8$_2$	(GGWGG)$_2$	2+2	5.0×10^5	-24.2	16.3
7•Q8$_3$	(GGWGG)$_3$	3+3	4.7×10^6	-39.4	30.2

ITC experiments in 10 mM sodium phosphate, pH 7.0, at 27 °C.[30]

The 30-300-fold increase in affinity is far below the optimal gains on the order of 10^4-10^8 fold based on additive energies. Some insight to this commonly observed suboptimal gain in affinity due to multivalency was found by analysis of the detailed thermodynamics. In this system, the enthalpy of binding was additive in comparing 1+1, 2+2, and 3+3 binding (Table 7), which indicated that the binding groups were able to bind simultaneously. The entropy of binding, however, was more than additively unfavorable as valency increased. This result was likely due to an increasing loss in conformational freedom in the oligo(Gly) backbone upon binding. The modularity of this system enables such a systematic analysis of the energetic effects of multivalency. The synthetic and analytical techniques involved in this

approach should allow for relatively rapid iteration in studying the structure-activity relationships in multivalent binding.

3.4. Benzobis(imidazolium) as an Auxiliary Guest for Peptide Recognition and Sensing.

Methyl viologen has proved useful for assisting in the binding and detection of peptides as the first guest in ternary Q8 complexes. MV is, however, unstable to mild reducing conditions and has no interesting optical properties of its own, and thus other first guests for Q8 have been explored. An interesting alternative to methyl viologen as a first guest for Q8 is tetramethylbenzobis(imidazolium) (MBBI),[12] which is similar in size and charge separation as MV (Figure 6). In this study, our group in collaboration with the Bielawski and Scherman groups demonstrated that Q8 binds to MBBI ($5.7 \times 10^5 \, M^{-1}$) with similar affinity as MV ($8.5 \times 10^5 \, M^{-1}$), and that the resulting Q8•MBBI complex binds to the Trp-containing peptides with identical affinity as the corresponding Q8•MV complex (Figure 12), thus showing sequence-specific binding. NMR spectra in D_2O of mixtures of Q8, MV and peptides showed upfield perturbations and broadening of the indole peaks, indicating simultaneous binding inside the cavity of the Q8.

The similarity in the binding of peptides by Q8•MV and Q8•MBBI points to the role of the first guest as stabilizing the second guest by filling space in the cavity and providing a surface for van der Waals contact, as opposed to stabilizing the complex by charge-transfer interactions. Given the similar binding properties of MBBI and

MV as first guests, the chief advantages of MBBI lie in its high stability to reducing

conditions and to heat (and thus more amenable to synthetic modification), and its

strong intrinsic fluorescence. Optical detection of Trp-containing peptides by the

quenching of MBBI fluorescence is more sensitive than by the quenching of indole

fluorescence, and the Q8•MBBI system is amenable to nonfluorescent analytes, thus

opening the possibility of sensing a broader range of peptides and other analytes.

3.5. Sequence Recognition and Peptide Dimerization by Cucurbit[8]uril

In characterizing the ternary complexes of Q8 with MV and Trp-containing

peptides, the question arose early whether peptides could bind directly to Q8 if MV

were not present as an auxiliary guest. If so, we anticipated that the Q8 cavity could

accommodate two peptides. Given that this phenomenon was not observed in the

presence of MV, it could be deduced that the binding of Q8 to two equivalents of

peptide (i.e., AAB ternary complex) is not as stable as the ABC ternary combination.

Nonetheless, further study was merited. Our group tested the binding of Q8 to

aromatic peptides of sequence **X**GG, G**X**G, and GG**X** (**X** = Trp, Phe, Tyr, His) by

ITC.[32] These experiments revealed several interesting phenomena: 1) out of twelve

peptides, only **W**GG and **F**GG were observed to bind to Q8; 2) the stoichiometry was

2:1 (peptide:Q8); and 3) **F**GG bound more tightly than **W**GG due to an enthalpic

advantage (Table 8). NMR spectra of **F**GG:Q8 mixtures at different guest:host ratios

indicated positive cooperativity due to the dominance of the 2:1 complex in

substoichiometric mixtures. The selectivity of Q8 for these two peptides is

remarkable because all peptides in the study had an aromatic sidechain as well as an

N-terminal ammonium group. Only when the two groups were positioned proximal to one another, however, was binding observed. The sequence selectivity in this system is large (>100-fold for **WGG** and >1000-fold for **FGG**) and perhaps unprecedented for a synthetic receptor. Moreover, the selectivity of Q8 alone was much greater than Q8•MV, which bound N-terminal Trp only 6-fold more tightly non-terminal Trp.

Table 8. Thermodynamic binding data for 2:1 Peptide:Q8 complexes.

Peptide	K (M^{-2})	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)
Phe-Gly-Gly	1.5×10^{11}	-29.6	14.2
Trp-Gly-Gly	3.6×10^{9}	-22.8	9.7

ITC experiments in 10 mM sodium phosphate, pH 7.0, at 27 °C.[32]

This study provided the first structural basis for peptide recognition by Qn macrocycles.[32] Sub-Ångstrom-resolution crystal structures of Q8•**WGG** and Q8•(**FGG**)$_2$ (one Q8 bound to two **FGG** molecules) were reported, showing in both cases the inclusion of the aromatic sidechain within the Q8 cavity and close interaction of the N-terminal nitrogen with the carbonyl oxygens of Q8 (Figure 13). These interactions support the observed selectivity of Q8 for the N-terminal aromatic residue. In addition, close dipole-dipole interactions between the peptide N-H groups and proximal carbonyl oxygens of Q8 were observed. In the case of the **FGG** dimer structure, the phenyl rings of the two respective Phe residues were stacked in a staggered face-to-face orientation within the Q8 cavity. These favorable interactions may explain the enthalpic advantages of this complex and the selectivity for **FGG** versus **WGG**.

3.6. Peptide Binding by Cucurbit[7]uril

Soon after this study was published, Kim, Inoue and coworkers reported a study of the binding of peptides containing N-terminal Phe residues by the smaller macrocycle Q7.[33] Q7 is large enough to bind only one aromatic guest, and 1:1 Q7:peptide binding was indeed observed in all cases. This study focused on the ability of Q7 to discriminate diastereomers of dipeptides such as L-Phe-L-Pro vs. L-Phe-D-Pro, and compared the effects of zwitterionic peptides (amino and carboxylate termini) to cationic peptides (amino and amide termini) using ITC (Table 9). In the series of zwitterionic peptides, L-Phe-L-Ala and L-Phe-L-Pro bound Q7 with slightly higher affinity (1.7-fold) than the respective L-D diastereomers, and the Phe-Ala peptides bound more tightly (25-30-fold) than the Phe-Pro peptides. The authors suggested that the diminished conformational flexibility of the proline residue negatively impacts the binding geometry, resulting in fewer stabilizing interactions, as evidenced by the reduction in exothermicity for binding of the Phe-Pro peptides. In the cationic peptide series, an interesting reversal of diastereomeric selectivity (e.g., K_{LL} vs. K_{DL}) was observed. The C-terminal carboxamide-containing peptides L-Phe-L-Ala-CONH$_2$ and L-Phe-L-Leu-CONH$_2$ bound to Q7 4-8-times more tightly than the respective L-D diastereomers. On the basis of these data, the authors suggested that recognition of the zwitterionic peptides is driven by long-range repulsion of the negatively charged carboxylate terminus with the Q7 oxygens, whereas recognition of the cationic peptides is driven by short-range van der Waals interactions. The lack of an NOE between one of the Phe β-CH protons of L-Phe and the outer surface protons of Q7 in

the ROESY spectrum of the Q7•L-Phe-L-Leu-CONH$_2$ complex was used to rationalize a deeper insertion of the Phe residue within the Q7 cavity and thus stronger intermolecular interactions and greater exothermicity of binding. It was suggested that the origin of diastereomeric recognition in this system lies in the optimization of van der Waals contacts between the peptide and host, and thus the achiral nature of the host should play little if any role in recognition.

Table 9. Equilibrium binding data for Q7 with peptides.

Peptide	K (M^{-1})
Phe-Ala	7.9 x 10^5 [a]
D-Phe-Ala	1.3 x 10^6 [a]
Phe-Pro	2.9 x 10^4 [a]
D-Phe-Pro	5.0 x 10^4 [a]
Phe-Phe-CONH$_2$	5.3 x 10^6 [a]
D-Phe-Phe-CONH$_2$	1.3 x 10^6 [a]
Phe-Leu-CONH$_2$	1.4 x 10^7 [a]
Phe-D-Leu-CONH$_2$	1.7 x 10^6 [a]
Phe-Gly	3.0 x 10^7 [b]
Gly-Phe	1.3 x 10^3 [b]
Tyr-Gly	3.6 x 10^6 [b]
Gly-Tyr	2.0 x 10^2 [b]
Trp-Gly	5.6 x 10^5 [b]
Gly-Trp	2.8 x 10^2 [b]

[a] ITC experiments in 0.1 M NaCl, pH 7.0, at 25 °C.[33] [b] ITC experiments in water at 25 °C.[34]

Kim and Inoue have also studied the binding of Q7 with peptides containing Trp and Tyr residues.[34] Using ITC in pure water they showed that Q7 binds selectively to **X**-Gly vs. Gly-**X** (**X** = Phe, Tyr, Trp) with selectivities for the N-terminal aromatic peptides of 2000-23000-fold (Table 9). This result is consistent with the prior work on recognition of N-terminal aromatic peptides by Q8[32] and the Q8•MV complex,[10] The strong and selective binding of Tyr-Lys vs. Lys-Tyr by Q7 and the weak but selective binding of Lys-Tyr vs. Tyr-Lys by the smaller Q6 was used to show that

when these four molecules are mixed and analyzed by NMR, they "self-sort"[2i] selectively into the Q7•Tyr-Lys and Q6•Lys-Tyr complexes.

5. Protein Binding by Cucurbit[n]urils

From a molecular recognition standpoint proteins are among the most attractive targets for drug design, medical diagnostics, and separations because they are involved in so many biological processes and so much has yet to be learned about their chemical and biological properties. From a practical standpoint, proteins differ from peptides in that they are larger, are more likely to be folded stably, and many of their sidechains are likely to be buried in the interior of the folded macromolecule. The larger size means there are more possible binding sites. The folded structure means that not all residues are available for binding, and that the exact surface of the folded protein and the surrounding water molecules is extraordinarily difficult to predict without high-resolution structural information. The majority of this review has focused on the binding of amino acids and peptides because this fundamental work is critical to understanding the detailed interactions of these molecules with the cucurbit[n]urils. With the information gained from these studies, several groups have begun to explore the binding of Qn receptors to proteins as well as the application of other cucurbit[n]uril chemistry to address problems in protein science.

Brunsveld and coworkers applied the Q8•(FGG)$_2$ interaction described in Section 3.5 to the dimerization of proteins.[35] They used intein-based protein expression to place the FGG sequence at the N-terminus of yellow fluorescent protein (YFP) and of

cyan fluorescent protein (CFP). Upon adding Q8 to a homogeneous sample of FGG-YFP they observed noncovalent dimerization of the protein by Förster resonance energy transfer (FRET), isothermal titration calorimetry, mass spectrometry, and size-exclusion chromatography. Experiments with a mixture of FGG-terminated YFP and CFP also showed a strong FRET. The protein dimerization was reversible upon addition of MV as a competitive ligand for Q8.

Wang and coworkers demonstrated the use of Q8 to mediate the binding of a photodynamic therapy sensitizer to bovine serum albumin (BSA).[36] The fluorescence of Trp residues on BSA was quenched to a modest extent in the presence of the sensitizer 5,10,15,20-tetrakis(1-methyl-4-pyridinio)porphyrin (TPP). Addition of Q8 to the mixture was shown to strongly enhance the extent of quenching, whereas Q7 produced no effect. NMR studies also showed substantial upfield perturbation to the aromatic peaks of the pyridinium and indole peaks in the presence of Q8. These effects are therefore likely due to the simultaneous inclusion of a pyridinium group of TPP and the indole sidechain of tryptophan in the cavity of Q8, as observed in the studies of Q8•MV with tryptophan derivatives and Trp-containing peptides.[10] Dynamic light scattering and atomic force microscopy experiments showed that addition of Q8 and TPP to a BSA solution induced the formation of BSA aggregates. They also reported that addition of Q8 and TPP significantly enhanced the light-induced cleavage of BSA, likely due to the observed increase in the triplet excited state lifetime of TPP.

Bhasikuttan, Nau and coworkers reported the use of Q7 to enhance the fluorescence and binding of a sensitizing dye, Brilliant Green, to BSA.[37] They

observed enhancement of fluorescence by the dye in the presence of Q7 *or* protien

and further enhancement in the presence of Q7 *and* protein, indicating the formation

of a Q7•dye•protein ternary complex. This finding was further supported by an

observed 10-fold increase in affinity of the dye for the protein in the presence of Q7

(K_a = 3.9 x 10^6 M^{-1}) versus in the absence of Q7 (K_a = 3.2 x 10^4 M^{-1}). Subsequently,

Mohanty and Pal showed that increasing concentrations of NaCl induced the transfer

of a dye, Neutral Red, from the binding pocket of Q7 to that of BSA by competitive

binding.[38]

Our group has applied the N-terminal recognition concept to a natural protein,

human insulin, which has an N-terminal Phe residue on the B-chain (i.e., Phe^{B1}).[39]

ITC experiments showed that Q7 binds to insulin with a K_a value of 1.5 x 10^6 M^{-1}

(Table 10). A variant with Glu at the B1 position showed no binding by ITC and thus

a loss of more than three orders of magnitude in binding affinity. **F**GG was tested as a

positive control under these conditions and found to bind with similar affinity as

insulin. G**F**G and G**Y**G peptides were tested as models for the non-terminal Phe and

Tyr residues on the surface of insulin and found to bind 100- and 1000-fold more

weakly than the N-terminal sequence isomer. The Glu variant also had non-terminal

Phe and Tyr but did not bind to Q7, suggesting that steric hindrance may be a

problem for protein surface binding at non-terminal sites. A crystal structure of the

Q7•insulin complex (Figure 14) revealed binding of Q7 at the Phe^{B1} position and

showed that the last few residues of the N-terminus of the B-chain unravel, like a ball

of string, to accommodate Q7 binding. These results suggest that the terminus of a

protein is well suited for targeting by a synthetic receptor because the terminal

residue is a unique chemical epitope that cannot exist elsewhere on the protein, and because the terminus can unfold more easily than other sites on the protein in order to accommodate binding. A fluorescent indicator displacement assay[40] was used to show that Q7 binds selectively to insulin versus a series of other blood proteins and short peptides that lack an aromatic N-terminus.

Table 10. Equilibrium Binding Data for Q7

Protein or Peptide	K_a (M^{-1})
native insulin	1.5×10^6
variant insulin	$< 10^3$
FGG	2.8×10^6
GFG	2.2×10^4
GYG	2.7×10^3

ITC experiments at 27 °C in 10 mM sodium phosphate, pH 7.0.[39]

6. Protein-Related Applications of Cucurbit[n]urils

Numerous applications of cucurbit[n]urils in protein science and technology have been reported. The high selectivities and wide-ranging affinities of Qn receptors provide a tool kit for building assemblies with controlled structures and properties.

Directed protein assembly and capture is a major theme. Kim and coworkers reported a very unusual observation that an amphiphilic Q6 derivative assembles into vesicles that can be modified noncovalently at their surfaces via specific interactions with alkylammonium guests.[41] Using an alkylammonium-modified mannose, they

showed that the vesicles selectively assemble the lectin Concanavalin A in a multivalent fashion.

In another breakthrough application of cucurbit[7]uril,[42] Kim and coworkers reported the selective capture of a protein using the extraordinarily stable complex between Q7 and ferrocenemethylammonium (AFc) ($K_a \sim 10^{12}$ M^{-1}) as an intermediary.[2j] A self-assembled monolayer of alkanethiolates covalently linked at their ω-termini to Q7 was prepared and shown by surface plasmon resonance spectroscopy to selectively capture molecules of glucose oxidase that had been conjugated to numerous equivalents of ferrocenemethylammonium groups. The captured enzymes were demonstrated to be active for glucose oxidation. Remarkably, this technique was recently applied to the capture of labeled cell-surface receptors by Q7-coated beads.[3c] Therefore, the highly stable Q7•AFc can be an effective small-molecule replacement for biotin-avidin in surface-immobilization and target-capture experiments. The Q7•AFc-mediated capture of proteins has also been used to print well aligned monolayers of proteins.[43] Brunsveld, Jonkheijm and coworkers used expressed protein ligation to conjugate a single equivalent of ferrocenemethyl-ammonium group to yellow fluorescent protein. The protein-AFc conjugates were deposited onto a Q7 monolayer that had been formed by spontaneous adsorption of Q7 on Au and shown to reach a stable and densely packed layer of protein. They further demonstrated the compatibility of this approach with microcontact printing from a patterned poly(dimethylsiloxane) stamp. Interestingly, treatment of the surface with free AFc ligand resulted in quantitative displacement of the protein, and thus printing was shown to be reversible. Very recently,[44] Scherman and coworkers used

the Q8•MV•HN interaction[8] to selectively and reversibly conjugate BSA to a poly(ethylene glycol) chain. A major difference between the weaker binding Q8•MV•HN vs. the Q7•ferrocene system for linking molecules to proteins is the reversibility of the Q8•MV•HN under practical conditions.

Nau and coworkers demonstrated that Q7 can effectively inhibit the action of a protease by binding to the substrate.[45] Trypsin and leucine aminopeptidase (LAP) activity toward certain peptide substrates was inhibited in the presence of Q7. Activity was not inhibited in the presence of other substrates, and thus inhibition was mediated by the binding of Q7 to the substrate and not to the enzyme. Although the exact sites of Q7 binding were not determined, knowledge of the chemical mechanism of the proteases was used to deduce likely binding sites.

Ghosh and Isaacs used "two-faced" compounds to regulate enzyme activity via interplay with Q7.[46] Compounds presenting an enzyme-binding group (arylsulfonamide for carbonic anhydrase or tacrine for acetylcholinesterase) linked to a Q7-binding group (aminoadamantyl, tetraalkylsilyl, or pentyl) were prepared, and enzyme activity was measured in the presence of these compounds with and without Q7. In the case of carbonic anhydrase, the arylsulfonamide-containing compounds inhibited enzyme activity, and addition of Q7 restored activity by sequestering the inhibitor from the active site. This process was reversed repeatedly by addition of a high-affinity competitor for Q7, and then addition of more Q7. In the case of acetylcholineasterase, addition of Q7 to the enzyme inhibitor complex did not sequester the inhibitor, but rather formed a stable ternary enzyme•inhibitor•Q7 complex. This study highlights the interesting possibilities for Qn-mediated enzyme

control and the challenges with respect to different types of enzymes. The careful choice of conditions is important, however, as high concentrations (>0.1 mM) of host can inhibit enzyme activity by effectively sequestering all of the substrate.[47]

Q7 was used to enhance the readout of enzyme-linked immunosorbent assays (ELISA).[48] Au nanoparticles have been shown to aggregate via binding to Q7 on the Au surface.[43, 49] de la Rica and Velders used UV-visible spectroscopy and dynamic light scattering experiments to demonstrate that urease, which produces ammonium ions upon reaction with urea, induces dissociation of Q7-nanoparticle assemblies via the competition of ammonia for binding to Q7. This activity was then applied to an ELISA for the detection of mouse immunoglobulin G.

An interesting application of the sequence-selectivity of Qn receptors for aromatic peptides was demonstrated recently.[50] The metalloendopeptidase thermolysin selectively cleaves the amide bond to the nitrogen side of Phe residues in substrate peptides, thus producing product peptides containing an N-terminal Phe. Q7 binds weakly to non-terminal Phe residues ($K_a = 10^4$ M^{-1}) but strongly to N-terminal Phe ($K_a > 10^6$ M^{-1}). Using the supramolecular tandem enzyme assay approach described in Section 2.4 for amino acid decarboxylases,[19, 21] our group in collaboration with Nau and coworkers developed an enzyme assay for proteases (Figure 15). The reporter pair in this assay comprised Q7 and acridine orange, which binds to Q7 with an affinity ($K_a = 2$ x 10^5 M^{-1}) that is between that of the substrate (e.g., Thr-Gly-Ala-Phe-Met-CONH$_2$) and product (e.g., Phe-Met-CONH$_2$).[51] By measuring the rate of hydrolysis as a function of the substrate concentration, Michaelis-Menten k_{cat}/K_M values (Table 11) were obtained for a series of enkephalin-type peptides that

corresponded well to previously determined values using fluorescently labeled peptides. This assay was used to accurately determine sequence specificity (e.g., Ser vs. Ala), stereoselectivity (e.g., L-Ala vs. D-Ala), and endo vs. exopeptidase activity. By using a known protease inhibitor, phosphoramidon, the assay was validated for the measurement of inhibitory constants.

Table 11. Equilibrium Binding Data for Q7.

Peptide sequence	K_a $(M^{-1})^a$	k_{cat}/K_M $(10^4\ s^{-1}\ M^{-1})^b$
TGAFM-CONH$_2$	1.3×10^4	14
TGDAFM-CONH$_2$	2.6×10^4	≤ 0.005
TGAFL-CONH$_2$	3.5×10^3	3.2
TGSFM-CONH$_2$	1.9×10^4	6.9
TGGFM-CONH$_2$	1.4×10^4	2.3
TGAFL	1.8×10^3	1.2
FM-CONH$_2$	1.5×10^6	c
FL-CONH$_2$	2.7×10^6	c
FL	2.1×10^6	c
Phe	2.0×10^4	c

[a] Competitive fluorescence titration experiments in 10 mM ammonium phosphate, pH 7.2, at 37 °C. [b] Determined by supramolecular tandem enzyme assay. [c] No conversion detected due to N-terminal Phe.[50]

7. Summary and Outlook

This review focused on the interactions of cucurbut[n]uril receptors with amino acids, peptides, and proteins, with an eye for structure-activity relationships as determined by comparative binding studies, as well as recent applications in biochemistry and biotechnology. It is clear from this body of work that this direction is promising for a

number of reasons. First, all of these binding studies were carried out in aqueous solution. While this may seem like a given, the fact is that few classes of synthetic receptors function well in water, due either to poor solubility or to competition with water itself. Second, the affinities observed for Qn binding to amino acids, peptides, and proteins has been demonstrated to be as high as 10^7 M^{-1}, and commonly in the range 10^5-10^6 M^{-1}. High affinity is critical for use in bio-related applications because target peptides and proteins are often present at low concentrations. While the affinities demonstrated in the Qn family are as good as or better than other known synthetic receptors, *the critical challenge for this field* lies in pushing affinities up to and beyond the 10^9 M^{-1} mark so that binding becomes effective at and below nanomolar concentrations. Compounding this challenge is the fact that salt competes for binding to the Qn portals. Na^+ cations are present at >100 mM concentrations in biological systems, and although Na^+ binds weakly, at such high concentrations it can substantially reduce the effective binding constant of target analytes. Therefore, it is of critical importance that binding studies be carried out in the presence of significant salt concentrations. Third, the selectivities observed for the binding of Qn receptors to peptides and proteins is impressive. In particular, the sequence-specific recognition of peptides at an N-terminal aromatic residue is unparalleled in the field and promises a number of applications, including the sensing and separating of peptides and proteins on the basis of the identity of the terminal residue, and the placement of labels (e.g., spectroscopic, redox) and reactive groups at a single site on the surface of target proteins.

In addition to the direct binding of Qn receptors to peptides and proteins, a remarkable feature of the papers described herein is the creativity employed to merge the properties of Qn receptors with proteins when direct binding is not involved. In particular, the tandem enzyme assays developed by Nau and coworkers appear to have enormous potential for measuring the activity of enzymes that produce or destroy compounds that bind tightly to Qn receptors, which may extend well beyond this family of receptors. Also, the use of the high-affinity interaction between Q7 and ferrocene derivatives for protein capture, as developed by Kim and coworkers, presages a replacement of biotin-avidin technology with the considerably smaller yet similarly stable and selective Q7•ferrocene complex. The major challenge of this area involves the difficult synthetic chemistry involved in making singly modified Qn derivatives for the purpose of conjugation. Such an approach would broaden the utility of these receptors to such a great extent, not just for protein-related applications but for all applications of the Qn family, that we encourage those involved in this pursuit to more fully develop this technology for the benefit of the entire community.

References

[1]   a) P. B. Dervan, *Bioorg. Med. Chem.* **2000**, *9*, 2215-2235; b) M. W. Peczuh, A. D. Hamilton, *Chem. Rev.* **2000**, *100*, 2479-2493; c) H. Gohlke, G. Klebe, *Angew. Chem. Int. Ed.* **2002**, *41*, 2644-2676; d) K. N. Houk, A. G. Leach, S. P. Kim, X. Zhang, *Angew. Chem. Int. Ed.* **2003**, *42*, 4872-4897; e) E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem. Int. Ed.* **2003**, *42*, 1210-1250; f) D. H. Williams, E. Stephens, D. O'Brien, M. Zhou, *Angew. Chem. Int. Ed.* **2004**, *43*, 6596-6616; g) H. Yin, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2005**, *44*, 4130-4163; h) H.-J. Schneider, *Angew. Chem. Int. Ed.* **2009**, *48*, 3924-3977.

[2]   a) R. Behrend, E. Meyer, F. Rusche, *Liebigs Ann. Chem.* **1905**, *339*, 1-37; b) W. L. Mock, N.-Y. Shih, *J. Org. Chem.* **1983**, *48*, 3618-3619; c) W. L. Mock, N.-Y. Shih, *J. Org. Chem.* **1986**, *51*, 4440-4446; d) W. L. Mock, N.-Y. Shih, *J. Am. Chem. Soc.* **1988**, *110*, 4706-4710; e) L. Isaacs, *Chem. Comm.* **2009**, 619-629; f) J. Kim, I.-S. Jung, S.-Y. Kim, E. Lee, J.-K. Kang, S. Sakamoto, K. Yamaguchi, K. Kim, *J. Am. Chem. Soc.* **2000**, *122*, 540-541; g) J. W. Lee, S. Samal, N. Selvapalam, H. Kim, K. Kim, *Acc. Chem. Res.* **2003**, *36*, 621-630; h) L. Lagona, P. Mukhopadhyay, S. Chakrabarti, L. Isaacs, *Angew. Chem. Int. Ed.* **2005**, *44*, 4844-4870; i) S. Liu, C. Ruspic, P. Mukhopadhyay, S. Chakrabarti, P. Y. Zavalij, L. Isaacs, *J. Am. Chem. Soc.* **2005**, *127*, 15959-15967; j) M. V. Rekharsky, T. Mori, C. Yang, H. K. Young, N. Selvapalam, H. Kim, D. Sobransingh, A. E. Kaifer, S. Liu, L. Isaacs, W. Chen, S. Moghaddam, M. K. Gilson, O. Kim, Y. Inoue, *Proc. Natl. Acad. Sci.* **2007**, *104*, 20737-20742.

[3]   a) M. del Pozo, L. Hernandez, C. Quintana, *Talanta* **2010**, *81*, 4-5; b) C. Kim, S. S. Agasti, Z. Zhu, L. Isaacs, V. M. Rotello, *Nature Chem.* **2010**, *2*, 962-966; c) D.-W. Lee, K. M. Park, M. Banerjee, S. H. Ha, T. Lee, K. Suh, S. Paul, H. Jung, J. Kim, N. Selvapalam, S. H. Ryu, K. Kim, *Nature Chem.* **2011**, *3*, 154-159; d) P. Montes-Nvajas, M. Gonzalez-Bejar, J. C. Scaiano, H. Garcia, *Photochem. Photobiol. Sci.* **2009**, *8*, 1743-1747; e) V. D. Uzunova, C. Cullinane, K. Brix, W. M. Nau, A. I. Day, *Org. Biomol. Chem.* **2010**, *8*, 2037-2042; f) S. Walker, R. Kaur, F. J. McInnes, N. J. Wheate, *Mol. Pharmac.* **2010**, *7*, 2166-2172; g) N. J. Wheate, *J. Inorg. Biochem.* **2008**, *102*, 2060-2066; h) N. J. Wheate, R. I. Taleb, A. M. Krause-Heuer, R. L. Cook, S. Wang, V. J. Higgins, J. R. Aldrich-Wright, *Dalton Trans.* **2007**, 5055-5064.

[4]   S. M. Liu, P. Y. Zavalij, L. Isaacs, *J. Am. Chem. Soc.* **2005**, *127*, 16798-16799.

[5]   H.-J. Bohm, G. Schneider, Wiley-VCH, Weinheim, **2003**.

[6]   a) C. Schmuck, *Coord. Chem. Rev.* **2006**, *250*, 3053-3067; b) C. Schmuck, P. Wich, *Angew. Chem. Int. Ed.* **2006**, *45*, 4277-4281; c) M. Wehner, D. Janssen, G. Shafer, T. Schrader, *Eur. J. Org. Chem.* **2006**, 138-153; d) S. Tashiro, M. Tominaga, M. Kawano, B. Therrien, T. Ozeki, M. Fujita, *J. Am. Chem. Soc.* **2005**, *127*, 4546-4547; e) A. T. Wright, M. J. Griffin, Z. L. Zhong, S. C. McCleskey, E. V. Anslyn, J. T. McDevitt, *Angew. Chem. Int. Ed.* **2005**, *44*, 6375-6378; f) H. Zhou, L. Baldini, J. Hong, A. J. Wilson, A. D. Hamilton, *J. Am. Chem. Soc.* **2006**, *128*, 2421-2425; g) P. L. Toogood, *J. Med. Chem.* **2002**, *45*, 1543-1558; h) T. Berg, *Angew. Chem. Int. Ed.* **2003**, *42*, 2462-2481; i) A. J. Wilson, *Chem. Soc. Rev.* **2009**, *38*, 3289-3300; j) C. G. Cummings, A. D. Hamilton, *Curr. Opin. Chem. Biol.* **2010**, *14*, 341-346; k) E. Giralt, M. W. Peczuh, X. Salvatella, John Wiley & Sons, West Sussex, **2011**.

[7]   a) H.-J. Buschmann, K. Jansen, E. Schollmeyer, *Thermochim. Acta* **1998**, *317*, 95-98; b) H.-J. Buschmann, E. Schollmeyer, L. Mutihac, *Thermochim. Acta* **2003**, *399*, 203-208.

[8]   H.-J. Kim, J. Heo, W. S. Jeon, E. Lee, J. Kim, S. Sakamoto, K. Yamaguchi, K. Kim, *Angew. Chem. Int. Ed.* **2001**, *40*, 1526-1529.

[9]   K. Kim, J. Kim, I.-S. Jung, S.-Y. Kim, E. Lee, J.-K. Kang, Pohang University of Science and Technology Foundation, USA, **2000**.

[10]   M. E. Bush, N. D. Bouley, A. R. Urbach, *J. Am. Chem. Soc.* **2005**, *127*, 14511-14517.

[11]   Y. Ling, W. Wang, A. E. Kaifer, *Chem. Comm.* **2007**, 610-612.

[12]   F. Biedermann, U. Ruwald, M. Cziferszky, K. A. Williams, L. D. Gann, B. Y. Guo, A. R. Urbach, C. W. Bielawski, O. A. Scherman, *Chem. Eur. J.* **2010**, *87*, 13716-13722.

[13]   P. Rajgariah, A. R. Urbach, *J. Incl. Phenom. Macrocyc. Chem.* **2008**, *62*, 251-254.

[14]   A. Radzicka, R. Wolfenden, *Biochem.* **1988**, *27*, 1664-1670.

[15]   C. Chothia, *J. Mol. Biol.* **1976**, *105*, 1-12.

[16]   L. Lagona, B. D. Wagner, L. Isaacs, *J. Org. Chem.* **2006**, *71*, 1181-1190.

[17]   W.-H. Huang, P. Y. Zavalij, L. Isaacs, *Angew. Chem. Int. Ed.* **2007**, *119*, 7569-7571.

[18]   H. Cong, L.-L. Tau, Y.-H. Yu, F. Yang, Y. Du, S.-F. Xue, Z. Tao, *Acta Chim. Sin.* **2006**, *64*, 989-996.

[19]   D. M. Bailey, A. Hennig, V. D. Uzunova, W. M. Nau, *Chem. Eur. J.* **2008**, *14*, 6069-6077.

[20]   J.-M. Yi, Y.-Q. Zhang, H. Cong, S.-F. Xue, Z. Tao, *J. Mol. Struct.* **2009**, *933*, 112-117.

[21]   A. Hennig, H. Bakirci, W. M. Nau, *Nat. Methods* **2007**, *4*, 629-632.

[22]   J. Lui, X. Du, X. Zhang, *Chem. Eur. J.* **2011**, *17*, 810-815.

[23]   H.-J. Buschmann, L. Mutihac, R. C. Mutihac, E. Schollmeyer, *Thermochim. Acta* **2005**, *430*, 79-82.

[24]   H. Zhang, M. Grabenauer, M. T. Bowers, D. V. Dearden, *J. Phys. Chem. A* **2009**, *113*, 1508-1517.

[25]   a) R. Breslow, Z. Yang, R. Ching, G. Trojandt, F. Odobel, *J. Am. Chem. Soc.*
         1998, *120*, 3536-3537; b) M. A. Hossain, H.-J. Schneider, *J. Am. Chem. Soc.*
         1998, *120*, 11208-11209; c) C. Schmuck, L. Geiger, *J. Am. Chem. Soc.* **2004**,
         126, 8898-8899; d) J. Shepherd, T. Gale, K. B. Jensen, J. D. Kilburn, *Chem.*
         Eur. J. **2006**, *12*, 713-720.

[26]   F. Tian, M. Cziferszky, D. Jiao, K. Wahlstrom, J. Geng, O. A. Scherman,
         Langmuir **2011**, *27*, 1387-1390.

[27]   a) J. D. Badjic, A. Nelson, S. J. Cantrill, W. B. Turnbull, J. F. Stoddart, *Acc.*
         Chem. Res. **2005**, *38*, 723-732; b) L. L. Kiessling, J. E. Gestwicki, L. E.
         Strong, *Angew. Chem. Int. Ed.* **2006**, *45*, 2348-2368; c) M. Mammen, S.-K.
         Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.* **1998**, *37*, 2754-2797; d) A.
         Mulder, J. Huskens, D. N. Reinhoudt, *Org. Biomol. Chem.* **2004**, *2*, 3409-
         3424.

[28]   a) J. Rao, J. Lahiri, L. Isaacs, R. M. Weis, G. M. Whitesides, *Science* **1998**,
         280, 708-711; b) J. Rao, J. Lahiri, R. M. Weis, G. M. Whitesides, *J. Am.*
         Chem. Soc. **2000**, *122*, 2698-2710.

[29]   S.-K. Choi, *Synthetic Multivalent Molecules: Concepts and Biomedical*
         Applications, John Wiley & Sons, Hoboken, NJ, **2004**.

[30]   J. J. Reczek, A. A. Kennedy, B. T. Halbert, A. R. Urbach, *J. Am. Chem. Soc.*
         2009, *131*, 2408-2415.

[31]   J. J. Reczek, E. Rebolini, A. R. Urbach, *J. Org. Chem.* **2010**, *75*, 2111-2114.

[32]   L. M. Heitmann, A. B. Taylor, P. J. Hart, A. R. Urbach, *J. Am. Chem. Soc.*
         2006, *128*, 12574-12581.

[33]   M. V. Rekharsky, H. Yamamura, C. Inoue, M. Kawai, I. Osaka, R. Arakawa,
         K. Shiba, A. Sato, Y. H. Ko, N. Selvapalam, K. Kim, Y. Inoue, *J. Am. Chem.*
         Soc. **2006**, *128*, 14871-14880.

[34]   M. V. Rekharsky, H. Yamamura, Y. H. Ko, N. Selvapalam, K. Kim, Y. Inoue,
         Chem. Comm. **2008**, 2236-2238.

[35]   H. D. Nguyen, D. T. Dang, J. L. J. van Dongen, L. Brunsveld, *Angew. Chem.*
         Int. Ed. **2010**, *49*, 895-898.

[36]   W. Lei, G. Jiang, Q. Zhou, B. Zhang, X. Wang, *Phys. Chem. Chem. Phys.*
         2010, *12*, 13255-13260.

[37]   A. C. Bhasikuttan, J. Mohanty, W. M. Nau, H. Pal, *Angew. Chem. Int. Ed.*
         2007, *46*, 4120-4122.

[38]   M. Shaikh, J. Mohanty, A. C. Bhasikuttan, V. D. Uzunova, W. M. Nau, H. Pal, *Chem. Comm.* **2008**, 3681-3683.

[39]   J. M. Chinai, A. B. Taylor, L. M. Ryno, N. D. Hargreaves, C. A. Morris, P. J. Hart, A. R. Urbach, *J. Am. Chem. Soc.* **2011**, *133*, in press.

[40]   B. T. Nguyen, E. V. Anslyn, *Coord. Chem. Rev.* **2006**, *250*, 3118-3127.

[41]   H.-K. Lee, K. M. Park, Y. J. Jeon, D. Kim, D. H. Oh, H. S. Kim, C. K. Park, K. Kim, *J. Am. Chem. Soc.* **2005**, *127*, 5006-5007.

[42]   I. Hwang, K. Baek, M. Jung, Y. Kim, K. M. Park, D.-W. Lee, N. Selvapalam, K. Kim, *J. Am. Chem. Soc.* **2007**, *129*, 4170-4171.

[43]   J. F. Young, H. D. Nguyen, L. Yang, J. Huskens, P. Jonkheijm, L. Brunsveld, *Chembiochem* **2010**, *11*, 180-183.

[44]   F. Biedermann, U. Rauwald, J. M. Zayed, O. A. Scherman, *Chem. Sci.* **2011**, *2*, 279-286.

[45]   A. Hennig, G. Ghale, W. M. Nau, *Chem. Comm.* **2007**, 1614-1616.

[46]   S. Ghosh, L. Isaacs, *J. Am. Chem. Soc.* **2010**, *132*, 4445-4454.

[47]   W. M. Nau, G. Ghale, A. Hennig, H. Bakirci, D. M. Bailey, *J. Am. Chem. Soc.* **2009**, *131*, 11558-11570.

[48]   R. de la Rica, A. H. Velders, *Small* **2011**, *7*, 66-69.

[49]   T. C. Lee, O. A. Scherman, *Chem. Comm.* **2010**, 2438-2440.

[50]   G. Ghale, V. Ramalingam, A. R. Urbach, W. M. Nau, *J. Am. Chem. Soc.* **2011**, *133*, in press.

[51]   M. Shaikh, J. Mohanty, P. K. Singh, W. M. Nau, H. Pal, *Photochem. Photobiol. Sci.* **2008**, *7*, 408-414.

Figure Captions

Figure 1. Structure of cucurbit[n]uril (Qn) molecular containers.

Figure 2. Q8 forms a heteroternary complex via a selective, stepwise association of methyl viologen (MV) followed by a second guest such as dihydroxynaphthalene (HN). The structure in the lower right was derived from the coordinates of the crystal structure reported by Kim and coworkers.[8]

Figure 3. Comparison of the sidechain surface area of the 20 amino acids versus the free energy of transfer from cyclohexane to aqueous solution. Trp, Phe, and Tyr lie in a unique region in the upper right corner.[14-15]

Figure 4. Chemical formulas of singly charged derivatives of tryptophan (Trp).[10]

Figure 5. Energy minimized computer model of Q8•MV•Trp showing both aromatic groups stacked face-to-face inside the Q8 cavity, and the N-terminal ammonium group interacting with Q8 carbonyl oxygens. Nitrogens are blue. Oxygens are red.

Figure 6. Chemical formulas of DPT and MBBI.

Figure 7. Concept of the supramolecular tandem enzyme assay as applied to the measurement of lysine carboxylase activity.[21] The cadaverine product binds more tightly than lysine and the dye to Q7, and thus the dye is displaced as the reaction proceeds, yielding a loss in fluorescence intensity.

Figure 8. Structural congruence between singly charged amino acid derivatives, and the terminal residue of a peptide chain. This critical concept enabled the leap from amino acid binding to sequence-specific peptide recognition.[10]

Figure 9. Schematic of the driving forces involved in the selective recognition of an N-terminal aromatic residue.

Figure 10. Schematic of the concept of self-assembled multivalent receptors for peptides. A scaffold presenting two viologen groups binds to two equivalents of Q8, and the resulting divalent receptor binds simultaneously to two Trp groups on a divalent peptide.[30]

Figure 11. Solid-phase synthesis of peptide-viologen conjugates yielding mono-, di-, and trivalent scaffolds scaffolds **5-7**.[30-31]

Figure 12. Plot showing the remarkably similar affinities of Q8•MV and Q8•MBBI for N-terminal, C-terminal, and non-terminal Trp.[12]

Figure 13. Crystal structures of (left) Q8 bound to one equivalent of **W**GG and (right) Q8 bound to two equivalents of **F**GG. All aromatic sidechains are bound within the Q8 cavity, and all N-terminal ammonium groups are associated with carbonyl oxygens on Q8. In the dimer structure, the two phenyl groups are pi-stacked in a staggered face-to-face arrangement.[32] Nitrogens are blue. Oxygens are red.

Figure 14. Crystal structure of Q7 (grey) bound to the N-terminal Phe of the B-chain of human insulin (green). The first few residues of the B-chain unfold from the surface of the macromolecule to accommodate Q7 binding.[39] Nitrogens are blue. Oxygens are red.

Figure 15. Schematic of a tandem enzyme assay for the protease thermolysin. The product contains an N-terminal Phe, which binds tightly to Q7 and outcompetes the dye, thus yielding a loss in fluorescence intensity as the reaction proceeds.[50]