Sequence recognition and self-sorting of a dipeptide by cucurbit[6]uril and cucurbit[7]uril†

Mikhail V. Rekharsky,[a] Hatsuo Yamamura,[b] Young Ho Ko,[c] N. Selvapalam,[c] Kimoon Kim*[c] and Yoshihisa Inoue*[ad]

Received (in Cambridge, UK) 2nd January 2008, Accepted 11th February 2008
First published as an Advance Article on the web 10th March 2008
DOI: 10.1039/b719902c

Cucurbit[7]uril forms very strong complex with zwitterionic dipeptide Phe–Gly with affinity exceeding 10^7 M^{-1} and effectively recognizes peptide sequence of Phe–Gly over Gly–Phe as well as Tyr–Gly over Gly–Tyr and Trp–Gly over Gly–Trp with relative affinities of 23 000, 18 000 and 2000, respectively.

Peptide sequence recognition is a key to the various functions of enzymes and antibodies, and hence enormous efforts have been devoted to mimic these natural systems.[1–3] Previously, we observed that zwitterionic Phe exhibits a high affinity (1.8×10^6 M^{-1})[2c] toward cucurbit[7]uril (CB[7]; Fig. 1), while CB[6] forms complexes with aliphatic amino acid derivatives, *e.g.* Leu-amide.[3b] Here, we report our results of calorimetric (ITC: isothermal titration calorimetry) and NMR spectral investigations on the sequence recognition and self-sorting of dipeptides by CB[7] and CB[6] cavities.

Our ITC study on the complexation of CB[7] with several zwitterionic dipeptides containing a Phe, Tyr or Trp residue, which are listed in Table 1, revealed that the affinity of dipeptides carrying an aromatic residue at the *N*-terminus is consistently much higher than that of the corresponding dipeptides with an aromatic residue at the *C*-terminus, giving relative affinities of 23 000, 18 000 and 2000 for Phe–Gly over Gly–Phe, Tyr–Gly over Gly–Tyr and Trp–Gly over Gly–Trp, respectively (Table 1). Fig. 2 allows us to clearly visualize the large difference in affinities between Phe–Gly and Gly–Phe, as the ITC titration curve is much steeper in the former case than in the latter. The present and previous results[4,5] confirm that, in general, CBs are quite effective at recognizing the peptide sequence. As can be seen from the thermodynamic parameters obtained in aqueous solution (Table 1), complexation of these dipeptides is driven exclusively by enthalpy with an accompanying negative (unfavorable) entropy change.

Fig. 3 illustrates how such unusually high peptide–sequence recognition occurs. Zwitterionic CB[7] forms a highly stable complex with Aryl–Gly (Aryl = Phe or Tyr) by efficiently including the guest's aromatic group in the CB cavity, and by the electrostatic attraction between its ammonium group and the carbonyl oxygens of the CB portal (more negative ΔH°), with accompanying extensive desolvation (less negative ΔS°), while avoiding electrostatic repulsion between the carbonyl oxygens and the guest's carboxylate (Fig. 3(a)). In contrast, none of these mechanisms operate for the complexation of a Gly–Aryl guest with CB[7], where deep penetration of the aromatic moiety into the CB[7] cavity is prevented by electrostatic repulsion between the anionic carboxylate moiety of the guest and the CB's carbonyl oxygens at the portal, and the ammonium group of Gly is not allowed to efficiently interact with the CB portal (Fig. 3(b)). These factors jointly discourage the complexation of a Gly–Aryl with CB[7] by a factor of 2000–23 000 in comparison with an Aryl–Gly guest (Table 1).

The host–guest interactions proposed in Fig. 3 have been confirmed by ^{1}H NMR experiments (Fig. 4). The NMR spectra shown in Fig. 4 allow us to elucidate molecular events occurring between CB[7] (Fig. 4(a)) and Phe–Gly$^{\pm}$ (Fig. 4(b)) or Gly–Phe$^{\pm}$ (Fig. 4(d)). Upon interaction of zwitterionic Phe–Gly with CB[7], all guest aromatic protons are shifted (Fig. 4(c)), revealing deep insertion of the phenyl ring inside the cavity. At the same time, the β-CH_2 protons of Phe are split, indicating very different environments around these two protons, and thus suggesting that they located at the edge of the cavity. The α-CH_2 protons of Gly exhibit the opposite shift compared to the aromatic protons of Phe. This observation suggests that the Gly residue is located outside the cavity. The NMR spectrum (Fig. 4(c)) can be taken as experimental evidence of strong complexation between CB[7] and Phe–Gly, with an optimal van der Waals interaction between the inner walls of the cavity and the inserted Phe residue (affording a large negative ΔH°), and the optimal charge–dipole interaction

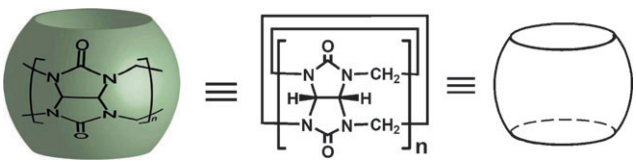

Fig. 1 Chemical structures of cucurbit[*n*]uril macrocycles (*n* = 6–8).

[a] *Entropy Control Project, ICORP, JST, 2-1 Yamada-oka, Suita 565-0871, Japan. E-mail: mvr@chem.eng.osaka-u.ac.jp; Fax: +81 6-6879-7923; Tel: +81 6-6879-7920*
[b] *Graduate School of Technology, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan. E-mail: yamamura.hatsuo@nitech.ac.jp; Fax: +81 052-735-5246; Tel: +81 052-735-5246*
[c] *National Creative Research Initiative Center for Smart Supramolecules and Department of Chemistry, Pohang University of Science and Technology, San 31 Hyojadong, Pohang 790-784, Korea. E-mail: kkim@postech.ac.kr*
[d] *Department of Applied Chemistry, Osaka University, 2-1 Yamada-oka, Suita 565-0871, Japan. E-mail: inoue@chem.eng.osaka-u.ac.jp; Fax: +81 6-6879-7923; Tel: +81 6-6879-7920*

† Electronic supplementary information (ESI) available: Thermodynamic data on complexation of CB[6] and CB[7] with several dipeptides, NMR spectra of [CB[7]·Tyr–Lys] and [CB[6]·Lys–Tyr] complexes. See DOI: 10.1039/b719902c

Table 1 Stability constants (K), standard enthalpies (ΔH°) and entropy changes (ΔS°) for the complexation of selected dipeptides with cucurbit[7]uril (CB[7]) in H_2O at $T = 298.15$ K

Dipeptide	K/M^{-1}	ΔH°/kJ mol^{-1}	$T\Delta S^\circ$/kJ mol^{-1}
Phe–Gly	$(3.0 \pm 0.4) \times 10^7$	-47.4 ± 0.5	-4.7 ± 0.5
Gly–**Phe**	1300 ± 200	-29.9 ± 0.3	-12.2 ± 0.4
Tyr–Gly	$(3.6 \pm 0.2) \times 10^6$	-44.1 ± 0.4	-6.7 ± 0.4
Gly–**Tyr**	200 ± 20	-23.1 ± 0.2	-10.2 ± 0.3
Trp–Gly	$(5.6 \pm 0.2) \times 10^5$	-44.6 ± 0.4	-11.9 ± 0.4
Gly–**Trp**	280 ± 25	-18.1 ± 0.2	-4.3 ± 0.3

between the guest's positively-charged ammonium cation and the host's carbonyl oxygens, with accompanying extensive dehydration (leading to a small negative or even positive ΔS°), both of which strengthen the complex's stability.

The NMR spectrum of a mixture of CB[7] and Gly–Phe (Fig. 4e) is very different. As shown in Fig. 4(e), the signals of both free and complexed guests are simultaneously observed and are very broad, indicating slow exchange. It is deduced that deep penetration of the aromatic moiety is prevented by the electrostatic repulsion of the carboxylate moiety against the CB's carbonyl oxygens (Fig. 3(b)). Such a complex structure cannot lead to strong van der Waals contacts inside the cavity, and naturally is associated with a low–moderate enthalpy. Another unfavorable aspect of CB[7] interaction with Gly–Phe$^{\pm}$ is the restriction of guest conformational upon complexation. Indeed, originally located on the opposite side of the peptide chain, the Phe residue and the ammonium cation are forced to come closer to each other upon complex formation, leading to conformational restriction of the guest and a large negative entropy. This simple rationalization,

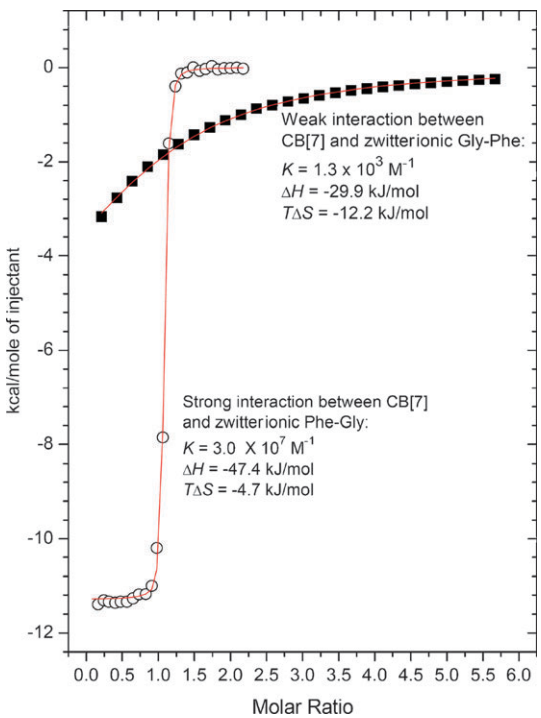

Fig. 2 ITC titration curves for interaction of CB[7] with zwitterionic Gly–Phe (solid squares) and zwitterionic Phe–Gly (open circles).

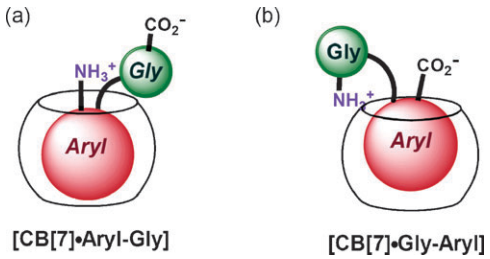

Fig. 3 Selective complexation of CB[7] with (a) Aryl–Gly, which forms much stronger complexes than (b) Gly–Aryl (see the main text for a detailed explanation).

arising from the NMR study, is nicely compatible with the results of the ITC experiments (Table 1).

CB[6] forms only weak complexes with dipeptides having an aromatic residue at the *N*-terminus, but interacts more strongly with those having an aliphatic residue, particularly Lys, at the *N*-terminus (see ESI†). This contrasting complexation behavior of CB[6] *vs.* CB[7] prompted us to study an effective self-sorting system (as it was defined previously[6a]) comprising of a dipeptide mixture, CB[6] and CB[7]. As mentioned above, dipeptides possessing an aromatic residue at the *N*-terminus interact strongly with CB[7], while dipeptides having a Lys residue at the *N*-terminus preferentially form a complex with CB[6]. To check this possibility, Tyr–Lys

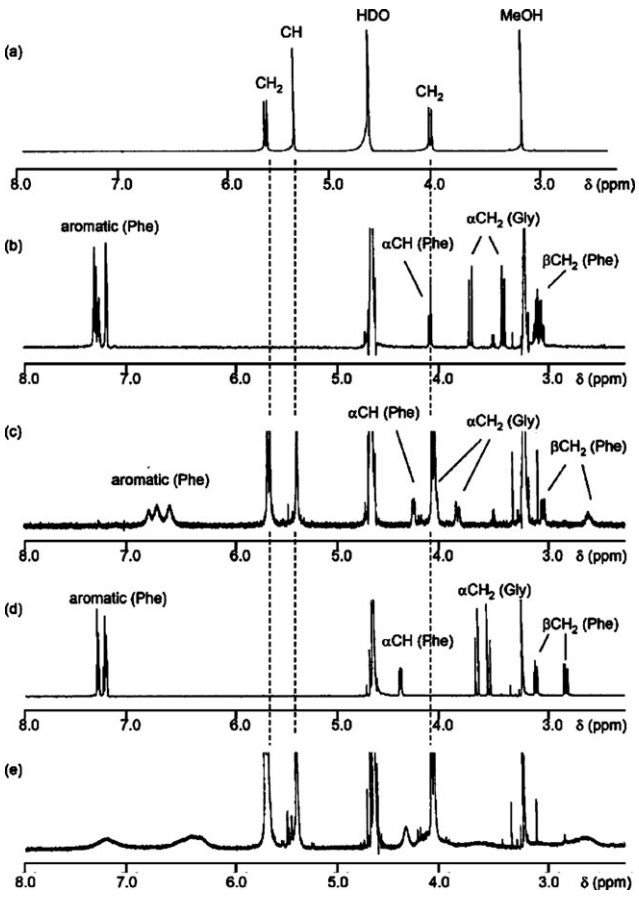

Fig. 4 NMR spectra of (a) CB[7], (b) zwitterionic Phe–Gly, (c) a mixture of CB[7] and Phe–Gly, (d) zwitterionic Gly–Phe, and (e) a mixture of CB[7] and Gly–Phe in D_2O.

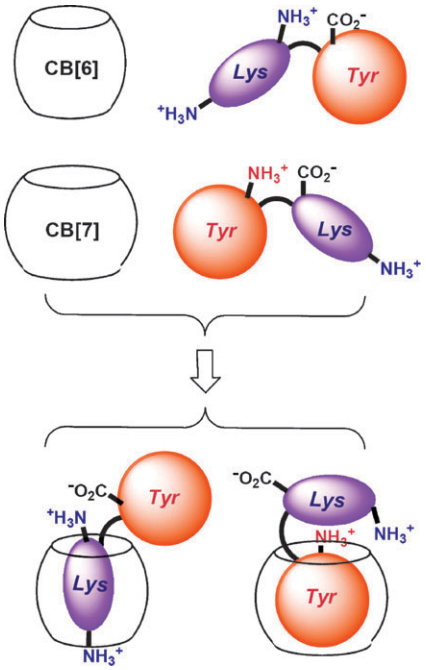

Fig. 5 Self-sorting of Tyr–Lys and Lys–Tyr in an aqueous solution containing CB[6] and CB[7].

and Lys–Tyr were subjected to an NMR spectral investigation. We chose Tyr, since Tyr–Gly reveals moderate affinity (affinity order toward CB[7] = Phe–Gly > Tyr–Gly > Trp–Gly (Table 1)) and the experimental results obtained for the Tyr-containing dipeptide could be readily propagated to the relevant Phe- and Trp-containing dipeptides.

Indeed, our NMR study confirmed that such spontaneous self-sorting does occur. Thus, upon preparation of the solution containing two guests, *i.e.* Tyr–Lys and Lys–Tyr, and two hosts, *i.e.* CB[6] and CB[7], the reaction mixture is clearly separated into two distinguishable supramolecular complexes, *i.e.* [CB[6]·Lys–Tyr] and [CB[7]·Tyr–Lys] (see ESI†). Due to the low stability of the [CB[6]·Lys–Tyr] complex, we applied a three-fold excess of CB[6] to drive the complexation of the Lys–Tyr peptide by this host (see ESI for details†). Supramolecular self-sorting was achieved previously with CBs for various guests,[6] but to the best of our knowledge, this is the first self-sorting of a pair of dipeptides with reverse amino acid sequences by CB or any other hosts.

The self-sorting of Tyr–Lys *vs.* Lys–Tyr, as illustrated in Fig. 5, mimics the biologically-important sequence-selective recognition of various peptides by natural enzymes. For instance, the active site of chymotripsin preferentially includes aromatic residues, while trypsin-like enzymes interact with positively-charged aliphatic residues such as Lys or Arg.

In conclusion, we have confirmed the remarkable ability of CBs to recognize an amino acid sequence of short peptides and to self-sort closely-related peptides. Our results suggest that CBs may serve as a simple yet effective models to mimic the recognition of small peptides by enzymes and antibodies.

Notes and references

1 M. W. Peczuh and A. D. Hamilton, *Chem. Rev.*, 2000, **100**, 2479–2493 and references cited therein.

2 (*a*) A. Ojida, Y. Mito-oka, M. Inoue and I. Hamachi, *J. Am. Chem. Soc.*, 2002, **124**, 6256–6258; (*b*) A. T. Wright and E. V. Anslyn, *Org. Lett.*, 2004, **6**, 1341–1344; (*c*) W. S. Jeon, K. Moon, S. H. Park, H. Chun, Y. H. Ko, J. Y. Lee, E. S. Lee, S. Samal, N. Selvapalam, M. V. Rekharsky, V. Sindelar, D. Sobransingh, Y. Inoue, A. E. Kaifer and K. Kim, *J. Am. Chem. Soc.*, 2005, **127**, 12984–12989; (*d*) S. Tashiro, M. Tominaga, M. Kawano, B. Therrien, T. Ozeki and M. Fujita, *J. Am. Chem. Soc.*, 2005, **127**, 4546–4547; (*e*) M. E. Bush, N. D. Bouley and A. R. Urbach, *J. Am. Chem. Soc.*, 2005, **127**, 14511–14517; (*f*) C. Schmuck, M. Heil, J. Scheiber and K. Baumann, *Angew. Chem., Int. Ed.*, 2005, **44**, 7208–7212; (*g*) H.-J. Buschmann, L. Mutihac, R.-C. Mutihac and E. Schollmeyer, *Thermochim. Acta*, 2005, **430**, 79–82.

3 For recent examples of the supramolecular recognition of peptides and related compounds, see: (*a*) J. Shepherd, T. Gale, K. B. Jensen and J. D. Kilburn, *Chem.–Eur. J.*, 2006, **12**, 713–720; (*b*) M. V. Rekharsky, H. Yamamura, C. Inoue, M. Kawai, I. Osaka, R. Arakawa, K. Shiba, A. Sato, Y. H. Ko, N. Selvapalam, K. Kim and Y. Inoue, *J. Am. Chem. Soc.*, 2006, **128**, 14871–14880; (*c*) S. Tashiro and M. Fujita, *Bull. Chem. Soc. Jpn.*, 2006, **79**, 833–837; (*d*) S. Tashiro, M. Kobayashi and M. Fujita, *J. Am. Chem. Soc.*, 2006, **128**, 9280–9281; (*e*) L. M. Heitmann, A. B. Taylor, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.*, 2006, **128**, 12574–12581; (*f*) S. Tashiro, M. Tominaga, Y. Yamaguchi, K. Kato and M. Fujita, *Angew. Chem., Int. Ed.*, 2006, **45**, 241–244; (*g*) S. Tashiro, M. Tominaga, Y. Yamaguchi, K. Kato and M. Fujita, *Chem.–Eur. J.*, 2006, **12**, 3211–3217; (*h*) S. Liu, P. Y. Zavalij, Y.-F. Lam and L. Isaacs, *J. Am. Chem. Soc.*, 2007, **129**, 11232–11241; (*i*) S. Kobayashi, K. Morino and E. Yashima, *Chem. Commun.*, 2007, 2351–2353; (*j*) Y. Liu, X.-Y. Li, H.-Y. Zhang, H.-Y. Li and F. Ding, *J. Org. Chem.*, 2007, **72**, 3640–3645; (*k*) Y. Ling, W. Wang and A. E. Kaifer, *Chem. Commun.*, 2007, 610–612; (*l*) V. Sindelar and A. E. Kaifer, *Chem. Commun.*, 2007, 2185–2187.

4 W. S. Jeon, H.-J. Kim, C. Lee and K. Kim, *Chem. Commun.*, 2002, 1828–1829.

5 (*a*) W. L. Mock, in *Comprehensive Supramolecular Chemistry*, ed. F. Vögtle, Pergamon, Oxford, 1996, vol. 2, pp. 477–493; (*b*) I.-S. Kim, J. Jung, S.-Y. Kim, E. Lee, J.-K. Kang, S. Sakamoto, K. Yamaguchi and K. Kim, *J. Am. Chem. Soc.*, 2000, **122**, 540–541; (*c*) J. W. Lee, S. Samal, N. Selvapalam, H.-J. Kim and K. Kim, *Acc. Chem. Res.*, 2003, **36**, 621–630; (*d*) J. Lagona, P. Mukhopadhyay, S. Chakrabarti and L. Isaacs, *Angew. Chem., Int. Ed.*, 2005, **44**, 4844–4870; (*e*) M. V. Rekharsky, Y. H. Ko, N. Selvapalam, K. Kim and Y. Inoue, *Supramol. Chem.*, 2007, **19**, 39–46.

6 (*a*) A. Wu and L. Isaacs, *J. Am. Chem. Soc.*, 2003, **125**, 4831–4835; (*b*) S. Liu, C. Ruspic, P. Mukhopadhyay, S. Chakrabarti, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2005, **127**, 15959–15967 and references cited therein; (*c*) P. Mukhapadhyay, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2006, **128**, 14093–14102.