COMMUNICATION

A synthetic transcription factor pair mimic for precise recruitment of an epigenetic modifier to the targeted DNA locus

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Zutao Yu[a,b], Mengting Ai[a], Soumen K. Samanta[c], Fumitaka Hashiya[a], Junichi Taniguchi[a], Sefan Asamitsu[a], Shuji Ikeda[a], Kaori Hashiya[a], Toshikazu Bando[a], Ganesh N. Pandian[*b], Lyle Isaacs[*c], Hiroshi Sugiyama[*a,b]

We developed an epigenetically active, cooperative DNA binding transcription factor platform assisted by cucurbit[7]uril (CB7) host-guest modules. This new type of molecule termed ePIP–HoGu not only mimics the operation of transcription factors as a pair but also recruits the epigenetic modifier to a particular DNA locus.

The cooperative DNA binding and subsequent transcriptional modulation are ubiquitous in natural gene regulatory systems, especially by transcription factors (TFs). In mammals, 50–70% of TFs operate in pairs (and clusters) to orchestrate accurate spatiotemporal gene expression[1]. Therefore, there is a need to develop a synthetic mimic that encompasses both the capability to undergo cooperative DNA binding and epigenetic modulation.

Cooperative DNA-binding systems using a peptide as a DNA binder or cooperation domain have been explored previously (Table S1)[2-5]. Pyrrole-imidazole polyamides (PIPs) are a class of well-characterized small-molecule DNA minor-groove binders[6-8]. Recently, we reported two synthetic cooperative DNA-binding systems, i) PIPs conjugated to either a host-guest assembly (PIP–HoGu)[9] and ii) a nucleic acid-based cooperation system (PIP–NaCo)[10] to provide exemplary models for mimicking DNA binding of TF pairs using small molecules (Table S1). Apart from covalent PIP dimers showing high binding affinity and affinity to fixed binding sites[11, 12], noncovalent cooperative systems can apply versatile binding modes, including different spacings and orientations of two individual DNA motifs, and, has the potential to constitute precise gene regulation via an amenable paternal conjugate.

Simple DNA binding using a cooperative system provokes biological effects through disruption of TF pair binding[9]. However, to achieve a higher level of cellular efficacy and more eminent biological applications such as gene activation, the next challenge is the installation of an epigenetic modulator (epi-drug) to advance them as a robust cooperative DNA-binding system[13, 14]. Here, we report an epigenetically active cucurbit[7]uril-assisted DNA-binding system, termed ePIP–HoGu that mimic the cooperative function of a TF pair and is capable of precisely recruiting epigenetic modifiers to the target DNA sites (Fig. 1).

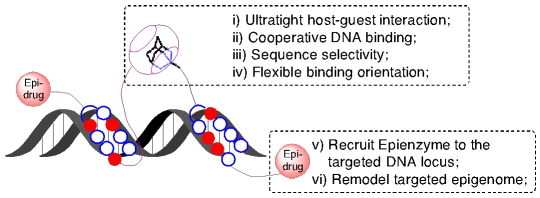

Fig. 1 Illustration of ePip-HoGu system.

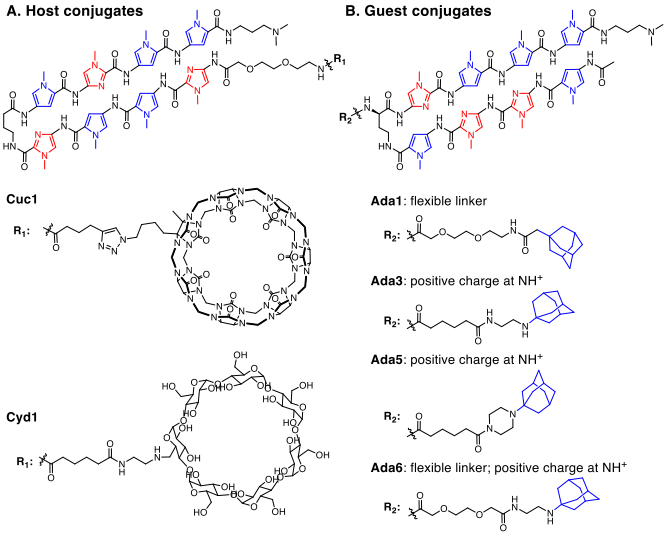

Fig. 2 Chemical structures of host conjugates CB7–PIP and Cyd–PIP (A), and guest conjugates Ada–PIP (B).

[a] Department of Chemistry, Graduate School of Science, Kyoto University,
[b] Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502 (Japan).
[c] Department of Chemistry and Biochemistry, University of Maryland, College Park, Institute for Biomolecular Design and Discovery (IBD), Maryland 20742 (United States).
E-mail: ganesh@kuchem.kyoto-u.ac.jp, lisaacs@umd.edu, hs@kuchem.kyoto-u.ac.jp

†Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

We first upgraded the cooperation domain in the PIP-HoGu system by replacing cyclodextrin (Cyd) with CB7, because an advanced DNA-binding system such as ePIP–HoGu requires a very tight host–guest interaction[15, 16]. As a heptameric member of CB[*n*], CB7 has received considerable attention because it is cell-permeable, non-toxic, readily soluble in water (20 mM)[15, 17]. Also, it has been widely explored in biosensing, separation, catalysis, and drug-delivery applications[18-23]. Significantly, CB7 exhibits an ultratight binding affinity to adamantane (Ada) (K_D routinely in the 10^{-9}–10^{-12} M^{-1} range), whereas Cyd–Ada has a relatively weak host–guest interaction (K_D of ca. 10^{-5} M^{-1})[16]. Using host **Cyd1** (5′-WGWCGW-3′) as a template[9], CB7-PIP conjugate **Cuc1** was synthesized by using click chemistry between PIP–alkyne and CB7–azide[24]. The synthesized guest derivatives **Ada1**–**6** vary structurally in linker length, linker type, and positive charge (Fig. 2A, B).

The CB7-assisted cooperative binding system was evaluated by using a thermal stabilization assay and closely compared with the Cyd-assisted system[9, 25]. As expected, when paired with **Ada1**–**6**, **Cuc1** exhibited notably higher thermal stability than **Cyd1** with ΔT_m values varying from 0.6 to 2.2 °C in the presence of ODNs with a spacing of 2 bp (Table S3). Moreover, electrostatic potential profiles revealed that, unlike Cyd that has a nearly neutral charge of portal and cavity, CB[*n*] displays a strongly negative charge around the entrance carbonyl oxygen atoms and the inner surface and promote the formation of complexes with positively charged guest species (especially ammonium ions)[15, 22]. Alkyl chain linkers have recently been shown to act as a chaperone in strengthening host–guest interactions[26]. Indeed, **Ada3**, with an ethyldiamino residue[15] and alkyl chain, showed the most prominent stabilization effect (Table S3). The piperazine moiety in **Ada5** has a deleterious impact on cooperation, which needs to be obviated.

Surface plasmon resonance (SPR) assays shed further light on the ultratight host–guest interaction and can reveal the binding dynamics. Previously, we showed that **Cyd1**–**Ada1** exhibited an association rate constant (k_a) of 2.6×10^5 $M^{-1}s^{-1}$ using an SPR assay by immobilizing dsDNA-biotin on a chip[9]. Here, in the absence of dsDNA, immobilized **Ada3-biotin** interacted with **Cyd1** with a k_a of 1.4×10^5 $M^{-1}s^{-1}$ (Fig. 3A, S1B). Therefore, the kinetic DNA binding mechanism of Cyd-assisted PIP–HoGu could occur either by the pair first binding to DNA followed by the host-guest interaction or by the procession of these two steps at a similar rate. Significantly, **Cuc1** exhibited a k_a of 4.1×10^5 $M^{-1}s^{-1}$ and did not further dissociate even by a series washing with harsh buffers ($K_D < 1.83 \times 10^{-11}$), which in turn demonstrates a remarkable binding potency of the system that is comparable to the irreversible binding of an antibody (Fig. 3B, S1)[22]. Thus, it is plausible to conclude that **Cuc1** first binds the partner guest and is followed by synergic DNA binding.

Electrophoretic mobility shift assay (EMSA) was conducted to investigate the influence of spacing and binding orientation on cooperation. **Cuc1**–**Ada3** assembly was applied in positive (Mode A) and negative binding modes (Mode B) (ODNs are listed in Table S2). Diverging from the **Cyd1**-system showing cooperativity when the spacing was limited to 0–5 bp[9], **Cuc1**–**Ada3** exhibited high complex formation potency at spacings of 0–5 bp, 8 bp, and, surprisingly, in Mode C (8 bp with partially reversed orientation) to suggest a potent binding affinity (Fig. 3C, D). The difference in band-shift behaviour of ODNs with the spacings of 6 bp and 8 bp could be explained by the combinatorial effects of DNA twist angle, the distance between the two PIP-binding sites, and the linker length of the two conjugates. Inserting a spacer between two PIP-binding sites not only shifts the linear range but also rotates the sites from their original position. While, for the spacing of 8 bp, host–guest moieties could meet through crossing the DNA major groove[10]. The finding supports that **Ada4**, with a longer linker, exhibited a robust band-shift at a spacing of 6 bp in Mode A (Fig. S2).

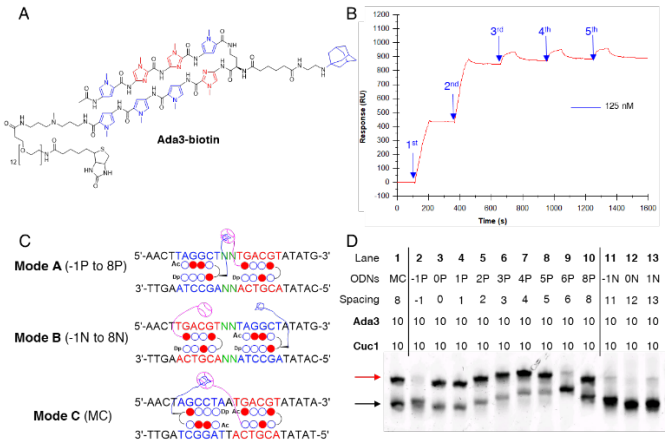

Fig. 3 (A, B) **Cuc1** binds **Ada3** irreversibly in the absence of DNA in an SPR assay. (A) Chemical structure of **Ada3-biotin**. (B) SPR sensorgram of **Cuc1** (125 nM) with multiple rounds of standard injection. One standard injection consisted of 180 s sample injection, followed by 180 s elution at 20 μL/min. (C, D) EMSA illustrating the cooperativity of the CB7-assisted DNA-binding system. (C) Three binding modes. Positive binding mode (Mode A) contains series dsDNA (–1P to 8P) with a gap distance (N) ranging from –1 to 8 bp. Similarly, negative binding mode (Mode B) includes dsDNA (–1N to 8N) with gap distance of –1 to 8 bp. (D) The gel-shift behavior of Modes A, B, and C with **Ada3**–**Cuc1**. ODN concentrations: 1.0 μM. Compound concentrations: 10.0 μM. Black arrow: ODNs. Red arrow: ODNs/**Cuc1**/**Ada3**.

The cooperation strength is altered not only by spacing but also by the length of the PIPs. Specifically, a weak host–guest force is presumed to be saturated for the synergic short PIPs binding because of slow PIP–DNA association (**Ada1** with a k_a of 1.1×10^4 $M^{-1}s^{-1}$ and **Cyd1** with a k_a of 5.9×10^4 $M^{-1}s^{-1}$) is the rate-limiting step[9], while PIPs with longer length require a stronger host–guest system[27]. To verify this notion, **Ada7**, with two extended bp-binding sites corresponding to parental **Ada3**, was prepared. Consistent with the results obtained with the EMSA assay (Fig. S4), the T_m assay revealed that at a spacing of 2 bp, **Cuc1**–**Ada7** strongly stabilized dsDNA compared with **Cyd1**–**Ada7**, with a ΔT_m of 2.1 °C.

Taken together, we optimized the PIP–HoGu system by introducing host CB7 and ethyldiamino-Ada as a guest molecule, which could serve as a reference design for developing advanced DNA-binding systems with longer spacing, longer PIPs length, mismatch recognition (Fig. S6), and a flexible binding

orientation (Fig. S3). As a model for the next-generation TF pair system, we also installed an epigenetically active domain in the reference design of PIP-HoGu.

Similar to studies in other laboratories[2, 14, 28, 29], our group has been making steadfast progress in developing small-molecule, gene-specific activators by conjugating PIPs with epigenetic modulators, such as histone deacetylase inhibitor (SAHA)[30], p300 activator (CTB)[30] and inhibitor (C646)[25], and bromodomain inhibitor (JQ1 and Bi)[13, 14]. However, despite recent progress, there major roadblocks such as high rate of nonspecific binding and the requirement for enriched repeat DNA-binding sites remain. In particular, PIPs designed to be enriched at an expanded DNA repeat in a disease model suggests that such systems could have versatile therapeutic applications[14]. The inclusion of a cooperative, gene-specific modifier that can target a DNA repeat locus would potentially overcome the existing roadblocks; however, there is no report of this achievement to date. To this end, our notion is to tether an epi-drug to the PIP–HoGu and construct an advanced synthetic transcription factor mimic termed ePIP–HoGu. This construct is expected to be capable of cooperatively recruiting the epigenetic modifiers to the predetermined DNA locus and nearby nucleosome. Histone acetylation is a significant epigenetic mark that is critical for gene activation. We previously established a biochemical assay in which sequence-selective histone acetylation could be quantified by combining reconstituted nucleosomes, HAT reaction, and chromatin immunoprecipitation (ChIP) with histone acetylation antibody and qPCR[13, 31, 32].

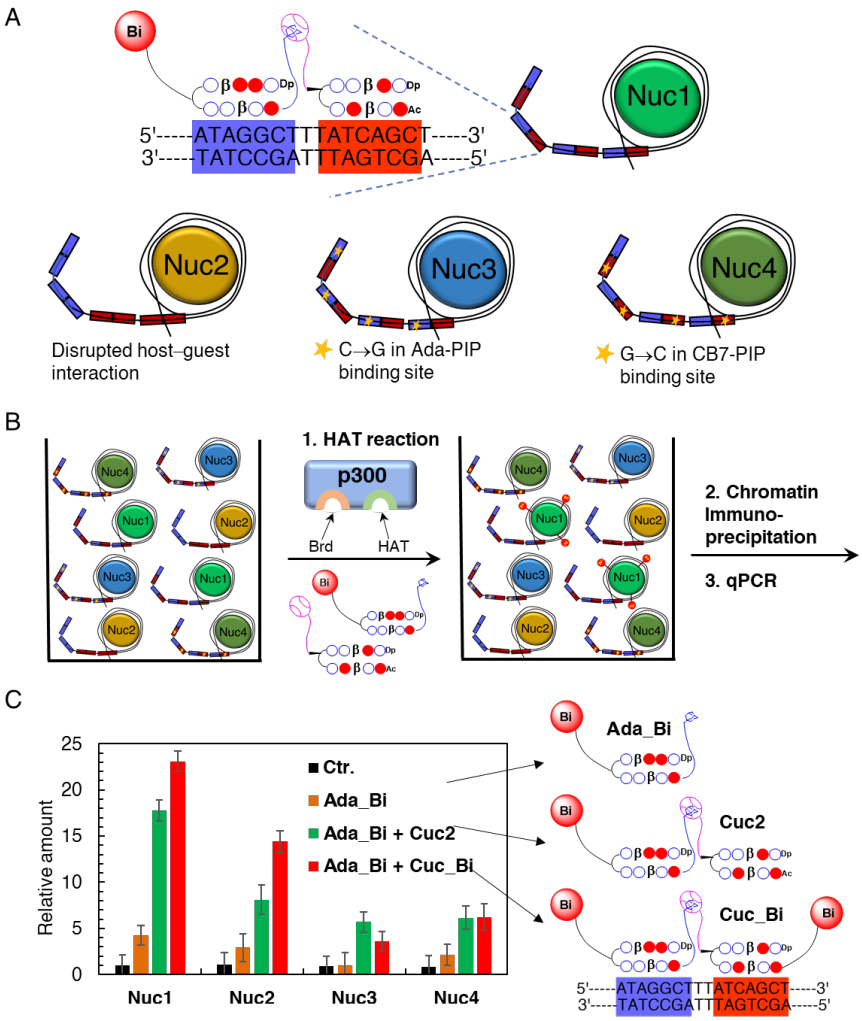

Fig. 4 ePIP–HoGu synergistically recruits an epigenetic modifier to the target DNA repeat locus. (A) Schematic illustration of four kinds of nucleosomes with different DNA templates. Nuc1 contains four-matched repeat sequence of PIP–HoGu binding. Nuc2 has two homodimeric binding sites of Ada-PIP and CB7-PIP separately, which cannot form a host–guest interaction (Nuc2 has potential synergic binding partially between site 2 and 3, because of the short distance between them). One-mismatch bp localizes in the binding site of Ada-PIP for Nuc3 and CB7-PIP for Nuc4. (B) The workflow of the *in vitro* HAT assay. The HAT reaction was conducted in 15 μL HAT buffer, with the addition of four nucleosomes (each concentration was 25 nM), 10 μM Ac-CoA, 15 nM recombinant human P300, 250 nM of each compound. The reaction was conducted for 1 h at 30 °C in HAT buffer (50 mM Tris-HCl, 0.1 mM EDTA, 10% glycerol, 1 mM DTT, pH 8.0). (C) Results of the *in vitro* HAT-ChIP-qPCR assay. Compound treatment in three groups compared with control (DMSO), i.e., **Ada_Bi1**, **Ada_Bi1** + **Cuc2**, and **Ada_Bi1** + **Cuc_Bi1**.

To explore the synergic effect of recruiting recombinant human P300 (965−1810 aa, containing HAT and Brd domains) and the ensuing histone acetylation[33], four types of DNA templates containing a Widom 601 sequence and distinct PIP-binding sites were constructed and reconstituted to form the nucleosome[31]. Nuc1 includes four tandem repeats of the cooperative binding site with a separation of 2 bp, in which ePIP–HoGu was expected to form tetrameric cooperative complexes. To verify the magnitude of cooperation, Nuc2–4 were prepared as control systems (Fig. 4A, S7). Nuc1–4 were mixed together before the *in vitro* HAT reaction. Meanwhile, three PIP conjugates (**Ada_Bi1**, **Cuc2**, **Cuc_Bi1**) were designed to match these DNA-targeting sites, which were (i) either tethered with the guest Ada or host CB7, and (ii) with or without the covalent linkage with the Brd inhibitor, Bi (Fig. S5). The sequence selectivity of the conjugates was firstly confirmed by EMSA that was consistent with the design (Fig. S6). It showed 2–3 folds and > 20 folds selectivity to the sequence with 1 bp

and 2 bp mismatch respectively. These compounds were evaluated in three groups. The Bi-PIP conjugate was anticipated to recruit epigenetic enzyme to the proximate histone tail wrapped by the template DNA containing multiple matched PIP-binding sites[13].

The *in vitro* HAT-ChIP-qPCR assay showed that, in the absence of PIPs, all four reconstituted nucleosomes showed similar, low levels of acetylation, suggesting a minimal influence of partial DNA sequence variation on histone acetylation (Fig. 4C). **Ada_Bi1** induced a similar level of histone acetylation for Nuc1, 2, and 4 with a ratio of 3–5-fold, but not for Nuc3 because of one mismatch insert at the binding sites. In contrast, co-treatment of **Ada_Bi1** and **Cuc2** hugely increased the acetylation level nearly 20-fold for the fully matched Nuc1; however, there was only 5–7-fold enhancement for Nuc2–4. Moreover, **Cuc_Bi1** further enhanced the acetylation level in Nuc1 (to 23.5-fold), which is almost 4–6-fold higher than that of Nuc3 and Nuc4. It would be reasonable to assume a further divergence in acetylation levels after an increase of mismatch frequency at the PIP-binding sites (Fig. S6). Thus, these results validate the favourable sequence-selective and synergic recruitment of functional enzymes augmented by ePIP–HoGu, suggesting their use for biological regulation.

In summary, for the first time, a small-molecule-based system has been developed to closely mimic natural TF pairs that contain a DNA binding domain, an interaction domain, and a gene regulatory domain. A CB7-assisted PIP-HoGu system complexed with ethyldiamino-Ada-PIPs has been shown to exhibit host–guest interactions that are superior to those of the CyD-system[9], which is established as a reference model. Furthermore, the incorporation of a cooperative dimer system into PIPs–epi-drug conjugates increases the DNA recognition length, reinforces reasonable sequence selectivity, and allows versatile binding modes. As a proof-of-concept study, the ePIP–HoGu system is shown to be adept at synergistically augmenting proximate histone acetylation with valuable efficiency and selectivity. The ePIP–HoGu system could thus evolve further into a chemical alternative to protein-based systems such as dCas9 and ZFs that deliver high efficiency and selectivity[34]. Further efforts on the optimization of the epi-drug and assay platform will fast-track the application of this synthetic tool to cell fate control and, ultimately, as therapeutic drugs.

L.I. thanks the National Science Foundation (CHE-1807486) for financial support.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 A. Jolma, Y. Yin, K. R. Nitta, K. Dave, A. Popov, M. Taipale, M. Enge, T. Kivioja, E. Morgunova and J. Taipale, *Nature*, 2015, **527**, 384-388.

2 A. K. Mapp, A. Z. Ansari, M. Ptashne and P. B. Dervan, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 3930-3935.

3 M. Ueno, A. Murakami, K. Makino and T. Morii, *J. Am. Chem. Soc.*, 1993, **115**, 12575-12576.

4 J. Mosquera, A. Jimenez-Balsa, V. I. Dodero, M. E. Vazquez and J. L. Mascarenas, *Nat. Commun.*, 2013, **4**, 1874.

5 J. B. Blanco, V. I. Dodero, M. E. Vázquez, M. Mosquera, L. Castedo and J. L. Mascareñas, *Angew. Chem. Int. Ed. Engl.*, 2006, **45**, 8210-8214.

6 P. B. Dervan, *Bioorg. Med. Chem.*, 2001, **9**, 2215-2235.

7 Z. Yu, G. N. Pandian, T. Hidaka and H. Sugiyama, *Adv. Drug Deliv. Rev.*, 2019, **147**, 66-85.

8 C. L. Kielkopf, S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan and D. C. Rees, *Science*, 1998, **282**, 111-115.

9 Z. Yu, C. Guo, Y. Wei, K. Hashiya, T. Bando and H. Sugiyama, *J. Am. Chem. Soc.*, 2018, **140**, 2426-2429.

10 Z. Yu, W. C. Hsieh, S. Asamitsu, K. Hashiya, T. Bando, D. H. Ly and H. Sugiyama, *Chem. Eur. J.*, 2018, **24**, 14183-14188.

11 P. Weyermann and P. B. Dervan, *J. Am. Chem. Soc.*, 2002, **124**, 6872-6878.

12 A. T. Poulin-Kerstien and P. B. Dervan, *J. Am. Chem. Soc.*, 2003, **125**, 15811-15821.

13 J. Taniguchi, Y. Feng, G. N. Pandian, F. Hashiya, T. Hidaka, K. Hashiya, S. Park, T. Bando, S. Ito and H. Sugiyama, *J. Am. Chem. Soc.*, 2018, **140**, 7108-7115.

14 G. S. Erwin, M. P. Grieshop, A. Ali, J. Qi, M. Lawlor, D. Kumar, I. Ahmad, A. McNally, N. Teider, K. Worringer, R. Sivasankaran, D. N. Syed, A. Eguchi, M. Ashraf, J. Jeffery, M. Xu, P. M. C. Park, H. Mukhtar, A. K. Srivastava, M. Faruq, J. E. Bradner and A. Z. Ansari, *Science*, 2017, **358**, 1617-1622.

15 S. Liu, C. Ruspic, P. Mukhopadhyay, S. Chakrabarti, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2005, **127**, 15959-15967.

16 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.

17 J. Kim, I.-S. Jung, S.-Y. Kim, E. Lee, J.-K. Kang, S. Sakamoto, K. Yamaguchi and K. Kim, *J. Am. Chem. Soc.*, 2000, **122**, 540-541.

18 B. Gong, B.-K. Choi, J.-Y. Kim, D. Shetty, Y. H. Ko, N. Selvapalam, N. K. Lee and K. Kim, *J. Am. Chem. Soc.*, 2015, **137**, 8908-8911.

19 X. Zhou, X. Su, P. Pathak, R. Vik, B. Vinciguerra, L. Isaacs and J. Jayawickramarajah, *J. Am. Chem. Soc.*, 2017, **139**, 13916-13921.

20 M. J. Webber, E. A. Appel, B. Vinciguerra, A. B. Cortinas, L. S. Thapa, S. Jhunjhunwala, L. Isaacs, R. Langer and D. G. Anderson, *Proc. Natl. Acad. Sci. U.S.A.*, 2016, **113**, 14189-14194.

21 K. M. Park, J. Murray and K. Kim, *Acc. Chem. Res.*, 2017, **50**, 644-646.

22 D. Shetty, J. K. Khedkar, K. M. Park and K. Kim, *Chem. Soc. Rev.*, 2015, **44**, 8747-8761.

23 K. L. Kim, G. Sung, J. Sim, J. Murray, M. Li, A. Lee, A. Shrinidhi, K. M. Park and K. Kim, *Nature Commun.*, 2018, **9**, 1712.

24 B. Vinciguerra, L. Cao, J. R. Cannon, P. Y. Zavalij, C. Fenselau and L. Isaacs, *J. Am. Chem. Soc.*, 2012, **134**, 13133-13140.

25 Z. Yu, J. Taniguchi, Y. Wei, G. N. Pandian, K. Hashiya, T. Bando and H. Sugiyama, *Eur. J. Med. Chem.*, 2017, **138**, 320-327.

26 S. Pandey, D. V. D. W. Kankanamalage, X. Zhou, C. Hu, L. Isaacs, J. Jayawickramarajah and H. Mao, *J. Am. Chem. Soc.*, 2019, DOI: 10.1021/jacs.9b09019.

27 G. Yu, K. Jie and F. Huang, *Chem. Rev.*, 2015, **115**, 7240-7303.

28 X. Xiangshu, Y. Peng, L. Hyun‐Suk, S. Devanjan and K. Thomas, *Angew. Chem. Int. Ed. Engl.*, 2007, **46**, 2865-2868.

29 Y. Kwon, H.-D. Arndt, Q. Mao, Y. Choi, Y. Kawazoe, P. B. Dervan and M. Uesugi, *J. Am. Chem. Soc.*, 2004, **126**, 15940-15941.

30 L. Han, G. N. Pandian, A. Chandran, S. Sato, J. Taniguchi, G. Kashiwazaki, Y. Sawatani, K. Hashiya, T. Bando, Y. Xu, X. Qian and H. Sugiyama, *Angew. Chem. Int. Ed. Engl.*, 2015, **54**, 8700-8703.

31 T. Zou, F. Hashiya, Y. Wei, Z. Yu, G. N. Pandian and H. Sugiyama, *Chem. Eur. J.*, 2018, **24**, 15998-16002.

32 U. T. T. Nguyen, L. Bittova, M. M. Müller, B. Fierz, Y. David, B. Houck-Loomis, V. Feng, G. P. Dann and T. W. Muir, *Nat. Meth.*, 2014, **11**, 834.

33 M. Delvecchio, J. Gaucher, C. Aguilar-Gurrieri, E. Ortega and D. Panne, *Nat. Struct. Mol. Biol.*, 2013, **20**, 1040-1046.

34 A. Chavez, M. Tuttle, B. W. Pruitt, B. Ewen-Campen, R. Chari, D. Ter-Ovanesyan, S. J. Haque, R. J. Cecchi, E. J. K. Kowal, J. Buchthal, B. E. Housden, N. Perrimon, J. J. Collins and G. Church, *Nat. Meth.*, 2016, **13**, 563.