Advanced Drug Delivery Reviews 147 (2019) 66–85

Contents lists available at ScienceDirect

Advanced Drug Delivery Reviews

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

Therapeutic gene regulation using pyrrole–imidazole polyamides

Zutao Yu [a,b], Ganesh N. Pandian [b,*], Takuya Hidaka [a], Hiroshi Sugiyama [a,b,*]

[a] Department of Chemistry, Graduate School of Science, Kyoto University, Kitashirakawa-Oiwakecho, Sakyo-Ku, Kyoto 606-8502, Japan
[b] Institute for Integrated Cell-Material Sciences (WPI-iCeMS), Kyoto University, Yoshida-Ushinomaecho, Sakyo-Ku, Kyoto 606-8501, Japan

ARTICLE INFO

Article history:
Received 27 September 2018
Received in revised form 22 November 2018
Accepted 4 February 2019
Available online 10 February 2019

Keywords:
Artificial genetic switches
Designer ligands
Pyrrole–imidazole polyamides
Fluorescent probes
Therapeutic gene modulation

ABSTRACT

Recent innovations in cutting-edge sequencing platforms have allowed the rapid identification of genes associated with communicable, noncommunicable and rare diseases. Exploitation of this collected biological information has facilitated the development of nonviral gene therapy strategies and the design of several proteins capable of editing specific DNA sequences for disease control. Small molecule-based targeted therapeutic approaches have gained increasing attention because of their suggested clinical benefits, ease of control and lower costs. Pyrrole–imidazole polyamides (PIPs) are a major class of DNA minor groove-binding small molecules that can be predesigned to recognize specific DNA sequences. This programmability of PIPs allows the on-demand design of artificial genetic switches and fluorescent probes. In this review, we detail the progress in the development of PIP-based designer ligands and their prospects as advanced DNA-based small-molecule drugs for therapeutic gene modulation.

Contents

Abbreviations: PIPs, pyrrole–imidazole polyamides; TFs, transcription factors; γ-turn, γ-aminobutyric acid; Hp, 3-hydroxypyrrole; RNAP, RNA polymerase; ATFs, artificial transcription factors; Brd, bromodomain; AR, androgen receptor; HIF, hypoxia-inducible factor; ER, estrogen receptor; GR, glucocorticoid receptor; ARE, androgen-response element; EVI1, human ectopic viral integration site 1; MMP9, matrix metalloprotein 9; hiPS, human induced pluripotent stem (hiPS) cells; HDACs, histone deacetylases; HATs, histone acetyltransferases; SAHA, suberanilohydroxamic acid; MEFs, mouse embryonic fibroblasts; ChIP, chromatin immune precipitation (ChIP) analysis; HDFs, human dermal fibroblasts; IPA, isophthalic acid; CTB, [*N*-(4-chloro-3-trifluoromethyl-phenyl)-2-ethoxy-benzamide]; Syn-TEF1, sequence-specific synthetic transcription elongation factor 1; Duo, durocarmycin; Chb, chlorambucil; CBI, *seco*-1-chloromethyl-5-hydroxy-1,2-dihydro-3*H*-benz[e]indole; ICL, interstrand crosslinks; RUNX1, Runt-related transcription factor 1; AML1, acute myeloid leukemia 1 protein; SOS, son of sevenless homolog (SOS) family; CPI, cyclopropylindole; DM1, myotonic dystrophy type 1; Ht, hoechst analogues; EBV, Epstein—Barr virus; PNA, peptide nucleic acid; pcPNA, pseudo-complementary (pcPNA) strand; G4, G-quadruplex; CID, cooperative interaction domain; Cyd, cyclodextrin; Ada, adamantane; Pip–HoGu, PIPs conjugated to host–guest assemblies; Pip–NaCo, PIPs conjugated with nucleic acid-based cooperation system; FISH, fluorescence *in situ* hybridization; mtDNA, mitochondria possess their own DNA; TFAM, mitochondrial transcription factor A; LSP, light strand promoter.

* Corresponding authors at: Institute for Integrated Cell-Material Sciences (WPI-iCeMS), Kyoto University, Yoshida-Ushinomaecho, Sakyo-Ku, Kyoto 606-8501, Japan.
E-mail addresses: ganesh@kuchem.kyoto-u.ac.jp (G.N. Pandian), hs@kuchem.kyoto-u.ac.jp (H. Sugiyama).

https://doi.org/10.1016/j.addr.2019.02.001
0169-409X/

1. Introduction

Artificial control of spatiotemporally variable gene expression has been the goal of generations of scientists. The current availability of open-access, user-friendly platforms enables us to track any part of the genetic information in the whole genome. Even though the abundant available information about the genome can be harnessed to identify well-defined targets for drug design, there is still a lack of promising small molecule-based designer drug capable of achieving targeted modulation of therapeutically important genes. Therefore, there is a need to explore the potential of programmable, sequence-specific DNA binders and advance them as designer drugs for achieving on-demand therapeutic gene modulation.

Pyrrole–imidazole polyamides (PIPs) represent the best-characterized class of small molecules that can be modified to bind to any predetermined DNA sequence and regulate gene expression patterns in a transgene-free and cost-effective manner [1,2]. Simple binary code has been elucidated to correlate the target DNA sequence with the side-by-side antiparallel pairing between pyrrole (Py) and imidazole (Im) under physiological conditions, i.e., Py/Im recognizes C/G pairs and Py/Py recognizes degenerate A/T and T/A pairs [3].

PIPs have been extensively applied to regulate specific gene expression because they can bind to the cognate DNA sequences and alter their interaction with natural transcription factors (TFs), which in turn modulate the transcriptional networks and the epigenetic landscapes that govern cell fate. In eukaryotes, TFs play an essential role in the spatiotemporal regulation of gene transcription, which decide the cell fate by modulating the genetic program associated with a particular phenotype. More than 1000 TFs recognize around 200–300 short DNA motifs in promoter and enhancer regions, recruiting gene coactivators or corepressors, and transforming that genetic information to allow an ON/OFF status [4]. Specific TFs are responsible for a particular biological function, and blocking the binding of such TFs to a particular DNA motif can perturb their biological function and have a profound physiological effect. For example, SOX2, a Yamanaka factor, is essential for stem cell pluripotency, and disrupting its binding using a sequence-specific DNA binder could alter the fate of stem cell differentiation and deflect it into a particular lineage [5].

More significantly, PIPs serve as a valuable DNA binding platform to achieve higher order biological functions in a programmable manner. First, tethering epi-drugs to DNA binders constrains the pharmacology of these drugs to the target locus, providing a robust effect on the target gene with minimal off-target effects. Second, PIP-alkylating agents are capable of blocking RNAP activity in the coding region of target genes to sequence-specifically downregulate target gene expression, and preclinical studies show that they have promising anticancer efficacy. Moreover, the development of covalent and especially noncovalent conjugates of PIPs with additional DNA binders has emerged as the new research field that has helped to overcome several limitations of conventional PIPs. Lastly, PIP-based fluorescent probes are an important application for sequence-specific DNA imaging in living cells.

Given the well-defined methods of working used in the past, recent breakthroughs in gene-specific modulation by biomolecules, and several recent significant innovations in PIPs, we recapitulate recent advances in this field including the development of PIPs as artificial genetic switches and fluorescent probes, and outline the difficulties and future challenges of their use as designer drugs.

2. Brief history of pyrrole–imidazole polyamides

The precursors of cell-permeable, programmable, DNA minor groove-binding PIPs are the natural antibiotics netropsin and distamycin A, a class of Py dimer/trimer molecules linked with an amide bond (Fig. 1) [6,7]. From 1982, researchers made considerable efforts to decipher, modify and optimize these crescent-shaped chemical entities, which enabled the development of an easy-to-use DNA binder [8,9]. In particular, the clarification of the crystal structure of cooperative distamycin–DNA 2:1 complexes and the incorporation of Im building blocks established the basis for the subsequent development of PIPs [10,11].

The Dervan group first developed hairpin PIPs by incorporating aliphatic γ-aminobutyric acid (γ-turn), which nicely combined and fixed two linear binding chains [12]. Hairpin PIPs are the most commonly used chemical architecture, which can be attributed to their ease of synthesis, unambiguous pairing rules, high binding affinity, and excellent sequence selectivity. Results from X-ray crystal studies manifested that the Im/Py pair recognizes G/C pairs and the Py/Py pair recognizes A/T or T/A; this defines the essence of their sequence specificity and binding site programmability (Fig. 2a) [13]. The precise hydrogen bond interactions included in G:Im are 2-NH_2·3′-N and 3-N·amide-NH_2; for C:Py is 2-CO·amide-NH_2; for A:Py is 3-N·amide-NH_2; and for T:Py is 2-CO·amide-NH_2 (Fig. 2b). Meanwhile, γ-turn and terminal *N*, *N*′-dimethyl-1,3-propanediamine prefer to bind A/T pairs. High torsional rigidity significantly reduce the binding affinity and specificity when there are over four continuous rings owing to the exaggerated curvature. A β-alanine linker used as a replacement for the Py moiety in the more extended PIP sequences aides relaxing torsional rigidity and provides the possibility of enhanced cellular uptake. β-alanine also “springs” the aliphatic amino acid residues in together with γ-turn [14]. To resolve the degeneracy of the Py/Py pair binding, 3-hydroxypyrrole (Hp) confers the best distinguishing ability, but its chemical instability limits its widespread application (Fig. 2b) [15]. Modification of the N-terminus from Im to 5-alkyl thiazole building blocks results in an increase in dsDNA binding affinity and causes a more significant degree of major groove compression of the target dsDNA sequence than its Im counterpart [16]. PIPs exhibit strong

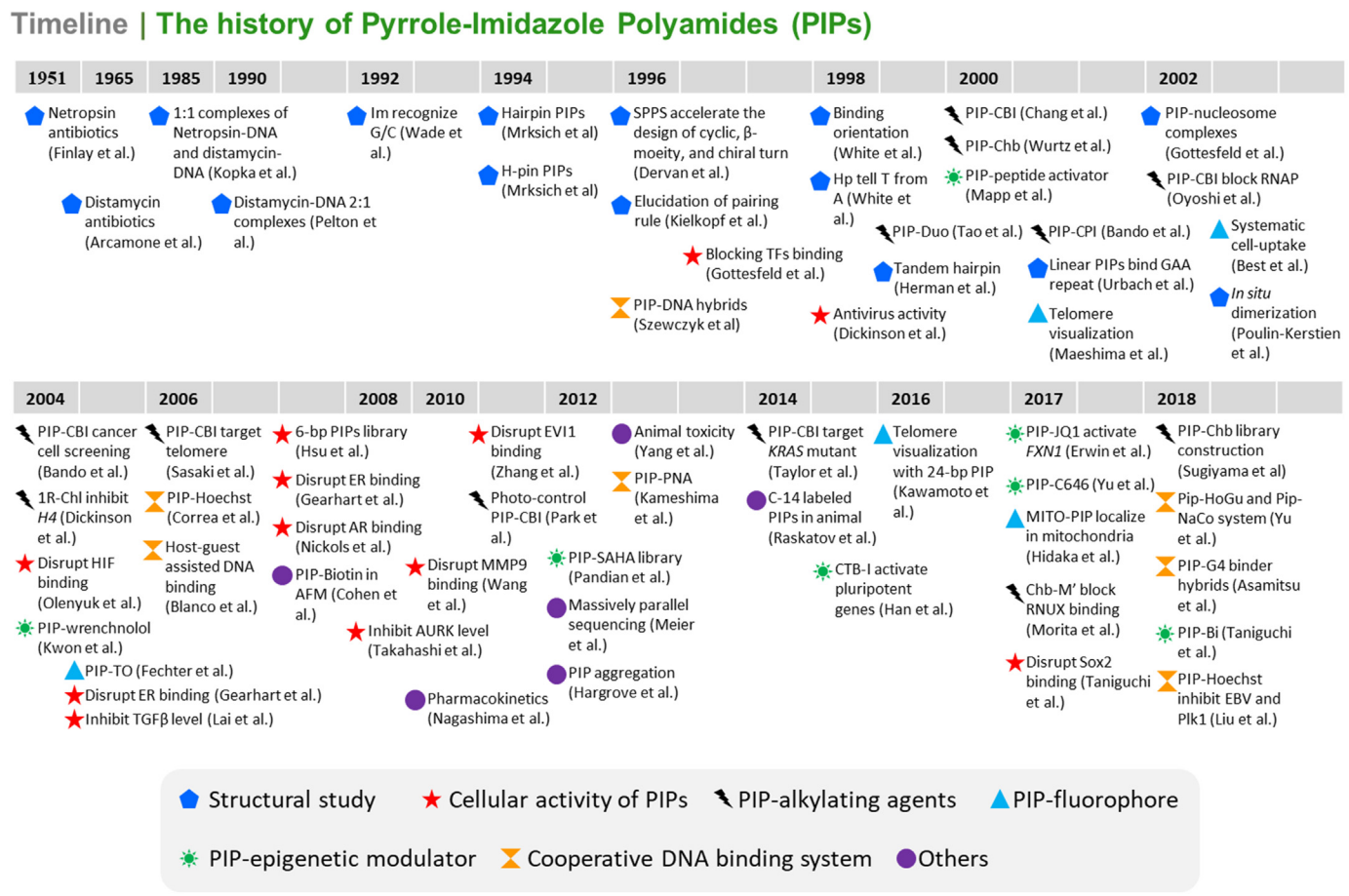


Fig. 1. The timeline of pyrrole-imidazole polyamides (PIPs) research.

binding affinity to their matching sequence. For example, 6-bp-binding PIPs showed binding affinities with K_d values at the nanomolar or even picomolar level. Moreover, PIPs can distinguish dsDNA from ssDNA, ssRNA and dsRNA with high selectivity [17].

The development of solid-phase peptide synthesis resulted in a surge in the design and characterization of various intricate molecular architectures (Fig. 2c) [18,19]. Cyclic PIPs display higher binding affinities (40-fold enhancement for 5'-TGTTA-3') than their counterpart hairpin PIPs [20]. Despite their small reduction in binding affinity, H-pin and U-pin PIPs extend the targeting sequence to the terminal G/C-rich DNA sequence [21,22]. Tandem hairpin structures, including head-to-head and head-to-tail styles, are applicable to the design of molecules targeting more extended sequences, such as a PIP targeting the 24-bp telomere repeat [23]. Polypurine DNA has a narrow minor groove, and a linear PIP with a 1:1 binding mode was successfully applied to bind to a GA-rich sequence [24]. The replacement at γ-turn showed significant effect on DNA binding affinity [25], binding orientation, sequence selectivity, solubility [26], cell uptake [27] and toxicity [28], including γ-aminobutyric acid (γ-turn), (*R*)- and (*S*)-β-amino-γ-turn, (*R*)-α-amino-γ-turn, α-diaminobutyric acid, and the respectively acetylated form at amino group.

The demonstration that PIPs are capable of disrupting TF–DNA interactions to modulate RNA polymerase (RNAP) II and III functions led to their broad applications as sequence-specific gene regulators (described in Section 3) [2]. PIPs also serve as an exquisitely specific DNA binding platform. The proof-of-concept that they are sequence-specific gene activators was provided by tethering peptide-based activation domains [29], especially the recently identified epigenetic modulators [30], and offers fresh possibilities of artificial transcription factors (ATFs), especially PIP–bromodomain (Brd) inhibitor conjugates (described in Section 4). Sequence-specific DNA alkylating agents targeting a coding region to block RNAP II/III showed promising efficacy under both *in vitro* and *in vivo* conditions (described in Section 5) [31]. The development of covalent and especially noncovalent conjugates of PIPs with additional DNA binders has emerged as a new research strategy that has helped to overcome some limitations of conventional PIPs (described in Section 6). Several potential therapeutic applications of PIPs supplemented with various functional tags are outlined in Section 7.

3. Therapeutic gene regulation using designer PIPs

In eukaryotes, more than 1000 TFs recognize around 200–300 short DNA motifs in the promoter and enhancer regions of genes, recruiting gene coactivators or corepressors, which in turn transcribe the biological information to code for ON/OFF states [4]. Distinct TFs are responsible for defined biological functions and hindering the binding of such TFs blocks the transcription machinery. PIPs possess potent DNA binding affinity capable of blocking TF–DNA interactions. Accordingly, the primary therapeutic strategy using PIPs has been applied to block the interaction of DNA-binding proteins with DNA and trigger the modulation of RNAP II/III activity, which is followed by downstream gene perturbation.

3.1. Short PIPs

Dervan, Gottesfeld and colleagues first demonstrated the selectivity and efficiency of PIPs as regulators of gene expression. In living cells, 8-ring hairpin PIPs were shown to interfere with 5S RNA expression by occupying specific regions of TFIII binding sites and suppressing RNAP III activity [2]. Moreover, PIPs targeting the flanking regions of the binding

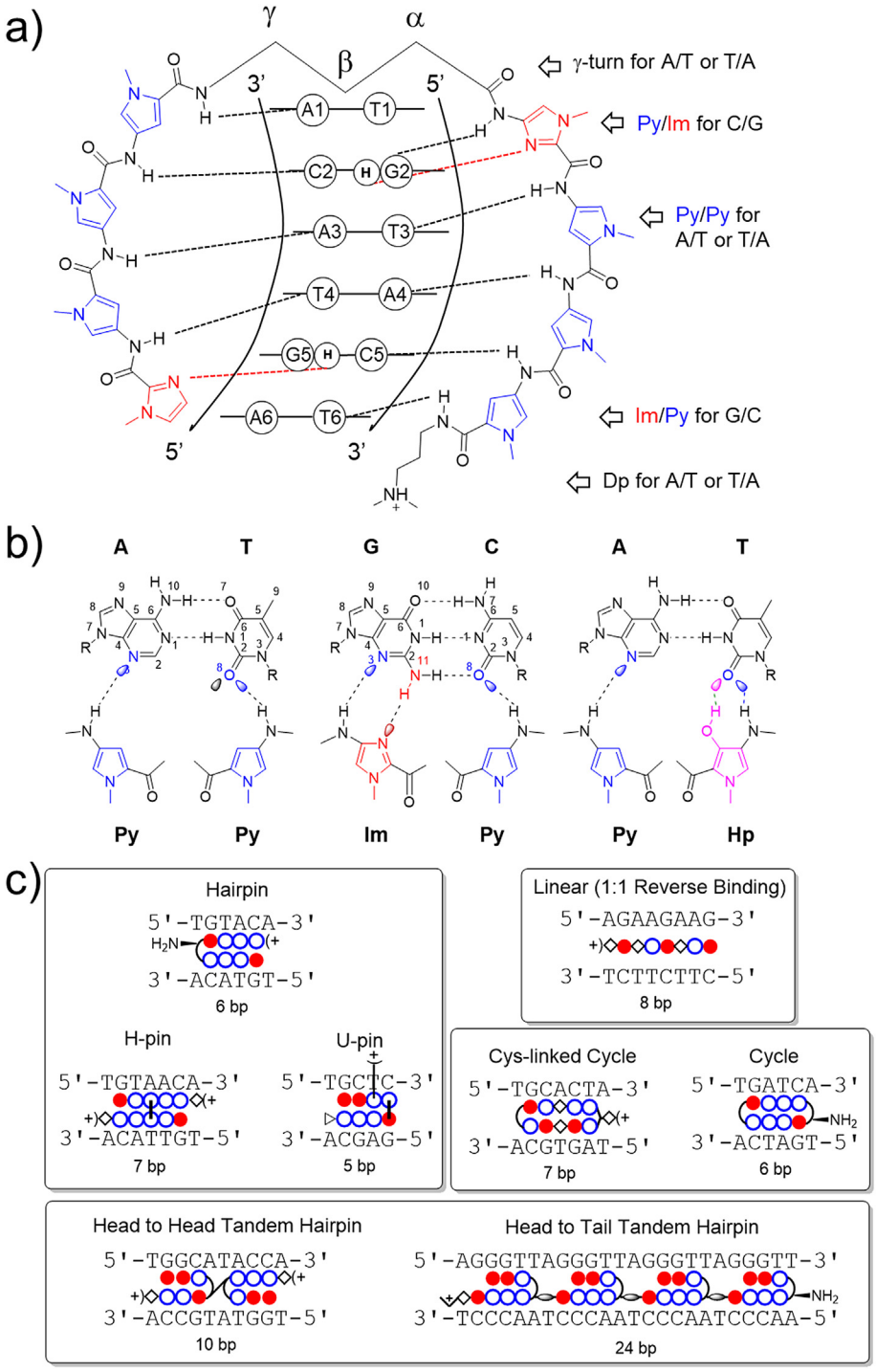


Fig. 2. (A) Detailed hydrogen bond interaction and pairing rules of PIP/DNA interactions. (B) Schematic hydrogen interaction map of PIP building blocks with four bases. (C) Different type of PIPs design and its binding to DNA. Main symbols include ○: Py. ●: Im; c: γ-aminobutyric acid (γ-turn). (+: *N,N′*-Dimethyl-1,3-propanediamine. ◇: β-alanine. (Modified from [187]).

sites of Ets-1, LEF-1 and TBP in the HIV-1 promoter/enhancer suppressed RNAP II, thereby inhibiting virus replication in human lymphocytes by >99% [32]. A parallel study in human cytomegalovirus demonstrated the use of PIPs in activating RNAP II and inducing transcription by blocking the DNA sequence where a repressor TF was bound [33]. Short PIPs with a second positive charge at the C-terminal exhibited antifungal activity against clinically relevant fungal strains under *in vitro* conditions and a mouse model [34]. In addition, crystallographic structural studies revealed allosteric perturbation of the DNA helix as the molecular basis by which minor groove-targeting PIPs disrupted the major groove-binding TFs [35]. Systematic cellular uptake studies using PIPs provided direct evidence of their effective cell permeability and nuclear localization properties to suggest their potential as clinical treatments [36]. Consequently, PIPs have been explored for their broader applications: structural optimization studies were performed to identify useful cellular applications and a selected list of reported PIPs targeting TFs and related information (Table 1).

Dervan and colleagues reported a series of results regarding the biological evaluation of 8-ring PIPs in cell assays and mouse models. For example, androgen-response element (ARE)-targeting PIPs (5′-WGWWCW-3′) in the prostate-specific antigen promoter disrupted androgen receptor–ARE interactions and inhibited the expression of

Table 1
Designer PIPs mediate TFs disruption.

TFs affected	PIP-binding site (5' to 3')	Target genes	Biological activity	Ref.
Short PIPs				
Androgen receptor (AR)	WGWWCW	AR target genes	Disruption of AR–ARE interactions to inhibit the expression of AR downstream genes; antitumor activity in the mice model of VCaP xenografts	[37,38]
Hypoxia-inducible factor (HIF)	WTWCGW	HIF target genes	Reduction in the expression of proangiogenic and pro-metastatic factors; inhibition of the formation of new tumor blood vessels and tumor growth	[39]
		HER	Decrease in the HER2 expression and inhibition of the cellular proliferation in various cancer cells	[40,41]
		VEGF	Disruption of a subset of hypoxia-induced genes and demonstration of consistency of biological effect with the binding site preferences of PIPs	[42]
Estrogen receptor (ER)	WGGWCW	ER target genes	Reduction of ER-driven luciferase expression in xenograft tumors	[43]
NFκB	WGGWWW	NF-κB	Binding to NFκB sites and reduction of the expression of various NFκB–driven genes including *IL6* and *IL8* by qRT-PCR	[44]
EVI1	WCWGWW	EVI1 reporter	Blocking of EVI1-responsive reporter activity	[45]
Glucocorticoid receptor (GR)	WGWWCW	GILZ	Reduction of GR occupancy at the GILZ promoter and inhibition of GILZ expression and several other known GR target genes	[46]
TFIIIA	WGWWCW	5S RNA	Target and recognition of TFIIIA binding site and interfering with *5S RNA* expression	[2]
Long PIPs				
RBPJ	GGWWWGWW	HES1	Sequence-specific suppression of neurogenesis-associated HES1 and its downstream genes; neural differentiation	[47]
SOX2	CWWWGWW	SOX2	Downregulation of SOX2 expression and mesoderm induction	[5]
Mitochondrial TFA	WGWCWGWW	LSP	Targeted transcriptional inhibition of ND6 gene in the mitochondria light strand	[48]
REL/ELK1 binding site	WWWWCCWW	EVI1	Inhibition of EVI1 in MDA-MB-231 cells; inhibit breast cancer cell migration	[49]
c-MYC	CWCGWGW	eIF4G1, CCND1, and CDK4	Inhibition of MYC binding at the promoter and downregulation at the mRNA level and protein expression of target genes	[50]
AP1	WCWCWCCW	TGFβ	Inhibition of hypertrophic scar formation at 35 days post-incision and prohibition of cellular infiltration, TGF-β1 and vimentin staining, and epidermal thickness	[51]
NFκB	WWWWCCW	MMP9	Reduction in *MMP9* gene expression and the enzymatic activity of MMP9 proteins; inhibition of migration and invasion in SaOS-2 cells	[52]
E4TF1 (GABPB)	WCCWCWW	AURKA	Inhibition of the promoter activities, mRNA expression, and protein levels of AURKA and AURKB; antiproliferative synergy to HeLa cells	[53]
Ets	WGWWWWW	HER2	Disruption of endogenous Ets-mediated HER2/neu preinitiation complexes	[54]

androgen downstream genes in LNCaP and VCaP cancer cells. Intensive structural modifications revealed that the oxime linkage between the PIPs and an aromatic functionality on the C-terminus generated a significant increase (~20-fold) in their cellular uptake and gene regulation activity [55]. Cyclic PIPs showed similar cellular activity and higher DNA binding affinity ($\Delta\Delta T_m = 5.3°C$) than their hairpin counterparts [56]. An aryl modification at the site of γ-turn enhanced their cellular uptake [27]. In a study of their anticancer mechanism, PIPs were shown to disrupt RNAP II localization at the transcription start sites and activated p53 signaling in cancer cell lines. Furthermore, PIP inhibition of DNA helicases leading to the inhibition of S-phase, topoisomerase–DNA binding and disruption of DNA ligase function could contribute to their anticancer activity [38,57,58]. PIPs did not cause detectable DNA damage but had potent antitumor activity ($IC_{50} = 6.2 \pm 2.6$ μM) in LNCaP prostate tumor cells, and in an enzlutamide-resistant prostate tumor xenograft model, PIPs had anticancer effects with limited host toxicity [37,59]. Because of the similarity of TF-binding sequences, ARE-targeting PIPs showed off-target effects and could also bind to the glucocorticoid response element to inhibit the glucocorticoid receptor–DNA interaction, thereby perturbing a cluster of glucocorticoid-induced and repressed genes. Together, obstacles including their limited bioactivity at the target DNA sequences of several TFs, their off-target recognition of flanking sequences and nonspecific TFs limit the use of short PIPs that target a cognate DNA sequence in a gene promoter region.

3.2. Long PIPs

Therefore, our group and collaborators focuse on the construction of longer designer PIPs. Such hairpin PIPs targeting longer sequence of therapeutically important genes have successfully altered the endogenous expression in different cell types of target genes including aurora kinases, TGF-β1, human lectin-like oxidized low-density lipoprotein, HER2 and EBNA1 [53,60–63]. For example, a designer PIP was shown to localize inside human breast adenocarcinoma (MDA-MB-231) cells and recognize the specific sequence of the REL/ELK1 binding site in the promoter region of the human ectopic viral integration site 1 (EVI1). EVI1 is a well-known oncogenic TF associated with a various type of cancers, and microarray studies revealed that the designer PIP inhibited the expression of genes related to breast cancer cell migration [45,49]. Likewise, a PIP targeting the activator protein-1-binding site in the promoter region of matrix metalloprotein 9 (MMP9), a tumor invasion and metastasis-associated factor, inhibited MMP9 expression and altered the migration and invasion of MDA-MB-231 cells [64]. Pharmacokinetic studies after intravenous administration of this PIP to rats suggested that it had notable effects on tumor metastasis [65,66]. Designer PIPs were also shown to bind to methylated 5′-CpG-3′ sequences and demonstrated a higher binding affinity to the unmethylated counterpart (3-fold) [67].

Recently, a ground-breaking application of PIPs was demonstrated in human induced pluripotent stem (hiPS) cells. Targeted differentiation of hiPS cells using only chemicals were reported to have the potential to generate clinically useful differentiated cell types such as cardiomyocytes [68]. Exploiting the available knowledge of the intertwined transcription machinery associated with cardiomyocyte differentiation, a designer PIP termed PIP-S2 was constructed [5]. PIP-S2 recognized the sequence 5′-CTTTGTT-3′ and inhibited the SOX2–DNA minor groove interaction. PIP-S2 (2 μM) caused promoter-specific suppression of transcription of SOX2 downstream genes and inhibited their protein expressions. Whole transcriptome analyses suggested that PIP-S2 specifically targeted and altered SOX2-associated gene regulatory networks associated with mesoderm induction. More significantly, the sequential treatment of PIP-S2 (2 μM treatment from day 1-5) and then Wnt/β-catenin inhibitor (day 6-7) effectively generated spontaneously contracting cardiomyocytes at day 7, which opens up new vistas of opportunity for differentiating stem cells into desired

cell types [5]. Similarly, we designed PIP–RBPJ-1, which caused the promoter-specific suppression of neurogenesis-associated HES1 (showed 20% inhibition at the concentration of 2 μM for 24 h treatment) and its downstream genes. PIP–RBPJ-1 alone altered the expression of neural system-associated Notch-signaling factors and triggered the neurogenesis gene program with an efficiency comparable to the conventional approach using small molecules targeting the protein–protein interactions [47].

Over the past 20-year on the application of PIPs-mediated TFs–DNA disruption, various commonly applicable guide rules and preferences got proposed for the construction of designer PIPs for a cognate DNA sequence in the specific study. First, the anticipated targeting DNA motif of PIP is preferred to have multiple binding sites in the promoter of target genes. The bioefficacy of PIPs is high when it is designed to target DNA sites, such as the highly enriched TF binding sites identified from TFs-based ChIP-Seq assay. Second, a TF playing a critical role in a signal pathway network with a positive feedback manner is a preferable target. For example, PIP-S2 blocks SOX2 binding to all core pluripotency genes thereby altering the global expression profile of the pluripotency network and shifts them to a specific terminal cell lineage. Third, PIP directly inhibiting the DNA minor groove-binding TFs is preferred. Because the inhibition of major groove-binding TFs through allosteric effect is relatively less potent. The critical factor in PIP design is the length while the short PIPs generally exert high rate of off-target binding. Also, the bioefficacy of short PIPs can be highly-cell type dependent; however, the screening studies using microarray could aid in identifying and characterizing the hit. Inside the allowance of TF-common binding sequence, PIP length needs to be as long as possible. Moreover, structural optimization such as T/A selective moiety, γ-turn modification, and cyclic vs linear PIP needs to be considered.

4. Therapeutic gene regulation using designer epigenetic switches

PIPs can be engineered as a DNA binder platform to design secondary gene-modulating architecture. In this section, we describe the progress in producing ATFs by tethering epi-drugs (modulators of epigenetic enzymes) to the PIP platform.

The Dervan and Ansari groups did pioneering work in the development of PIP-based ATFs [69,70]. Utilizing PIPs as the DNA binding platform, they covalently conjugated them to short peptide-based gene activation domains, including the Gal4 activation domain, VP16, polyproline helical peptide and Exd recruiting peptides [203]. In addition, the Uesugi group reported the conjugation of a small molecule wrenchnolol-based gene activation domain, a Sur-2 recruiting agent, to PIPs [204]. A hybrid generated by tethering a synthetic catalyst DMAP-SH to PIPs could actively transfer various acyl groups to the lysine residue of the proximate nucleosome [71]. These ATFs successfully activated sequence-specific gene transcription in cell-free assays.

Those producing ATFs often overlook the essential role of epigenetic enzymes such as histone deacetylases (HDACs), histone acetyltransferases (HATs) and Brd, which control the acetylation status of the histone proteins and therefore, the accessibility of the DNA to the transcription machinery (Fig. 3a). Epi-drugs are a group of promising small molecule modulators targeting epigenetic enzymes [72,73]. Epi-drugs function by reversing epigenetic patterns across the whole genome. However, the lack of gene selectivity causes side effects. By tethering an epi-drug to a DNA binder, its pharmacology is constrained to the target locus, providing a robust effect on the target gene with minimal off-target effects (Table 2).

4.1. PIP–HDAC inhibitors

Inhibiting the "eraser" of histone acetylation at sequence-specific sites of target genes reveals distinct gene activation patterns. Our group reported a type of ATF termed SAHA–PIPs, in which 6-bp-binding PIPs and suberanilohydroxamic acid (SAHA, a clinically used HDAC inhibitor) constituted the DNA-binding and functional modules, respectively (Fig. 3b). Our group initially confirmed that a designer SAHA–PIP specifically triggered histone H3 Lys9 acetylation in the promoter region of the p16 tumor suppressor gene and transformed the morphological features of HeLa cells [30]. Because targeted histone acetylation could be applied to reprogram somatic cells into pluripotent stem cells; it would be interesting if SAHA–PIPs could regulate endogenous expression of Yamanaka factors, i.e., *Oct4, Sox2, Klf4* and *c-Myc* [81]. Therefore, we screened a library of 32 SAHA–PIPs (A to Φ) in mouse embryonic fibroblasts (MEFs). Quantitative RT-PCR studies showed that the SAHA–PIPs D, E, J, O and δ could specifically induce *c-Myc, Nanog, Sox2* and *Klf4*, respectively. Chromatin immune precipitation (ChIP) analysis validated that SAHA–PIP induced gene expression by establishing transcriptionally permissive chromatin modifications in histone H3, including the acetylation of Lys9 and Lys14 and trimethylation of Lys4 [82]. Notably, SAHA-δ at nanomolar concentrations had the potent ability to switch ON endogenous expression of *Oct3/4* and *Nanog* by about 30-fold in mouse fibroblasts. In just 24 h, SAHA-δ overcame the rate-limiting step of the dedifferentiation process by shifting the transcriptional network from the fibroblast state to the mesenchymal–epithelial transition state [83].

Considering the notable effect of SAHA-δ in mouse MEFs, we screened the SAHA–PIPs library (A to Φ) in human dermal fibroblasts (HDFs). Bioinformatic analyses of microarray data validated the remarkable capacity of SAHA–PIPs to trigger transcriptional activation of particular clusters of therapeutically important genes and noncoding RNAs [74]. For example, SAHA-K was identified as the first ever reported small molecule capable of inducing unusual transcriptional activation of germ cell genes in human somatic cells (Fig. 3b) [84]. Interestingly, the SAHA-K-induced *MOV10L1* and *piRNA* factors are essential in maintaining postmeiotic genome integrity and are not expressed in somatic cells. Similarly, SAHA-I displayed distinctive bioactivity and switched ON a core pluripotency network including the miR-302 family [85]. Prolonged incubation with SAHA-I for 21 days generated partially reprogrammed alkaline-phosphatase-positive cells with an induction efficiency of 0.06%. SAHA-X triggered the transcriptional activation of retinal cell-specific genes associated with ocular disorders including *CERKL*, *PAX6*, *RS1*, *USH2A*, *CRYBB3* and *STRA6* [86], while SAHA-L was characterized as the first ever reported multitarget small molecule that epigenetically induced neural system and brain synapse TFs in human foreskin fibroblasts (BJ cells) and iPS cells. It successfully differentiated iPS cells into a neural progenitor state to suggest their potential for acquiring desired neural cell types by targeting the corresponding TFs [76].

To identify the optimal chemical architecture for enhanced bioefficacy, the number of β-alanine linkers in the architecture of SAHA-E was varied, and the presence of three β-alanine linkers moderately increased *OCT4* expression levels [87]. Alterations of the structure of SAHA in SAHA-δ generated a new type of ATFs termed JAHA-δ, (JAHA is a derivative of SAHA lacking its surface-recognition domain). Bioassays demonstrated that JAHA-δ had higher HDAC8 inhibitory activity (~6-fold) than SAHA-δ and activated the skull morphogenesis-governing *Otx2* and *Lhx1* [88]. Similarly, the incorporation of an isophthalic acid (IPA) at the C-terminus of SAHA-δ improved its aqueous solubility and overcame the issue of aggregation. Bind-n-seq analysis provided vital information about sequence bias and aided in the construction of next-generation SAHA–PIPs with improved binding specificity and bioefficacy [89,90]. Overall, these results demonstrate that SAHA–PIPs can be tailored to improve their bioefficacy. While targeting a specific site is a potential strategy to improve potency, it is challenging because it is not straightforward to identify the actual target site in the genome and clarify the mechanism of gene modulation [91]. Future studies are warranted that pay attention to the interpretation of this "black box" by applying an operable assay system [92], in-depth data analysis method and alternative targets with a similar mechanism [80].

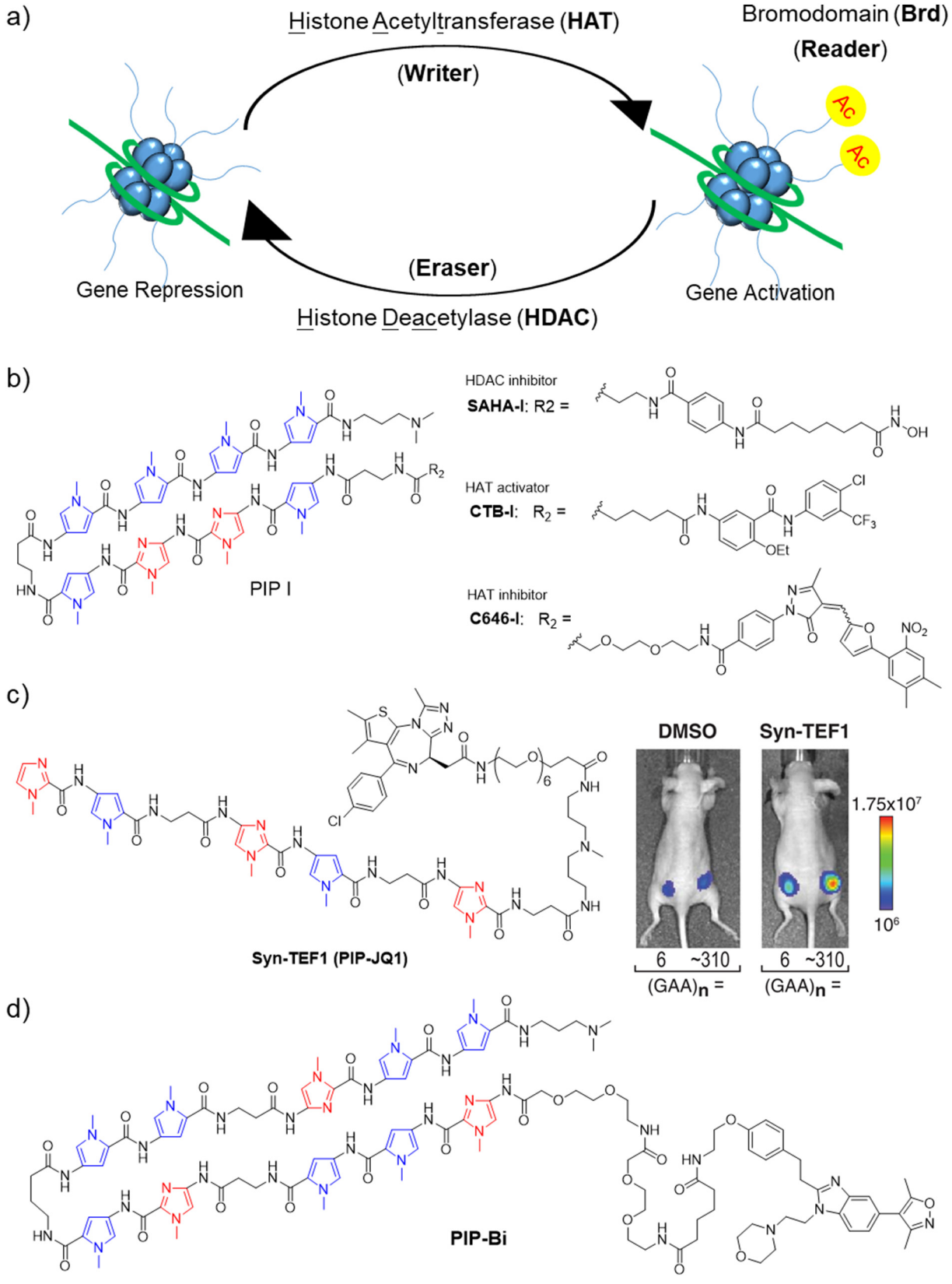
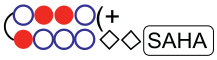
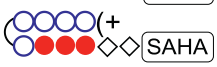
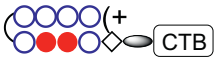
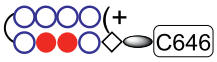


Fig. 3. (a) Schematic illustration of the key process of histone acetylation. (b) Chemical structures of SAHA-I, CTB-I, and C646-I. (c) Chemical structure of PIP–JQ1; bioluminescent images of two representative mice harboring xenografts with six and ~310 GAA repeats after 22 h treatment. The picture is reproduced from [78]. (d) Chemical structure of PIP–Bi.

Table 2
Biological functions of sequence-specific PIP-based Epi-drugs.

Comp.	Structure	Target genes	Regulated gene networks	Ref.
SAHA-K		Germ cell genes	Induce *MOV10L1* and *piRNA* factors	[74]
SAHA-I		Pluripotent genes	Activate *SOX2*, *POU5F1* and miR-302 family	[75]
SAHA-X		Retinal genes	Activate *CERKL*, *PAX6*, *RS1*, *USH2A*, *CRYBB3*, and *STRA6*	[74]
SAHA-L		Neural genes	Induce *NESTIN*, *PAX6*, and *NGN2*	[76]
CTB-I		Pluripotent genes	Activate *SOX2*, *POU5F1* and *NANOG*	[75]
C646-I		P53 targeting genes	Induce *SPATA18*, *EDA2R*, *MDM2*, *BTG2*, and *FDXR*	[77]
PIP-JQ1		*FXN1*	Specifically activate *FXN1* in FRDA cells	[78]
PIP-Bi		Sequence-selective genes	Sequence-specific relevant activation, such as *NTS* gene	[79]
PIP-NCD38		Epigenome changes	Region-selective alteration of epigenetic status	[80]

4.2. PIP–HAT activators and PIP–HAT inhibitors

A more direct way to activate gene(s) of interest than by blocking the epigenetic suppressor HDACs is to create a transcriptionally permissive chromatin architecture by recruiting epigenetic activators such as HATs, "writers" of histone acetylation. However, unlike the widely studied HDAC inhibitors, there is a considerable gap in application of knowledge between HATs function and testified HAT modulators. We substituted the SAHA in SAHA-I with a known HAT activator [*N*-(4-chloro-3-trifluoromethyl-phenyl)-2-ethoxy-benzamide] (CTB) and generated CTB–PIP I (CTB-I) (Fig. 3b). Transcriptome analysis of SAHA-I and CTB-I-induced genes indicated that CTB-I remarkably boosted the bioefficacy of CTB for activating a similar set of genes to those activated by SAHA-I in HDFs. ChIP-Seq analysis revealed that the importance of putative specific sequences in the promoter region was that this was the location where the DNA binding domain of both CTB-I and SAHA-I bound and triggered transcriptionally permissive chromatin marks. This proof-of-concept study indicated that distinct epigenetic modulators could be transformed upon conjugation with a PIP to have the same bioactivity [75]. Integration of CTB, a designer PIP and other functional modules on a nanoparticle resulted in a new type of gene regulation platform termed NanoScript. A NanoScript targeting SOX9, a master TF in chondrogenic differentiation, successfully activated endogenous *SOX9* expression (5-fold at day 7) in adipose-derived mesenchymal stem cells and differentiated them into chondrocytes with enhanced bioefficacy [93]. We also conjugated PIPs with a HAT inhibitor termed C646 to construct sequence-specific HAT inhibitors that could induce antiproliferative and apoptotic activity with IC50 of 1.8–2.6 μM in cancer cell lines (A549 and MV4-11) (Fig. 3b). Microarray analysis and studies in mutant cell lines revealed that this sequence-specific HAT inhibitor operated by the initiation of the tumor suppressor p53-dependent apoptosis pathway [77].

4.3. PIP–Brd inhibitors

As the "reader" of acetylated histones, Brd plays a pivotal role in transcription initiation and elongation and recruiting it to a specific gene locus has demonstrated therapeutic potential. In the recently outstanding study, Ansari group created a sequence-specific synthetic transcription elongation factor 1 (Syn-TEF1) to stimulate the transcription elongation machinery (Fig. 3c). PIP was flexibly tethered to the Brd4 inhibitor JQ1. This Syn-TEF1 (PIP–JQ1) actively enabled transcription across the repressive GAA repeats that silence frataxin expression in Friedreich's ataxia (FRDA) [78]. This novel design defined a promising therapeutic regimen for FRDA patients, a terminal neurodegenerative disease with no effective therapy. More specifically, PIP–JQ1 (1 μM, 24 h) bound to the GAA sites in an FRDA patient-derived cell line to specifically activate *FXN* level (about 100-fold) and restored frataxin level to the level observed in healthy human cells.

Syn-TEF1 works efficiently to accelerate transcription elongation rather than transcription initiation. Recently, a bifunctional molecule PIP–Bi, which was generated by tethering Bi (a p300/CBP-selective Brd inhibitor) to PIPs, recruited p300 to the nucleosomes via its target DNA sequence and markedly accelerated the site-specific acetylation with ~30-fold enhancement *in vitro* (Fig. 3d) [79]. PIP–Bi also induced endogenous expression inside living cells of those genes that have the corresponding cognate DNA sequence. Further works should give direct evidence of transcription initiation by promoter-binding assisted and sequence-specific manner.

4.4. Biomolecule-based ATFs

Until now, it has not been straightforward to use the current PIP-based ATFs and deliberately modulate gene expression in a sequence-specific and target-specific manner. Recently, substantial results with biomolecule-based DNA binding platforms such as dCas9 [94], TALEs [95] and ZNF [96] demonstrated the feasibility of specifically modulating the expression of a target gene and an exclusive set of genes [97]. For example, simple binding of dCas9 in the specific coding region blocked RNAP-mediated elongation to cause a significant decrease in downstream gene expression [98]. Fusion protein dCas9–core p300 catalyzed acetylation of histone H3K27 at its gene locus, leading to robust transcriptional activation of target genes with a ~10,000-fold induction [99]. By targeting a predesigned OCT4 enhancer, TALE–VP64 induced epigenetic changes, reactivated OCT4 expression and substituted for exogenous OCT4 in reprogramming MEFs to induced pluripotent stem cells [100]. Similarly, binding of a single dCas9–Suntag–VP64 system at the SOX2 promoter produced iPS cells in the absence of the conventional four Yamanaka factors [94]. However, the clinical utility of the generated cells was hampered because of barriers including the large molecular size of the complex, the requirement for transfection, and the involvement of genetic materials. Nevertheless, the success of these biomolecule-based tools offers a blueprint for future design and optimization of PIP-based ATFs. Significantly, with substantial optimizations in PIP binding specificity, the PIP-based ATFs could minimize the existing off-target issues of the protein-targeting epi-drugs and improve their bioactivity. PIP-based ATFs could be a promising strategy for

selective and robust epigenetic regulation of gene networks and aid in treating complex incurable disorders like neurodegenerative diseases and cancers.

5. Alkylating PIPs for gene regulation

After the cornerstone human genome project and subsequent clarification of abnormal DNA sequences as potential drug targets in terminal diseases, chemists turned their interest to sequence-specific DNA alkylators targeting longer and more accurate DNA targeting sites. Since our reports two decades ago on describing PIP–duocarmycin (Duo) hybrids [31,101], substantial PIP–alkylator researches have led to the improvements in the precision of this concept. Currently, PIP-alkylating hybrids targeting protein coding regions are progressing toward preclinical research because they can cause specific silencing of oncogenes at nanomolar concentrations [102–104]. Four alkylating agents have been proposed for conjugation with DNA binders, including Duo, Chlorambucil (Chb), *seco*-1-chloromethyl-5-hydroxy-1,2-dihydro-3*H*-benz[e]indole (CBI) and α-bromoacrylamide (Fig. 4a). Here, we detail three critical issues associated with the requirements of PIP–alkylator hybrids.

First, although PIPs function adequately to regulate gene expression by blocking DNA binding of natural TFs [2], some barriers impede the broad application of conventional PIPs. The major barrier is that minor groove-binding PIPs are far less competent to interfere with the elongation function of RNAP in the coding region and with DNA polymerase in

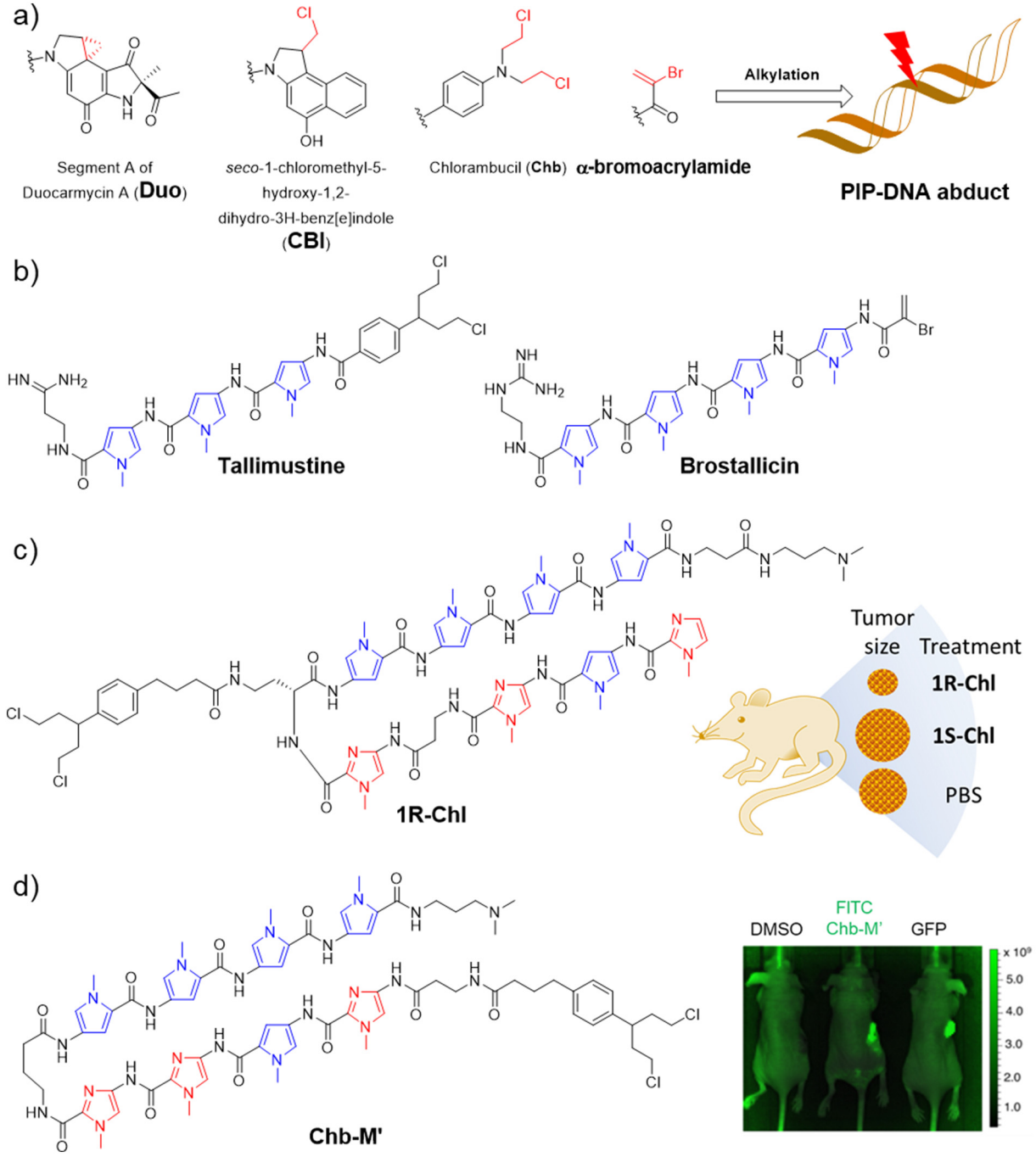


Fig. 4. (a) Chemical structures of four alkylating moieties (in square box, alkylating core modules highlighted in red). (b) Chemical structures of Tallimustine and Brostallicin. (c) Chemical structure of 1R-Chl and representative images of 1R-Chl, 1S-Chl and PBS treated animal with K562 xenograph. (d) Chemical structure of Chb-M′; Live biofluorescence images at 48 hours after treatment with FITC-Chb-M′ in MKN45-transplanted mice (GFP shows tumor sites). These pictures are reproduced from [118] and [104], respectively.

DNA replication [59,105]. By supplementing the PIPs with a DNA-alkylating moiety to produce bifunctional PIP conjugates that attach irreversibly to the DNA minor groove in the coding region could inhibit polymerase-mediated elongation [106–108].

A second key question to consider is the influence of PIPs on the level of alkylation activity, the alkylation efficacy of the monomeric alkylators and whether the conjugates offer a therapeutic advantage over monomeric alkylators. For example, PIP–Chb hybrids exhibit greater alkylating activity and more potent antitumor activity than monomeric Chb, with their IC_{50} decreased from millimolar to micromolar or nanomolar levels [109]. However, monomeric CPI produces extremely high alkylation activity in both cancer and normal cells at picomolar concentrations [110]. The clinical use of monomeric alkylators is limited by their narrow therapeutic window and their conjugation with PIPs resulted in a considerable reduction in their cytotoxicity for normal mammalian cells (~3 orders of magnitude). This suggested that the conjugates have the potential to overcome the toxicity barrier of monomeric alkylators in clinical applications [110].

Finally, PIP is the core component responsible for sequence specificity, and its length and sequence determine the biological targets and functions. Conventional DNA alkylating agents remain among the most common drugs used clinically for cancer treatment. However, they alkylate DNA sites methodically and cannot discriminate abnormal from normal cells, which causes severe toxicity [111]. At the early stage of the study of DNA binder–alkylator conjugates, tallimustine (a tripyrrole–Chb hybrid) [112] and brostallincin (a tetrapyrrole-α-bromoacrylamide conjugate) [113] showed potent anticancer activity and reached phase II clinical trial (Fig. 4b) [114,115]. However, the priority of these two studies was to identify a superior alkylating agent and the DNA binding length and sequence were not examined in detail. More specifically, such hybrids with a short DNA recognition sequence and imprecise DNA binding rules have limited capacity to discriminate between cancer and normal cells [116,117]. Currently, the strategy of producing hairpin PIP–alkylator hybrids has been demonstrated to be a promising tool to tackle this performance issue by endowing the hybrids with higher sequence selectivity, precise DNA binding rules and higher binding affinity [104,118,119]. Therefore, the design of the length and sequence of DNA binders should be considered thoroughly for fulfilling this specific research target.

5.1. PIP–Chb hybrids

Chb and CBI are the two most widely applied alkylators for generating sequence-specific alkylators [109,116,120]. Studies of their alkylating mechanism have shown that hairpin PIP–Chb conjugates operate through sequence-specific alkylation of adenine at N3 in the DNA minor groove and cause interstrand crosslinks (ICL) through their two reactive arms [121]. PIP–Chb covalently binds to both DNA strands and prevents their separation, thereby affecting the vital processes of polymerase-mediated elongation in replication and transcription. Moreover, PIP–Chb conjugates improve DNA alkylation efficiency more than 100-fold compared with that of Chb, suggesting the importance of the efficient and more rapid delivery of Chb alkylating groups to DNA-reactive sites by PIPs [109].

5.1.1. 1R-Chl

The Dervan group did substantial work in the characterization, identification and optimization of hairpin PIP–Chb hybrids. For example, after intensive structure-activity optimization of the turn unit in the hairpin PIPs structure, they concluded that Chb substitution at the site of the (*R*)-α-DABA turn displayed moderate binding affinity and enhanced alkylation specificity with a preference toward a forward hairpin binding conformation under *in vitro* conditions and a cell culture model [122,123]. In a screening of their small library of PIP–Chb hybrids, Dervan and colleagues singled out an interesting compound, 1R–Chl (Fig. 4c). 1R–Chl, which recognizes 5′-WGGWGW-3′ (where W = A or T), caused cells to arrest in the G2/M stage of the cell cycle without apparent cytotoxicity in several normal cell lines (0.5 μM for 4 days) and shown efficacy in preventing tumor growth in several mouse xenograft models (Fig. 4c) [124]. Subsequent studies demonstrated that 1R–Chl alkylated target dsDNA efficiently within the coding region of several histone *H4* gene families, in which it was more significant in a cancer cell line to a greater extent than in a cell line of noncancerous origin [118]. More specifically, their differential effect on the chromatin status, histone acetylation and expression of *H4* genes in cancerous and noncancerous cell lines demonstrated the potential of the application of 1R–Chl [103].

5.1.2. PIP–Chb library construction

Screening a PIP-Chb library is a simple approach to identifying the drugs with the highest potential when there is a specific disease phenotype, a convenient screening method and a clear experimental endpoint. Especially, screening a PIP–Chb library on a plate are expected to be rapid and reliable, but because of its short DNA recognition length, we recognize that it is challenging to clarify the details of the site of alkylation and elucidate a functional mechanism. However, because of the accumulative evidences that these compounds have meaningful therapeutic values, these issues are gradually being addressed with the help of various analytical techniques and “omics” tools.

Encouraged by the results from 1R–Chl and the success in identifying hits from screening PIP–SAHA library, our group started to construct an in-house library of PIP–Chb hybrids in which Chb is covalently attached at the N-terminal of the PIPs instead of at the α-DABA turn. This change enhanced the alkylating activity with a slight decrease in alkylation specificity, improved the efficiency of synthesis and avoided potential conformational complexity [125]. The PIP part recognizes an 8-bp DNA sequence, and the library includes a total of 81 PIPs in every possible combination to cover all possible DNA sequences. The PIP–Chb hybrids in this library are small, with an average molecular weight of 1500 Da, and can be efficiently incorporated into cells without the assistance of a transfection reagent. After the construction of the library, our lab prepared precoated 96-well plates to accelerate the screening efficiency. This combinatorial approach to screening the PIP–Chb library for diverse biological purposes is underway and is expected to accelerate the identification of potential drugs.

5.1.3. Chb–M′

While screening our PIP–HDAC inhibitor library, our laboratory identified that SAHA-M enhanced RUNX expression. Runt-related transcription factor 1 (RUNX1) known as acute myeloid leukemia 1 protein (AML1) is a master transcription factor essential for the differentiation of hematopoietic stem cells. A growing body of evidence suggests that it has pro-oncogenic properties in AML through a RUNX1–p53–core binding factor-β regulatory loop [126]. Such results motivated us to substitute the gene activator with an alkylator. In an *in vitro* assay, Chb–M′ was shown to specifically alkylate A/G of the predesigned consensus RUNX-binding sequence (5′-TGTGGT-3′) (Fig. 4d).

Based on this, we hypothesized that Chb–M′ efficiently blocked RUNX1 target gene expression, and induced the p53-dependent apoptotic pathway. By switching OFF the RUNX cluster, Chb–M′ functioned effectively in AML cell lines at the nanomolar level and against several solid tumors with poor prognosis in a xenograft mouse model of AML, without causing any notable adverse events to the concentration of 10 μM [104]. Consistent with these results, the RUNX1 depletion-mediated antileukemic effect required functional p53, and Chb–M′ lacked growth-suppressive capacity against cancer cells where p53 was mutated or downregulated. The results for cellular uptake and tissue distribution further highlight the bioefficacy of Chb–M′. When fluorescein isothiocyanate-labeled Chb–M′ was administered intravenously *in vivo*, it accumulated at much higher concentrations in tumor cells than in normal tissue cells, potentially creating a therapeutic window and minimizing the side effects (Fig. 4d). Subsequent studies

demonstrated the mechanism of RUNX-mediated inhibition of the engraftment of AML cells: through downregulating E-selectin expression in the vascular niche, Chb–M′ inhibited the engraftment of AML cells in the bone marrow, extending the overall survival of leukemic mice [127].

The antitumor mechanism of Chb–M′ has also been studied in other cancer cell lines, including gastric cancer [128], malignant rhabdoid tumor, acute lymphoblastic leukemia and neuroblastoma, in which the RUNX family plays a pivotal role in tumor survival. For example, in gastric cancer, the ErbB2/HER2 signaling pathway is frequently upregulated and is associated with the maintenance of the cancers, partly through the son of sevenless (SOS) homolog family. RUNX1 interacts with the RUNX1-binding sequence located in the SOS1 promoter and positively regulates its gene expression. Chb–M′, in a similar way as the RUNX1 knockdown, led to the decreased expression of SOS1 and dephosphorylation of ErbB2/HER2, which in turn suppressed the proliferation of gastric cancer cells [128].

5.2. PIP–CBI hybrids

The class of cyclopropylindole (CPI) alkylating agents include natural products ((+)-CC-1065, Duo and Duo-SA) and synthetic derivatives (CI, CBI, CBQ and CPI). To date, CBI has been the most widely studied alkylating agent tethered to PIPs, because it is conveniently synthesized and has enhanced activity. PIP–CBI conjugates predominantly alkylate adenine at N3, causing RNAP blockade, DNA damage and inhibition of replication. The PIP–CBI conjugate is much less cytotoxic (~3-order of magnitude) to mammalian cells than its parent. High-throughput sequencing identified the intact and natural binding sequence preference in a large DNA pool. The results show significant sequence-specific DNA alkylation by PIP–CBI conjugates corresponding to the proposed DNA binding rule, and their targeting pattern is similar in both a full sequence of free DNA and a chromatin-complexed genome [129]. Microarray analysis has been applied to study gene regulatory patterns across the whole genome. Based on screening of a small library of PIP–CBI conjugates, we identified that upregulated and commonly downregulated genes shared by different PIPs typically reflect genes involved in apoptosis, growth arrest and DNA damage. The antitumor activity of PIP–CBI conjugates is attributable to p53-dependent antitumor effects [130].

5.2.1. PIP–CBI hybrids targeting KRAS mutant locus

When targeting gene mutations, a PIP-based alkylating system can be preferable to a gene-editing methodology because its effects can be reversed by withdrawing the compound. The capability of Crispr-cas9 to edit mutant oncogenes is attractive, but the off-target effects carry a high risk and may cause irreversible damage to the human genome. KR12, a PIP–CBI conjugate, selectively recognized oncogenic mutations in codon 12 of *KRAS* (Fig. 5a, b). Around 90% of pancreatic cancers and 50% of colorectal cancers contain a *KRAS* mutation in codon 12 or 13, which results in dysfunction of the downstream MAPK pathway [131,132]. In a thermally induced strand cleavage assay, KR12 was shown to bind the *KRAS* mutation sequence 5′-ACGCCA**W**CA-3′ (W = A/T, T for G12D and A for G12V) with exceptionally high affinity and selectivity (10–15-fold). Bind-n-Seq enabled high-throughput sequencing of PIP binding sites in a pool of DNA sequences and the results revealed that hit-binding sequences matched KRAS mutation sequences with high enrichment (89-fold and 69-fold for G12D and G12V, respectively) [133]. KR12 causes strand cleavage and growth suppression in the human colon cancer cells with G12D or G12V mutations at the concentration of 50 nM for 48 h, thus inducing senescence and apoptosis. Moreover, the treatment significantly downregulated expression of total KRAS and the mutated *KRAS* allele but had no effect in cell lines with wild-type KRAS. In xenograft models in mice, KR12 induced significant tumor growth suppression of LS180 cells (heterozygous G12D/WT) and SW480 cells (homozygous G12V), but not in the wild-type tumor HT19 (Fig. 5b) [119].

Similarly, specific targeting of the *KRAS* codon 13 mutation was also achieved with PIP–CBI conjugates (Fig. 5a). The designed conjugate showed potent specific alkylation activity for the codon 13 mutation containing the 5′-ACGTC**A**CCA-3′ sequence, with a K_D of 2.2×10^{-7} M^{-1} [134]. The current PIP-alkylator approach could also be applied to other mutations driving oncogenesis in human tumors. Moreover, such agents targeting a specific mutation might be a promising approach to overcome drug resistance and side effects in cancer treatments [132].

5.2.2. Other applications

The conjugation of two CBI moieties to PIPs, termed PIP–(*S*)-biCBI conjugates, also allowed the efficient formation of ICL complexes with high DNA sequence selectivity and high reactivity (Fig. 5c). A higher alkylating potency of PIP–(*S*)-biCBI conjugates (~80-fold) than the corresponding (*R*)-form was observed, which can be explained by computational analysis showing that cyclopropane in the (*S*)-form displays proximity and a favorable steric angle to the N3 of adenine [135]. Secondly, photolabile groups were shown to be a useful monitoring system to protect reactive functional groups in drug delivery systems (Fig. 5d) [136]. After protecting a PIP–CBI conjugate with an electron-rich nitrobenzyl group as a photocleavable group, it was activated by UV irradiation for 5 min at 365 nm both *in vitro* and in a cell-based assay [137]. Sequence-specific alkylating PIPs containing photolabile linkers offer a useful approach for developing novel chemical- or enzyme-activated anticancer agents and may facilitate the spatiotemporal control of gene expression. In other applications, PIP–CBI hybrids were demonstrated to penetrate virions efficiently, and such agents would be advantageous compared with conventional CPI drugs because they are expected to possess stronger antiviral activity. PIP–CBI hybrids also show significantly lower cytotoxicity toward uninfected cells than monomeric CPI and therefore provide a possible therapeutic advantage for treating virus infections [138].

5.3. Targeting tandem repeat sequences with PIP–Chb and PIP–CBI hybrids

The telomere, a 5′-TTAGG-3′ repeat at the chromosome terminal, functions as a protective cap for the chromosome, and its length is a critical indicator of aging and tumorigenesis [139]. Two short hairpin PIP–CBI conjugates were shown to specifically alkylate the 3′-A of 5′-ACCC TA-3′ and 5′-AGGGTTA-3′ (5-bp and 7-bp respectively) and displayed potent antiproliferative activity in 39 cancer cell lines with average IC_{50} values of 110.0 nM and 57.5 nM respectively (Fig. 5e) [140]. Interestingly, in the presence of cooperative binding motifs, heterotrimeric PIP–CBI and PIP–Chb conjugates recognized 10- and 11-bp, respectively, in telomere repeats and triggered cooperative alkylation [141,142]. Despite the complexity of their synthesis, PIP–CBI conjugates encompassing tandem hairpin motifs showed higher alkylating efficacy and better affinity toward the 5′-TTAGG-3′ repeat sequence in the telomere regions.

Trinucleotide repeat sequences exist widely in genomic DNA, and the expansion of these repeat sequences often causes neurological disorders involving mechanisms of endogenous protein sequestration and abnormal protein aggregation with a poly Gly/Ala tail [143]. For example, the expansion of CAG/CTG trinucleotide repeats (from 36 to >100 repeats) within the first exon of the *Huntingtin* gene responsible for Huntington disease yields a full-length protein harboring expanded polyglutamine tracts, in which gain of function leads to the selective disruption of neuronal protein function. The dysexpansion of CTG within the 3′-UTR region of the *DMPK* gene generates CUG tracts to sequester MBNL protein, which is responsible for myotonic dystrophy type 1 (DM1) [144]. The strategy of PIP–alkylating hybrids could be applied to target these DNA repeat sequences to downregulate abnormal gene expression. A hairpin PIP–Chb conjugate with an 11-bp recognition sequence demonstrated sequence-specific alkylation of both DNA strands at the N3 of adenine or guanine in CAG/CTG repeats, causing specific

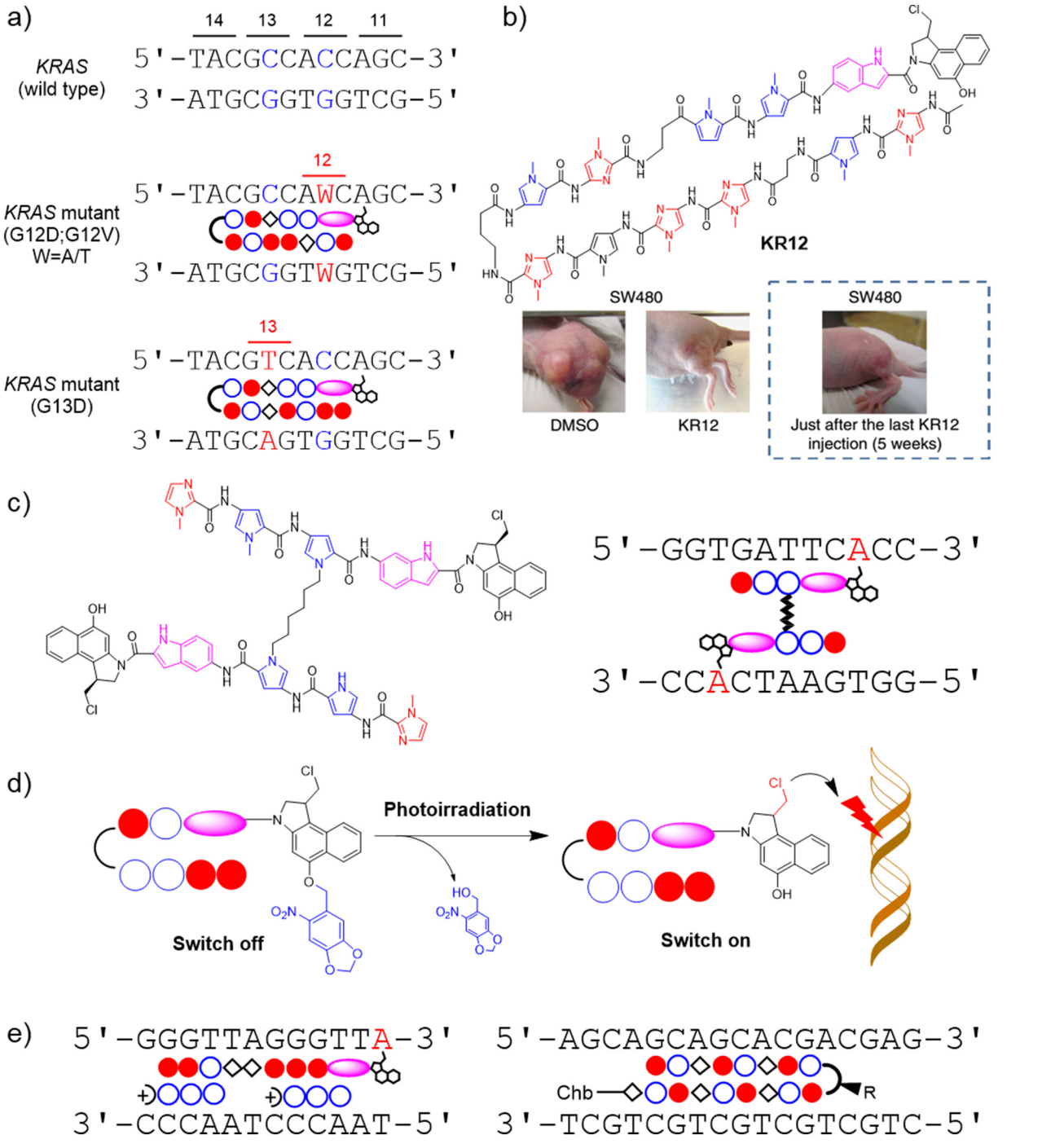


Fig. 5. (a) Schematic illustration of PIP–CBI hybrids targeting KRAS mutants. (b) Chemical structure of KR12; Mice images of SW480 xenograft with KR12 treatment. The picture is reproduced from [119]. (c) Chemical structure of PIP-(*S*)-biCBI and schematic representation of ICL formation. (d) Photo-controllable sequence-specific DNA alkylation with PIP–CBI conjugates. (e) Schematic representation of PIP–alkylator hybrids targeting telomere and CAG repeat sequence respectively.

DNA damage and transcriptional inhibition at the alkylating sites (Fig. 5e) [107]. Further biological evaluation in neural cells and mouse models is underway and is expected to validate their therapeutic usefulness.

6. Cooperative DNA binders for gene regulation

6.1. Covalent conjugates

In the steady progress in achieving higher DNA binding affinity and specificity, the covalent linkage of PIPs with other DNA binders such as small molecule DNA binders, DNA binding peptides, nucleic acid analogues and noncanonical DNA binders, has been extensively explored in an effort to target longer DNA sequences (Fig. 6a).

Small-molecule DNA binding analogues of Hoechst (Ht), a fluorescent dye for DNA, exhibit strong selectivity against A- and T-rich DNA sequences in the DNA minor groove. PIPs supplemented with Ht (PIP–Ht) recognized longer DNA sequences, had enhanced cellular permeability and exhibited appropriate fluorescence properties for easy detection of cellular localization [145]. Linear Py trimer–Ht conjugates bound dsDNA by forming a 1:1 complex with a recognition length of up to 9-bp [146]. Hairpin PIP–Ht conjugates targeting the *Plk1* promoter

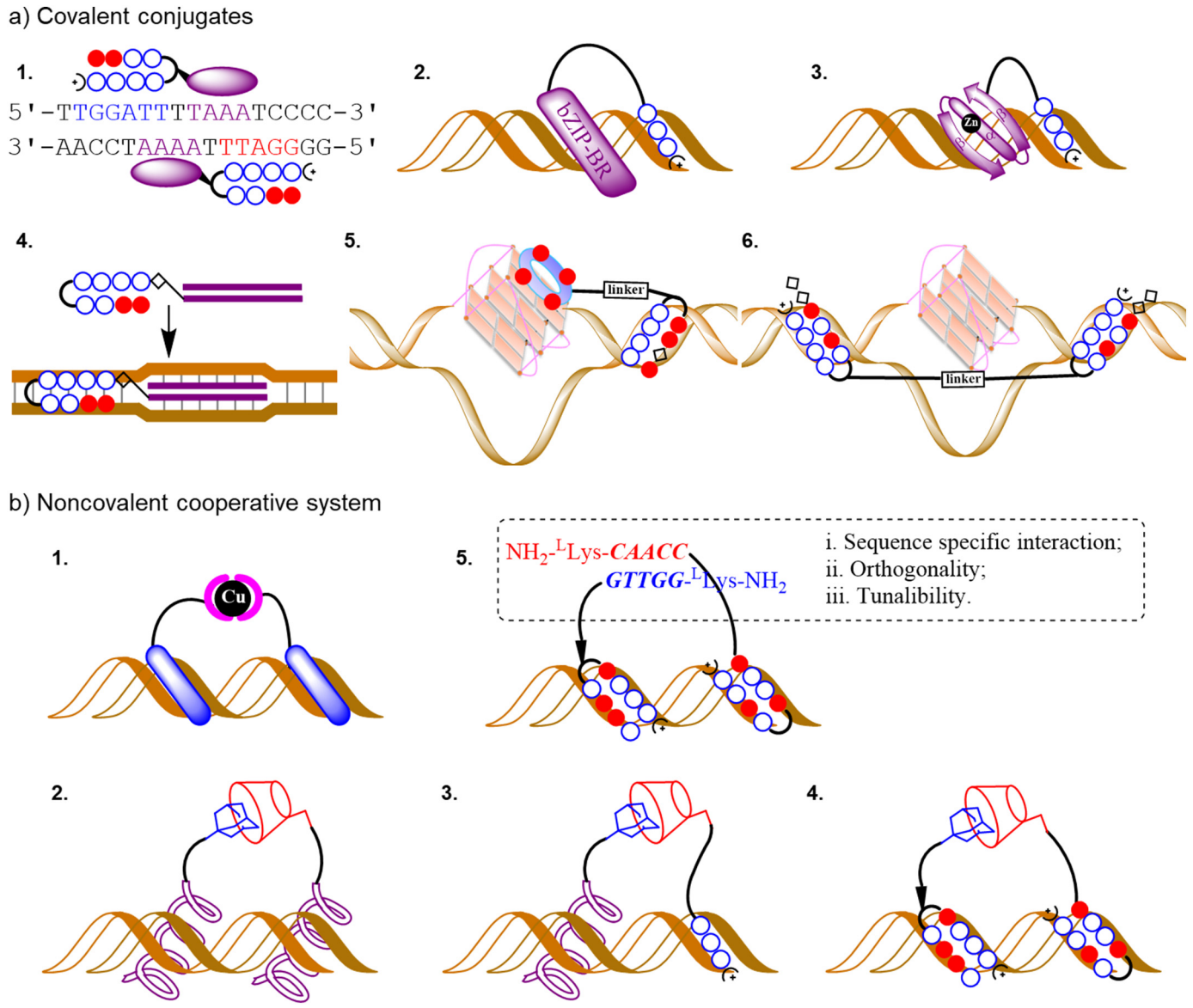


Fig. 6. (a) Schematic illustration of covalent conjugates-based DNA binding modes. (a1) PIP–Ht conjugates targeting *PLK1* promoter. (a2 and a3) PIP–peptide conjugates binding both DNA minor and major grooves simultaneously. (a4) PIP–PNA conjugates. (a5) PIP–G4 binder hybrids binding G-quadruplex (G4) and its proximal dsDNA simultaneously. (a6) G4 induction by a hybrid PIPs binding two adjacent dsDNA sites. (b) Noncovalent cooperative DNA binding modes. (b1) Schematic illustration of DNA binding with DNA minor groove binders supplemented with metal ion-ligand interaction. Host-guest interaction assisted cooperative DNA binding with DNA binding peptides (b2), PIP–peptide conjugates (b3), and PIPs-based Pip–HoGu system (b4). (b5) Schematic illustration of Pip–NaCo system.

specifically downregulated cell cycle-regulated Plk1 expression and consequently retarded tumor cell growth at the concentration of 10 μM for 48 h treatment, while leaving non-cancerous cells undamaged (Fig. 6a1). *In vivo* studies of human cancer xenografts in mice showed that subcutaneous injection of PIP–Ht conjugates led to significant tumor growth suppression with low host toxicity and 50% to 70% reduction in Plk1 expression [147]. Recently, PIP–Ht conjugates were successfully applied to suppress Epstein—Barr virus (EBV) replication. By eradicating the EBV episome in EBV-positive cells, the compound selectively inhibited EBV-positive cell proliferation and slowed down EBV-positive tumor growth [148]. PIP–Ht conjugates currently represent one of the candidates for DNA-based targeted therapeutics. Meanwhile, the meticulous *in vitro* and *in vivo* studies of their target specificity and potential cytotoxicity are warranted to substantiate their therapeutic value.

Although DNA minor groove-binding PIPs can potently disrupt major groove TF binding through allosteric changes to DNA, DNA binders capable of synergistically contacting both minor and major grooves might have additional benefits for intervening TF–DNA interactions [145]. The simple introduction of a protonated substituent of N-propylamine at the PIP backbone affords an electrostatic interaction with the DNA phosphodiester linkages at the interface between minor and major grooves, resulting in an augmentation of three orders of magnitude in the inhibitory activity of major groove TF–DNA interactions [145,149,150]. Mascarenas' group reported PIP–DNA binding peptide conjugates, where distamycin was covalently linked to the DNA binding domains of TFs GCN4 and GAGA (Fig. 6a2, 3). Such conjugates simultaneously binding to DNA major and minor grooves exhibited considerably higher affinity for their designated target DNA sequence than that their respective peptide parts (~60-fold) [151,152]. Oligonucleotide-directed triple helix formation is well-established as a nonbiological approached for sequence-specific DNA recognition and was first utilized to enhance PIP binding affinity, recognition length and sequence specificity [153]. Peptide nucleic acid (PNA) is a well-characterized DNA binder that has high binding affinity through the formation of triplex or double-strand intercalations. In the presence of a

pseudocomplementary (pcPNA) strand, covalent PNA–PIP conjugates (at the concentration of 0.5 μM) were shown to bind and cleave DNA by a process in which PNA bound to DNA through double-duplex strand intercalation (Fig. 6a4). The binding affinity and cleavage efficiency decreased in the absence of a pcPNA strand or truncated conjugates lacking PNA or PIP moieties [154]. However, several key issues, such as chemical synthesis, cellular uptake and biological stability need to be addressed before preclinical studies and their application in cells and *in vivo*.

Recently, PIP-based covalent hybrids were successfully applied to induction and stabilization of DNA G-quadruplex (G4). G4 is a noncanonical nucleic acid secondary structure adopted by guanine-rich sequences, and is believed to regulate various biological and pharmacological processes [155]. Despite the accumulative reports of G4 ligands, the target specificity of an individual G4 is challenging because of the topological similarity of the skeleton of diverse G4s [156]. By utilizing a specific sequence of adjacent duplex DNA, a covalent hybrid molecule generated by tethering cIKP (a G4 binder) to PIP enabled concerted recognition of G4 and its proximal duplex (Fig. 6a5). The synergistic effect of each binding component resulted in higher binding affinity and modest sequence specificity for the *c-myc* promoter [157]. In a parallel study, instead of directly binding G4, an alternative strategy by targeting two local adjacent dsDNA sequences further increased the recognition specificity of the individual G4 in the larger context. A covalent PIP dimer with appropriate linker was demonstrated to have a potent ability to draw the two dsDNAs closer and trigger the conformational collapse of the central G-rich duplex DNA followed by G4 formation with an efficiency of ~31% at the concentration of 1 μM (Fig. 6a6) [158]. As a proof-of-concept, the results highlighted that PIP dimers could be applied to induce G4 structures in native sequences and validated the strategy of using PIP dimers to cause changes in higher-level DNA structures. Although further refinement of the design is necessary to obtain higher potency and specificity, these two strategies established by our group provide an opportunity to control the biological role of individual G4s at specific genome locations. Further studies on validating the functional potency in living cells could substantiate their clinical potential in altering the G4-related cancer genes, such as *c-MYC*, *RAS*, *BCL-2* and *KIT*.

6.2. Cooperative interaction domain (CID)

In prokaryotes, separate binding of individual TFs and recognition of short DNA motifs are sufficient to define the specific function and precise regulation of gene(s). In higher organisms with larger genomes, however, the recognition sequences of individual TFs are too short to define unique genomic expression patterns. Therefore, TFs prefer to operate cooperatively, allowing both definition of unique genomic positions in large genomes, and complex information processing at the level of individual regulatory elements [159].

6.2.1. Natural TF pair systems

The observation that TF binding occurs in dense clusters in higher organisms suggests that most TFs coordinate widespread cooperative interactions. Indeed, among 1000 commonly expressed TFs in mammal cells, 50–70% of them work cooperatively through direct or indirect TF–TF interactions, which allow specific manipulation of downstream gene expression [159]. Moreover, substantial studies have established that, in addition to classic TF–TF structural interactions, DNA-facilitated TF–TF cooperation and DNA-mediated TF cooperation are equally prominent [160].

To mimic or synergistically inhibit the TF pair–DNA interaction, three major points must be considered. First, the TF pair–DNA interaction contains two TF-binding sites in which the binding sequences are predetermined. However, gap sequences are not conserved, making the conventional strategy of covalent conjugation of two individual PIPs ineffective for targeting a TF pair [161]. Second, the gap distances are flexible and usually vary from –1 to 5 bp [159]. Lastly, the most important point to consider concerns the use of two separate short PIPs for targeting the two binding sites. However, it is now known that if a TF pair partner is displaced, the TFs might generate divergent biological functions. For example, the SOX2–OCT4 pair functions to maintain pluripotent gene circuits in stem cells. However, when SOX2 changes its partner to PAX6, they are functionally responsible for lens development [162,163].

Therefore, to manipulate spatiotemporal gene expression in high resolution, conventional PIP design through a covalent linker could not achieve the purpose of intervening in TF pair–DNA interactions, suggesting the demand of new strategies. Several reported cooperative systems might be helpful in this context, including host-guest systems, nucleic acid analogues and others (Fig. 6b) [164–167].

6.2.2. Metal ion–ligand interaction systems

The complexing ability of metal ion–ligand interactions could be an approch to achieve noncovalent, cooperative DNA binding. Schepartz's group first reported Fe^{2+}-assisted peptide–DNA binding affinity and orientation by finely controlling the peptide conjugation site [168]. Similarly, Ihara's group reported that the Cu^{2+}-mediated L-glutamic acid interaction significantly stabilized polypurine and polypyridine triplex DNA formation with an almost 165-fold increase in binding affinity [169]. Sasaki's group applied Cu^{2+}-bipyridine complexes with a small molecule DNA binder system and demonstrated their efficacy in the dimeric binder and multimeric formats (Fig. 6b1) [170,171]. Recently, multivalent interactions of Mg^{2+}-assisted chromomycin A3 were shown to potently protect DNA from endonuclease cleavage upon the formation of multiple binding complexes at the concentration of 10–50 μM [172]. These studies have not yet progressed to cellular applications. Future studies on a cooperative PIP system assisted by metal ion–ligand interactions need to consider taking advantage of the natural abundance of metal ions (such as K^+ and Mg^{2+}) in the nucleus and cell-specific metal ions (such as Ca^{2+} and Fe^{2+}) to maximize cumulative cooperativity via multivalent interactions.

6.2.3. Host–guest systems

Host-guest systems (e.g., cyclodextrin (Cyd), cucurbit[n]uril, and carcerands with respective guests) have been widely applied *in vitro* and in cell-based systems [164,173]. Among these host–guest systems, Cyd–adamantane (Ada) has been extensively studied [174,175]. By replacing a leucine-zipper dimerization domain with Cyd/Ada, Morii's group designed an artificial system in which the cooperative Cyd–Ada interaction strongly stabilized the interaction of DNA with the DNA binding peptide of GCN4 homodimer (Fig. 6b2) [176,177]. Mascarenas' group generated DNA binding peptide–distamycin conjugates by noncovalently binding Cyd–Ada modules, which represents a step forward in the development of smaller, selective and ligand-responsive systems (Fig. 6b3) [178].

Recently, we reported the development of PIPs conjugated to Host-Guest assemblies, termed Pip–HoGu, which mimic the cooperation between natural TF pairs (Fig. 6b4) [179]. By incorporating Cyd and Ada separately, Ada1 (PIP1–Ada) and Cyd1 (PIP2–Cyd) were designed and synthesized. The results consistently demonstrated that the Pip–HoGu system formed stable, noncovalent, cooperative complexes, thereby meeting the critical criteria to mimic a natural TF pair. The chemical architecture also encompassed a longer recognition sequence (two PIP-binding intervals with gap distance), favorable sequence selectivity, higher binding affinity and importantly, a flexible gap distance. For example, they showed thermal stability of 7.2 °C and a minimum free energy of interaction of $-2.32\ kcal \cdot mol^{-1}$ with a targeting length of 14-bp. Significantly, cell-based evaluation validated the capacity of Pip–HoGu to exhibit potent cooperative inhibitory effects on gene expression under physiological conditions by disrupting TF pair–DNA function [43]. It is conceivable that the modular design of Pip–HoGu could

serve as a proof-of-concept and define a general framework for mimicking naturally occurring cooperative TF pair–DNA interactions.

Future studies could define the nature of several features that may underline the efficacy of the Pip–HoGu system. For example, the system is not practical for spacing sequences >5 nucleotides and there is a need for another host–guest system with higher interacting ability to overcome this limitation. In addition, the cooperative binding energy of the host-guest system could not be finely tuned independently. Essentially, the interactions between host-guest moieties are electrostatic and hydrophobic, rather than residue specific [180].

6.2.4. Nucleic acid-based systems

Unlike the host-guest system, the interaction between nucleic acids is sequence specific and can be finely modulated. The addition of discrete dimerization domains enhanced the cooperativity of oligonucleotide-directed triple helix formation. Devan's group were the first to report the binding properties of oligonucleotides that cause dimerization via Watson–Crick hydrogen bonds and bind neighboring sites on double-helical DNA by triple helix formation [181]. An outward-facing 8-mer dimerization domain can enhance binding to 11-bp and neighboring 15-bp triplex strands with cooperative energy of binding of –2.3 kcal·mol^{-1}, and the overall binding affinity increased by 10–15-fold in the presence of a partner strand [181,182]. Similarly, Winssinger's group reported PNA-based templated reactions, which proved to have higher efficiency because the PNA duplex has higher binding affinity than natural nucleic acids [183].

In addition to the importance of sequence-specific interactions, nucleic acid-based CIDs should avoid interference with natural DNA and RNA, i.e., have bio-orthogonality. More specifically, the recognition partner of a CID must be limited to binding its corresponding artificial nucleic acids, but not interact with endogenous natural nucleic acids. Recently, our group expanded the cooperation partners to include left-handed (LH) γPNA, i.e., PIPs conjugated with a nucleic acid-based cooperation system (Pip–NaCo) (Fig. 6b5) [184]. The γPNA duplex is parallel to dsDNA. LH γPNA was chosen because of its bio-orthogonality, and its high binding affinity and sequence specificity for the partner strand [185]. The cooperativity of its concerted DNA recognition is highly comparable to that of the natural system, with a minimum energetics of cooperation of –3.26 kcal mol^{-1} [184]. Moreover, the vertical binding mode, in which γPNA duplex is vertical to dsDNA, demonstrated further enhancement of cooperativity. Furthermore, by changing the linker conjugation site, binding mode, and γPNA sequence and length, the cooperative energetics of Pip–NaCo can be tuned independently and rationally. The current orthogonal Pip–NaCo platform has the potential to provide triple to multiple heterobinding systems to target further complicated natural TF cluster networks.

Taken together with metal ion-assisted- and Pip-HoGu cooperative system, those cooperative DNA binding systems provide novel approaches in regenerative medicine to precisely manipulate gene expression machinery of biological processes in stem cell research and neuron differentiation.

7. Other applications

PIP-based DNA binding domain system has been widely explored for various applications, and several excellent reviews have summarized them [186,187]. Here, we highlight two emerging research fields including DNA visualization probes and mitochondrial targeting PIPs, which may have the therapeutic potential for clinical applications in the future.

7.1. In vitro visualization of specific DNA sequences

In addition to their therapeutic potential, PIPs could also be programmed to be harnessed as bioimaging probes. By conjugating fluorescent molecules to PIPs, it is possible to label DNA in a sequence-specific manner. TAMRA–polypyrrole, reported in 2018, is a compound designed to label A/T-rich sequences in DNA (Fig. 7a) [188]. The advantage of TAMRA–polypyrrole is that there is no fluorophore-mediated photocleavage and no structural deformation, which can be a problem with YOYO-1, because TAMRA–polypyrrole does not intercalate into DNA. Furthermore, while popular DNA staining reagents such as 4′,6-diamidino-2-phenylindole (DAPI) and Hoechst need ultraviolet light to be excited, TAMRA–polypyrrole is excited at a longer wavelength (~550 nm), which reduces damage to DNA during observation. When extended DNA fragments fixed to the nanochannel surface of a flow cell in flowing buffer were stained with TAMRA–polypyrrole, each fragment showed a specific staining pattern depending on its A/T frequency. By comparing this staining pattern with an *in silico* map generated from the original DNA sequence, it is possible to identify the position of each fragment in the whole genome. This gives a simple method to determine the sequence of DNA fragments using only microscopic observation when the sequence of the original DNA is available.

7.2. Tandem-type PIPs for telomere visualization

PIP-based fluorescent probes can also be applied to the visualization of specific DNA sequences in cells. Fluorescence *in situ* hybridization (FISH) is one method to achieve DNA sequence visualization using oligonucleotide probes, but this technique requires harsh conditions to denature the dsDNA and anneal the probes to their target sequence [189]. Because PIPs can bind to DNA without denaturing its structure, DNA can be labeled with PIPs under mild conditions.

Telomeres are located at the ends of chromosomes and play an essential role in the stability of the chromosome and DNA replication. Human telomeres have repetitive sequences of 5′-$(TTAGGG)_n$-3′ and contain a duplex region and a single-stranded 3′ overhang. To target the repetitive sequence, short PIP units that recognize one repeat sequence can be used to develop tandem-type PIPs to recognize longer sequences. The production of fluorescent-labeled tandem-type PIPs that contain two, three and four PIP units targeting telomeres have been reported [23,190–192]. Of these, a tandem tetramer that can recognize 24-bp within the telomere repeat sequence showed the best signal-to-noise ratios in fluorescence imaging of telomeric foci at the concentration of 75 nM (Fig. 7b) [23]. Hence, a longer recognition sequence may bestow better selectivity for the target sequence. These PIP probes can stain telomeres in fixed cells and tissue sections after a simple incubation at room temperature, and there is no need to introduce foreign DNA to express a fusion protein to label telomeres [193]. Future work could explore live cell imaging to study chromosomal dynamics using cell-uptake reagents or covalently incorporating cell-penetrating tags.

7.3. PIPs targeting mitochondrial DNA

Mitochondria possess their own DNA (mtDNA), and some of the mutations in mtDNA are known to cause mitochondrial diseases [194]. A promising approach to treat these diseases is to use PIPs, because PIPs can control transcription or replication based on DNA sequence. However, default accumulation of PIPs in nuclei can be a problem for this purpose. To overcome this, a PIP conjugated to mitochondrial penetrating peptide (MITO–PIP) was developed (Fig. 7c) [48,195]. The peptide contains hydrophobic cyclohexylalanine residues to reduce the energy barrier during mitochondrial membrane penetration and basic arginine residues that generate the driving force to penetrate the mitochondrial membrane in a membrane potential-dependent manner. Fluorescence microscopy showed that TAMRA-labeled MITO–PIP accumulated in mitochondria, and MITO–PIP targeting the binding sequence of mitochondrial transcription factor A (TFAM) in the light-strand promoter (LSP) region of mtDNA repressed expression of *ND6* (90% inhibition at the concentration of 10 μM), a gene downstream of LSP, in HeLa cells. Hence, the introduction of a simple peptide can switch the target of PIPs from nuclei to mitochondria without impairing its sequence

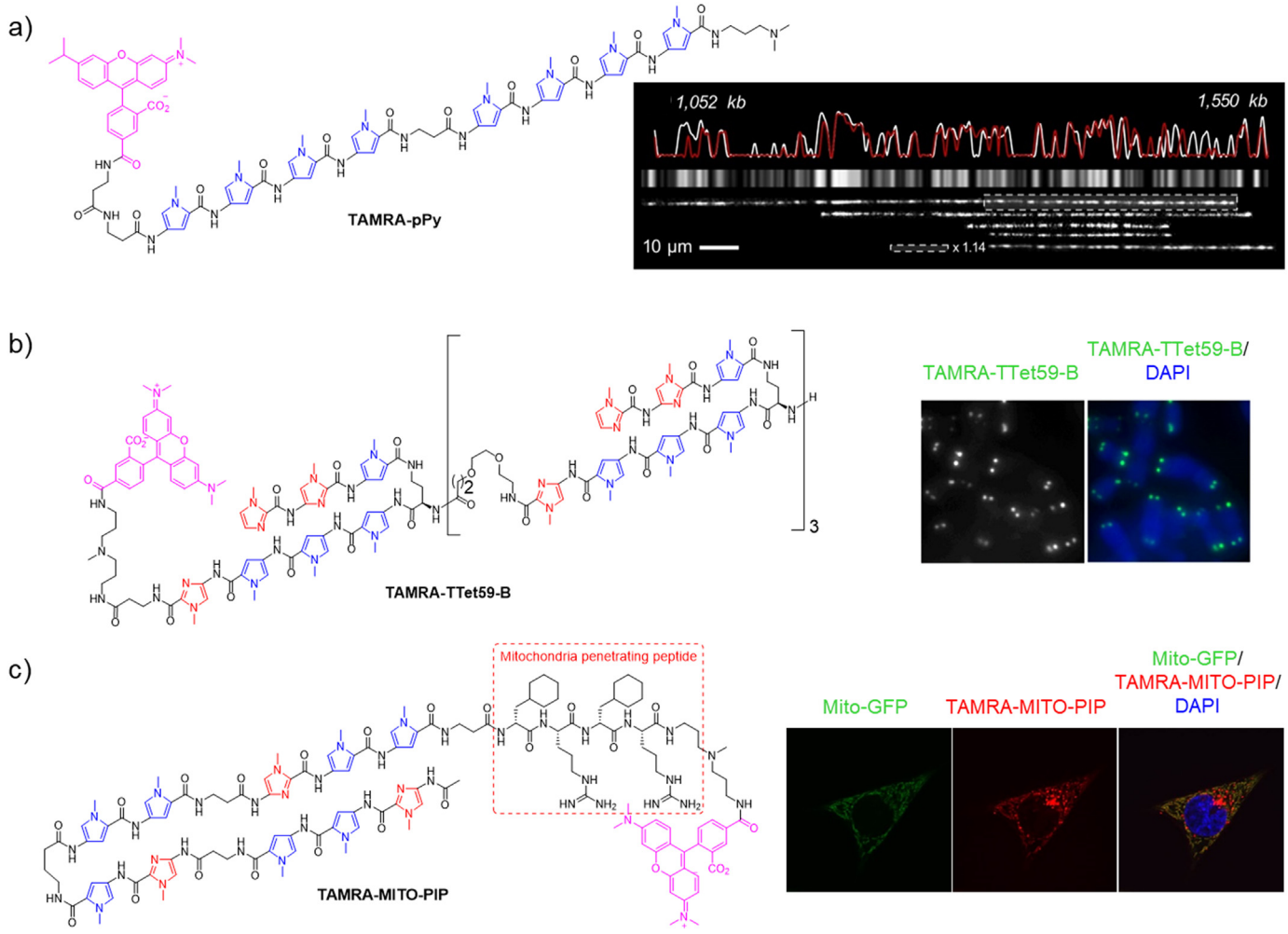


Fig. 7. (a) Chemical structure of TAMRA-pPy and sequence alignment of 498 kb bacterial genomic fragment. (b) Chemical structure of tandem tetramer TAMRA-TTet59-B and telomere staining with TAMRA-TTet59-B (green). Nuclei were visualized by DAPI (blue). (c) TAMRA-MITO-PIP structure and mitochondrial colocalization image of Mito-GFP (green) and TAMRA-MITO-PIP (red). These pictures are reproduced from [188], [23] and [48], respectively.

selectivity. The MITO-PIP study provides a fresh platform to alter either a normal or a mutated mitochondrial genome in a sequence-specific manner. The tunability of PIPs bestows a huge potential to this strategy to treat rare mitochondrial diseases.

8. Challenges and perspectives

PIPs can enter live mammalian cells, localize inside the nucleus, read their cognate DNA sequences and alter basal transcription with no constraint from transfection agents [196]. Hence, they remain at the forefront of promising small molecule-based designer drugs. While PIP-based systems show potential as sophisticated drugs, there is a need to perform substantial characterizations and optimizations to overcome their shortcomings.

To stimulate the global changes to the transcriptome required to change cell fate, it may be advantageous that the designed factor can more closely mimic natural TFs. While binding to one specific genome site increaser selectivity, this effect can be weakened in perturbing local transcriptional status and further diminished in transcriptome modulation. Therefore, PIP-based ATFs recognizing many sites in the genome while maintaining a binding affinity that rivals natural TFs are promising candidates for inducing changes in cell fate.

Short PIPs with a short recognition DNA sequence are synthesized easily. Short PIPs and its conjugates could be widely evaluated either using a library screening method or by their targeting of repetitive DNA short sequences. Extensive library screenings are expected to identify potential chemical agents with novel functional mechanisms [83]. For example, after several rounds of screening of a 6-bp binding PIP–SAHA library, we identified certain PIP–SAHA conjugates capable of activating the pluripotent gene network in somatic cells.

However, to achieve higher specificity with a low rate of off-target effects during the predesign of PIPs, a longer target length is necessary. Even though the molecular weight of PIPs is smaller than those of nucleic acids and biomolecules, cellular uptake is still a key issue for their application as pharmaceutical agents. More specifically, design of PIPs capable of reading more than 10-bp could have a detrimental effect on their cell uptake efficiency [36,190]. Dervan's group proposed an approach to resolving this problem by applying in situ click chemistry: two short PIPs can form a covalent bond when they bind to adjacent sites, although there is an issue with low reaction efficiency [197]. Cooperative systems such as Pip–HoGu and Pip–NaCo showed huge potential for *in vitro* and cell-based assays, and further optimization would let them feasible for versatile application [179,184]. The incorporation of an IPA moiety at the C-terminal and γ-turn substituents led to significant enhancement of cell uptake and biological activity [27,198]. The incorporation of NLS, CPP and methoxypolyethylene glycol (PEG) is a good option for the enhancement of cell uptake, while the attachment of mitochondrial penetrating moiety shifts the localization to mitochondria. Moreover, PIPs aggregation caused a dramatic loss in solubility and cellular uptake, and this issue could be addressed by the addition of carbohydrate aggregate-solubilizing agents, especially 2-hydroxypropyl-β-cyclodextrin [26]. Cell-uptake enhancing reagents such as Endo-porter

and liposome operated efficiently to transport neutrally-charged long PIPs and the relevant conjugates [79,130,179].

Since several PIP-based therapeutics exhibits clinical potential, such as Chb-M' in AML and PIP-JQ1 in Friedreich's ataxia A, extensive preclinical researches and useful research tools are necessary [78,104]. A minor modification such as ^{14}C- and ^{18}F-labeled radioactive PIPs enabled *in vivo* observation of localization and convenient determination of parameters in ADMET studies in animals [199,200]. Toxicity and side-effects always impede the further development of clinical therapy, and PIP-originated toxicity and off-effects should be closely examined. Most of the DNA minor groove binders are substrates of P-glycoprotein. Resistance to DNA binders is often related to transcriptional mechanisms and DNA repair pathways, particularly defects in transcription-coupled nucleotide excision repair [201].

The strategy of precise gene manipulation with small molecule-based, locus-specific gene regulatory hybrids has been pursued for decades. Since our first report on DNA sequence-selective gene activator in cell assays a decade ago, steadfast progress has been made where PIPs conferred sequence specificity to various epigenetic modulators like HDAC inhibitor, HAT activator and inhibitor, bromodomain inhibitor and histone demethylase inhibitor. However, the current emerging platform should recognize their intrinsic limitations originating from both PIP part and functional groups, thereby warranting the need for novel strategies. For example, targeting repeat sequences could theoretically increase the success rate. Both longer PIPs and more potent epidrugs should be substantially exploited. A rational approach to the choice of modulating moiety i.e., structural optimization of PIPs or functional group, is needed. A convenient cell-based evaluation system is highly urged to accelerate such research [92]. Furthermore, the recent achievements of gene-specific modulation by dCas9, TALEs and ZNFs fused to functional protein demonstrate the possibility of using the abovementioned strategy in DNA-based small molecules [94]. Future research on PIP-based ATFs should also take cues from the dCas9-, TALEs- and ZNFs-based systems.

The employment of TFs to regulate gene networks and reprogram cell states has had increasing success [202]. However, recognizing a minimal set of natural TFs driving the cell fate alterations remain labor intensive and challenging. Modern emerging techniques in ATF design have advanced upon decades of research in natural DNA-binding proteins, allowing them to be custom tailored for a variety of purposes. Nonprotein-based methods such as the PIP-based systems can be rationally designed to act as ATFs allowing for temporal regulation of genes without the introduction of genetic material. Further intensive optimization and new designs of DNA binding domains, FD and CIDs of ATFs are highly required before they can be used clinically as the drugs. Together, the rational and knowledge-based design could aid in improving and developing the PIPs as drugs for disease treatment.

Acknowledgements

This work was supported by JSPS [16H06356 to H.S.]; Suzuken Memorial Foundation [16-077 to G.N.P.]; Kyoto University SPIRITS grants [to G.N.P.].

References

[1] J.W. Trauger, E.E. Baird, P.B. Dervan, Recognition of DNA by designed ligands at subnanomolar concentrations, Nature 382 (1996) 559–561.
[2] J.M. Gottesfeld, L. Neely, J.W. Trauger, et al., Regulation of gene expression by small molecules, Nature 387 (1997) 202–205.
[3] P.B. Dervan, Molecular recognition of DNA by small molecules, Bioorg. Med. Chem. 9 (2001) 2215–2235.
[4] A. Jolma, J. Yan, T. Whitington, et al., DNA-binding specificities of human transcription factors, Cell 152 (2013) 327–339.
[5] J. Taniguchi, G.N. Pandian, T. Hidaka, et al., A synthetic DNA-binding inhibitor of SOX2 guides human induced pluripotent stem cells to differentiate into mesoderm, Nucleic Acids Res. 45 (2017) 9219–9228.
[6] A.C. Finlay, F.A. Hochstein, B.A. Sobin, et al., Netropsin, a new antibiotic produced by a streptomyces, J. Am. Chem. Soc. 73 (1951) 341–343.
[7] F. Arcamone, S. Penco, P. Orezzi, et al., Structure and synthesis of distamycin A, Nature 203 (1964) 1064.
[8] M.L. Kopka, C. Yoon, D. Goodsell, et al., The molecular origin of DNA-drug specificity in netropsin and distamycin, Proc. Natl. Acad. Sci. U.S.A. 82 (1985) 1376–1380.
[9] M.W. Van Dyke, R.P. Hertzberg, P.B. Dervan, Map of distamycin, netropsin, and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA.Fe(II), Proc. Natl. Acad. Sci. U.S.A. 79 (1982) 5470–5474.
[10] J.G. Pelton, D.E. Wemmer, Structural characterization of a 2:1 distamycin A.d(CGCA AATTGGC) complex by two-dimensional NMR, Proc. Natl. Acad. Sci. U.S.A. 86 (1989) 5723–5727.
[11] W.S. Wade, M. Mrksich, P.B. Dervan, Binding affinities of synthetic peptides, pyridine-2-carboxamidonetropsin and 1-methylimidazole-2-carboxamidonetropsin, that form 2:1 complexes in the minor groove of double-helical DNA, Biochemistry 32 (1993) 11385–11389.
[12] M. Mrksich, M.E. Parks, P.B. Dervan, Hairpin peptide motif. A new class of oligopeptides for sequence-specific recognition in the minor groove of double-helical DNA, J. Am. Chem. Soc. 116 (1994) 7983–7988.
[13] C.L. Kielkopf, E.E. Baird, P.B. Dervan, et al., Structural basis for G.C recognition in the DNA minor groove, Nat. Struct. Biol. 5 (1998) 104–109.
[14] J.M. Turner, S.E. Swalley, E.E. Baird, et al., Aliphatic/aromatic amino acid pairings for polyamide recognition in the minor groove of DNA, J. Am. Chem. Soc. 120 (1998) 6219–6226.
[15] S. White, J.W. Szewczyk, J.M. Turner, et al., Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands, Nature 391 (1998) 468–471.
[16] G. Padroni, J.A. Parkinson, K.R. Fox, et al., Structural basis of DNA duplex distortion induced by thiazole-containing hairpin polyamides, Nucleic Acids Res. (2017) 1211.
[17] A. Iguchi, N. Fukuda, T. Takahashi, et al., RNA binding properties of novel gene silencing pyrrole-imidazole polyamides, Biol. Pharm. Bull. 36 (2013) 1152–1158.
[18] E.E. Baird, P.B. Dervan, Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids, J. Am. Chem. Soc. 118 (1996) 6141–6146.
[19] N.R. Wurtz, J.M. Turner, E.E. Baird, et al., Fmoc solid phase synthesis of polyamides containing pyrrole and imidazole amino acids, Org. Lett. 3 (2001) 1201–1203.
[20] J. Cho, M.E. Parks, P.B. Dervan, Cyclic polyamides for recognition in the minor groove of DNA, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 10389–10392.
[21] M. Mrksich, P.B. Dervan, Design of a covalent peptide heterodimer for sequence-specific recognition in the minor groove of double-helical DNA, J. Am. Chem. Soc. 116 (1994) 3663–3664.
[22] A. Heckel, P.B. Dervan, U-pin polyamide motif for recognition of the DNA minor groove, Chem. Eur. J. 9 (2003) 3353–3366.
[23] Y. Kawamoto, A. Sasaki, A. Chandran, et al., Targeting 24 bp within telomere repeat sequences with tandem tetramer pyrrole–imidazole polyamide probes, J. Am. Chem. Soc. 138 (2016) 14100–14107.
[24] A.R. Urbach, P.B. Dervan, Toward rules for 1:1 polyamide:DNA recognition, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 4343–4348.
[25] C. Dose, M.E. Farkas, D.M. Chenoweth, et al., Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit, J. Am. Chem. Soc. 130 (2008) 6859–6866.
[26] A.E. Hargrove, J.A. Raskatov, J.L. Meier, et al., Characterization and solubilization of pyrrole-imidazole polyamide aggregates, J. Med. Chem. 55 (2012) 5425–5432.
[27] J.L. Meier, D.C. Montgomery, P.B. Dervan, Enhancing the cellular uptake of Py-Im polyamides through next-generation aryl turns, Nucleic Acids Res. 40 (2012) 2345–2356.
[28] F. Yang, N.G. Nickols, B.C. Li, et al., Animal toxicity of hairpin pyrrole-imidazole polyamides varies with the turn unit, J. Med. Chem. 56 (2013) 7449–7457.
[29] A.K. Mapp, A.Z. Ansari, M. Ptashne, et al., Activation of gene expression by small molecule transcription factors, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 3930–3935.
[30] A. Ohtsuki, M.T. Kimura, M. Minoshima, et al., Synthesis and properties of PI polyamide–SAHA conjugate, Tetrahedron Lett. 50 (2009) 7288–7292.
[31] Z.-F. Tao, T. Fujiwara, I. Saito, et al., Rational design of sequence-specific DNA alkylating agents based on duocarmycin A and pyrrole−imidazole hairpin polyamides, J. Am. Chem. Soc. 121 (1999) 4961–4967.
[32] L.A. Dickinson, J.W. Trauger, E.E. Baird, et al., Inhibition of Ets-1 DNA binding and ternary complex formation between Ets-1, NF-kappaB, and DNA by a designed DNA-binding ligand, J. Biol. Chem. 274 (1999) 12765–12773.
[33] L.A. Dickinson, J.W. Trauger, E.E. Baird, et al., Anti-repression of RNA polymerase II transcription by pyrrole-imidazole polyamides, Biochemistry 38 (1999) 10801–10807.
[34] N.J. Marini, R. Baliga, M.J. Taylor, et al., DNA binding hairpin polyamides with antifungal activity, Chem. Biol. 10 (2003) 635–644.
[35] D.M. Chenoweth, P.B. Dervan, Structural basis for cyclic Py-Im polyamide allosteric inhibition of nuclear receptor binding, J. Am. Chem. Soc. 132 (2010) 14521–14529.
[36] T.P. Best, B.S. Edelson, N.G. Nickols, et al., Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 12063–12068.
[37] A.A. Kurmis, F. Yang, T.R. Welch, et al., A pyrrole-imidazole polyamide is active against enzalutamide-resistant prostate cancer, Cancer Res. 77 (2017) 2207–2212.
[38] A.E. Hargrove, T.F. Martinez, A.A. Hare, et al., Tumor repression of VCaP xenografts by a pyrrole-imidazole polyamide, PLoS One 10 (2015), e0143161.
[39] J.O. Szablowski, J.A. Raskatov, P.B. Dervan, An HRE-binding Py-Im polyamide impairs hypoxic signaling in tumors, Mol. Cancer Ther. 15 (2016) 608–617.
[40] V.S. Mysore, J. Szablowski, P.B. Dervan, et al., A DNA-binding molecule targeting the adaptive hypoxic response in multiple myeloma has potent antitumor activity, Mol. Cancer Res. 14 (2016) 253–266.

[41] B.Z. Olenyuk, G.J. Zhang, J.M. Klco, et al., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 16768–16773.
[42] N.G. Nickols, C.S. Jacobs, M.E. Farkas, et al., Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface, ACS Chem. Biol. 2 (2007) 561–571.
[43] N.G. Nickols, J.O. Szablowski, A.E. Hargrove, et al., Activity of a Py-Im polyamide targeted to the estrogen response element, Mol. Cancer Ther. 12 (2013) 675–684.
[44] J.A. Raskatov, J.L. Meier, J.W. Puckett, et al., Modulation of NF-kappaB-dependent gene transcription using programmable DNA minor groove binders, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 1023–1028.
[45] Y. Zhang, G. Sicot, X. Cui, et al., Targeting a DNA binding motif of the EVI1 protein by a pyrrole-imidazole polyamide, Biochemistry 50 (2011) 10431–10441.
[46] K.A. Muzikar, N.G. Nickols, P.B. Dervan, Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 16598–16603.
[47] Y. Wei, G.N. Pandian, Z. Yu, et al., Synthetic DNA-binding inhibitor of HES1 alters the notch signaling pathway and induces neuronal differentiation, ACS Omega 3 (2018) 3608–3616.
[48] T. Hidaka, G.N. Pandian, J. Taniguchi, et al., Creation of a synthetic ligand for mitochondrial DNA sequence recognition and promoter-specific transcription suppression, J. Am. Chem. Soc. 139 (2017) 8444–8447.
[49] J. Syed, G.N. Pandian, S. Sato, et al., Targeted suppression of EVI1 oncogene expression by sequence-specific pyrrole-imidazole polyamide, Chem. Biol. 21 (2014) 1370–1380.
[50] M. Taniguchi, K. Fujiwara, Y. Nakai, et al., Inhibition of malignant phenotypes of human osteosarcoma cells by a gene silencer, a pyrrole–imidazole polyamide, which targets an E-box motif, FEBS Open Bio. 4 (2014) 328–334.
[51] J. Igarashi, N. Fukuda, T. Inoue, et al., Preclinical study of novel gene silencer pyrrole-imidazole polyamide targeting human TGF-beta1 promoter for hypertrophic scars in a common marmoset primate model, PLoS One 10 (2015), e0125295.
[52] T. Kojima, X. Wang, K. Fujiwara, et al., Inhibition of human osteosarcoma cell migration and invasion by a gene silencer, pyrrole-imidazole polyamide, targeted at the human MMP9 NF-kappaB binding site, Biol. Pharm. Bull. 37 (2014) 1460–1465.
[53] T. Takahashi, Y. Asami, E. Kitamura, et al., Development of pyrrole-imidazole polyamide for specific regulation of human aurora kinase-A and -B gene expression, Chem. Biol. 15 (2008) 829–841.
[54] S.Y. Chiang, R.W. Burli, C.C. Benz, et al., Targeting the ets binding site of the HER2/neu promoter with pyrrole-imidazole polyamides, J. Biol. Chem. 275 (2000) 24246–24254.
[55] C.S. Jacobs, P.B. Dervan, Modifications at the C-terminus to improve pyrrole-imidazole polyamide activity in cell culture, J. Med. Chem. 52 (2009) 7380–7388.
[56] D.M. Chenoweth, D.A. Harki, J.W. Phillips, et al., Cyclic pyrrole-imidazole polyamides targeted to the androgen response element, J. Am. Chem. Soc. 131 (2009) 7182–7188.
[57] S. Diaz-Perez, N. Kane, A.A. Kurmis, et al., Interference with DNA repair after ionizing radiation by a pyrrole-imidazole polyamide, PLOS ONE 13 (2018), e0196803.
[58] T.F. Martinez, J.W. Phillips, K.K. Karanja, et al., Replication stress by Py-Im polyamides induces a non-canonical ATR-dependent checkpoint response, Nucleic Acids Res. 42 (2014) 11546–11559.
[59] F. Yang, N.G. Nickols, B.C. Li, et al., Antitumor activity of a pyrrole-imidazole polyamide, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 1863–1868.
[60] T. Ueno, N. Fukuda, A. Tsunemi, et al., A novel gene silencer, pyrrole-imidazole polyamide targeting human lectin-like oxidized low-density lipoprotein receptor-1 gene improves endothelial cell function, J. Hypertens. 27 (2009) 508–516.
[61] T. Suzuki, Y. Asami, T. Takahashi, et al., Development of a molecule-recognized promoter DNA sequence for inhibition of HER2 expression, J. Antibiot. 62 (2009) 339–341.
[62] H. Matsuda, N. Fukuda, T. Ueno, et al., Transcriptional inhibition of progressive renal disease by gene silencing pyrrole-imidazole polyamide targeting of the transforming growth factor-beta1 promoter, Kidney Int. 79 (2011) 46–56.
[63] A. Yasuda, K. Noguchi, M. Minoshima, et al., DNA ligand designed to antagonize EBNA1 represses Epstein-Barr virus-induced immortalization, Cancer Sci. 102 (2011) 2221–2230.
[64] X. Wang, H. Nagase, T. Watanabe, et al., Inhibition of MMP-9 transcription and suppression of tumor metastasis by pyrrole-imidazole polyamide, Cancer Sci. 101 (2010) 759–766.
[65] T. Nagashima, T. Aoyama, T. Yokoe, et al., Pharmacokinetic modeling and prediction of plasma pyrrole-imidazole polyamide concentration in rats using simultaneous urinary and biliary excretion data, Biol. Pharm. Bull. 32 (2009) 921–927.
[66] T.W. Synold, B. Xi, J. Wu, et al., Single-dose pharmacokinetic and toxicity analysis of pyrrole-imidazole polyamides in mice, Cancer Chemother. Pharmacol. 70 (2012) 617–625.
[67] M. Minoshima, T. Bando, S. Sasaki, et al., Pyrrole-imidazole hairpin polyamides with high affinity at 5'-CGCG-3' DNA sequence; influence of cytosine methylation on binding, Nucleic Acids Res. 36 (2008) 2889–2894.
[68] P.W. Burridge, E. Matsa, P. Shukla, et al., Chemically defined generation of human cardiomyocytes, Nat. Meth. 11 (2014) 855–860.
[69] A. Eguchi, Garrett O. Lee, F. Wan, et al., Controlling gene networks and cell fate with precision-targeted DNA-binding proteins and small-molecule-based genome readers, Biochem. J. 462 (2014) 397–413.
[70] A.K. Mapp, A.Z. Ansari, A TAD further: exogenous control of gene activation, ACS Chem. Biol. 2 (2007) 62–75.
[71] Y. Amamoto, Y. Aoi, N. Nagashima, et al., Synthetic posttranslational modifications: chemical catalyst-driven regioselective histone acylation of native chromatin, J. Am. Chem. Soc. 139 (2017) 7568–7576.
[72] A.P. Feinberg, M.A. Koldobskiy, A. Gondor, Epigenetic modulators, modifiers and mediators in cancer aetiology and progression, Nat. Rev. Genet. 17 (2016) 284–299.
[73] G.N. Pandian, J. Taniguchi, H. Sugiyama, Cellular reprogramming for pancreatic β-cell regeneration: clinical potential of small molecule control, Clin. Transl. Med. 3 (2014) 6.
[74] G.N. Pandian, J. Taniguchi, S. Junetha, et al., Distinct DNA-based epigenetic switches trigger transcriptional activation of silent genes in human dermal fibroblasts, Sci. Rep. 4 (2014), e3843.
[75] L. Han, G.N. Pandian, A. Chandran, et al., A synthetic DNA-binding domain guides distinct chromatin-modifying small molecules to activate an identical gene network, Angew. Chem. Int. Ed. Engl. 54 (2015) 8700–8703.
[76] Y. Wei, G.N. Pandian, T. Zou, et al., A multi-target small molecule for targeted transcriptional activation of therapeutically significant nervous system genes, ChemistryOpen 5 (2016) 517–521.
[77] Z. Yu, J. Taniguchi, Y. Wei, et al., Antiproliferative and apoptotic activities of sequence-specific histone acetyltransferase inhibitors, Eur. J. Med. Chem. 138 (2017) 320–327.
[78] G.S. Erwin, M.P. Grieshop, A. Ali, et al., Synthetic transcription elongation factors license transcription across repressive chromatin, Science 358 (2017) 1617–1622.
[79] J. Taniguchi, Y. Feng, G.N. Pandian, et al., Biomimetic artificial epigenetic code for targeted acetylation of histones, J. Am. Chem. Soc. 140 (2018) 7108–7115.
[80] K. Alagarswamy, K.I. Shinohara, S. Takayanagi, et al., Region-specific alteration of histone modification by LSD1 inhibitor conjugated with pyrrole-imidazole polyamide, Oncotarget 9 (2018) 29316–29335.
[81] P. Hou, Y. Li, X. Zhang, et al., Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds, Science 341 (2013) 651–654.
[82] G.N. Pandian, K. Shinohara, A. Ohtsuki, et al., Synthetic small molecules for epigenetic activation of pluripotency genes in mouse embryonic fibroblasts, Chembiochem 12 (2011) 2822–2828.
[83] G.N. Pandian, Y. Nakano, S. Sato, et al., A synthetic small molecule for rapid induction of multiple pluripotency genes in mouse embryonic fibroblasts, Sci. Rep. 2 (2012) e544.
[84] L. Han, G.N. Pandian, S. Junetha, et al., A synthetic small molecule for targeted transcriptional activation of germ cell genes in a human somatic cell, Angew. Chem. Int. Ed. 52 (2013) 13410–13413.
[85] G.N. Pandian, S. Sato, C. Anandhakumar, et al., Identification of a small molecule that turns ON the pluripotency gene circuitry in human fibroblasts, ACS Chem. Biol. 9 (2014) 2729–2736.
[86] J. Syed, A. Chandran, G.N. Pandian, et al., A synthetic transcriptional activator of genes associated with the retina in human dermal fibroblasts, Chembiochem 16 (2015) 1497–1501.
[87] G.N. Pandian, A. Ohtsuki, T. Bando, et al., Development of programmable small DNA-binding molecules with epigenetic activity for induction of core pluripotency genes, Bioorg. Med. Chem. 20 (2012) 2656–2660.
[88] A. Saha, G.N. Pandian, S. Sato, et al., Synthesis and biological evaluation of a targeted DNA-binding transcriptional activator with HDAC8 inhibitory activity, Bioorg. Med. Chem. 21 (2013) 4201–4209.
[89] C. Anandhakumar, Y. Li, S. Kizaki, et al., Next-generation sequencing studies guide the design of pyrrole-imidazole polyamides with improved binding specificity by the addition of beta-alanine, Chembiochem 15 (2014) 2647–2651.
[90] C. Anandhakumar, S. Kizaki, T. Bando, et al., Advancing small-molecule-based chemical biology with next-generation sequencing technologies, Chembiochem 16 (2015) 20–38.
[91] G.N. Pandian, H. Sugiyama, Strategies to modulate heritable epigenetic defects in cellular machinery: lessons from nature, Pharmaceuticals 6 (2012) 1–24.
[92] K.V. Butler, A.M. Chiarella, J. Jin, et al., Targeted gene repression using novel bifunctional molecules to harness endogenous histone deacetylation activity, ACS Synth. Biol. 7 (2018) 38–45.
[93] S. Patel, T. Pongkulapa, P.T. Yin, et al., Integrating epigenetic modulators into NanoScript for enhanced chondrogenesis of stem cells, J. Am. Chem. Soc. 137 (2015) 4598–4601.
[94] P. Liu, M. Chen, Y. Liu, et al., CRISPR-based chromatin remodeling of the endogenous Oct4 or Sox2 locus enables reprogramming to pluripotency, Cell Stem Cell 22 (2018) 252–261.
[95] M.L. Maeder, J.F. Angstman, M.E. Richardson, et al., Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins, Nat. Biotechnol. 31 (2013) 1137–1142.
[96] A. Eguchi, M.J. Wleklinski, M.C. Spurgat, et al., Reprogramming cell fate with a genome-scale library of artificial transcription factors, Proc. Natl. Acad. Sci. U.S.A. 113 (2016) 8257–8266.
[97] E.A. Heiderscheit, A. Eguchi, M.C. Spurgat, et al., Reprogramming cell fate with artificial transcription factors, FEBS Lett. 592 (2018) 888–900.
[98] L.S. Qi, M.H. Larson, L.A. Gilbert, J.A. Doudna, J.S. Weissman, A.P. Arkin, W.A. Lim, Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression, Cell 152 (2013) 1173–1183.
[99] I.B. Hilton, A.M. D'Ippolito, C.M. Vockley, et al., Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers, Nat. Biotechnol. 33 (2015) 510.
[100] X. Gao, J. Yang, J.C.H. Tsang, et al., Reprogramming to pluripotency using designer tale transcription factors targeting enhancers, Stem Cell Reports 1 (2013) 183–197.
[101] Z.-F. Tao, T. Fujiwara, I. Saito, et al., Sequence-specific DNA alkylation by hybrid molecules between segment A of duocarmycin A and pyrrole/imidazole diamide, Angew. Chem. Int. Ed. Engl. 38 (1999) 650–653.

[102] T. Bando, H. Sugiyama, Synthesis and biological properties of sequence-specific DNA-alkylating pyrrole-imidazole polyamides, Acc. Chem. Res. 39 (2006) 935–944.
[103] C. Jespersen, E. Soragni, C. James Chou, et al., Chromatin structure determines accessibility of a hairpin polyamide-chlorambucil conjugate at histone H4 genes in pancreatic cancer cells, Bioorganic Med. Chem. Lett. 22 (2012) 4068–4071.
[104] K. Morita, K. Suzuki, S. Maeda, et al., Genetic regulation of the RUNX transcription factor family has antitumor effects, J. Clin. Invest. 127 (2017) 2815–2828.
[105] J. Xu, I. Lahiri, W. Wang, et al., Structural basis for the initiation of eukaryotic transcription-coupled DNA repair, Nature 551 (2017) 653–657.
[106] T. Oyoshi, W. Kawakami, A. Narita, et al., Inhibition of transcription at a coding sequence by alkylating polyamide, J. Am. Chem. Soc. 125 (2003) 4752–4754.
[107] S. Asamitsu, Y. Kawamoto, F. Hashiya, et al., Sequence-specific DNA alkylation and transcriptional inhibition by long-chain hairpin pyrrole–imidazole polyamide–chlorambucil conjugates targeting CAG/CTG trinucleotide repeats, Bioorg. Med. Chem. 22 (2014) 4646–4657.
[108] K. Shinohara, S. Sasaki, M. Minoshima, et al., Alkylation of template strand of coding region causes effective gene silencing, Nucleic Acids Res. 34 (2006) 1189–1195.
[109] N.R. Wurtz, P.B. Dervan, Sequence specific alkylation of DNA by hairpin pyrrole-imidazole polyamide conjugates, Chem. Biol. 7 (2000) 153–161.
[110] K. Shinohara, T. Bando, H. Sugiyama, Anticancer activities of alkylating pyrrole-imidazole polyamides with specific sequence recognition, Anticancer Drugs 21 (2010) 228–242.
[111] B.A. Chabner, T.G. Roberts Jr., Chemotherapy and the war on cancer, Nat. Rev. Cancer 5 (2005) 65–72.
[112] F.M. Arcamone, F. Animati, B. Barbieri, et al., Synthesis, DNA-binding properties, and antitumor activity of novel distamycin derivatives, J. Med. Chem. 32 (1989) 774–778.
[113] A. Fedier, C. Fowst, J. Tursi, et al., Brostallicin (PNU-166196)—a new DNA minor groove binder that retains sensitivity in DNA mismatch repair-deficient tumour cells, Br. J. Cancer 89 (2003) 1559–1565.
[114] J. Viallet, D. Stewart, F. Shepherd, et al., Tallimustine is inactive in patients with previously treated small cell lung cancer. A phase II trial of the National Cancer Institute of Canada Clinical Trials Group, Lung Cancer 15 (1996) 367–373.
[115] H. Gelderblom, J.Y. Blay, B.M. Seddon, et al., Brostallicin versus doxorubicin as first-line chemotherapy in patients with advanced or metastatic soft tissue sarcoma: an European Organisation for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group randomised phase II and pharmacogenetic study, Eur. J. Cancer 50 (2014) 388–396.
[116] G. Pezzoni, M. Grandi, G. Biasoli, et al., Biological profile of FCE 24517, a novel benzoyl mustard analogue of distamycin A, Br. J. Cancer 64 (1991) 1047.
[117] C. Geroni, S. Marchini, P. Cozzi, et al., Brostallicin, a novel anticancer agent whose activity is enhanced upon binding to glutathione, Cancer Res. 62 (2002) 2332–2336.
[118] C.J. Chou, M.E. Farkas, S.M. Tsai, et al., Small molecules targeting histone H4 as potential therapeutics for chronic myelogenous leukemia, Mol. Cancer Ther. 7 (2008) 769–778.
[119] K. Hiraoka, T. Inoue, R.D. Taylor, et al., Inhibition of KRAS codon 12 mutants using a novel DNA-alkylating pyrrole-imidazole polyamide conjugate, Nat. Commun. 6 (2015) 6706.
[120] A.Y. Chang, P.B. Dervan, Strand selective cleavage of DNA by diastereomers of hairpin polyamide-seco-CBI conjugates, J. Am. Chem. Soc. 122 (2000) 4856–4864.
[121] Y.D. Wang, J. Dziegielewski, N.R. Wurtz, et al., DNA crosslinking and biological activity of a hairpin polyamide-chlorambucil conjugate, Nucleic Acids Res. 31 (2003) 1208–1215.
[122] M.E. Farkas, S.M. Tsai, P.B. Dervan, Alpha-diaminobutyric acid-linked hairpin polyamides, Bioorg. Med. Chem. 15 (2007) 6927–6936.
[123] S.M. Tsai, M.E. Farkas, C.J. Chou, et al., Unanticipated differences between alpha- and gamma-diaminobutyric acid-linked hairpin polyamide-alkylator conjugates, Nucleic Acids Res. 35 (2007) 307–316.
[124] L.A. Dickinson, R. Burnett, C. Melander, et al., Arresting cancer proliferation by small-molecule gene regulation, Chem. Biol. 11 (2004) 1583–1594.
[125] M. Minoshima, T. Bando, K. Shinohara, et al., Comparative analysis of DNA alkylation by conjugates between pyrrole-imidazole hairpin polyamides and chlorambucil or seco-CBI, Bioorg. Med. Chem. 18 (2010) 1236–1243.
[126] K. Morita, M. Noura, C. Tokushige, et al., Autonomous feedback loop of RUNX1-p53-CBFB in acute myeloid leukemia cells, Sci. Rep. 7 (2017), 16604.
[127] K. Morita, C. Tokushige, S. Maeda, et al., RUNX transcription factors potentially control E-selectin expression in the bone marrow vascular niche in mice, Blood Adv. 2 (2018) 509–515.
[128] Y. Mitsuda, K. Morita, G. Kashiwazaki, et al., RUNX1 positively regulates the ErbB2/HER2 signaling pathway through modulating SOS1 expression in gastric cancer cells, Sci. Rep. 8 (2018) 6423.
[129] A. Chandran, J. Syed, R.D. Taylor, et al., Deciphering the genomic targets of alkylating polyamide conjugates using high-throughput sequencing, Nucleic Acids Res. 44 (2016) 4014–4024.
[130] G. Kashiwazaki, T. Bando, T. Yoshidome, et al., Synthesis and biological properties of highly sequence-specific-AlkylatingN-methylpyrrole–N-methylimidazole polyamide conjugates, J. Med. Chem. 55 (2012) 2057–2066.
[131] C. Ambrogio, J. Kohler, Z.W. Zhou, et al., KRAS dimerization impacts MEK inhibitor sensitivity and oncogenic activity of mutant KRAS, Cell, 172 e815 (2018) 857–868.
[132] Z. Yu, S. Ye, G. Hu, et al., The RAF-MEK-ERK pathway: targeting ERK to overcome obstacles to effective cancer therapy, Future Med. Chem. 7 (2015) 269–289.
[133] R.D. Taylor, A. Chandran, G. Kashiwazaki, et al., Selective targeting of the KRAS codon 12 mutation sequence by pyrrole-imidazole polyamide seco-CBI conjugates, Chemistry 21 (2015) 14996–15003.
[134] R.D. Taylor, S. Asamitsu, T. Takenaka, et al., Sequence-specific DNA alkylation targeting for Kras codon 13 mutation by pyrrole-imidazole polyamide seco-CBI conjugates, Chem. Eur. J. 20 (2014) 1310–1317.
[135] C. Guo, S. Asamitsu, G. Kashiwazaki, et al., DNA interstrand crosslinks by H-pin polyamide (S)-seco-CBI conjugates, ChemBioChem 18 (2017) 166–170.
[136] E. Pazos, J. Mosquera, M.E. Vázquez, et al., DNA recognition by synthetic constructs, ChemBioChem 12 (2011) 1958–1973.
[137] S. Park, T. Bando, K. Shinohara, et al., Photocontrollable sequence-specific DNA alkylation by a pyrrole-imidazole polyamide seco-CBI conjugate, Bioconjugate Chem. 22 (2011) 120–124.
[138] B.J. Philips, A.Y. Chang, P.B. Dervan, et al., DNA damage effects of a polyamide-CBI conjugate in SV40 virions, Mol. Pharmacol. 67 (2005) 877–882.
[139] E. Varela, M.A. Muñoz-Lorente, A.M. Tejera, et al., Generation of mice with longer and better preserved telomeres in the absence of genetic manipulations, Nat. Commun. 7 (2016), 11739.
[140] S. Sasaki, T. Bando, M. Minoshima, et al., Sequence-specific alkylation of double-strand human telomere repeat sequence by pyrrole-imidazole polyamides with indole linkers, J. Am. Chem. Soc. 128 (2006) 12162–12168.
[141] G. Kashiwazaki, T. Bando, K. Shinohara, et al., Cooperative alkylation of double-strand human telomere repeat sequences by PI polyamides with 11-base-pair recognition based on a heterotrimeric design, Bioorg. Med. Chem. 17 (2009) 1393–1397.
[142] G. Kashiwazaki, T. Bando, K. Shinohara, et al., Alkylation of a human telomere sequence by heterotrimeric chlorambucil PI polyamide conjugates, Bioorg. Med. Chem. 18 (2010) 2887–2893.
[143] S.M. Mirkin, Expandable DNA repeats and human disease, Nature 447 (2007) 932–940.
[144] B.H.M. Mooers, J.S. Logue, J.A. Berglund, The structural basis of myotonic dystrophy from the crystal structure of CUG repeats, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 16626–16631.
[145] C.M. White, A.L. Satz, T.C. Bruice, et al., Inhibition of transcription factor-DNA complexes and gene expression by a microgonotropen, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 10590–10595.
[146] P.M. Reddy, P.T. Jindra, A.L. Satz, et al., Sequence selective recognition in the minor groove of dsDNA by pyrrole, imidazole-substituted bis-benzimidazole conjugates, J. Am. Chem. Soc. 125 (2003) 7843–7848.
[147] K. Liu, L. Fang, H. Sun, et al., Targeting polo-like kinase 1 by a Novel pyrrole-imidazole polyamide-Hoechst conjugate suppresses tumor growth in vivo, Mol. Cancer Ther. 17 (2018) 988–1002.
[148] Z. Cheng, W. Wang, C. Wu, et al., Novel pyrrole-imidazole polyamide Hoechst conjugate suppresses epstein-barr virus replication and virus-positive tumor growth, J. Med. Chem. 61 (2018) 6674–6684.
[149] K.A. Browne, G.X. He, T.C. Bruice, Chemistry of phosphodiesters, DNA and models. 4. Microgonotropens and their interactions with DNA. 2. Quantitative evaluation of equilibrium constants for 1:1 and 2:1 binding of dien-microgonotropen-a, -b, and -c as well as distamycin and Hoechst 33258 to d(GGCGCAAATTTGGCGG)/d(CCGCCAAATTTGCGCC), J. Am. Chem. Soc. 115 (1993) 7072–7079.
[150] S. Chavda, B. Babu, P. Patil, et al., Design, synthesis, and DNA binding characteristics of a group of orthogonally positioned diamino, N-formamido, pyrrole- and imidazole-containing polyamides, Bioorg. Med. Chem. 21 (2013) 3907–3918.
[151] M.E. Vázquez, A.M. Caamaño, J. Martínez-Costas, et al., Design and synthesis of a peptide that binds specific DNA sequences through simultaneous interaction in the major and in the minor groove, Angew. Chem. Int. Ed. Engl. 40 (2001) 4723–4725.
[152] V. Olalla, V.M. Eugenio, J.B. Blanco, et al., Specific DNA recognition by a synthetic, monomeric Cys2His2 zinc-finger peptide conjugated to a minor-groove binder, Angew. Chem. Int. Ed. Engl. 46 (2007) 6886–6890.
[153] J.W. Szewczyk, E.E. Baird, P.B. Dervan, Sequence-specific recognition of DNA by a major and minor groove binding ligand, Angew. Chem. Int. Ed. Engl. 35 (1996) 1487–1489.
[154] W. Kameshima, T. Ishizuka, M. Minoshima, et al., Conjugation of peptide nucleic acid with a pyrrole/imidazole polyamide to specifically recognize and cleave DNA, Angew. Chem. Int. Ed. Engl. 52 (2013) 13681–13684.
[155] G. Biffi, D. Tannahill, J. McCafferty, et al., Quantitative visualization of DNA G-quadruplex structures in human cells, Nat. Chem. 5 (2013) 182–186.
[156] S. Asamitsu, Z. Obata, T. Yu, H. Bando, Sugiyama, Recent Progress of Targeted G-Quadruplex-Preferred Ligands Toward Cancer Therapy, Molecules 24 (2019) 429.
[157] S. Asamitsu, S. Obata, A.T. Phan, et al., Simultaneous binding of hybrid molecules constructed with dual DNA-binding components to a G-quadruplex and its proximal duplex, Chem. Eur. J. 24 (2018) 4428–4435.
[158] S. Obata, S. Asamitsu, K. Hashiya, et al., G-quadruplex induction by the hairpin pyrrole-imidazole polyamide dimer, Biochemistry 57 (2018) 498–502.
[159] A. Jolma, Y. Yin, K.R. Nitta, et al., DNA-dependent formation of transcription factor pairs alters their binding specificity, Nature 527 (2015) 384–388.
[160] B. Deplancke, D. Alpern, V. Gardeux, The genetics of transcription factor DNA binding variation, Cell 166 (2016) 538–554.
[161] A. Jolma, J. Yan, T. Whitington, et al., DNA-binding specificities of human transcription factors, Cell 152 (2013) 327–339.
[162] N. Tapia, C. MacCarthy, D. Esch, et al., Dissecting the role of distinct OCT4-SOX2 heterodimer configurations in pluripotency, Sci. Rep. 5 (2015), 13533.
[163] Y. Kamachi, M. Uchikawa, A. Tanouchi, et al., Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development, Genes Dev. 15 (2001) 1272–1286.
[164] G. Yu, K. Jie, F. Huang, Supramolecular amphiphiles based on host–guest molecular recognition motifs, Chem. Rev. 115 (2015) 7240–7303.

[165] Y. Azuma, M. Imanishi, T. Yoshimura, et al., Cobalt(II)-responsive DNA binding of a GCN4-bZIP protein containing cysteine residues functionalized with iminodiacetic acid, Angew. Chem. Int. Ed. Engl. 121 (2009) 6985–6988.
[166] T. Ihara, A. Uemura, A. Futamura, et al., Cooperative DNA probing using a β-cyclodextrin—DNA conjugate and a nucleobase-specific fluorescent ligand, J. Am. Chem. Soc. 131 (2009) 1386–1387.
[167] X. Zhou, X. Su, P. Pathak, et al., Host–guest tethered DNA transducer: ATP fueled release of a protein inhibitor from Cucurbit[7]uril, J. Am. Chem. Soc. 139 (2017) 13916–13921.
[168] B. Cuenoud, A. Schepartz, Altered specificity of DNA-binding proteins with transition metal dimerization domains, Science 259 (1993) 510–513.
[169] T. Ihara, Y. Takeda, A. Jyo, Metal ion-directed cooperative triple helix formation of glutamic acid—oligonucleotide conjugate, J. Am. Chem. Soc. 123 (2001) 1772–1773.
[170] M. Tanada, S. Tsujita, T. Kataoka, et al., Cu2+-mediated assembly of the minor groove binders on the DNA template with sequence selectivity, Org. Lett. 8 (2006) 2475–2478.
[171] H. Koda, J.A. Brazier, I. Onishi, et al., Strong positive cooperativity in binding to the A3T3 repeat by Hoechst 33258 derivatives attaching the quinoline units at the end of a branched linker, Bioorganic Med. Chem. 23 (2015) 4583–4590.
[172] H. Murase, T. Noguchi, S. Sasaki, Evaluation of simultaneous binding of Chromomycin A3 to the multiple sites of DNA by the new restriction enzyme assay, Bioorganic Med. Chem. 28 (2018) 1832–1835.
[173] B. Vinciguerra, L. Cao, J.R. Cannon, et al., Synthesis and self-assembly processes of monofunctionalized Cucurbit[7]uril, J. Am. Chem. Soc. 134 (2012) 13133–13140.
[174] J. Lai, B.P. Shah, E. Garfunkel, et al., Versatile fluorescence resonance energy transfer-based mesoporous silica nanoparticles for real-time monitoring of drug release, ACS Nano 7 (2013) 2741–2750.
[175] J. Zhang, P.X. Ma, Cyclodextrin-based supramolecular systems for drug delivery: recent progress and future perspective, Adv. Drug Deliv. Rev. 65 (2013) 1215–1233.
[176] M. Ueno, A. Murakami, K. Makino, et al., Arranging quaternary structure of peptides by cyclodextrin-guest inclusion complex: sequence-specific DNA binding by a peptide dimer with artificial dimerization module, J. Am. Chem. Soc. 115 (1993) 12575–12576.
[177] Y. Aizawa, Y. Sugiura, M. Ueno, et al., Stability of the dimerization domain effects the cooperative DNA binding of short peptides, Biochemistry 38 (1999) 4008–4017.
[178] J.B. Blanco, V.I. Dodero, M.E. Vázquez, et al., Sequence-specific DNA binding by noncovalent peptide–tripyrrole conjugates, Angew. Chem. Int. Ed. Engl. 45 (2006) 8210–8214.
[179] Z. Yu, C. Guo, Y. Wei, et al., Pip-HoGu: an artificial assembly with cooperative DNA recognition capable of mimicking transcription factor pairs, J. Am. Chem. Soc. 140 (2018) 2426–2429.
[180] X. Ma, Y. Zhao, Biomedical applications of supramolecular systems based on host–guest interactions, Chem. Rev. 115 (2015) 7794–7839.
[181] M.D. Distefano, J.A. Shin, P.B. Dervan, Cooperative binding of oligonucleotides to DNA by triple helix formation: dimerization via Watson-Crick hydrogen bonds, J. Am. Chem. Soc. 113 (1991) 5901–5902.
[182] S.A. Strobel, P.B. Dervan, Cooperative site specific binding of oligonucleotides to duplex DNA, J. Am. Chem. Soc. 111 (1989) 7286–7287.
[183] K. Gorska, N. Winssinger, Reactions templated by nucleic acids: more ways to translate oligonucleotide-based instructions into emerging function, Angew. Chem. Int. Ed. Engl. 52 (2013) 6820–6843.
[184] Z. Yu, W.C. Hsieh, S. Asamitsu, et al., Orthogonal gammaPNA dimerization domains empower DNA binders with cooperativity and versatility mimicking that of transcription factor pairs, Chemistry 24 (2018) 14183–14188.
[185] I. Sacui, W.-C. Hsieh, A. Manna, et al., Gamma peptide nucleic acids: as orthogonal nucleic acid recognition codes for organizing molecular self-assembly, J. Am. Chem. Soc. 137 (2015) 8603–8610.
[186] T. Vaijayanthi, T. Bando, G.N. Pandian, et al., Progress and prospects of pyrrole-imidazole polyamide-fluorophore conjugates as sequence-selective DNA probes, Chembiochem 13 (2012) 2170–2185.
[187] Y. Kawamoto, T. Bando, H. Sugiyama, Sequence-specific DNA binding pyrrole-imidazole polyamides and their applications, Bioorg. Med. Chem. 26 (2018) 1393–1411.
[188] S. Lee, Y. Kawamoto, T. Vaijayanthi, et al., TAMRA-polypyrrole for A/T sequence visualization on DNA molecules, Nucleic Acids Res. (2018) 531.
[189] G.T. Rudkin, B. Stollar, High resolution detection of DNA–RNA hybrids in situ by indirect immunofluorescence, Nature 265 (1977) 472.
[190] Y. Kawamoto, T. Bando, F. Kamada, et al., Development of a new method for synthesis of tandem hairpin pyrrole-imidazole polyamide probes targeting human telomeres, J. Am. Chem. Soc. 135 (2013) 16468–16477.
[191] Y. Kawamoto, A. Sasaki, K. Hashiya, et al., Tandem trimer pyrrole-imidazole polyamide probes targeting 18 base pairs in human telomere sequences, Chem. Sci. 6 (2015) 2307–2312.
[192] K. Maeshima, S. Janssen, U.K. Laemmli, Specific targeting of insect and vertebrate telomeres with pyrrole and imidazole polyamides, EMBO J. 20 (2001) 3218–3228.
[193] A. Sasaki, S. Ide, Y. Kawamoto, et al., Telomere visualization in tissue sections using pyrrole-imidazole polyamide probes, Sci. Rep. 6 (2016), 29261.
[194] G.S. Gorman, P.F. Chinnery, S. DiMauro, et al., Mitochondrial diseases, Nat. Rev. Dis. Primers 2 (2016), 16080.
[195] K.L. Horton, K.M. Stewart, S.B. Fonseca, et al., Mitochondria-penetrating peptides, Chem. Biol. 15 (2008) 375–382.
[196] G.N. Pandian, H. Sugiyama, Nature-inspired design of smart biomaterials using the chemical biology of nucleic acids, Bull. Chem. Soc. Jpn. 89 (2016) 843–868.
[197] A.T. Poulin-Kerstien, P.B. Dervan, DNA-templated dimerization of hairpin polyamides, J. Am. Chem. Soc. 125 (2003) 15811–15821.
[198] A. Saha, G.N. Pandian, S. Sato, et al., Chemical modification of a synthetic small molecule boosts its biological efficacy against pluripotency genes in mouse fibroblasts, ChemMedChem 9 (2014) 2374–2380.
[199] J.A. Raskatov, J.O. Szablowski, P.B. Dervan, Tumor xenograft uptake of a pyrrole-imidazole (Py-Im) polyamide varies as a function of cell line grafted, J. Med. Chem. 57 (2014) 8471–8476.
[200] D.A. Harki, N. Satyamurthy, D.B. Stout, et al., In vivo imaging of pyrrole-imidazole polyamides with positron emission tomography, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 13039–13044.
[201] B. Colmegna, S. Uboldi, E. Erba, et al., Resistance to minor groove binders, Drug Discov. Today Technol. 11 (2014) 73–79.
[202] K. Takahashi, S. Yamanaka, Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors, Cell 126 (2006) 663–676.
[203] X. Xiangshu, Y. Peng, L. Hyun-Suk, S. Devanjan, K. Thomas, A Cell-Permeable Synthetic Transcription Factor Mimic, Angew. Chem. Int. Ed. Engl. 46 (2007) 2865–2868.
[204] Y. Kwon, H.-D. Arndt, Q. Mao, Y. Choi, Y. Kawazoe, P.B. Dervan, M. Uesugi, Small Molecule Transcription Factor Mimic, J. Am. Chem. Soc. 126 (2004) 15940–15941.