Chinese Chemical Letters xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Chinese Chemical Letters

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

Review

Programmable pyrrole-imidazole polyamides: A potent tool for DNA targeting

Chunlei Wu, Wei Wang, Lijing Fang*, Wu Su*

Guangdong Key Laboratory of Nanomedicine, Institute of Biomedicine and Biotechnology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Py-Im polyamides
DNA binding
Gene modulation
Targeting ligands

ABSTRACT

Hairpin pyrrole-imidazole (Py-Im) polyamides are a class of programmable minor-groove binders that recognize pre-determined DNA double helixes with high affinity and specificity. They are capable of regulating gene expression by modulating the activity of transcription factors. To date, Py-Im polyamides have been successfully applied as a potent tool to disturb DNA functions and considered as a group of promising candidates for the clinical applications. Herein, this review will focus on summarizing the recent advances of Py-Im polyamides from their synthesis to applications *via* various modifications at the molecular level.

1. Introduction

DNA is widely involved in many functional roles including replication, and transcription. A myriad of devastating health diseases such as cancer and some hereditary conditions are related to the aberrant expression of DNA [1]. DNA has become the prime target for some related diseases. Many synthetic small molecules targeting DNA *via* binding to the grooves, intercalation, cross-linking, or DNA strand-scission, have been discovered to disturb DNA functions [2]. *N*-Methylpyrrole (Py)-*N*-methylimidazole (Im) polyamides (PIPs), originally derived from distamycin A, are a class of programmable minor-groove binders that recognize pre-determined DNA sequences with nanomolar binding affinity [3,4]. They are consisting of two polyamide strands with a γ-aminobutyric acid (γ-turn) linkage, and form hairpin-like structures. Their specificity is achieved by forming hydrogen bonds with double-strand DNA (dsDNA) based on the Watson-Crick rules, that is antiparallel Im/Py pair discriminates G/C base pair from C/G, and *N*-methyl-3-hydroxypyrrole (Hp)/Py shows specificity for T/A over A/T, whereas Py/Py pair recognizes both A/T and T/A base pairs. This binding mode is oriented N→C with respect to the 5′→3′ direction of the adjacent DNA in the minor groove (Fig. 1). Once Py-Im polyamides bind with matched dsDNA, DNA-binding proteins will be competitively displaced, so that the expression of corresponding genes is inhibited or promoted. In recent years, various Py-Im polyamides and their conjugates have been developed to modulate endogenous expression of nuclear genes and proteins. Py-Im polyamides have also been used as targeting ligands for DNA recognition and bioimaging. To date, there have been many excellent reviews introducing the development of Py-Im polyamides in this field [3–7]. In this review, we summarized the new synthetic methods and the structure-activity relationship of Py-Im polyamides, and outlined the recent advances of their applications as genetic modulators and targeting ligands.

2. New synthetic methods for Py-Im polyamides

Initial Py-Im polyamide syntheses started from long and difficult solution phase. A significant advance was the solid-phase synthesis of Py-Im polyamides developed by Baird and Dervan, which reduced the synthetic timescale for a single polyamide from months to days or weeks [8,9]. Moreover, multifunctional polyamides were produced by using Kaiser oxime or safety-catch hydrazine resin followed by nucleophilic cleavage with various compounds [10]. However, the low yields and lengthy durations were still the crucial problems in the synthesis of Py-Im polyamides [11]. Hence, Dervan *et al.* developed a convergent synthesis of Py-Im polyamides on gram-scale with minimal use of chromatography by using a [3 + 2+4]-peptide fragment-coupling strategy in solution phase [12]. In recent years, microwave assisted solid-phase synthesis was utilized to prepare Py-Im polyamides, which significantly enhanced the coupling yields and reduced the

* Corresponding authors.
E-mail addresses: lj.fang@siat.ac.cn (L. Fang), wu.su@siat.ac.cn (W. Su).

https://doi.org/10.1016/j.cclet.2018.05.025

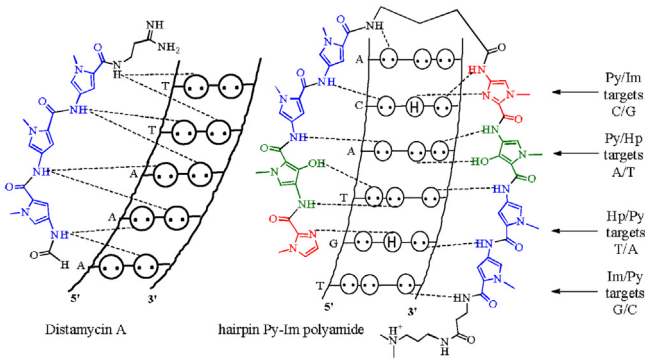

Fig. 1. Binding models of distamycin A and hairpin Py-Im polyamide with dsDNA.

coupling times [13]. Burley *et al.* developed a convergent synthesis of Py-Im polyamides with superior yields and purities, which involved the Boc-based solid-phase peptide synthesis (SPPS) of the C-terminal fragment and the solution phase synthesis of the N-terminal fragment [14]. We previously reported a facile and highly efficient solid-phase synthesis of Py-Im polyamides using the cost-effective triphosgene (BTC) as activating agent [15]. BTC converts carboxylic acids into highly electrophilic acid chlorides, which readily acylate the sterically hindered and electronically deactivated amines. However, the strong exothermic nature and the large amount of precipitates generated in the process of BTC-mediated activation were unfavorable to automated synthesis. Therefore, development of a milder and solid free activation method was crucial for the automation of the BTC strategy. We found that the insoluble collidine hydrochloride salts produced by the preactivation of the carboxyl acid with BTC and collidine in THF could be converted to the soluble *N*,*N*-diisopropylethylamine (DIEA) hydrochloride by the addition of a DIEA/DMF solution. So we used this special activation procedure to avoid the potential clogging of the peptide synthesizer pipelines, and realized the fully automated solid-phase synthesis of Py-Im polyamides, which represented a significant step forward for the multiple, parallel production of Py-Im polyamides (Fig. 2) [16].

In addition to linear and hairpin Py-Im polyamides, tandem hairpin structures have been designed and synthesized. Tandem hairpin Py-Im polyamides are comprised of two or more hairpin Py-Im polyamides linked with a flexible hinge segment, thus can recognize up to 10 base pairs. However, the synthesis of this kind of polyamides suffered from low yields because multiple steps were required to remove Boc groups or Fmoc groups in Fmoc SPPS or Boc SPPS, respectively. By preparing a new tetraamide unit in solution phase and introducing it to Fmoc SPPS, Sugiyama *et al.* developed an efficient synthetic route for creating tandem hairpin Py-Im polyamides [17]. Using this method, they synthesized a series of tandem hairpin Py-Im polyamides including tandem dimer [18], trimer [19] and tetramer [20].

3. Structure-activity relationship of Py-Im polyamides

Structures of Py-Im polyamides influence their activity. Compared to natural distamycin A (interacted with DNA in a 1:1 stoichiometry) (Fig. 1), hairpin Py-Im polyamides, in which a γ-turn linker connects two polyamide strands, display approximately 100-fold higher affinity [21]. A number of molecular recognition researches were conducted to evaluate the structure-activity relationship of Py-Im polyamides. For example, previous reports have demonstrated the cell uptake and nuclear localization of Py-Im polyamides were related to Py/Im content, number and location of positive and negative charges, the modification of C-terminus and N-terminus, choice of fluorophore, linker composition and attachment [22–25]. It was noted that hairpin Py-Im polyamides with eight or more rings became over-curved and would no longer match the minor groove shape of dsDNA. In replacement of Py, introduction of a flexible β-alanine (β) to longer Py-Im polyamides was needed to compensate the structural incompatibility. But β substitution could sometimes decrease the binding affinity of Py-Im polyamides with DNA. Wilson *et al.* found the number and position of β substitutions affected the affinity and specificity of Py-Im polyamides [26,27]. They summarized when Py between two Im was replaced by β, the binding affinity tended to go up (such as KA2034 in Table 1), while if β appeared on a lower strand consisting of four Py, the binding affinity tended to go down (such as KA2041 and KA2114 in Table 1). They reasoned that Im heterocycles were more rigid than Py rings, and displacement Py with β increased the flexibility of the whole molecule. For the former, increase of the molecular flexibility adjusted Py-Im polyamides to the curvature of DNA more easily. Inversely for the latter, this increase reduced van der Waals interaction between Py-Im polyamides and dsDNA, leading to slack the binding.

According to well-defined pairing rules, Py-Im polyamides retain the orientation preferences of aligning N to C terminus with respect to the 5′–3′ direction of the DNA (referred to as the forward orientation). However, for some Py-Im polyamides, reversed binding (a C to N terminus alignment of Py-Im polyamides with respect to the 5′–3′ direction of DNA) has been observed as the dominant orientation. Dervan *et al.* reported β-containing polyamides with different γ-turn units (α-amino substituted, β-amino substituted or unsubstituted) displayed unanticipated binding motifs with DNA by Bind-N-Seq [28]. They discovered that conformationally flexible, β-containing polyamides with a β-amino-γ-turn linker preferred a reverse orientation, and the forward orientation was recovered when β-amino-γ-turn linker was replaced by α-amino-γ-turn linker. However, for the polyamide sequence PyImβIm-γ-PyImβIm with a reversed binging mode,

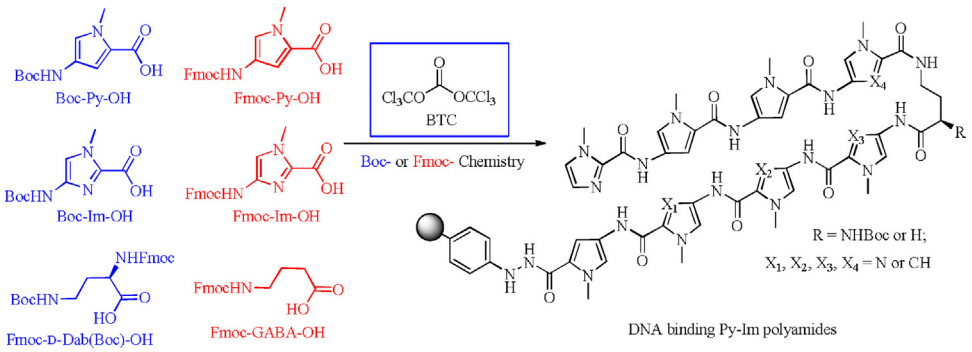

Fig. 2. Synthesis of Py-Im polyamides *via* BTC activation.

Please cite this article in press as: C. Wu, et al., Programmable pyrrole-imidazole polyamides: A potent tool for DNA targeting, Chin. Chem. Lett. (2018), https://doi.org/10.1016/j.cclet.2018.05.025