

library in numerous other assays, including adipocyte differentiation, muscle differentiation, chondrocyte differentiation, inflammation, proliferation of primary hemopoietic cells, and cell cycle regulation in primary endothelial cells. All of these screens led to the identification of both known and novel regulators. We are currently analyzing the novel hits that were isolated from these screens.

As reported in the above article by Elferink, we have a >120 000 human placental library available and are generating additional human cDNA libraries from fetal liver and brain. The identity and length of the cDNAs in these random 'knock-in' libraries is unknown and is established only after a phenotype is observed in our screens. However, we are also generating other, defined and focused libraries of classical drugable genes (GPCRs, kinases, phosphatases, proteases etc.) to identify and validate disease-related candidate genes for further drug discovery. The first of these libraries was established through a collaboration with Incyte Pharmaceuticals (Palo Alto, CA, USA) and contains >1000 individual full-length cDNAs. A second, 'knock-in' library will contain arrayed cDNAs representing all known drugable genes. A third library, aimed at down-modulation (knock-down) of expression of the drugable genes, will consist of adenoviral vectors that express small pan-handle RNAs that target specific mRNAs through a process known as RNA interference. The main advantages of these defined libraries are that they are comparatively small (30 000 maximum) and

not disease specific, thus they can be used in all therapeutic areas. The relatively small size will reduce the costs of screening these libraries. Furthermore, it allows the construction of these libraries in different adenoviral vectors by using the capsid variants mentioned above so as to increase the direct infectability and decrease the amount of virus.

All in all, the adenoviral PhenoSelect™ libraries have the potential to contribute to the discovery of novel therapeutic targets at the genome scale. The libraries are produced in an arrayed adenoviral format that enables high-throughput screening. The great benefit of this library format is the ability to perform versatile, functional assays with a wide variety of human cell types, including primary cells. Galapagos Genomics (Leiden, The Netherlands; Mechelen, Belgium) will use these libraries to isolate novel therapeutic targets within its main disease areas, namely Alzheimer's disease, bone disorders and psoriasis. In addition, we are setting up screens aimed at the identification of regulators for IgE-dependant mast cell degranulation with mast cells derived from human cord blood progenitors.

#### References

- 1 Horwood, N.J. *et al.* (2002) High-efficiency gene transfer into non-transformed cells: utility for studying gene regulation and analysis of potential therapeutic targets. *Arthritis Res.* 4, S215–S225

0167-7799/03/\$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.  
doi:10.1016/S0167-7799(03)00027-1

#### Research Focus

## Earrings and padlocks for the double helix: topological labeling of duplex DNA

Vadim V. Demidov

Center for Advanced Biotechnology, Boston University, 36 Cummington St., Boston, MA 02215, USA

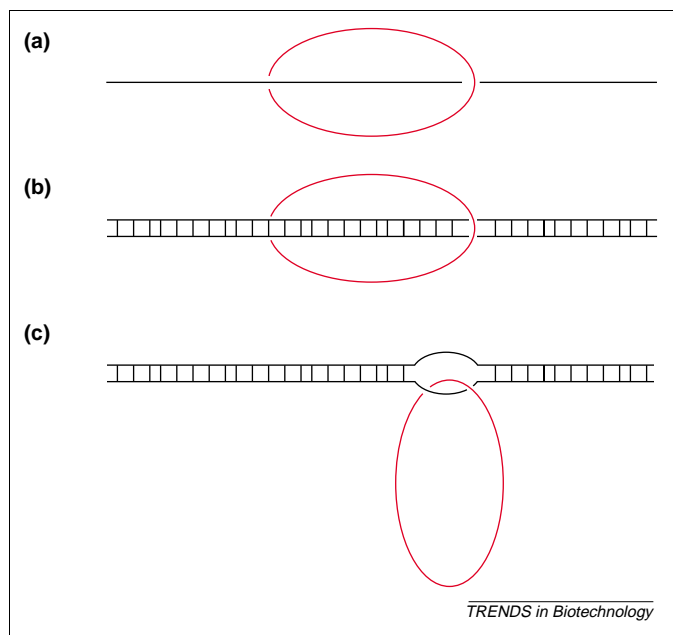
**Concatenation of hybridization probe with DNA target is crucial for highly localized detection of targeted sequences and might also be used in various gene-therapy applications. Several approaches based on the attachment of a circular oligonucleotide to designated DNA sites have been proposed. Recently, earring-like probes provide a true topological linkage between a probe and the target, thus allowing the DNA labeling by essentially immobile tags. The latest development in this direction takes advantage of oligonucleotide uptake by supercoiled DNA and is an important step forward.**

Site-specific tagging of DNA with various nucleobase oligomers, such as regular oligonucleotides or their analogs and peptide nucleic acids (PNAs), underlies many biotechnological DNA manipulations, including

isolation and detection of chosen DNA sequences. To be widely applicable, DNA labels should have high sequence selectivity and high stability, and be without serious sequence limitations. The first two requirements assume that the labeling approach will yield firm, ideally immovable, attachment of tags to the designated DNA sites only, and the last requirement means that a wide variety of DNA target sites can be used.

Methods for bonding of a circular oligodeoxyribonucleotide (cODN) with the desired DNA sequence via pseudotopological or topological links (Fig. 1) are promising [1]. These methods provide extra stability for labeling owing to the probe-to-target catenation. The *in situ* circularized probes have also an increased specificity because multiple-coincidence hybridization/circularization processes are involved. Labeling by cODNs offers significant improvement for DNA detection via isothermal rolling-circle amplification (RCA) [2].

Corresponding author: Vadim V. Demidov (vvd@bu.edu).



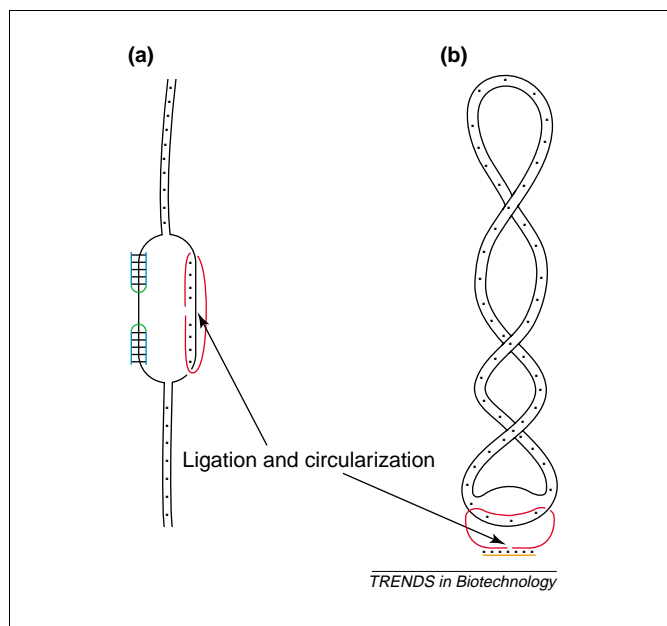
**Fig. 1.** A variety of pseudorotaxane-like topological links between DNA (black lines) and cODN (red ring): (a) padlock probe formed on ssDNA; (b) padlock probe or sliding clamp formed on dsDNA; (c) earring probe for dsDNA.

### DNA padlocks

The first step in the development of topological DNA labeling was the design of padlock-type probes consisting of two target-complementary segments connected by a target-unrelated linker sequence [3]. The DNA padlock (Fig. 1a) can be formed by enzymatic [3] or non-enzymatic [4] circularization of a linear oligonucleotide probe following the Watson-Crick pairing of its termini with the complementary target on single-stranded (ss) DNA. In the hybridized state, the probe's termini are juxtaposed and can therefore be covalently joined by DNA ligase or chemical autoligation to form the padlock similar to assemblies shown in Figure 2.

Normally, the hybridized part of a circular probe is ~20 nucleotides long (to target the unique DNA sequences) making about two turns around the ssDNA target. Thus, the padlock probe and target strand are intertwined twice (for simplicity, probes are drawn with only one link). Such linkage between a closed endless probe and the ssDNA target imposes, in a manner similar to regular padlocks, certain topological restrictions on their separation, as compared with essentially free, topologically unrestrained dissociation of target-bound linear probes. Besides padlocks, one can imagine the long stick encircled by a small ring in the middle: there is no possibility for that ring to leave the stick other than to slide all the way long to either of its termini and only then freely slip away.

The requirement for two separate probe segments to simultaneously react with the target sequence and each other results in a higher selectivity of target recognition by padlock probes. What's more, the topologically stabilized probes resist extreme washing conditions that further reduces the nonspecific probe binding. Owing to exquisite specificity, padlock probes have found application in several genetic assays, including the chromosomal



**Fig. 2.** Schematics of the two types of earring assemblies on dsDNA (black lines): the post-hybridization closure of circularizing probes (red) with DNA ligase results in the probe's threading between the two strands of duplex DNA. (a) The peptide nucleic acid (PNA)-assisted target-directed earring assembly on a linear DNA fragment, which can be obtained via enzymatic digest of longer genomic DNA. A pair of *bis*PNA openers is shown in blue. (b) The supercoiling-assisted template-directed earring assembly on a covalently closed circular plasmid DNA. A template oligonucleotide is shown in light brown. Note that in the supercoiling-based approach an earring can be linked with either of the two DNA strands.

localization and real-time detection of particular DNA sequences [5–8].

The original design of padlocks has two limitations. First, these probes can be formed this way on the ssDNA target only. Given that double-stranded (ds) DNA is the major DNA form, prior denaturation of DNA is necessary. Second, in spite of the movement restrictions, the *pseudo* topological link featured by padlock-type probes still allows the cODN label to move along the targeted DNA, thus preventing precise spatial positioning of hybridized probe. Therefore, the target site cannot be located with precision better than a few hundred nucleotides [3].

Triplex-based approaches to target dsDNA sites with cODNs have recently been proposed (Fig. 1b) [9–12]. Initially, a prior circularized triplex-forming cODN, called sliding clamp, was threaded on the specific dsDNA fragment [9]. This approach can be applied to a relatively short linear dsDNA fragment with free, unobstructed ends. Circular plasmids, as well as linear chromosomes, the ends of which are blocked by telomeric complexes, are not accessible for sliding clamps. Later, the triplex-directed padlock-like approach for the *in situ* probe circularization on dsDNA target was demonstrated [10–12]. An oligonucleotide probe was circularized around the dsDNA target after triplex formation by using DNA ligase and supplementary template oligonucleotide (similar to that shown in Fig. 2b). The resulting structure looks like the sliding clamp in Fig. 1b yet it can be successfully assembled on circular plasmids [11,12] and presumably on DNA within intact chromosomes.

Note that DNA ligases operate mostly with dsDNA and rarely with ssDNA but these enzymes do not work with

DNA triplexes. Therefore, the immediate target-directed enzymatic ligation cannot be used to make dsDNA padlocks *in situ*. Hence, triplex-forming probe oligonucleotide should either be in a circular form or be circularized with DNA ligase outside the target site. In both cases, a single probing segment participates in the target recognition, which might compromise the detection selectivity, as compared with bipartite probes. The padlock can probably be latched straight away on dsDNA in a highly sequence-specific way by non-enzymatic target-directed autoligation of a linear precursor similar to the established case with ssDNA [4]. Still, it has to be proved whether this can be accomplished within triplexes taking into account the conformational differences between double- and triple-helical complexes, as well as certain sequence limitations of autoligation reaction [4].

Thus, circular probe can be loaded onto dsDNA via triple helix formation. However, sequence restrictions intrinsic in triplexes limit this method because mostly long oligopurine DNA sites can be targeted through Hoogsteen-type base pairing required by triplexes [13]. Also, owing to the pseudotopological character of the probe–target link thus formed, the problem with probe slippage away from target site remains, let alone the possible complications with sequence selectivity. Although padlock movement along the dsDNA can be constrained by triplex-stabilizing ligands [11], other ligand-free approaches would be desirable to trap cODN probe at a precise position. Indeed, triplex-stabilizing ligand might interfere with DNA–protein interactions, which is disadvantageous for some applications, for example for performing the RCA reaction with DNA polymerase [2]. Conversely, insufficient stabilization might result in the displacement of circular probes by proteins that translocate on dsDNA.

### DNA earrings

The solutions of aforementioned problems came again from topology: two research groups had closely related ideas on how to assemble truly topological, and hence absolutely immovable, links of cODNs and dsDNA sites. They decided to locally open the DNA double helix at designated position thus making the DNA strand(s) of a tiny DNA segment readily accessible for common Watson-Crick-type hybridization. In the final structure, which resembles an earring (Fig. 1c), a segment of the circular label is threaded sequence-specifically between complementary strands of dsDNA. Therefore, in this construct, an oligonucleotide probe is firmly fastened to chosen dsDNA site and cannot move along the targeted DNA unless strand separation in the flanking DNA sequences occurs.

Our group used special biomolecular nano-tools, peptide nucleic acid (PNA) openers [14], which selectively invade DNA duplex thus prying it apart (Fig. 2a). Accordingly, the oligonucleotide probe was circularized by means of target-directed enzymatic ligation to yield first earring labels on both linear and closed circular dsDNAs [15]. Subsequent studies demonstrated the exceptional sequence specificity and stability of PNA-assisted target-directed earrings, the possibility of their tagging by streptavidin for ensuing plasmid functionalization, the workability of in-gel earrings

assembly on long DNAs embedded in agarose plugs, and that the RCA reaction can be performed on earring probes by some DNA polymerases [1,16,17].

The PNA-assisted formation of earrings includes two more, as compared with the target-directed padlocks, independent recognition elements – a pair of PNA openers (Fig. 2a). Because only the concerted sequence-specific binding of all four participants, two PNA openers and two termini of the oligonucleotide probe, to the dsDNA target results in the formation of earrings, their assembly should be even more discriminative than the very selective target-directed formation of padlock probes. Hence, one can expect practically zero tolerance of earrings to mismatches between a cODN and the dsDNA target, as observed [1,16]. Note that despite the PNA-assisted earrings being formed on mixed-base dsDNA sites, the involvement into the assembly of PNA openers targeted to short neighboring oligopurine DNA sites (Fig. 2a) limits the applicability of this technique by purine-rich sequences (although it is less sequence restricted than triplex-based approaches).

Very recently, Potaman *et al.* managed to exploit the energy of DNA supercoiling as a driving force for dsDNA local opening and subsequent formation of earrings on closed circular dsDNAs. Specifically, they made use of oligonucleotide uptake by plasmid DNA under negative superhelical tension [18], a factor that promotes DNA unwinding. This enabled the assembly of earring probes on supercoiled dsDNA without any engagement of PNA or some other openers into the process (Fig. 2b). The supercoiling-assisted earrings can apparently be formed on a fully mixed sequence of DNA nucleobases, therefore expanding the repertoire of DNA sites available for topological dsDNA labeling. Nevertheless, one sequence limitation is still intrinsic in this approach: the targeting DNA site must be AT-rich to facilitate its unwinding. Elevated temperatures or reduced salt concentrations might soften the prerequisite need in AT-rich sites because other sequences are known to be exposed by DNA supercoiling with increase of temperature and decrease in saline contents of solution.

Thus designed earrings comprise a single recognition segment being circularized by DNA ligase with the aid of supplementary template oligonucleotide. In principle, the supercoiling-assisted earrings can be assembled in a target-directed, and therefore more sequence-specific, way if two separate end-located probe segments simultaneously bind the target sequence (see Fig. 2a). However, the efficiency of this mode of assembly was significantly reduced compared with the template-directed approach (Potaman V.N., personal communication), which is not surprising given that the targeted dsDNA site is AT-rich and binding of shorter probe segments should be less stable. Besides, in case of the PNA-assisted earring assembly, the PNA openers keep the non-target displaced DNA strand away from the targeted site. However, nothing restrains this strand from interference with probe binding in case of the supercoiling-assisted formation of earrings.

In spite of these limitations, the supercoiling-based approach is well suited for plasmid labeling. Indeed, most cloning vectors carry a lengthy AT-rich site near the

$\beta$ -lactamase gene, which could serve as a unique target for earring labeling. It is highly unlikely to meet the same site, as well as its mismatched variants, within cloned recombinants. Using the common pUC-derived cloning vector and atomic force microscopy (AFM), Potaman *et al.* showed that the earring they formed is stably localized at the designated target, an unwinding element neighboring the  $\beta$ -lactamase gene [18]. This proves that earring labels might serve as a reference point on circular plasmids. Accordingly, microscopy visualization techniques can be used to map the protein binding sites, identify local unusual structures in supercoiled DNA and monitor conformational dynamics of dsDNA molecules in real time.

Importantly, the AFM images revealed that labeling of dsDNA with earring results in a sharp bend of the DNA double helix, which tends to occupy an apical position on a superhelix [18]. Such an artificially induced site-specific DNA bending can be used for modifying the action of DNA-processing proteins as recently demonstrated with a conversion of restriction enzyme into the DNA nickase using the PNA-induced bend [19]. Formation of DNA earrings might serve as a new approach to modulate the extent of DNA curvature that has important roles in numerous cellular processes involving DNA.

### The bottom line

DNA topological labeling with motional-restricted cyclization probes, padlocks and earrings, offers significant improvements to common hybridization approaches, including the extra stability and increased specificity of labeling and capacity for immense amplification by RCA at a single temperature. Through triplex formation, padlock probes can be loaded directly onto duplex DNA—the primary DNA form. By comparison with DNA padlocks, formation of earring probes results in much more stable and strictly localized labeling of dsDNA. Considering that DNA earrings can be fastened to a wider diversity of target DNA sequences and could be even more sequence-specific than padlocks, the former might be of substantial value in a variety of biotechnological, nanotechnological and gene-therapeutic applications. Potential uses are: the SNP detection directly in genomic DNA, protein-mediated gene delivery, localization of certain markers within intact chromosomes, blocking specific DNA sequences, further design of higher-order supramolecular DNA assemblies, manipulating DNA libraries and monitoring of intracellular distribution of foreign DNA.

New advances might extend the applications of topological probes. First, given the reagent-free format, high sequence fidelity and significant level of self-amplification observed for autoligation reaction [4,20], the workability of non-enzymatic target-directed probe circularization within triplexes (to form DNA padlock) and locally exposed duplexes (to form DNA earring) has to be studied. Second, the development of techniques based on covalent but readily reversible probe closing (e.g. via connection of two SH-groups located at the probe ends through S–S bond [21]) as well as non-covalent probe ‘self-ligation’ (e.g. via biotin–streptavidin–biotin linkage of a pair of terminal biotin residues [22]) could add extra

pliability to enzyme-free labeling approaches. Finally, the use of mixed-base pseudocomplementary PNA openers [23] along with the pseudocomplementary oligonucleotide probes [24,25] can further relax the sequence limitations on the PNA-assisted design of earrings [1,14].

### References

- 1 Demidov, V.V. *et al.* (2001) Peptide nucleic acid-assisted topological labeling of duplex DNA. *Methods* 23, 123–131
- 2 Demidov, V.V. (2002) Rolling-circle amplification in DNA diagnostics – the power of simplicity. *Expert Rev. Mol. Diagn.* 2, 542–548
- 3 Nilsson, M. *et al.* (1994) Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science* 265, 2085–2088
- 4 Xu, Y. and Kool, E.T. (1999) High sequence fidelity in a non-enzymatic DNA autoligation reaction. *Nucleic Acids Res.* 27, 875–881
- 5 Isaksson, A. and Landegren, U. (1999) Accessing genomic information: alternatives to PCR. *Curr. Opin. Biotechnol.* 10, 11–15
- 6 Baner, J. *et al.* (2001) More keys to padlock probes: mechanisms for high-throughput nucleic acid analysis. *Curr. Opin. Biotechnol.* 12, 11–15
- 7 Schweitzer, B. and Kingsmore, S. (2001) Combining nucleic acid amplification and detection. *Curr. Opin. Biotechnol.* 12, 21–27
- 8 Nilsson, M. *et al.* (2002) Making ends meet in genetic analysis using padlock probes. *Hum. Mutat.* 19, 410–415
- 9 Ryan, K. and Kool, E.T. (1998) Triplex-directed self-assembly of an artificial sliding clamp on duplex DNA. *Chem. Biol.* 5, 59–67
- 10 Escude, C. Garestier, T. and Helene, C. (1999) Padlock oligonucleotides for duplex DNA based on sequence-specific triple helix formation. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10603–10607
- 11 Roulon, T. Helene, C. and Escude, C. (2001) A ligand-modulated padlock oligonucleotide for supercoiled plasmids. *Angew. Chem. Int. Ed.* 40, 1523–1526
- 12 Roulon, T. *et al.* (2002) Padlock oligonucleotides as a tool for labeling superhelical DNA. *Nucleic Acids Res.* 30, e12
- 13 Gowers, D.M. and Fox, K.R. (1999) Towards mixed sequence recognition by triple helix formation. *Nucleic Acids Res.* 27, 1569–1577
- 14 Demidov, V.V. (2001) PD-loop technology: PNA openers at work. *Expert Rev. Mol. Diagn.* 1, 343–351
- 15 Kuhn, H. Demidov, V.V. and Frank-Kamenetskii, M.D. (1999) Topological links between duplex DNA and a circular DNA single strand. *Angew. Chem. Int. Ed.* 38, 1446–1449
- 16 Kuhn, H. *et al.* (2000) An earring for the double helix: assembly of topological links comprising duplex DNA and a circular oligodeoxynucleotide. *J. Biomol. Struct. Dyn. Sp. Iss.* S2, 22–225
- 17 Kuhn, H. *et al.* (2002) Rolling-circle amplification under topological constraints. *Nucleic Acids Res.* 30, 574–580
- 18 Potaman, V.N. *et al.* (2002) Site-specific labeling of supercoiled DNA at the A + T rich sequences. *Biochemistry* 41, 13198–13206
- 19 Protozanova, E. *et al.* (2002) Tailoring the activity of restriction endonuclease *Ple I* by PNA-induced DNA looping. *EMBO Rep.* 3, 956–961
- 20 Xu, Y. Karalkar, N.B. and Kool, E.T. (2001) Nonenzymatic autoligation in direct three-color detection of RNA and DNA point mutations. *Nat. Biotechnol.* 19, 148–152
- 21 Gilbert, H.F. (1995) Thiol/disulfide exchange equilibria and disulfide bond stability. *Methods Enzymol.* 251, 8–28
- 22 Niemeyer, C.M. (2002) The developments of semisynthetic DNA-protein conjugates. *Trends Biotechnol.* 20, 395–401
- 23 Demidov, V.V. *et al.* (2002) Kinetics and mechanism of the DNA double helix invasion by pseudocomplementary peptide nucleic acids. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5953–5958
- 24 Kutvavin, I.V. *et al.* (1996) Oligonucleotides containing 2-aminoadenine and 2-thiothymine act as a selectively binding complementary agents. *Biochemistry* 35, 11170–11176
- 25 Woo, J. *et al.* (1996) G/C-modified oligodeoxynucleotides with selective complementarity: synthesis and hybridization properties. *Nucleic Acids Res.* 24, 2470–2475