COMMUNICATION

ChemComm
www.rsc.org/chemcomm

Fluoroaromatic universal bases in peptide nucleic acids

Kathryn A. Frey and Stephen A. Woski*
Department of Chemistry, The University of Alabama, Box 870336, Tuscaloosa, AL 35487, USA.
E-mail: swoski@bama.ua.edu; Fax: +1 205 348 9104; Tel: +1 205 348 0454

Received (in Cambridge, UK) 14th June 2002, Accepted 9th August 2002
First published as an Advance Article on the web 30th August 2002

Fluoroaromatic peptide nucleic acid residues were found to possess little base discrimination when incorporated into PNA·DNA double helices

The increasing availability of genetic information for numerous organisms has promoted the continued development of oligonucleotide-based probe molecules. Such probes are potentially valuable for the development of array-based assays, techniques based on primer extensions and *in situ* hybridizations, and antisense chemotherapeutics.[1] While in most cases these methods depend on the high specificity of Watson–Crick base pairing, situations arise where the sequence information is ambiguous or unknown. For example, the design of probes and primers based on the amino acid sequence of a protein is complicated by codon degeneracy. Sequence ambiguities can also arise from natural variations of single nucleotides. The most efficient means to overcome these difficulties is to incorporate universal residues into the probe molecules.[2] Ideally, these wildcard bases should indiscriminately pair with all four natural bases with an affinity at least equal to the native base pair.

While several universal base residues have been described for oligodeoxynucleotides,[2] relatively few have been investigated for peptide nucleic acid (PNA) probes.[3–5] PNA is a synthetic analogue of DNA that incorporates repeating *N*-(2-aminoethyl)-glycine units in its backbone.[6,7] Nucleobases are attached to this backbone *via* methylene carbonyl linkers. The unique, charge-neutral polypeptide backbone of PNA imparts numerous benefits to this system including tighter and salt-independent binding to complementary DNAs and a higher sensitivity to base mismatches.[7] Combined with the ability of PNA oligomers to displace homologous strands from DNA double helices (strand invasion), these characteristics make PNA a superior molecule for numerous biochemical applications.[7]

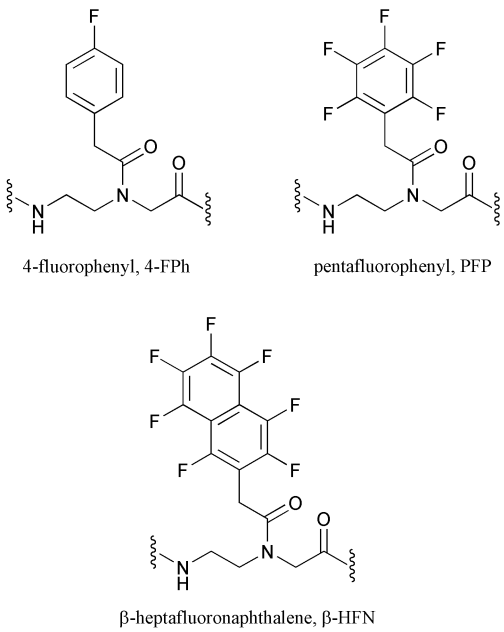
Preliminary investigations of potential universal bases in PNA have indicated that residues with considerable structural variation can still show base pairing degeneracy.[3,4] PNA 15-mers containing a 5-nitroindole,[3] 3-nitropyrrole,[3] phenyl,[4] or hydrogen (abasic)[4] residue showed remarkably uniform T_ms when hybridized with DNAs containing each of the four natural bases positioned opposite to the modified site. While the ranges of T_ms exhibited by these duplexes were surprisingly small, the average T_ms were depressed by varying degrees compared to a canonical PNA·DNA duplex. This suggests that other residues may yet be synthesized that combine base pairing degeneracy with no reduction in duplex stability.

As part of a search for such residues, we have undertaken the syntheses of several classes of modified PNA monomers. Described herein is a preliminary report on the hybridization properties of PNA oligomers containing fluoroaromatic residues.[8] The incorporation of fluorine substituents can perturb the electron distribution of arenes without providing sites for strong hydrogen bonding.[9] Furthermore, perfluorinated aryl systems feature an inversion of the normal quadrupole distribution of electrons; model systems have shown an enhancement of the face-to-face π-stacking interactions between perfluoroarenes and 'normal' arenes.[10] In order to examine effects of fluoroarene residues on the formation of PNA·DNA duplexes and their potential as universal bases, 4-fluorophenyl (4-FPh), pentafluorophenyl (PFP), and β-heptafluoronaphthalene (β-HFN) PNA monomers were synthesized and incorporated into oligomers (Fig. 1).

Syntheses of the fluoroarene PNA monomers was performed as previously described.[3,11] Thus, 4-fluorophenylacetic acid, pentafluorophenylacetic acid (obtained commercially) and β-heptafluoronaphthaleneacetic acid (prepared by the method of Vlasov and Yacobson[12]) were coupled with 2-*N*-Fmoc-2-aminoethylglycine *tert*-butyl ester[13] in good yields. Treatment with trifluoroacetic acid produced the corresponding carboxylic acids, which were then utilized for standard automated PNA synthesis with manual coupling of the novel residues. PNA 15-mers were purified by HPLC, and their identities were confirmed by MALDI-TOF mass spectrometry. Complementary DNA oligomers were made containing each of the four natural nucleobases opposite to the modified site on PNA.

Hybridization of the PNAs with each of the four complementary DNAs was performed in PES buffer (10 mM phosphate, 0.1 mM EDTA, 100 mM NaCl, pH 7). Thermal

A.



B.

PNA X = 4-FPh, PFP, β-HFP, T

H–TGT	ACG	X	CAC	AAC	TA–NH_2
3'–ACA	TGC	Y	GTG	TTG	AT–5'

DNA Y = A, C, G, T

Fig. 1 (A) Structures of fluoroaromatic PNA residues. (B) Sequences of PNA and DNA oligomers used in thermal denaturation assays.

DOI: 10.1039/b205734d

denaturation of the PNA·DNA complexes was followed using absorbance spectroscopy at 260 nm. All of the absorbance *vs.* temperature plots showed sigmoidal curves indicating a cooperative transition, and the data were fit to a two-state model using Meltwin 3.5 software.[14] As was the case with the PNA universal base candidates examined previously, each of the fluoroaromatic residues showed considerable base pairing degeneracy with the natural nucleobases (Table 1). For example, PNA·DNA duplexes with the 4-FPh residue showed T_m values ranging from 55.1 °C (*vs.* dA) to 53.3 °C (*vs.* dG), a ΔT_m of 1.8 °C. The duplexes bearing the perfluorophenyl residue behaved similarly. Compared to the complexes with the monocyclic residues, those with the larger β-HFN residue were approximately 2–4 °C more stable and showed a smaller range of T_m values *vs.* the four natural bases (ΔT_m = 1.3 °C).

The similarity in the behaviors of the 4-FPh and PFP residues is interesting. The parent arenes possess different charge distributions both on the faces and on the peripheries of the rings. The similar stabilities of duplexes containing either of these two residues suggest that the inversion of the facial charge distribution of the arenes (negative in center and positive along edges for 4-FPh; positive in center and negative along edges for PFP) does not significantly effect the aromatic stacking interactions within the double helical complexes. This result differs with conclusions from model studies that show a strong role for quadrupole interactions in aromatic stacking.[10] The similarity of base pairing specificity for these residues is also interesting in light of the different distribution of partial atomic charges along the peripheries of the arenes. Assuming that both residues maintain a coplanar arrangement with all of the natural bases, one might expect that the PFP residue (with electronegative fluorines at all positions) would interact differently than the 4-FPh residue which possesses a single fluorine. The fact that both 4-FPh and PFP formed their least stable pairs with dG is suggestive of a weak repulsive interaction between the 4-fluoro group and the guanine O-6 group. Experiments with other fluorobenzenes are underway to more fully examine these issues.

Finally, β-HFN proved to be the least destabilizing and least discriminating residue of the three examined. Interestingly, model building studies indicate that this residue cannot exist in a coplanar arrangement with either purine or pyrimidine partners while in an "*anti*-like" conformation. Either this residue flips to the "*syn*-like" conformation or it partially intercalates into the helix. Such intercalation has been proposed for large hydrophobic residues placed in internal sites of DNA·DNA[15] and DNA·PNA[16] duplexes; however, structural studies on this system will be required to definitively assign the position of the aryl group. If this residue does in fact intercalate, it will be interesting to examine the magnitudes of effects arising from potential interactions with the neighboring base pairs. Although the β-HFN-containing complexes were destabilized compared to a control duplex containing a canonical T–A base pair (T_m = 67.3 °C), the potential utility of the β-HFN residue as a universal base can, nevertheless, be easily illustrated by the following observation: the T_m values of all of the PNA·DNA duplexes containing β-HFN were ≥3.5 °C higher than the T_m reported for a 15-mer DNA·DNA duplex with same sequence (with an T–A pair at the variable site).[17] Furthermore, in the same sequence context, the β-HFN residue produced results similar to the best of the previously studied PNA universal bases, 5-nitroindole (5-Ni) and 3-nitropyrrole (3-Np).[3] Specifically, incorporation of β-HFN produced destabilization intermediate between 5-Ni and 3-Np while maintaining equivalent base pairing degeneracy.

Overall, the fluoroaromatic PNA residues showed little base pairing discrimination toward any of the four bases in a complementary DNA strand. Because of its higher stability, the β-HFN residue appears to be the most promising of these analogues for further universal base studies. The pursuit of universal base candidates, the examination of their characteristics, and the development of applications is ongoing in these laboratories.

The authors wish to thank Dr. S. Webb for measurements of mass spectra and Prof. A. J. Arduengo for helpful discussions. The National Institutes of Health and the University of Alabama are acknowledged for support of our research.

Table 1 Data for the thermal denaturation of fluoroaromatic-PNAs with complementary DNAs[a]

	Y	T_m/°C	$\Delta G_{298\,K}$/kcal mol^{-1} K^{-1}
X = 4-FPh	A	55.1	−15.9
	C	54.4	−15.8
	G	53.3	−15.5
	T	53.8	−15.6
X = PFP	A	54.3	−15.1
	C	54.7	−16.4
	G	52.5	−15.3
	T	53.9	−15.8
X = β-HFN	A	56.8	−17.0
	C	58.1	−17.7
	G	57.5	−17.4
	T	58.0	−17.8
X = T	A	67.3	−21.0

[a] Absorbance *vs.* temperature curves were measured at concentrations of 4.0 μM for each strand. T_m values and thermodynamic data were derived from Meltwin fits of triplicate runs. The errors in the reported T_m and ΔG values are estimated to be ±0.2 °C and ±2%, respectively.

Notes and references

1 For some recent reviews, see: R. Ekong and J. Wolfe, *Curr. Opin. Biotechnol.*, 1998, **9**, 19–24; S. C. Case-Green, K. U. Mir, C. E. Pritchard and E. M. Southern, *Curr. Opin. Chem. Biol.*, 1998, **2**, 404–410; C. M. Roth and M. L. Yarmush, *Annu. Rev. Biomed. Eng.*, 1999, **1**, 265–297; E. Uhlmann, *Curr. Opin. Drug Discovery Dev.*, 2000, **3**, 203–213.

2 For a review, see: D. Loakes, *Nucleic Acids Res.*, 2001, **29**, 2437–2447.

3 H. Challa, M. L. Styers and S. A. Woski, *Org. Lett.*, 1999, **1**, 1639–1641.

4 H. Challa and S. A. Woski, *Tetrahedron Lett.*, 1999, **40**, 8333–8336.

5 P. Zhang, M. Egholm, N. Paul, M. Pingle and D. E. Bergstrom, *Methods*, 2001, **23**, 132–140.

6 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497–1500.

7 For some recent reviews see: B. Hyrup and P. E. Nielsen, *Bioorg. Med. Chem.*, 1996, **4**, 5–23; R. Gambari, *Curr. Pharm. Des.*, 2001, **7**, 1839–1862.

8 Recently, several fluoroarenes have been studied as universal bases in a wobble position within a RNA·RNA duplex: J. Parsch and J. W. Engels, *J. Am. Chem. Soc.*, 2002, **124**, 5664–5672.

9 B. A. Schweitzer and E. T. Kool, *J. Am. Chem. Soc*, 1995, **117**, 1863–1872; X. Wang and K. N. Houk, *Chem. Commun.*, 1998, 2631–2632.

10 F. Cozzi, F. Ponzini, R. Annunziata, M. Cinquini and J. S. Siegel, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 1019–1020; F. Cozzi and J. S. Siegel, *Pure Appl. Chem.*, 1995, **67**, 683–689.

11 The synthesis of a Boc–protected PFP PNA monomer was recently reported: N. Shibata, B. K. Das, H. Honjo and Y. Takeuchi, *J. Chem. Soc., Perkin Trans. 1*, 2001, 1605–1611.

12 V. M. Vlasov and G. G. Yakobson, *J. Org. Chem. USSR (Engl. transl.)*, 1976, **12**, 2345–2351.

13 S. A. Thomson, J. A. Josey, R. Cadilla, M. D. Gaul, C. F. Hassman, M. J. Luzzio, A. J. Pipe, K. L. Reed, D. J. Ricca, R. W. Wiethe and S. A. Noble, *Tetrahedron*, 1995, **51**, 6179–6194.

14 J. A. McDowell and D. H. Turner, *Biochemistry*, 1996, **35**, 14077–14089.

15 T. J. Matray and E. T. Kool, *J. Am. Chem. Soc*, 1998, **120**, 6191.

16 K. A. Frey and S. A. Woski, manuscript in preparation.

17 M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566–568.