Twisted Intercalating Nucleic Acids (TINA) in Guanosine-rich Oligonucleotides

A thesis submitted in the partial fulfilment

of the requirements for the degree of

Doctor of Philosophy in Chemistry

Massey University,

Palmerston North, New Zealand

MASSEY UNIVERSITY

Osman Doluca

2013

Abstract

The main role in the structural diversity of DNA molecules belongs to guanosines due to their array of hydrogen bond donors and acceptors, large aromatic surface and ability to adopt *syn* or *anti* conformations. These properties lead to the formation of various DNA topologies such as triplexes or G-quadruplexes by guanosine-rich oligonucleotides. For a long time these secondary structures were mainly considered to be a fascinating phenomenon with little practical use; it was subsequently realised that these structures are likely to be formed under physiological conditions and therefore might be involved in many important biological processes, including genome recombination, telomere stability and regulation of gene expression. Thus, there is a growing interest in development and control of these non-traditional nucleic acid structures.

Although the secondary structures of nucleic acids can be controlled to a certain extent by the careful design of oligonucleotide sequence this strategy alone is not always sufficient. In this thesis we investigated how to control the assemblies of guanosine-rich oligonucleotides using a novel tool, twisted intercalating nucleic acids (TINAs). The incorporation of pyrene-containing TINA monomers into guanosine-rich oligonucleotides led to the formation of stable triplexes or G-quadruplexes depending on the position of TINA monomers. In the light of our results, we have established a set of rules that helps to create a desired structure of guanosine-rich oligonucleotides using TINA molecules.

In the second half of the thesis we focused on expanding the functionality of TINA conjugated oligonucleotides. In terms of fluorescence, we synthesised several fluorescently-silent triplex-forming oligonucleotides (TFOs) equipped with a dye at different positions in the DNA. Fluorescence properties were strongly dependent on the position of the dye. These fluorescently silent TFOs showed up to an 18-fold increase in fluorescent intensity upon triplex formation.

These findings lay the foundation for the future design of artificial DNA sequences for expanding the repertoire of DNA secondary structures and function.

Acknowledgements

I would like to thank my supervisors Professor Geoff Jameson and Dr Vyacheslav V. Filichev for their support and assistance throughout this project. Especially Vyacheslav for the enthusiasm he showed towards the project and the time taken to show me various techniques of DNA chemistry. I would like to thank Professor Alexandre Boutorine for his support to this project and hospitality.

Thanks to all former and present colleagues in IFS for their assistance over the years, especially Jamie Withers and Emad Al-imarah for their friendship and Pat Edwards for his help with NMR. I am also grateful for assistance of departmental staff over the years I have been at Massey University.

I would like to acknowledge important financial support from the Marsden grant administered by the Royal Society of New Zealand, the Dumont d'Urville grant from NZ/France Science and Technology Support Programme and a doctoral completion bursary from Massey University.

Many thanks to my dear parents, Saliha and Nevzat Doluca, for their infinite support reaching all the way to the other side of the world.

Above all I would like to thank my dear wife, Rose, for her patience and support during completion of this PhD and my kids, Leila and Musa, for making my life cOL*ou***r***f*uL at every step.

Table of Contents

List of Figures

List of Tables

Abbreviations

$(EG)_6$	hexaethylene-glycol
µL	microlitre
µM	micromole/litre
µmol	micromole
A	adenosine
Å	Ångström
ABL	Abelson murine leukemia viral oncogene
Abs	absorbance
ACN	acetonitrile
AcOH	acetic acid
aq	aqueous
ATR	attenuated total reflectance
BCR	breakpoint cluster region
bp	basepair
C	cytosine
CD	circular dichroism
COMBO-FISH	combinatorial oligonucleotide probes in fluorescence *in situ* hybridization
CPG	controlled pore glass
CTAB	cetyl trimethylammonium bromide
CuAAC	copper-assisted azide-alkyne cycloaddition
DCA	dichloroacetic acid
DCI	4,5-dicyanoimidazole
DCM	dichloromethane
ddH_2O	double-distilled water
dH_2O	distilled water
DIEA	di*iso*propylethylamine
DMA	*N*,*N*-dimethylacetamide
DMF	*N*,*N*-dimethylformamide
DMPA	*N*,*N*-dimethylaminopropylamine
DMSO	dimethylsulfoxide
DMT	4,4′-dimethoxytrityl
DNA	deoxyribonucleic acid

Dp	(*N*,*N*-dimethylamino)propylamide
dsDNA	double-stranded DNA
ε	extinction coefficient
EDTA	*N*,*N*-ethylenediaminetetraacetic acid
eq	equivalent
ESI	electrospray ionisation
EtBr	ethidium bromide
EtOAc	ethyl acetate
EtOH	ethyl alcohol
F_c	fluorescence intensity of DNA complex
FISH	fluorescence *in situ* hybridization
FRET	Förster resonance energy transfer
F_{ss}	fluorescence intensity of single-stranded DNA
FT-IR	Fourier transform infrared spectroscopy
G	guanosine
HATU	*O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	human immunodeficiency virus
HPLC	high pressure liquid chromatography
HRAS	Harvey rat sarcoma viral oncogene
Im	*N*-methylimidazole
INA	intercalating nucleic acid
IR	infrared
K_d	dissociation constant
KRAS	Kirsten rat sarcoma viral oncogene
L	litre
LNA	locked nucleic acid
M	mole/litre
MALDI	matrix-assisted laser desorption/ionization
MeOH	methyl alcohol
MGB	minor groove binder
min	minute
mL	millilitre
mM	millimole/litre
mmol	millimole
mRNA	mitochondrial ribonucleic acid

MYC	myelocytomatosis viral oncogene
NaOAc	sodium acetate
N-FISH	non-denaturing fluorescence *in situ* hybridization
nm	nanometre
NMP	*N*-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
ON	oligonucleotide
P	pyrene
p	terminal phosphate
PAGE	polyacrylamide gel electrophoresis
Ph	phenyl
PhSH	thiophenol
PNA	peptide nucleic acid
ppm	parts per million
Ps	psoralen
Py	*N*-methylpyrrole
qRT-PCR	quantitative real time polymerase chain reaction
RNA	ribonucleic acid
s	second
SNP	single nucleotide polymorphism
SRC	sarcoma viral oncogene
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
T	thymidine
$T_{½}$	mid-transition point
TBTA	tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine
TCA	trichloroacetic acid
TDS	thermal difference spectra
TEAA	triethylammonium acetate
TFA	trifluoroacetic acid
TFO	triplex-forming oligonucleotide
THF	tetrahydrofuran
TINA	twisted intercalating nucleic acid
TLC	thin-layer chromatography
T_m	melting temperature
TO	thiazole orange

TOF	time of flight
TSP	trimethylsilyl propionate
U	uridine
UV	ultraviolet
Vis	visible
β	β-alanine
γ	γ-aminobutyric acid
λ_{em}	emission wavelength
λ_{ex}	excitation wavelength
Φ_F	fluorescence quantum yield