



UNIVERSITY OF  
BIRMINGHAM

# Modification of DNA aptamers for sensing applications

Francia Allabush

A thesis submitted to the University of Birmingham for  
the degree of Doctor of Philosophy

School of Chemistry  
College of Engineering and Physical Sciences  
University of Birmingham  
September 2018

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

## Abstract

Current clinical assays employ a number of different molecular recognition elements, with the most prevalent using antibodies as a consequence of their high selectivity towards target analytes. However, antibody-based assays carry several limitations. They are highly sensitive to temperature as well as pH, irreversibly denatured outside of physiological conditions, costly, laborious to produce, have a limited shelf life, and cannot be designed to detect targets that are not immunogenic. In recent years, aptamers have been highlighted as suitable alternatives to antibodies. Aptamers are single stranded oligonucleotides which fold to bind to non-oligonucleotide based analytes with high affinities and selectivities. With similar target affinities to antibodies, aptamers are inherently more stable, reversibly denatured by temperature, easily amenable to chemical modifications, and can potentially be developed for any analyte. However, aptamers experience several limitations when it comes to their use in therapeutic and diagnostic applications, one major factor being their inability to withstand nuclease degradation. This makes them unsuitable for analysing clinical samples. Fortunately, such limitations can be rectified by incorporating modifications into the aptamer sequence. This thesis aims to contribute to the growing field of modified aptamer development by introducing two vastly different but innovative modifications for incorporating aptamers into biosensing applications.

The first approach aims to incorporate aptamers into molecularly imprinted polymers, or MIPs. Often noted as a more stable substitute to both antibodies and aptamers, MIPs are formed when the analyte of interest is imprinted into a polymer. Although there are already a handful of reports on aptamer

incorporated MIPs, there is no mention on how the flexibility of the aptamer affects the sensing capabilities of the surface coupled aptamer-MIP hybrids. This approach aims to address this by studying the placement of novel acrylamide modified thymidine residues into the thrombin binding aptamer sequence. The synthesis of the acrylamide modified thymidine molecule is described, and the feasibility of the molecule to polymerise and be incorporated into DNA *via* solid phase oligonucleotide synthesis is reported. Modified aptamer sequences with various placements of acrylamide groups are then synthesised, and their ability to polymerise and bind to thrombin/complementary targets is studied. These studies demonstrate that placement of the modification needs to be carefully considered in order to maximise aptamer-target interactions.

The second approach aims to utilise aptamers as target recognition elements in liposomes for biosensing applications. DNA-lipid conjugates comprise either steroidal or long hydrocarbon units which anchor the molecule into liposomes by partial insertion into the lipid bilayer. To date, a conjugate does not exist where the steroidal moiety spans the entire bilayer. Here, a novel lipid tag with two consecutive steroidal units is created, designed to anchor the DNA-lipid conjugate across the whole bilayer, providing more stabilization to the supramolecular structure. The design and synthesis of the bischolesterol molecule and its attachment to DNA to create transmembrane molecules is described *via* four different synthetic strategies, each one with varying success. Azide-alkyne cycloaddition and phosphoramidite coupling techniques are employed to couple the bischolesterol molecule to aptameric sequences of DNA in solution and on solid supports *via* automated DNA synthesis. Cy3 and Cy5 fluorophores are also attached to the bischolesterol unit as signalling components for the system.



An overview of the work described in this thesis is given below:

## Chapter 1- Introduction

This chapter describes the structure and function of oligonucleotides and aptamers. A review of aptamer modifications is also provided.

## Chapter 2- Techniques

This chapter gives a general overview of the techniques used to synthesise and study the unmodified and modified DNA aptamer sequences detailed in this thesis.

## Chapter 3- Acrylamide modified aptamers for incorporation into molecularly imprinted polymers

This chapter describes the first approach undertaken in modifying aptamers by creating a novel acrylamide moiety to allow incorporation of aptamers into molecularly imprinted polymers. Acrylamide modified thymidine was used to place acrylamide groups in various positions of the thrombin binding aptamer sequence. The acrylamide modified strands were successfully crosslinked into acrylamide-based polymers and were observed to bind to thrombin and complementary DNA targets.

## Chapter 4- Lipid-aptamer conjugates for biosensing with liposomes

This chapter describes the second approach undertaken in modifying aptamers by creating a novel bischolesterol-dye tag for inserting lipid-aptamer conjugates into liposomes. DNA-bischolesterol and DNA-bischolesterol-dye conjugates were successfully synthesised *via* three different synthetic

strategies, purified by reversed phase HPLC, and characterized by mass spectrometry. The constructs afforded have the potential to bind to target molecules and ultimately be adopted into biosensing systems with liposomes.

## Chapter 5- Experimental

This chapter provides a detailed description of the experimental procedures employed during the investigations performed in this thesis.

## Chapter 6- Appendix

This chapter contains additional data obtained during the investigations performed in this thesis.

## Acknowledgements

First of all, I would like to acknowledge my supervisors Professor Paula Mendes and Professor James Tucker for giving me the opportunity to undertake such a challenging but rewarding project. I am forever grateful for their advice, guidance, and constant positive attitude throughout my PhD.

I also wish to thank the members of the Tucker and Mendes groups, past and present, for their help in all aspects of my PhD- Huy, Jean-Louis, Purbani, Rosie, James, Gemma, Jon, David, Haydn, Holly, Media, Aysha, Klaudia, Ed, Georgina, Charlotte, Liyao, Jake, Alex, Aaron, Lewis, Chiranjib, Stefano, Eduardo, Marcos, Josh G, Josh N, Barbara G, Barbara S, Monika, Fatima, Lydia, Zarrar, Eleonora, Philippa, Kam Yazmin, James, Giuseppe, Rosalia, Pushpa, and Alice. Apologies if I've forgotten to include anyone else!

There are others from the department outside of my research groups who I also wish to thank for their friendship and assistance in what can occasionally be an isolating environment- Shani, Greg, Dennis, Owen, Michael, Phil, Liz, Marcus, Abhishek, Gary, Ash, Will, John, Siobhan, Nat, Sophie, Louise, Nick, and Russell.

I would also like to acknowledge the staff at the Analytical facility for their advice and support. I personally believe that I would not have been able to finish this programme without the help of Chi, Peter, Allen, Cécile, Neil and Louise.

Finally, I would like to thank my close friends and family for their continued support through the ups and downs of the past four years. To my friends

Crystal, Stephen, Begum, and Carl for putting up with my 'living too far away' and thesis writing excuses, and to my brothers Ronnie, Andrew, and Darron, and my cousin Fiifi for providing me with laughs and drama to distract me from my studies. Special thanks goes to my mother for her guidance (and relentless threats to disown me if I did not finish this PhD).

I hope you find reading this thesis as fun as I did writing it...

# Contents

## Abbreviations

1. Introduction	1
1.1 Structure of nucleic acids	1
1.1.1 Nucleotide structure	1
1.1.2 Oligonucleotide structure	2
1.2 Biological function of nucleic acids	5
1.3 Aptamers	7
1.3.1 Aptamer structure	7
1.3.2 Aptamer sequence selection (SELEX)	8
1.3.3 Aptamer function	10
1.4 Aptamer modifications	13
1.4.1 Nucleobase modifications	14
1.4.2 Sugar modifications	16
1.4.3 Phosphate backbone modifications	18
1.4.4 End strand modifications	19
1.4.5 Other modifications	20
1.5 Summary	22
1.6 Aim	23
1.7 List of references	24
2. Techniques	31
2.1 Automated solid phase DNA synthesis	32

2.1.1	Synthesis of nucleoside phosphoramidite monomers	34
2.1.2	Attachment of protected nucleosides to solid supports	36
2.1.3	DNA synthesis process	37
2.2	Purification of oligonucleotides	40
2.2.1	High performance liquid chromatography (HPLC)	40
2.2.2	Polyacrylamide gel electrophoresis (PAGE)	41
2.3	Characterisation of oligonucleotides by mass spectrometry	43
2.3.1	Ionisation methods	43
2.3.2	Mass analysers	44
2.3.3	Interpretation of mass spectra	45
2.4	Circular dichroism spectroscopy (CD)	47
2.5	Thermal melting ( $T_m$ )	48
2.6	Gel electromobility shift assay (gel EMSA)	50
2.7	Isothermal titration calorimetry (ITC)	51
2.8	Förster resonance energy transfer (FRET)	53
2.9	Contact angle	55
2.10	Ellipsometry	57
2.11	X-ray photoelectron spectroscopy (XPS)	59
2.12	Surface plasmon resonance (SPR)	60
2.13	List of references	61

3. Acrylamide modified aptamers for incorporation into molecularly imprinted polymers	64
3.1 Introduction	64
3.2 Aim	68
3.3 Synthesis of thymidine acrylate and thymidine acrylamide	72
3.4 <sup>1</sup> H NMR homo- and co-polymerisation experiments with thymidine acrylate and thymidine acrylamide	73
3.4.1 <sup>1</sup> H homopolymerisation experiments	73
3.4.2 <sup>1</sup> H copolymerisation experiments	74
3.5 Contact angle experiments with thymidine acrylate and thymidine acrylamide	78
3.6 Testing the stability of thymidine acrylamide towards the oligonucleotide deprotection process	81
3.7 Synthesis of acrylamide modified thrombin binding aptamer strands	86
3.7.1 Synthesis of thymidine acrylamide phosphoramidite	86
3.7.2 Synthesis of acrylamide modified aptamer strands	87
3.7.3 Purification of acrylamide modified aptamer strands	88
3.8 Circular dichroism of unmodified and acrylamide modified thrombin binding aptamer strands	94
3.9 Gel electromobility shift assay of unmodified and acrylamide modified thrombin binding aptamer strands with thrombin	96

3.10	Thermal melting experiments of unmodified and acrylamide modified thrombin binding aptamer strands with complementary sequence	98
3.11	DNA-acrylamide gel copolymerisation experiments	100
3.12	Surface coupling experiments with unmodified and acrylamide modified thrombin binding aptamer strands	102
3.12.1	Contact angle results	102
3.12.2	Ellipsometry results	104
3.12.3	XPS results	105
3.12.4	Summary	107
3.13	Optimisation of APS concentration used in acrylamide DNA coupling experiments	109
3.13.1	HPLC experiments	109
3.13.2	Contact angle experiments	112
3.13.3	Ellipsometry experiments	114
3.13.4	XPS experiments	114
3.13.5	Summary	115
3.14	Conclusions and future work	117
3.15	List of references	119
4.	Lipid-aptamer conjugates for biosensing with liposomes	122
4.1	Introduction	122
4.2	Aim	126
4.3	Synthesis of bischolesterol derivatives	131
4.3.1	Synthesis of symmetrical bischolesterol derivatives	131
4.3.2	Synthesis of asymmetrical bischolesterol derivatives	135



4.3.3	Synthesis of DMT protected bischolesterol phosphoramidite	136
4.4	Synthesis of cyanine dyes	138
4.4.1	Synthesis of cyanine dye precursors	138
4.4.2	Synthesis of Cy3 monoalkyne	140
4.4.3	Synthesis of Cy5 monoalkyne	142
4.5	Synthesis of bischolesterol-dye conjugates	145
4.5.1	Synthesis of bischolesterol dye azides	145
4.5.2	Synthesis of bischolesterol dye alcohols	147
4.5.3	Synthesis of bischolesterol dye phosphoramidites	149
4.6	DNA selection and synthesis	151
4.6.1	DNA sequence selection	151
4.6.2	Splitting the thrombin binding aptamer sequence for sandwich-based sensing	151
4.6.3	DNA synthesis for strategies 1-2	153
4.7	Synthesis and purification of transmembrane molecules	155
4.7.1	Strategy 1- coupling alkyne modified DNA to bischolesterol dye azide	155
4.7.2	Strategy 2- coupling DBCO modified DNA to bischolesterol dye azide	159
4.7.3	Strategy 3- coupling bischolesterol dye phosphoramidite to DNA	161
4.7.4	Strategy 4- using the DNA synthesiser to couple components	161
4.7.5	Coupling bischolesterol to DNA	164
4.7.6	Purification of transmembrane molecules	165
4.8	Conclusions and future work	171
4.9	List of references	172

5. Experimental	176
5.1 Materials and methods	176
5.2 Synthesis of thymidine acrylate and thymidine acrylamide	177
5.3 <sup>1</sup> H NMR homo- and co- polymerisation experiments with thymidine acrylate and thymidine acrylamide	179
5.3.1 Homopolymerisation experiments	179
5.3.2 Copolymerisation experiments	179
5.4 Contact angle experiments with thymidine acrylate and thymidine acrylamide	180
5.4.1 Fabrication of acrylamide SAMs	180
5.4.2 Copolymerisation of acrylamide SAMs with modified bases	180
5.4.3 Contact angle measurements	181
5.5 Testing the stability of thymidine acrylamide towards the oligonucleotide deprotection process	182
5.6 Synthesis of thymidine acrylamide derivatives	183
5.6.1 Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine	184
5.6.2 Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine	185
5.6.3 Synthesis of 5'-O-(4,4'-dimethoxytrityl)-3'-O-[2-cyanoethoxy-(N,N'-diisopropylamino)-phosphino]-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine	187
5.7 Synthesis and purification of oligonucleotides	189
5.7.1 Standard synthesis of unmodified and alkyne modified strands	189

5.7.2	Ultramild synthesis of acrylamide, DBCO, and bischolesterol modified strands, and strategy 4 derived transmembrane molecules	190
5.7.3	Purification of unmodified, acrylamide modified, and alkyne modified strands	191
5.7.4	Purification of DCBO modified strands	192
5.7.5	Oligonucleotide characterisation	192
5.8	Polyacrylamide gel experiments	193
5.8.1	DNA-acrylamide gel copolymerisation experiments	193
5.8.2	Gel electromobility shift assay of unmodified and acrylamide modified thrombin binding aptamer strands with thrombin	193
5.8.3	Gel electromobility shift assay of full and split thrombin binding aptamer strands with thrombin	194
5.9	Circular dichroism of unmodified and acrylamide modified thrombin binding aptamer strands	195
5.10	Thermal melting experiments of unmodified and acrylamide modified thrombin binding aptamer strands with complementary sequence	196
5.11	ITC experiments on the binding of full and split thrombin binding aptamer strands with thrombin	197
5.12	Surface coupling experiments with unmodified and acrylamide modified thrombin binding aptamer strands	198
5.12.1	Reaction of acrylamide SAMs with unmodified/ acrylamide unmodified aptamer strands at high APS concentrations	198

5.12.2	Reaction of acrylamide SAMs with unmodified/ acrylamide unmodified aptamer strands at low APS concentrations	198
5.12.3	Contact angle measurements	198
5.12.4	Ellipsometry measurements	199
5.12.5	XPS measurements	199
5.13	Optimisation of APS concentration used in acrylamide DNA coupling experiments by HPLC analysis	200
5.14	Synthesis of symmetrical bischolesterol derivatives	201
5.14.1	Synthesis of lithocholic acid methyl ester	202
5.14.2	Synthesis of bischolesterol diester	203
5.14.3	Synthesis of bischolesterol dialcohol	204
5.14.4	Synthesis of bischolesterol dibromide	205
5.14.5	Synthesis of bischolesterol diazide	206
5.14.6	Crystal structure determination of bischolesterol diester, bischolesterol dibromide, and bischolesterol diazide	208
5.15	Synthesis of asymmetrical bischolesterol derivatives	211
5.15.1	Synthesis of bischolesterol monoester/alcohol	212
5.15.2	Synthesis of bischolesterol monoester/bromide	213
5.15.3	Synthesis of bischolesterol monoester/azide	214
5.15.4	Synthesis of bischolesterol monoazide/alcohol	215
5.16	Synthesis of DMT protected bischolesterol phosphoramidite	217
5.16.1	Synthesis of mono DMT protected bischolesterol alcohol	218
5.16.2	Synthesis of DMT protected bischolesterol phosphoramidite	219
5.17	Synthesis of cyanine dyes	221
5.17.1	Synthesis of 1,2,3,3-tetramethyl-3H-indolinium iodide	222

5.17.2	Synthesis of 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide	223
5.17.3	Synthesis of malondialdehyde bis(phenylimine) monohydrochloride	224
5.17.4	Synthesis of 1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolinium iodide	225
5.17.5	Synthesis of Cy3 monoalkyne	226
5.17.6	Synthesis of Cy5 monoalkyne	228
5.18	Synthesis of bischolesterol dye conjugates	230
5.18.1	General procedure	231
5.18.2	HPLC purification of bischolesterol dye conjugates	231
5.18.3	Synthesis of bischolesterol Cy3 azide	232
5.18.4	Synthesis of bischolesterol Cy5 azide	233
5.18.5	Synthesis of bischolesterol Cy3 alcohol	234
5.18.6	Synthesis of bischolesterol Cy5 alcohol	235
5.19	Synthesis of bischolesterol dye phosphoramidites (failed)	237
5.20	Synthesis and purification of transmembrane molecules	239
5.20.1	Strategy 1- coupling alkyne modified DNA to bischolesterol azide derivatives with copper(I) iodotriethylphosphite	239
5.20.2	Strategy 2- coupling DBCO modified DNA to bischolesterol dye azides	239
5.20.3	Purification of transmembrane molecules	240
5.21	List of references	242
6.	Appendix	243
6.1	NMR data for thymidine acrylate, thymidine acrylamide, and thymidine acrylamide derivatives	243
6.1.1	Compound <b>1</b> - thymidine acrylate <sup>1</sup> H and <sup>13</sup> C NMR	243
6.1.2	Compound <b>2</b> - thymidine acrylamide <sup>1</sup> H and <sup>13</sup> C NMR	244

6.1.3	Compound <b>3</b> - 5'-O-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine <sup>1</sup> H and <sup>13</sup> C NMR	245
6.1.4	Compound <b>4</b> - 5'-O-(4,4'-dimethoxytrityl)-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine <sup>1</sup> H and <sup>13</sup> C NMR	246
6.1.5	Compound <b>5</b> - 5'-O-(4,4'-dimethoxytrityl)-3'-O-[2-cyanoethoxy-(N,N'-diisopropylamino)-phosphino]-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine <sup>1</sup> H, <sup>13</sup> C, and <sup>31</sup> P NMR	247
6.2	NMR data for bischolesterol derivatives	249
6.2.1	Compound <b>6</b> - lithocholic acid methyl ester <sup>1</sup> H and <sup>13</sup> C NMR	249
6.2.2	Compound <b>7</b> - bischolesterol diester <sup>1</sup> H and <sup>13</sup> C NMR	250
6.2.3	Compound <b>8</b> - bischolesterol dialcohol <sup>1</sup> H and <sup>13</sup> C NMR	251
6.2.4	Compound <b>9</b> - bischolesterol dibromide <sup>1</sup> H and <sup>13</sup> C NMR	252
6.2.5	Compound <b>10</b> - bischolesterol diazide <sup>1</sup> H and <sup>13</sup> C NMR	253
6.2.6	Compound <b>11</b> - bischolesterol monoester/alcohol <sup>1</sup> H and <sup>13</sup> C NMR	254
6.2.7	Compound <b>12</b> - bischolesterol monoester/bromide <sup>1</sup> H and <sup>13</sup> C NMR	255
6.2.8	Compound <b>13</b> - bischolesterol monoester/azide <sup>1</sup> H and <sup>13</sup> C NMR	256
6.2.9	Compound <b>14</b> - bischolesterol monoazide/alcohol <sup>1</sup> H and <sup>13</sup> C NMR	257
6.2.10	Compound <b>15</b> - mono DMT protected bischolesterol alcohol <sup>1</sup> H and <sup>13</sup> C NMR	258
6.2.11	Compound <b>16</b> - DMT protected bischolesterol phosphoramidite <sup>1</sup> H, <sup>13</sup> C, and <sup>31</sup> P NMR	259
6.3	NMR data for cyanine dye derivatives	261
6.3.1	Compound <b>17</b> - 1,2,3,3-tetramethyl-3H-indolinium iodide <sup>1</sup> H and <sup>13</sup> C NMR	261
6.3.2	Compound <b>18</b> - 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide <sup>1</sup> H and <sup>13</sup> C NMR	262

6.3.3	Compound <b>19</b> - malondialdehyde bis(phenylimine) monohydrochloride <sup>1</sup> H and <sup>13</sup> C NMR	263
6.3.4	Compound <b>20</b> - 1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolium iodide <sup>1</sup> H and <sup>13</sup> C NMR	264
6.3.5	Compound <b>21</b> - Cy3 monoalkyne <sup>1</sup> H and <sup>13</sup> C NMR	265
6.3.6	Compound <b>22</b> - Cy5 monoalkyne <sup>1</sup> H and <sup>13</sup> C NMR	266
6.4	NMR data for bischolesterol-dye conjugates	267
6.4.1	Compound <b>23</b> - bischolesterol Cy3 azide <sup>1</sup> H and <sup>13</sup> C NMR	267
6.4.2	Compound <b>24</b> - bischolesterol Cy5 azide <sup>1</sup> H and <sup>13</sup> C NMR	268
6.4.3	Compound <b>25</b> - bischolesterol Cy3 alcohol <sup>1</sup> H and <sup>13</sup> C NMR	269
6.4.4	Compound <b>26</b> - bischolesterol Cy5 alcohol <sup>1</sup> H and <sup>13</sup> C NMR	270
6.5	Crystal structure images of bischolesterol dibromide and bischolesterol diazide	271
6.5.1	Compound <b>9</b> - bischolesterol dibromide crystal structure image	271
6.5.2	Compound <b>10</b> - bischolesterol diazide crystal structure image	271
6.6	Mass spectrometry characterisation of oligonucleotides	272
6.7	Analytical HPLC chromatograms of oligonucleotides	273
6.7.1	Acryl-T16	273
6.7.2	Acryl-T7/T9	274
6.7.3	Acryl-T12/T13	275
6.7.4	Acryl-T3/T4	276
6.7.5	Acryl-T3/T12	277
6.7.6	Acryl-T4/T13	278
6.7.7	Acryl-T7/T9/T16	279
6.7.8	TBA15	280
6.7.9	TBA15 complementary	281
6.7.10	TBA29	282
6.7.11	TBA15-A	283
6.7.12	TBA15-B	284

6.7.13	TBA15-Aex	285
6.7.14	TBA15-Bex	286
6.7.15	TBA15-alkyne	287
6.7.16	TBA15-DBCO	288
6.7.17	TBA29-alkyne	289
6.7.18	TBA29-DBCO	290
6.7.19	nina-alkyne	291
6.7.20	nina-DBCO	292
6.8	Analytical HPLC chromatograms of bischolesterol dye azides	293
6.8.1	Compound <b>23</b> - bischolesterol Cy3 azide	293
6.8.2	Compound <b>24</b> - bischolesterol Cy5 azide	294
6.9	Mass spectrometry characterisation of transmembrane molecules	295
6.10	$T_m$ data for thermal melting experiments of acrylamide modified thrombin binding aptamer strands with complementary sequence	296
6.10.1	First thermal melting experiment	296
6.10.2	Second thermal melting experiment	296
6.10.3	Third thermal melting experiment	297
6.11	Data for control HPLC experiments with acrylamide DNA and BAC with/without TEMED and APS	298
6.11.1	BAC and Acryl-T16	298
6.11.2	BAC and Acryl- T7/T9	298
6.11.3	BAC and Acryl-T7/T9/T16	299
6.11.4	BAC and TEMED with increasing concentrations of APS	299
6.11.5	Acryl-T16 and TEMED with increasing concentrations of APS	300
6.11.6	Acryl- T7/T9 and TEMED with increasing concentrations of APS	300
6.11.7	Acryl-T7/T9/T16 and TEMED with increasing concentrations of APS	301



## Abbreviations

A	<u>A</u> denine
APS	<u>A</u> mmonium persulphate
BAC	N,N'-bis( <u>a</u> cryloyl) <u>c</u> ystamine
BE	<u>B</u> inding <u>e</u> nergy
C	<u>C</u> ytosine
CD	<u>C</u> ircular <u>d</u> ichroism
CPG	<u>C</u> ontrolled pore glass
DABCO	1,4- <u>d</u> iaza <u>b</u> icyclo[2.2.2] <u>o</u> ctane
DBCO	<u>D</u> ibenzoc <u>y</u> clo <u>o</u> ctyne
DCM	<u>D</u> ichloro <u>m</u> ethane
DIPEA	N,N- <u>d</u> iisopropylethylamine
DMAP	4- <u>D</u> imethylaminopyridine

DMF	<u>D</u> imethyl <u>f</u> ormamide
DMSO	<u>D</u> imethyl <u>s</u> ulfoxide
DMT	<u>D</u> imethoxy <u>t</u> rityl
DNA	<u>D</u> eoxyribo <u>n</u> ucleic <u>a</u> cid
EDTA	<u>E</u> thylene <u>d</u> iamine <u>t</u> etraacetic <u>a</u> cid
EMSA	<u>E</u> lectromobility <u>s</u> hift <u>a</u> ssay
ESI	<u>E</u> lectrospray ionisation
FANA	2'- <u>F</u> <u>a</u> rabinose <u>n</u> ucleic <u>a</u> cid
FRET	<u>F</u> örster <u>r</u> esonance <u>e</u> nergy <u>t</u> ransfer
G	<u>G</u> uanine
HMBC	<u>H</u> eteronuclear <u>m</u> ultiple <u>b</u> ond <u>c</u> orrelation
HNA	<u>H</u> exitol <u>n</u> ucleic <u>a</u> cid
HPLC	<u>H</u> igh- <u>p</u> erformance <u>l</u> iquid <u>c</u> hromatography

HSQC	<u>H</u> eteronuclear <u>s</u> ingle <u>q</u> uantum <u>c</u> oherence
HRMS	<u>H</u> igh <u>r</u> esolution <u>m</u> ass <u>s</u> pectrometry
IMFP	<u>I</u> nelastic <u>m</u> ean <u>f</u> ree <u>p</u> ath
IR	<u>I</u> nfrared
ITC	<u>I</u> sothermal <u>t</u> itration <u>c</u> alorimetry
KE	<u>K</u> inetic <u>e</u> nergy
LC	<u>L</u> iquid <u>c</u> hromatography
LNA	<u>L</u> ocked <u>n</u> ucleic <u>a</u> cid
LOC	<u>L</u> ipid- <u>o</u> ligonucleotide <u>c</u> onjugate
$m/z$	Mass ( <u>m</u> ) to charge ( <u>z</u> ) ratio
MALDI	<u>M</u> atrix <u>a</u> ssisted <u>l</u> aser <u>d</u> esorption <u>i</u> onisation
MIP	<u>M</u> olecularly <u>i</u> mprinted <u>p</u> olymer
mRNA	<u>M</u> essenger <u>r</u> ibon <u>u</u> cleic <u>a</u> cid

MS	<u>M</u> ass <u>s</u> pectrometry
NBS	<u>N</u> - <u>b</u> romo <u>s</u> uccinimide
NMR	<u>N</u> uclear <u>m</u> agnetic <u>r</u> esonance
PAGE	<u>P</u> oly <u>a</u> crylamide gel <u>e</u> lectrophoresis
PCR	<u>P</u> olymerase <u>c</u> hain <u>r</u> eaction
R <sub>f</sub>	<u>R</u> etention <u>f</u> actor
RNA	<u>R</u> ibo <u>n</u> ucleic <u>a</u> cid
RP	<u>R</u> eversed <u>p</u> hase
rRNA	<u>R</u> ibosomal <u>r</u> ibo <u>n</u> ucleic <u>a</u> cid
RT	<u>R</u> oom <u>t</u> emperature
SAM	<u>S</u> elf- <u>a</u> ssembled <u>m</u> onolayer
SELEX	<u>S</u> ystematic <u>e</u> volution of <u>l</u> igands by <u>e</u> xponential enrichment
SOMAmers	<u>S</u> low <u>o</u> ff-rate <u>m</u> odified <u>a</u> ptam <u>e</u> rs

SPR	<u>S</u> urface <u>p</u> lasmon <u>r</u> esonance
T	<u>T</u> hymine
TBA	<u>T</u> hrombin <u>b</u> inding <u>a</u> ptamer
TBA	<u>T</u> ributyl <u>a</u> mine
TBE	<u>T</u> ris- <u>b</u> orate- <u>e</u> thylenediaminetetraacetic acid
TBTA	<u>T</u> ris( <u>b</u> enzyl <u>t</u> riazolylmethyl) <u>a</u> mine
TCA	<u>T</u> richloro <u>a</u> cetic acid
TCEP	<u>T</u> ris(2- <u>c</u> arboxy <u>e</u> thyl) <u>p</u> hosphine
TEA	<u>T</u> riethyl <u>a</u> mine
TEAA	<u>T</u> riethyl <u>a</u> mmonium <u>a</u> cetate
TEG	<u>T</u> riethyleneglycol
TEMED	N,N,N',N'- <u>t</u> etram <u>e</u> thylethylene <u>d</u> iamine
TFA	<u>T</u> ri <u>f</u> luoro <u>a</u> cetic acid

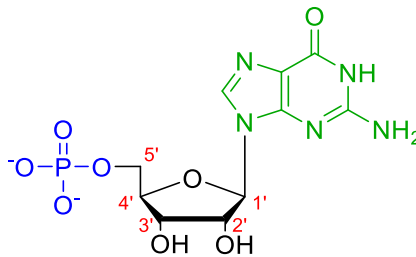
THF	<u>T</u> etra <u>h</u> ydro <u>f</u> uran
THPTA	<u>T</u> ris( <u>h</u> ydroxypropyl <u>t</u> riazolylmethyl) <u>a</u> mine
TLC	<u>T</u> hin <u>l</u> ayer <u>c</u> hromatography
$T_m$	<u>T</u> hermal <u>m</u> elting temperature
TMS	<u>T</u> etra <u>m</u> ethyl <u>s</u> ilane
TNA	<u>T</u> hreose <u>n</u> ucleic <u>a</u> cid
TOC	<u>T</u> otal <u>o</u> rganic <u>c</u> ontent
TOF	<u>T</u> ime <u>o</u> f <u>f</u> light
tRNA	<u>T</u> ransfer <u>r</u> ibo <u>n</u> ucleic <u>a</u> cid
U	<u>U</u> racil
UNA	<u>U</u> nlocked <u>n</u> ucleic <u>a</u> cid
UV/vis	<u>U</u> ltravio <u>l</u> et/ <u>v</u> isible
XPS	<u>X</u> -ray photoelectron <u>s</u> pectroscopy

# Chapter 1 Introduction

## 1.1 Structure of nucleic acids

### 1.1.1 Nucleotide structure

Nucleic acids are biopolymers essential for the existence and evolution of all living organisms. The monomers of these biopolymers are called nucleotides and consist of three groups- a phosphomonoester, a ribose sugar, and an N-heterocyclic base. The phosphomonoester is attached to the 5' position of the sugar and the N-heterocyclic base is attached to the 1' position of the sugar (Figure 1.1).<sup>1</sup>

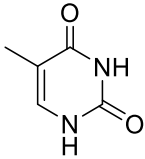
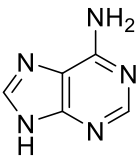
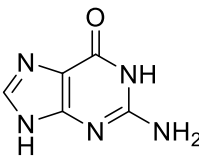
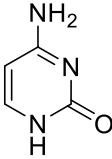
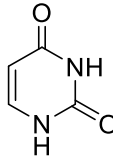
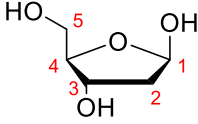
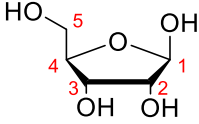


**Figure 1.1** Structure of a nucleotide. The deprotonated phosphomonoester is highlighted in blue, the ribose sugar in black (with numbered positions), and the N-heterocyclic group in green.

The phosphomonoester in Figure 1.1 is deprotonated in physiological conditions. This group is responsible for the acidity and negative charge of nucleic acids. The sugar of the nucleotide can either be ribose or deoxyribose (missing a hydroxyl group in the 2' position of the sugar). The name ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) is determined by the type of sugar used in the biopolymer. The N-heterocyclic base of nucleotides can either be monocyclic or bicyclic. The former are classed as pyrimidines and the latter

classed as purines. Naturally, nucleotides comprise of two purines and three pyrimidines, as shown in Table 1.1. Sequence labelling of nucleic acids denotes the first letter of each base for each nucleobase, hence A for adenine, G for guanine, C for cytosine, T for thymine, and U for uracil.<sup>1</sup> Thymine, uracil, cytosine, and guanine can exist as aromatic tautomers but are more stable in their amide-like ring forms.

**Table 1.1** Different N-heterocyclic bases and ribose sugars found in nucleic acids.

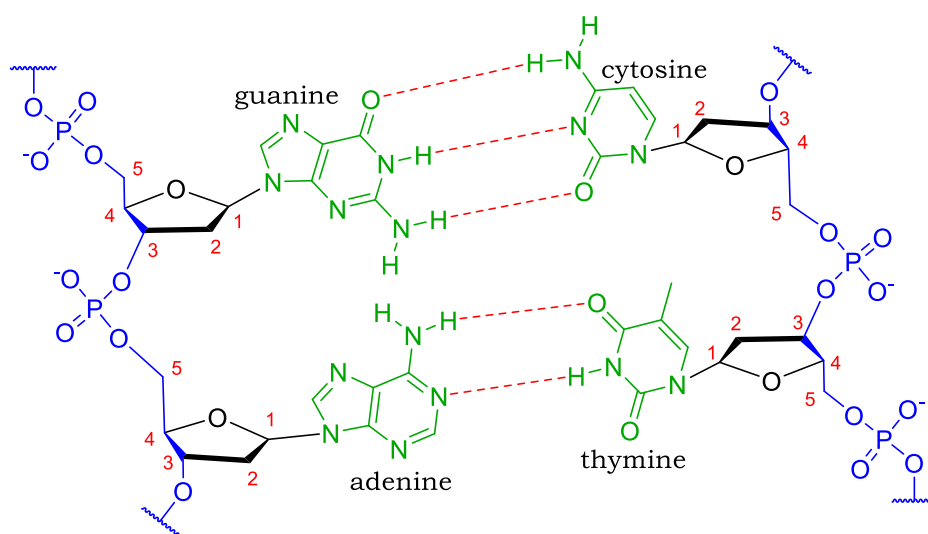
	DNA only	DNA and RNA			RNA only
bases	 thymine	 adenine	 guanine	 cytosine	 uracil
sugars	 2-deoxyribose				 ribose

### 1.1.2 Oligonucleotide structure

Nucleotides are linked by phosphodiester bonds to form oligonucleotides. Condensation polymerisation links the phosphate group of one nucleotide to the 3' position of the ribose sugar of another nucleotide. In living organisms, the linkages are formed among nucleotide triphosphates, eliminating pyrophosphate in the process. The primary structure of an oligonucleotide is a linear single strand. Hydrogen bonding between nucleobases either within the same strand or between different strands lead to higher ordered structures of oligonucleotides. Guanine pairs with cytosine with three hydrogen bonds, and adenine pairs with thymine (in DNA) or uracil (in RNA) with two hydrogen bonds



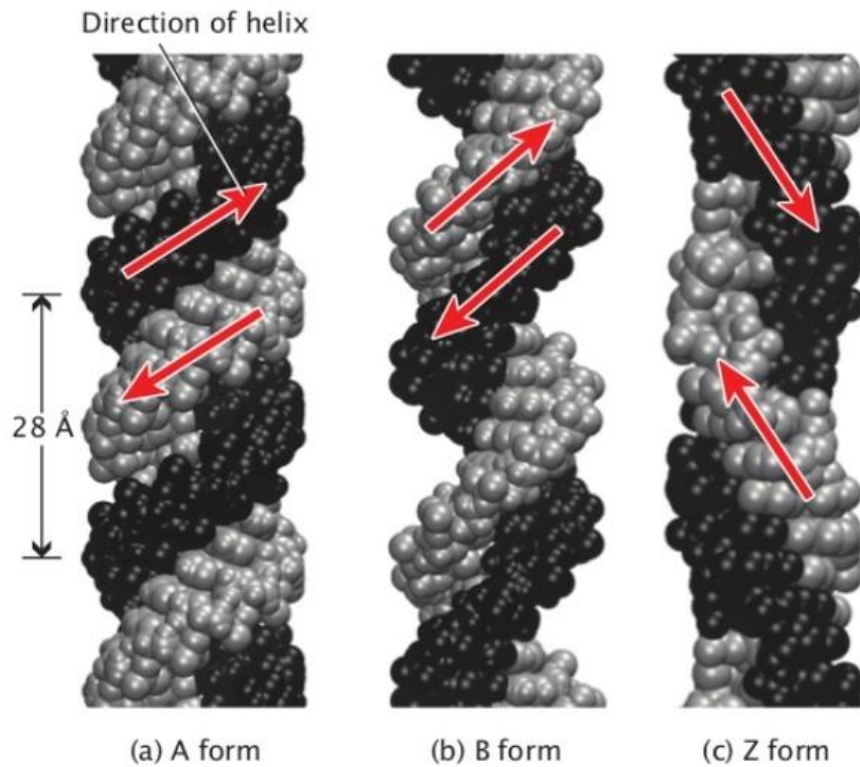
(Figure 1.2). The pairing of two strands of oligonucleotides with base sequences complementary to one another results in a double helical structure. Hydrophobic interactions and  $\pi$  stacking of the bases further stabilise the helix. In nature, this structure is more commonly observed in DNA than RNA, although double stranded RNA has been detected in some viruses. RNA in other living organisms exists as highly folded single oligonucleotide strands.<sup>2</sup>



**Figure 1.2** Hydrogen bonding of nucleobase pairs in DNA showing three hydrogen bonds between guanine and cytosine, and two hydrogen bonds between adenine and thymine. The phosphodiester backbone is shown in blue.

The structure of biological DNA was correctly identified as a double helix by Francis Crick and James Watson in the early 1950's.<sup>3</sup> The helix consists of complimentary oligonucleotide strands running antiparallel to one another, i.e. the 3' end of one strand is paired with the 5' end of the other and vice versa. It was later discovered that the dimensions and geometries of the helix can vary, but three main conformations are known (Figure 1.3). The predominant conformation for duplexes found in aqueous solutions is B-DNA; where the helix is wound in a right-handed direction. A-DNA is a shorter and fatter version of

the B form and is rarely found under physiological conditions. Z-DNA is a left-handed version of the helix and a transient form of DNA.<sup>4</sup>

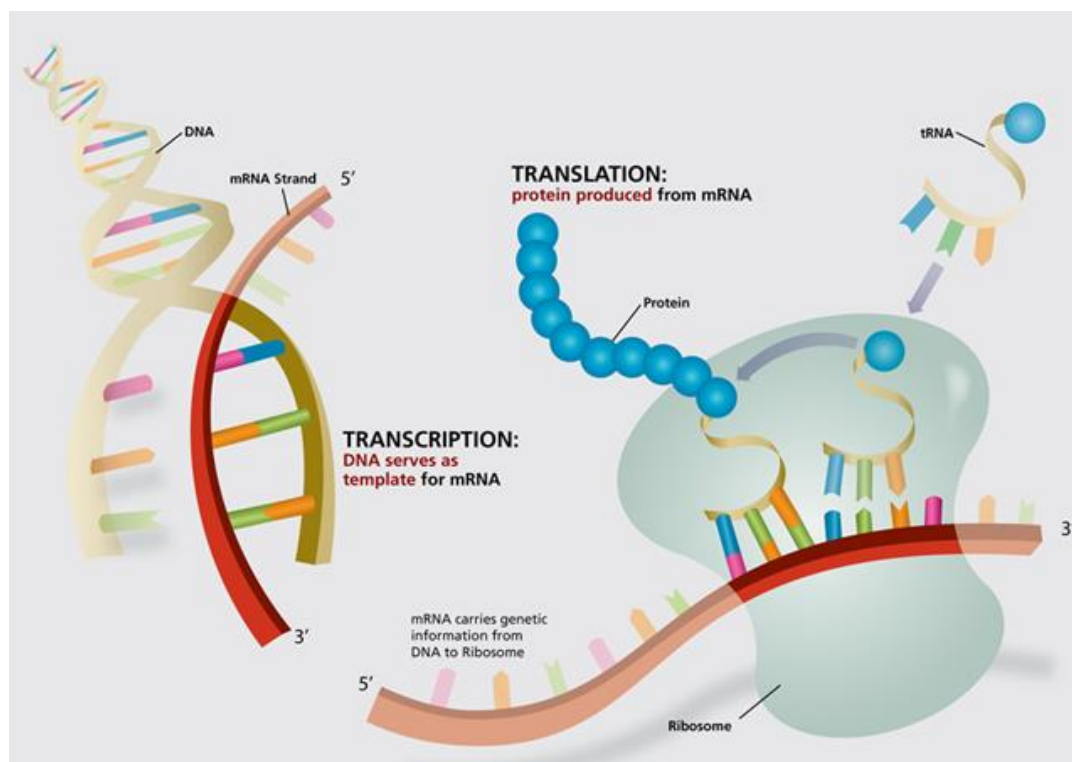


**Figure 1.3** The three different conformations of the DNA double helix.<sup>4</sup>

## 1.2 Biological function of nucleic acids

DNA is found in high molecular weights in the nuclei and mitochondria of eukaryotic cells, and the nucleoids of prokaryotic cells.<sup>5</sup> In eukaryotes, DNA is packed with proteins called histones to form chromosomes. In prokaryotes, various proteins are used to compact DNA. RNA is more abundant in living organisms but have much lower molecular weights than DNA. RNA predominantly exists as either ribosomal RNA (rRNA), messenger RNA (mRNA), or transfer RNA (tRNA). These three types of RNA serve different but relatable functions in cells.

Nucleic acids are the blueprint for the biosynthesis of proteins. As well as catalysing reactions in the form of enzymes, partaking in molecular recognition and signalling, and transporting small molecules; proteins also provide structural integrity to cells in the form of keratin or collagen.<sup>6</sup> Protein biosynthesis occurs *via* the transcription and translation of genes. Genes are DNA base sequences that correlate to the synthesis of specific polypeptides; the primary structure of proteins. Genes also contain hereditary information which is passed on from parents to offspring.<sup>7</sup> Transcription involves the transfer of genetic information from DNA to mRNA, and translation uses the mRNA, along with tRNA and rRNA in ribosomes, to generate polypeptide sequences. Both processes are assisted and catalysed by enzymes and other proteins (Figure 1.4).<sup>8</sup>



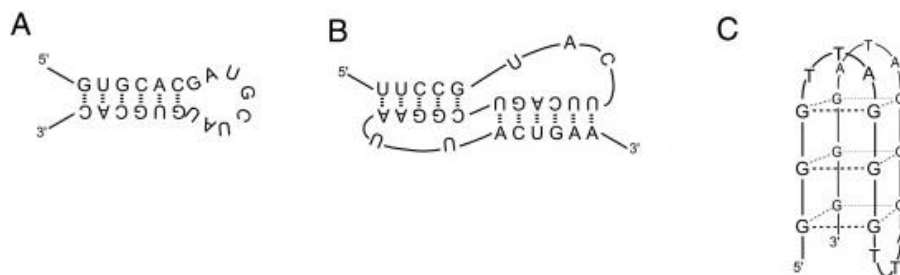
**Figure 1.4** Transcription and translation processes.<sup>9</sup>

Alongside storing genetic information for protein biosynthesis, nucleic acids also function in the catalysis of biological processes in the form of DNAzymes and RNAzymes.<sup>10</sup> RNA riboswitches control gene expression in prokaryotes by binding to small molecule metabolites with high affinity and selectivity, altering the synthesis of polypeptides encoded by mRNA in the process. There is growing interest in engineering catalytic nucleic acids and riboswitches to develop our understanding of the alternative functions of nucleic acids.<sup>11, 12</sup> There is also interest in using these oligonucleotides or chemically synthesised derivatives for gene expression, antisensing (the use of single stranded oligonucleotide probes to sense complementary sequences), and sensing of proteins and small molecule metabolites.<sup>13</sup> Automated solid phase chemical synthesis of oligonucleotides and in vitro amplification of nucleic acids by the polymerase chain reaction (PCR) has greatly enhanced the opportunities of finding nucleic acid sequences for such applications.<sup>14</sup>

## 1.3 Aptamers

### 1.3.1 Aptamer structure

Aptamers are classed as synthetically derived single stranded oligonucleotides that structurally conform to bind to specific targets.<sup>15</sup> The name ‘aptamer’ is derived from the Latin word *aptus* meaning ‘to join’, and the Greek word *meros*, meaning ‘parts’.<sup>16</sup> Aptamers can consist of either DNA or RNA sequences. Self-complementary regions within the sequence allow the aptamer to fold into secondary or three-dimensional structures that fit target molecules with high affinity and specificity (Figure 1.5). Targets include small molecules,<sup>17-21</sup> nucleic acids,<sup>22-24</sup> proteins,<sup>25-29</sup> viruses,<sup>30</sup> bacteria,<sup>31, 32</sup> and whole cells.<sup>33-35</sup> Depending on the structure of the target and the environmental conditions that are necessary, the interactions between aptamer and target are mainly non-covalent and so include van der Waals interactions, hydrogen bonding, and electrostatic forces. A number of secondary structure motifs for aptamers have been described, including G-quadruplexes,<sup>36</sup> hairpin structures,<sup>37</sup> and pseudoknots.<sup>38</sup> Three-dimensional structures of these biopolymers have not been as easy to define. However, a few 3D structures, mainly of the thrombin binding aptamer, have been determined by NMR spectroscopy<sup>39</sup> and X-ray crystallography.<sup>40</sup>



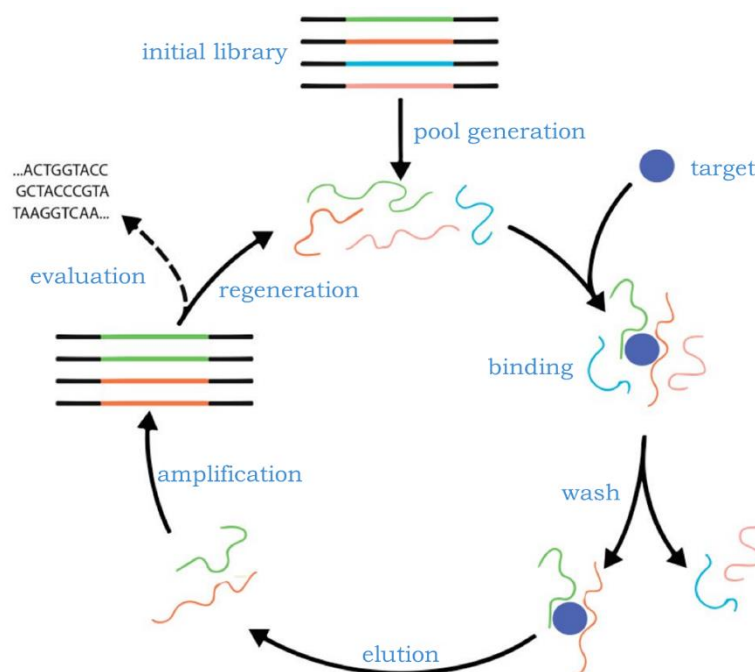
**Figure 1.5** Common secondary structures of aptamers (adapted from the review ‘Aptamers: molecules of great potential’<sup>15</sup>). A- hairpin, B- pseudoknot, C- G-quartet/quadruplex.

### 1.3.2 Aptamer sequence selection (SELEX)

Aptamer sequences are selected through a form of combinatorial chemistry called Systematic Evolution of Ligands by EXponential enrichment (SELEX). The process is termed so because it entails the removal of non-target binding sequences, making the process evolutionary. Exponential enrichment refers to the PCR amplification of selected target binding sequences.<sup>41</sup> It was first developed in the early nineties by two independent research groups; Tuerk and Gold selected RNA sequences showing high affinity to the bacteriophage T4 DNA polymerase gp43 from a collection of random RNA oligonucleotides,<sup>25</sup> whereas Ellington and Szostak used a similar procedure to target RNA aptamers for small organic dyes.<sup>42</sup>

The process starts with a large library or pool of chemically synthesised DNA or RNA oligonucleotides ( $\sim 10^{15}$  oligos)<sup>43</sup>. The sequences typically contain a region of 20-80 random nucleotides and are flanked with primer binding sites necessary for PCR amplification. The library is incubated with the chosen target and non-target binding oligos are partitioned from the mixture. The partitioning method implemented depends on the target selected. For large targets such as cells or proteins, membrane filters are used to filter off unbound oligos, whereas for smaller targets, the molecules are immobilised onto solid supports allowing for easy washing of unbound oligos. Target-bound oligos are isolated and amplified by PCR for DNA sequences, or reverse transcription PCR for RNA sequences. If symmetric PCR is used, the resulting double stranded DNA is transformed into a new oligonucleotide library by separating the relevant single stranded DNA or by in vitro transcription for RNA. This new library is then used for the next round of SELEX. The cycle is repeated until the initial

oligonucleotide pool is reduced to a small number of precise sequence motifs which show high affinity and specificity for the chosen target (Figure 1.6).<sup>44</sup>



**Figure 1.6** Aptamer selection process (SELEX).<sup>45</sup>

Usually, the stringency of each cycle is gradually increased by reducing the target concentration and/or by changing the binding and washing conditions (i.e. incubation time, buffer conditions) in later SELEX rounds. Negative and counter SELEX rounds are typically introduced to remove sequences showing affinity for filter membranes/solid supports and for molecules which are similar in structure to the target molecule. The success of the process is largely dependent on the design of the initial oligonucleotide library, target structure and concentration, ratio of target molecules to oligonucleotides, selection conditions (e.g. temperature, pH, salt), the efficiency of removing unbound oligos, and the elution of bound sequences.<sup>16</sup>

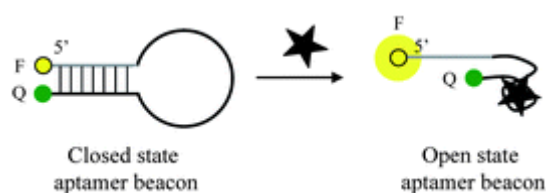
### 1.3.3 Aptamer function

The ability of aptamers to bind selectively to targets permits them to be utilised in a variety of therapeutic and diagnostic applications. Therapeutics is the treatment of disease. For therapeutic applications, aptamers can behave as agonists: species which interact with and activate cell receptors, as well as antagonists: species which inhibit protein-protein or receptor-ligand interactions. An example of this is the aptamer-based drug Macugen, which works as an antagonist to VEGF165, a protein involved in angiogenesis (the development of new blood vessels) and permeability (the leakage of blood vessels). It is used to treat age-related macular degeneration, an eye disease that can result in blindness if left untreated.<sup>46</sup> Also, aptamers specific for cell surface receptors are used to deliver drugs and other compounds to cells. There are many reports on aptamer mediated drug delivery of anticancer drugs to cancerous cells for therapeutic effects.<sup>47-50</sup>

In terms of diagnostics, aptamers have been incorporated into biosensors to detect a variety of biological targets. Biosensors are devices which utilise biological molecules for the detection of analytes. The device consists of a biologically derived recognition element (the aptamer in this case), which interacts with a target of interest, and a physiochemical detector, which translates the interaction to an observable signal. Aptamers are often considered as superior alternatives to popular antibody based systems as they exhibit similar target affinities but are more stable, reversibly denatured by temperature, easily amenable to chemical modifications, and can potentially be developed for any analyte.<sup>51</sup> Upon target binding, aptamers usually undergo conformational changes which can be detected by a number of techniques, including UV/vis analysis or circular dichroism.<sup>52, 53</sup> Mass based detection



techniques such as surface plasmon resonance are also used to identify aptamer-target binding.<sup>54</sup> Commonly, aptamer conformation changes in biosensors are detected using colorimetric, fluorescent or electrochemical techniques with oligonucleotides which are modified with fluorescent or electrochemical tags. Fluorescently labelled aptamers are occasionally called ‘apta-beacons’. Usually the fluorescence of these apta-beacons is switched off by a quencher group on a complementary sequence present either within the same strand or duplexed with a different strand. Recognition of the target by the aptamer results in separation of the fluorescent tag and the quencher group, allowing the fluorophore to emit light (Figure 1.7). Förster Resonance Energy Transfer (FRET) is a mechanism by which energy is transferred between two fluorophores. The amount of energy transferred is dependent on the distance between the groups. It can be used to correlate shifts in fluorophore emission wavelengths to aptamer conformational changes upon target binding<sup>15</sup> and is also used to determine structures of aptamers.<sup>55</sup>



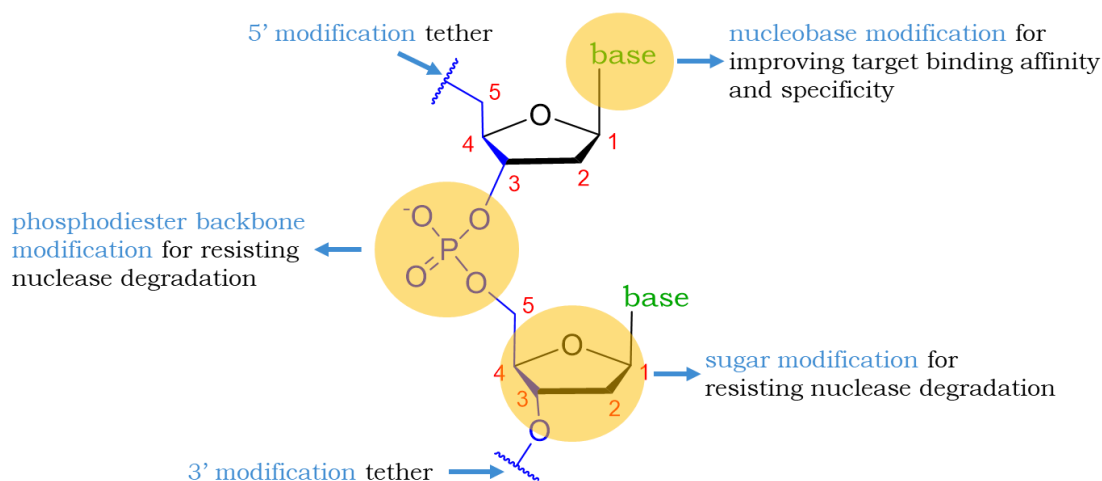
**Figure 1.7** Fluorescence signal emitted upon target recognition of a fluorescently tagged aptamer.<sup>56</sup>

Electrochemical detection using redox-tagged aptamers involves immobilising aptamers onto an electrode surface and measuring the effects of aptamer-target binding on the current flow through the system. The conformational change of the aptamer upon binding to the target affects the distance of the redox tag to the electrode surface, which in turn varies the rate of electron transfer.<sup>57</sup>

Alongside biosensing capabilities, aptamers can be used to separate targets from mixtures by immobilising the aptamer onto a solid support.<sup>58</sup> Oligonucleotides with aptameric tags showing high affinity to solid phase materials (e.g. cellulose) can be used to isolate sequences from mixtures of DNA or RNA.<sup>59</sup>

## 1.4 Aptamer modifications

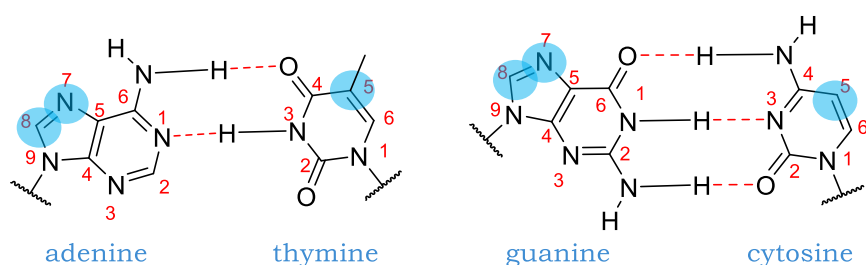
Unmodified aptamers have several limitations when it comes to their use in therapeutic and diagnostic applications. Their small size leads to rapid renal clearance and they are easily degraded by nuclease enzymes, making them unideal for therapeutics and for analysing clinical samples.<sup>60</sup> Fortunately, as aptamers are synthetically derived, various modifications can be applied to avoid these limitations. Alongside improving aptamer stability to renal clearance and nuclease degradation, these modifications are also adopted to enhance aptamer-target interactions and/or to incorporate transducing elements for biosensing applications. Sites of modification can take place on the nucleobase, sugar, phosphate backbone, or tethered to the end of the strand, and can be applied before, during, or after sequence selection (SELEX) (Figure 1.8). Most modifications are introduced post-SELEX due to the inability of polymerase enzymes to recognise the modified groups in the amplification stage of the SELEX process.<sup>61</sup> When added post-SELEX, structure activity relationship experiments are performed to determine the effect of the modification on aptamer-target interactions. A few examples of such modifications will now be discussed.



**Figure 1.8** Sites for aptamer modifications.

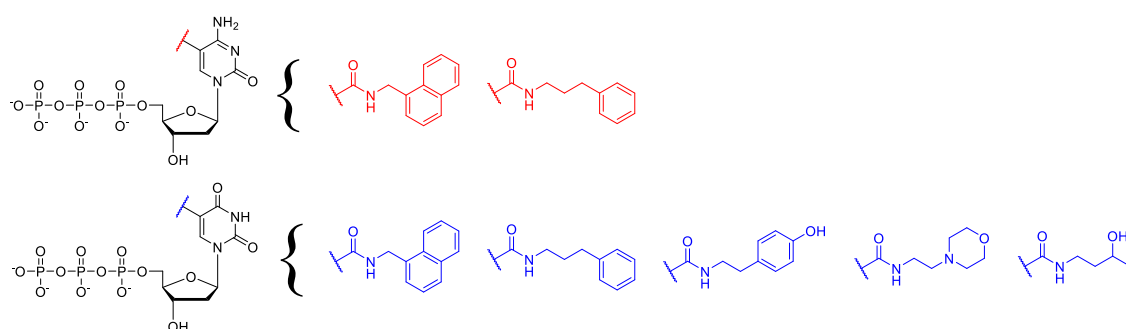
#### 1.4.1 Nucleobase modifications

Nucleobase modifications are usually applied to improve target binding affinity and selectivity, as the chemical diversity of unmodified aptamers is restricted to the five original nucleobases- A, T, C, G, and U.<sup>62</sup> The C5 site of pyrimidines and N7/C8 sites of purines are often chosen for modification as they are more synthetically accessible than other sites and do not interfere with base pairing and other hydrogen bonding interactions (Figure 1.9).<sup>63</sup> Affinities are enhanced *via* additional hydrophobic, hydrophilic, and charged interactions provided by the incorporated groups.



**Figure 1.9** Preferred modification sites on nucleobases (highlighted in blue).

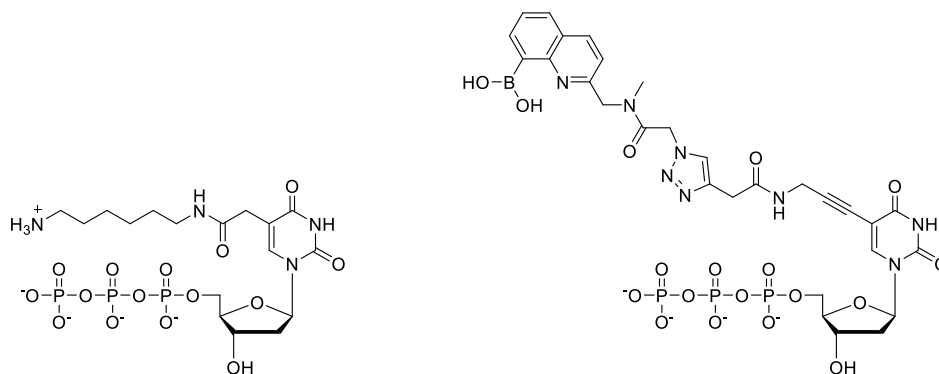
Functional groups placed at the C5 position of pyridines and the N7 position of purines by rigid alkyne, alkene or amide linkers are often tolerated by DNA polymerases and can thus be introduced into oligonucleotide libraries in SELEX rounds.<sup>64</sup> One notable example of such is the use of SOMAmers (Slow Off-rate Modified Aptamers) which are prepared from modified nucleoside triphosphates bearing hydrophobic amino acid side chains on the C5 position of cytosine and thymine (Figure 1.10).<sup>65</sup> These modified aptamers were found to increase affinities with protein targets by providing additional interactions with hydrophobic regions of the protein. The protein-aptamer complexes formed were observed to have significantly lower (10-1000 fold) dissociation rates than their unmodified counterparts, which led to the terminology of these modified sequences.



**Figure 1.10** Structures of some of the modified nucleoside triphosphates used to form SOMAmers.

As well as improving aptamer-protein interactions, modified nucleoside triphosphates have also been used to enhance aptamer-sugar binding (Figure 1.11). Sawai and coworkers modified thymidine residues by adding charged cationic groups to improve interactions with the oligosaccharide sialyllactose.<sup>66</sup> The theory behind this was that the cationic group could ionically interact with the anionic carboxylic acid group of sialic acid and enhance binding. Wang and

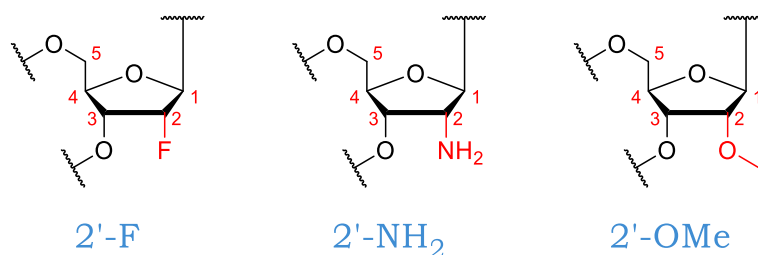
coworkers added boronic acid moieties to thymidine residues to direct DNA aptamer binding to the glycan part of the glycoprotein fibrinogen, as boronic acids are widely known to bind to diols and hydroxyl groups.<sup>67</sup> In both cases the modified aptamers showed higher affinities to selected targets than their unmodified counterparts.



**Figure 1.11** Structure of modified nucleoside triphosphates used to enhance aptamer-sugar interactions.

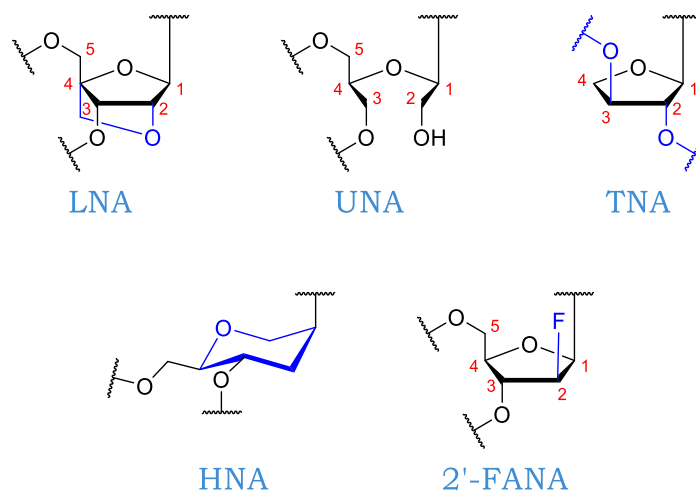
#### 1.4.2 Sugar modifications

Sugar modifications are adopted to increase the stability of aptamers to thermal and nuclease degradation. They are usually placed at the C2 or C2 and C4 positions of the sugar and are more often found on pyrimidine-based nucleosides as these are more susceptible to nuclease degradation.<sup>68</sup> Popular modifications of the C2 position include 2'-F, 2'-NH<sub>2</sub>, and 2'-OMe substituents (Figure 1.12). These substituents are compatible with certain engineered polymerases and thus can be used in SELEX rounds as modified nucleoside triphosphates.<sup>69</sup> The aptamer drug Macugen (mentioned in section 1.3.3) contains modified 2'-F pyrimidine residues which were incorporated into SELEX rounds and 2'-OMe purine residues which were added post-SELEX.<sup>46</sup>



**Figure 1.12** C2 sugar modifications used to increase stability of aptamers to nuclease degradation. Modifications are highlighted in red.

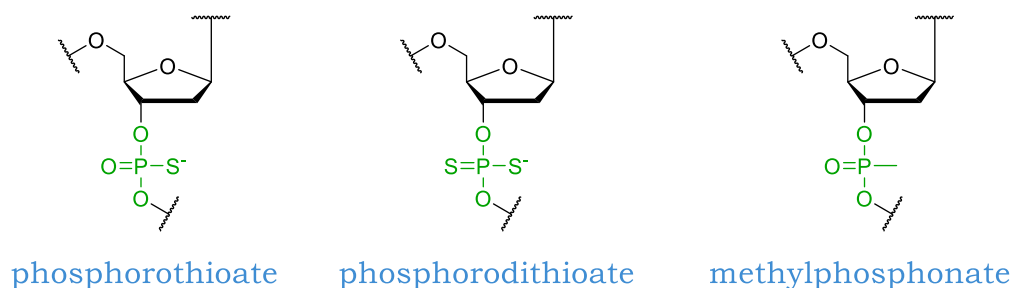
Locked nucleic acids (LNA) have a methylene bridge linking the 2'-oxygen to the C4 position of the sugar (Figure 1.13). These modifications have been incorporated into aptamer sequences post SELEX and were found to enhance both thermal stability and stability to nuclease digestion, albeit sometimes at the expense of target binding.<sup>70</sup> Jarstfer *et al* observed changes to aptamer-target affinities in the thrombin binding aptamer sequence depending on where the LNA sugar was placed.<sup>71</sup> Other sugar modifications utilised to enhance aptamer stability to nuclease degradation include unlocked nucleic acid (UNA),<sup>72</sup> threose nucleic acid (TNA),<sup>73</sup> hexitol nucleic acid (HNA),<sup>74</sup> and 2'-F arabinose nucleic acid (FANA)<sup>75</sup> (Figure 1.13).



**Figure 1.13** Different sugar modifications used to enhance stability of aptamers to nuclease degradation. Modifications are highlighted in blue.

### 1.4.3 Phosphate backbone modifications

As with sugar modifications, the main reasoning behind phosphate backbone modifications is to increase aptamer resistance to nuclease degradation. The most routinely used backbone modifications either replace one or both non-bridging oxygens with sulphur to give phosphorothioates and phosphorodithioates respectively (Figure 1.14).<sup>76</sup> Although effective, complete substitution for sulphur containing linkages can result in heightened non-specific interactions reducing target selectivity.<sup>77</sup> Another type of phosphate modification to use are methylphosphonates where one of the non-bridging oxygens is replaced with a methyl group, resulting in a neutrally charged linkage (Figure 1.14). Mergny and coworkers used methylphosphonate modifications to study the effect of charge loss on the structure of the thrombin binding aptamer.<sup>78</sup> They observed that replacing the negative phosphate backbone with neutral methylphosphonate groups destabilised the aptamer structure. Although novel methods are being developed to connect nucleosides *via* non-phosphate derived linkages for antisensing applications (e.g. triazole linked DNA)<sup>79</sup>, few examples of these technologies have so far been employed in aptamer research.<sup>80</sup>

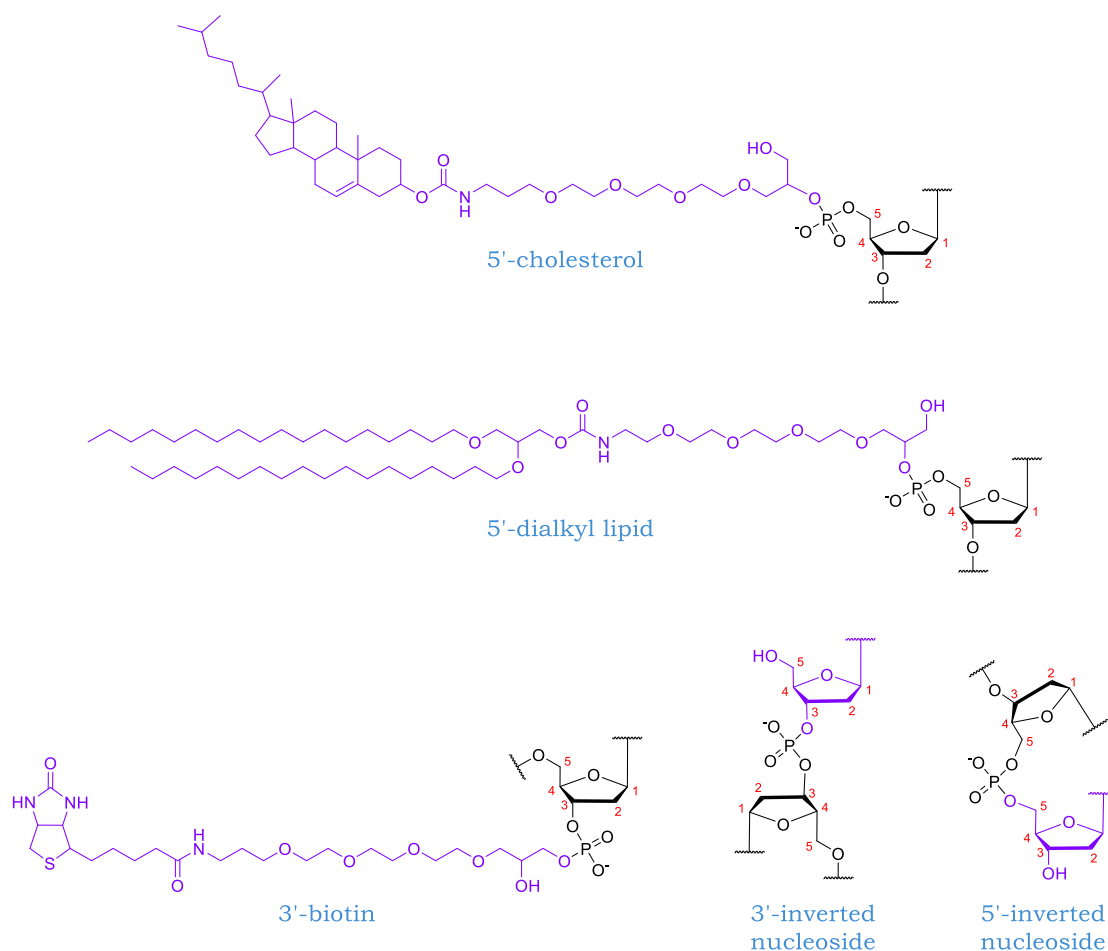


**Figure 1.14** Phosphate backbone modifications used to increase aptamer resistance to nuclease degradation. Modifications are highlighted in green.



#### 1.4.4 End strand modifications

Numerous end group modifications have been developed to overcome the limitations that arise for unmodified aptamers. These modifications are often introduced post SELEX as phosphoramidites during the solid phase oligonucleotide synthesis process. Examples of these include 5'-cholesterol<sup>81</sup> and 5'-dialkyl lipids<sup>82</sup> for enhancing cellular uptake of aptamers and resisting renal clearance, and also 3'-biotin<sup>83</sup> and 3'/5'-inverted nucleosides<sup>84</sup> for increasing aptamer stability to nuclease degradation (Figure 1.15). Alternatively, phosphoramidites of small functional handles such as amines, thiols, alkynes, and carboxylic acids are added instead which are then subsequently reacted to connect the aptamer to an array of substrates and transducing elements. Examples include attachment of strands to gold surfaces *via* thiol-gold connections,<sup>85</sup> and attachment of fluorescent and electrochemical tags *via* amide or triazole linkages.<sup>86-88</sup>

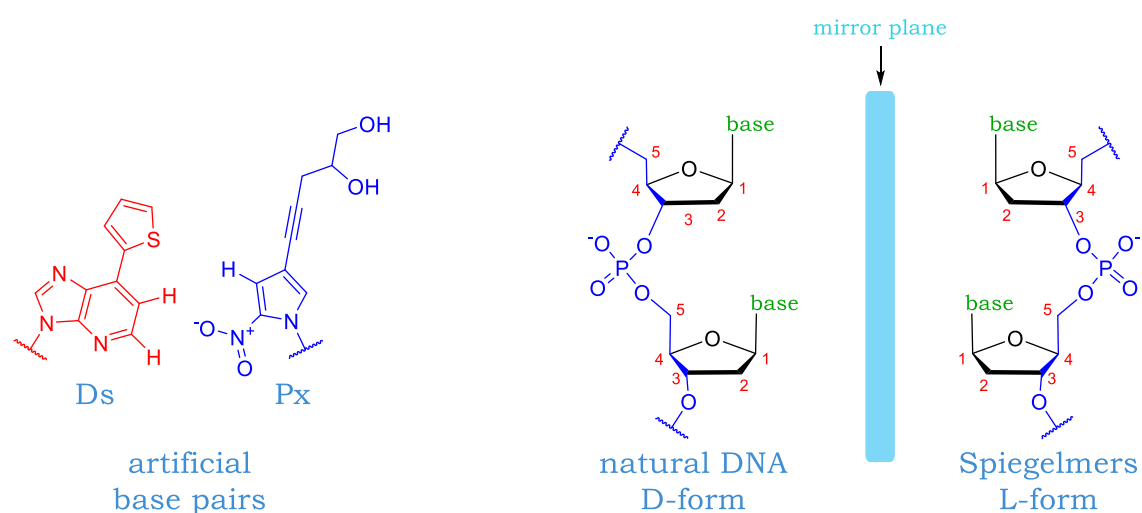


**Figure 1.15** Examples of 3' and 5' end group modifications. Modifications are highlighted in purple.

#### 1.4.5 Other modifications

In addition to the few examples listed above, other noticeable modifications include the use of artificial base pairs to enhance aptamer-target binding and specificity, and Spiegelmers to prevent nuclease degradation (Figure 1.16). Hirao *et al* incorporated hydrophobic unnatural base pairs termed Ds and Px into SELEX rounds to interact with a number of protein targets.<sup>89</sup> The modified aptamers yielded were found to have affinities that were >100 fold greater than sequences only containing natural bases. Spiegelmers are mirror images of natural oligonucleotides.<sup>90</sup> They exist as L-forms whereas natural oligonucleotides are D-forms. Nucleases do not recognise the L-form of

oligonucleotides, making Spiegelmers ideal candidates for preventing nuclease degradation. Spiegelmer based aptamers are chemically synthesised using sequences of D-form aptamers that express high affinities to the enantiomer of the desired target, as DNA polymerases do not recognise the L-forms of nucleoside triphosphates.



**Figure 1.16** Aptamer modifications using artificial base pairs and Spiegelmers.

## 1.5 Summary

Nucleic acids are an important class of biomolecules vital for the existence of living organisms. Alongside being the blueprint for the production of proteins, they also function in the catalysis of numerous biological processes in the form of DNAzymes and RNAzymes. Research into developing synthetic derivatives of these biomolecules particularly for use in sensing applications has grown significantly. One field which has spanned from this research is the study and use of aptamers- synthetically derived single stranded oligonucleotides that structurally conform to bind to specific targets. They are often championed as better alternatives for biosensing to commonly-used antibodies as they are inherently more stable, are easily amenable to chemical modifications, and can potentially be developed for any analyte whilst exhibiting similar if not greater affinities and selectivities. Although aptamers have some limitations, the major being their susceptibility to nuclease degradation, various modifications can be applied to avoid these limitations.

## 1.6 Aim

This thesis aims to contribute to the growing field of modified aptamer development by introducing two vastly different but innovative modifications for incorporating aptamers into biosensing applications. The first modification aims to incorporate aptamers into molecularly imprinted polymers, or MIPs. Often noted as a more stable substitute to both antibodies and aptamers, MIPs are formed when the analyte of interest is imprinted into a polymer. Although there are already a handful of reports on aptamer incorporated MIPs, there is no mention on how the flexibility of the aptamer affects the sensing capabilities of the aptamer-MIP hybrids. This approach aims to address this by studying how the placement of novel acrylamide modified thymidine residues into the thrombin binding aptamer sequence affects the ability of the modified strands to bind to their desired target. The second modification aims to utilise aptamers as target recognition elements in liposomes for biosensing applications. DNA-lipid conjugates comprise of either steroidal or long hydrocarbon units which anchor the molecule into liposomes by partial insertion into the lipid bilayer. At present, a conjugate does not exist where a steroidal moiety spans the entire bilayer. Here, a novel lipid tag with two steroidal units is created, designed to anchor the DNA-lipid conjugate across the whole bilayer, providing more stabilization to the supramolecular structure. Cy3 and Cy5 fluorophores are also attached to the bischolesterol unit as signalling components for the system. The constructs afforded have the potential to bind to target molecules and ultimately be adopted into biosensing systems with liposomes. Literature methods more specific to the two approaches are further explored in the introduction sections of Chapter 3 and Chapter 4.

## 1.7 List of references

1. J. Clayden, N. Greeves, S. Warren and P. Wothers, *Organic chemistry*, Oxford University Press, Oxford, 2000.
2. G. M. Blackburn and M. J. Gait, *Nucleic Acids in Chemistry and Biology*, Royal Society of Chemistry, Cambridge, 1990.
3. J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737-738.
4. L. A. Pray, *Nature Education*, 2008, **1**, 100.
5. A. Griswold, *Nature Education*, 2008, **1**, 57.
6. D. Hames and N. Hooper, *Biochemistry*, Taylor & Francis, New York, 2005.
7. L. E. Rosenberg, *Human Genes and Genomes: Science, Health, Society*, Elsevier Science, Burlington, 2012.
8. Nucleic Acids,  
<https://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/nucacids.htm>, (accessed 05/09/2018).
9. Transcription & Translation Illustration, <http://www.ignyc.com/our-work/portfolio/transcription-translation/>, (accessed 05/09/2018).
10. Z. J. Xu, L. F. Yang, L. Q. Sun and Y. Cao, *Chin. Sci. Bull.*, 2012, **57**, 3404-3408.
11. K. H. Link and R. R. Breaker, *Gene therapy*, 2009, **16**, 1189-1201.
12. J. Li, A. A. Green, H. Yan and C. Fan, *Nat. Chem.*, 2017, **9**, 1056-1067.
13. S. Ausländer and M. Fussenegger, *Curr. Opin. Biotechnol.*, 2017, **48**, 54-60.
14. F. Groher and B. Suess, *Biochim. Biophys. Acta, Gene Regul. Mech.*, 2014, **1839**, 964-973.
15. F. Radom, P. M. Jurek, M. P. Mazurek, J. Otlewski and F. Jeleń, *Biotechnol. Adv.*, 2013, **31**, 1260-1274.

16. R. Stoltenburg, C. Reinemann and B. Strehlitz, *Biomol. Eng*, 2007, **24**, 381-403.
17. D. E. Huizenga and J. W. Szostak, *Biochemistry*, 1995, **34**, 656-665.
18. D. Mann, C. Reinemann, R. Stoltenburg and B. Strehlitz, *Biochem. Biophys. Res. Commun.*, 2005, **338**, 1928-1934.
19. A. Shoji, M. Kuwahara, H. Ozaki and H. Sawai, *J. Am. Chem. Soc.*, 2007, **129**, 1456-1464.
20. Y. S. Kim, C. J. Hyun, I. A. Kim and M. B. Gu, *Bioorg. Med. Chem.*, 2010, **18**, 3467-3473.
21. K. M. Song, E. Jeong, W. Jeon, M. Cho and C. Ban, *Anal. Bioanal. Chem.*, 2012, **402**, 2153-2161.
22. C. Boiziau, E. Dausse, L. Yurchenko and J.-J. Toulmé, *J. Biol. Chem.*, 1999, **274**, 12730-12737.
23. F. Duconge, Eacute, Eacute, RIC and J.-J. Toulme, *RNA*, 1999, **5**, 1605-1614.
24. J. H. Ko, B. Cho, J. K. Ahn, Y. Lee and I. Park, *Bull. Korean Chem. Soc.*, 1999, **20**, 1335-1339.
25. C. Tuerk and L. Gold, *Science*, 1990, **249**, 505-510.
26. L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas and J. J. Toole, *Nature*, 1992, **355**, 564-566.
27. L. A. Jones, L. E. Clancy, W. D. Rawlinson and P. A. White, *Antimicrob. Agents Chemother.*, 2006, **50**, 3019-3027.
28. J. Ruckman, L. S. Green, J. Beeson, S. Waugh, W. L. Gillette, D. D. Henninger, L. Claesson-Welsh and N. Janjić, *J. Biol. Chem.*, 1998, **273**, 20556-20567.
29. F. Ylera, R. Lurz, V. A. Erdmann and J. P. Fürste, *Biochem. Biophys. Res. Commun.*, 2002, **290**, 1583-1588.

30. W. Pan, R. C. Craven, Q. Qiu, C. B. Wilson, J. W. Wills, S. Golovine and J. F. Wang, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 11509-11513.
31. F. Chen, J. Zhou, F. Luo, A.-B. Mohammed and X.-L. Zhang, *Biochem. Biophys. Res. Commun.*, 2007, **357**, 743-748.
32. C. L. A. Hamula, H. Zhang, L. L. Guan, X.-F. Li and X. C. Le, *Anal. Chem.*, 2008, **80**, 7812-7819.
33. Y. J. Lee and S. W. Lee, *J. Microbiol. Biotechnol.*, 2006, **16**, 1149-1153.
34. D. Shangguan, Y. Li, Z. Tang, Z. C. Cao, H. W. Chen, P. Mallikaratchy, K. Sefah, C. J. Yang and W. Tan, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 11838-11843.
35. C. Wang, M. Zhang, G. Yang, D. Zhang, H. Ding, H. Wang, M. Fan, B. Shen and N. Shao, *J. Biotechnol.*, 2003, **102**, 15-22.
36. R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe and J. Feigon, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 3745-3749.
37. P. Fan, A. K. Suri, R. Fiala, D. Live and D. J. Patel, *J. Mol. Biol.*, 1996, **258**, 480-500.
38. J. R. Lorsch and J. W. Szostak, *Biochemistry*, 1994, **33**, 973-982.
39. P. Schultze, R. F. Macaya and J. Feigon, *J. Mol. Biol.*, 1994, **235**, 1532-1547.
40. K. Padmanabhan, K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler and A. Tulinsky, *J. Biol. Chem.*, 1993, **268**, 17651-17654.
41. A. Cibiel, D. M. Dupont and F. Ducongé, *Pharmaceuticals*, 2011, **4**, 1216-1235.
42. A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818-822.
43. E. J. Cho, J.-W. Lee and A. D. Ellington, *Annu. Rev. Anal. Chem.*, 2009, **2**, 241-264.
44. S. Gopinath, *Anal. Bioanal. Chem.*, 2007, **387**, 171-182.



45. A. E. G. Cass and Y. Zhang, *Faraday Discuss.*, 2011, **149**, 49-61.
46. E. W. M. Ng, D. T. Shima, P. Calias, E. T. Cunningham, D. R. Guyer and A. P. Adamis, *Nat Rev Drug Discov*, 2006, **5**, 123-132.
47. E. Kim, Y. Jung, H. Choi, J. Yang, J.-S. Suh, Y.-M. Huh, K. Kim and S. Haam, *Biomaterials*, 2010, **31**, 4592-4599.
48. V. Bagalkot, O. C. Farokhzad, R. Langer and S. Jon, *Angew. Chem. Int. Ed.*, 2006, **45**, 8149-8152.
49. S. Dhar, F. X. Gu, R. Langer, O. C. Farokhzad and S. J. Lippard, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 17356-17361.
50. Y.-F. Huang, D. Shangguan, H. Liu, J. A. Phillips, X. Zhang, Y. Chen and W. Tan, *ChemBioChem*, 2009, **10**, 862-868.
51. W. Zhou, P.-J. Jimmy Huang, J. Ding and J. Liu, *Analyst*, 2014, **139**, 2627-2640.
52. Q. Yang, I. J. Goldstein, H. Y. Mei and D. R. Engelke, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 5462-5467.
53. S. Nagatoishi, Y. Tanaka and K. Tsumoto, *Biochem. Biophys. Res. Commun.*, 2007, **352**, 812-817.
54. M. d. Toro, R. Gargallo, R. Eritja and J. Jaumot, *Anal. Biochem.*, 2008, **379**, 8-15.
55. D. M. J. Lilley and T. J. Wilson, *Curr. Opin. Chem. Biol.*, 2000, **4**, 507-517.
56. A. Sassolas, L. J. Blum and B. D. Leca-Bouvier, *Analyst*, 2011, **136**, 257-274.
57. A. E. Radi, *Int. J. Electrochem.*, 2011, **2011**, 1-17.
58. Q. Deng, I. German, D. Buchanan and R. T. Kennedy, *Anal. Chem.*, 2001, **73**, 5415-5421.
59. B. J. Boese and R. R. Breaker, *Nucleic Acids Res.*, 2007, **35**, 6378-6388.

60. A. D. Keefe, S. Pai and A. Ellington, *Nat Rev Drug Discov*, 2010, **9**, 537-550.
61. S. Ni, H. Yao, L. Wang, J. Lu, F. Jiang, A. Lu and G. Zhang, *Int. J. Mol. Sci.*, 2017, **18**, 1683-1704.
62. M. R. Dunn, R. M. Jimenez and J. C. Chaput, *Nat. Rev. Chem.*, 2017, **1**, 0076.
63. A. Shivalingam and T. Brown, *Biochem. Soc. Trans.*, 2016, **44**, 709-715.
64. P. Röthlisberger and M. Hollenstein, *Adv. Drug Delivery Rev.*, 2018, DOI: 10.1016/j.addr.2018.1004.1007.
65. B. N. Gawande, J. C. Rohloff, J. D. Carter, I. von Carlowitz, C. Zhang, D. J. Schneider and N. Janjic, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 2898-2903.
66. M. M. Masud, M. Kuwahara, H. Ozaki and H. Sawai, *Biorg. Med. Chem.*, 2004, **12**, 1111-1120.
67. M. Li, N. Lin, Z. Huang, L. Du, C. Altier, H. Fang and B. Wang, *J. Am. Chem. Soc.*, 2008, **130**, 12636-12638.
68. T. K. Sharma, J. G. Bruno and A. Dhiman, *Biotechnol. Adv.*, 2017, **35**, 275-301.
69. T. Chen, N. Hongdilokkul, Z. Liu, R. Adhikary, S. S. Tsuen and F. E. Romesberg, *Nat Chem*, 2016, **8**, 556-562.
70. R. E. Wang, H. Wu, Y. Niu and J. Cai, *Curr Med Chem*, 2011, **18**, 4126-4138.
71. L. Bonifacio, F. Church and M. Jarstfer, *Int. J. Mol. Sci.*, 2008, **9**, 422-433.
72. T. B. Jensen, J. R. Henriksen, B. E. Rasmussen, L. M. Rasmussen, T. L. Andresen, J. Wengel and A. Pasternak, *Biorg. Med. Chem.*, 2011, **19**, 4739-4745.

73. H. Mei, J.-Y. Liao, R. M. Jimenez, Y. Wang, S. Bala, C. McCloskey, C. Switzer and J. C. Chaput, *J. Am. Chem. Soc.*, 2018, **140**, 5706-5713.
74. G. Kolb, S. Reigadas, C. Boiziau, A. van Aerschot, A. Arzumanov, M. J. Gait, P. Herdewijn and J.-J. Toulmé, *Biochemistry*, 2005, **44**, 2926-2933.
75. I. Alves Ferreira-Bravo, C. Cozens, P. Holliger and J. J. DeStefano, *Nucleic Acids Res.*, 2015, **43**, 9587-9599.
76. D. Volk and G. Lokesh, *Biomedicines*, 2017, **5**, 41-61.
77. S. Gao, X. Zheng, B. Jiao and L. Wang, *Anal. Bioanal. Chem.*, 2016, **408**, 4567-4573.
78. B. Saccà, L. Lacroix and J.-L. Mergny, *Nucleic Acids Res.*, 2005, **33**, 1182-1192.
79. P. Kumar, L. Truong, Y. R. Baker, A. H. El-Sagheer and T. Brown, *ACS Omega*, 2018, **3**, 6976-6987.
80. A. M. Varizhuk, V. B. Tsvetkov, O. N. Tatarinova, D. N. Kaluzhny, V. L. Florentiev, E. N. Timofeev, A. K. Shchylolkina, O. F. Borisova, I. P. Smirnov, S. L. Grokhovsky, A. V. Aseychev and G. E. Pozmogova, *Eur. J. Med. Chem.*, 2013, **67**, 90-97.
81. C. H. Lee, S.-H. Lee, J. H. Kim, Y.-H. Noh, G.-J. Noh and S.-W. Lee, *Mol. Ther.-Nucleic Acids*, 2015, **4**, e254.
82. M. C. Willis, B. D. Collins, T. Zhang, L. S. Green, D. P. Sebesta, C. Bell, E. Kellogg, S. C. Gill, A. Magallanez, S. Knauer, R. A. Bendele, P. S. Gill and N. Janjic, *Bioconjug Chem*, 1998, **9**, 573-582.
83. K. T. Shum and J. A. Tanner, *Chembiochem*, 2008, **9**, 3037-3045.
84. V. Esposito, M. Scuotto, A. Capuozzo, R. Santamaria, M. Varra, L. Mayol, A. Virgilio and A. Galeone, *Org. Biomol. Chem.*, 2014, **12**, 8840-8843.
85. S. Balamurugan, A. Obubuafo, R. L. McCarley, S. A. Soper and D. A. Spivak, *Anal. Chem.*, 2008, **80**, 9630-9634.

86. J. Tan, N. Yang, Z. Hu, J. Su, J. Zhong, Y. Yang, Y. Yu, J. Zhu, D. Xue, Y. Huang, Z. Lai, Y. Huang, X. Lu and Y. Zhao, *Nanoscale Res. Lett.*, 2016, **11**, 298-306.
87. S. Zhang, G. Zhou, X. Xu, L. Cao, G. Liang, H. Chen, B. Liu and J. Kong, *Electrochem. Commun.*, 2011, **13**, 928-931.
88. M. Oishi, S. Nakao and D. Kato, *Chem. Commun.*, 2014, **50**, 991-993.
89. M. Kimoto, R. Yamashige, K.-i. Matsunaga, S. Yokoyama and I. Hirao, *Nat. Biotechnol.*, 2013, **31**, 453-457.
90. L. Yatime, C. Maasch, K. Hoehlig, S. Klussmann, G. R. Andersen and A. Vater, *Nat. Commun.*, 2015, **6**, DOI: 10.1038/ncomms7481.

## Chapter 2 Techniques

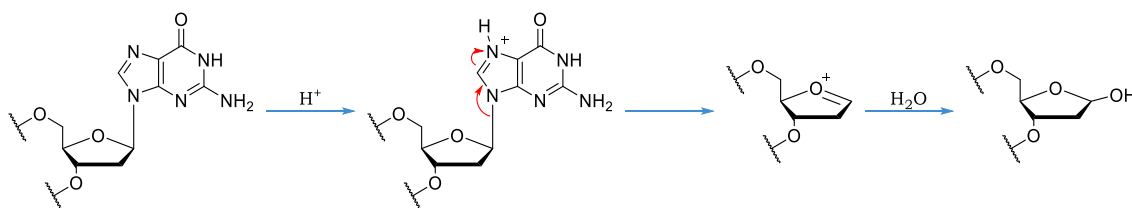
This chapter gives a general overview of the techniques used to synthesise and study the unmodified and modified DNA aptamer sequences detailed in this thesis. The production of the strands *via* automated solid phase DNA synthesis is first explained in detail. Afterwards, the methods behind the purification and characterisation of the strands is discussed. Then the analysis of oligonucleotide conformations using circular dichroism spectroscopy is next described. Subsequently, the solution-based techniques used for DNA-analyte interactions in the form of thermal melting, gel electrophoresis, and isothermal calorimetry is reported. Finally, the methods used to examine DNA reacted surfaces *via* contact angle, ellipsometry, and X-ray photoelectron spectroscopy are conveyed. Although not utilised in experiments described in this thesis, Förster resonance energy transfer and surface plasmon resonance spectroscopy are briefly touched upon as methods for future experiments.

## 2.1 Automated solid phase DNA synthesis

Solid phase DNA synthesis entails the heterogeneous coupling of a deoxyribonucleotide derivative in solution to another residue or growing oligonucleotide chain bound to a solid support (or resin). The technique is used to chemically synthesise single strands of oligonucleotides of up to 100 bases. Many types of solid supports can be used but the most common is controlled pore glass (CPG). CPG is rigid, non-swelling, and chemically inert to all reactions involved in DNA synthesis. It is manufactured with different particle and pore sizes with the most favoured pore sizes being 500Å and 1000Å. An advantage of solid phase DNA synthesis is that it permits the use of large excesses of reagents to force the coupling reaction to completion. Impurities and unreacted components are easily removed by washing the solid support, meaning no purification is required after each coupling step. The process is quicker and higher yielding than solution-based methods and is amenable to automation on computer-controlled DNA synthesisers. Several approaches to chemical synthesis of oligonucleotides have been developed to date, originating with the H-phosphonate method by Todd in the 1950s,<sup>1</sup> leading on to phosphodiester,<sup>2</sup> phosphotriester,<sup>3</sup> and phosphite triester<sup>4</sup> coupling chemistries. By far the most popular method for solid phase DNA synthesis is the phosphoramidite approach devised by Caruthers in 1981.<sup>5</sup>

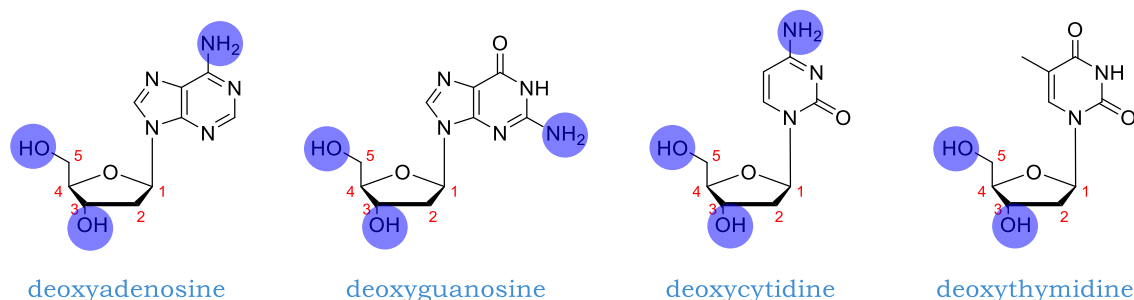
When designing the chemical synthesis of DNA oligonucleotides, one must be aware of the sensitivity of the nucleobases to acylation, alkylation, oxidation and reduction, and the sensitivity of the phosphodiester backbone to hydrolysis. There is also a risk of acid hydrolysis of the glycosidic C-N bond resulting in depurination of the nucleoside (Scheme 2.1). This effect is more prominent with

adenine and guanine nucleobases than with cytosine and thymine. This restricts the chemical reactions to mild alkaline hydrolysis, very mild acidic hydrolysis, mild nucleophilic displacement, base catalysed elimination reactions, and a few mild redox reactions.<sup>6</sup>



**Scheme 2.1** Depurination of a guanine nucleobase.

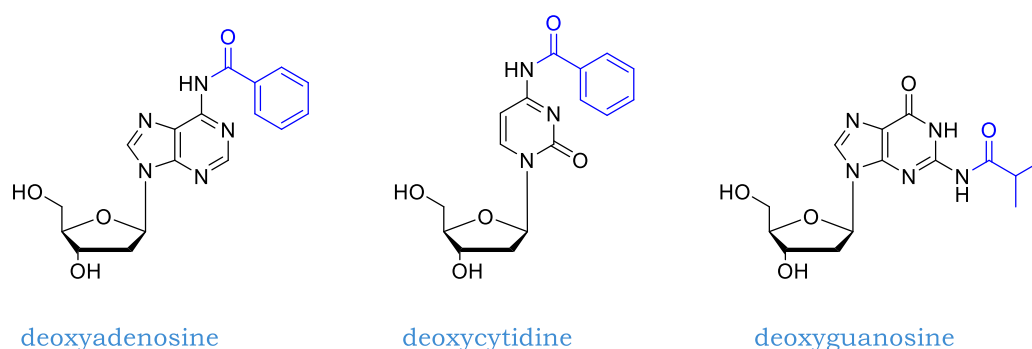
For the formation of phosphodiester connections between nucleosides *via* the phosphoramidite method, one sugar-based hydroxyl group must be phosphitylated with a phosphoramidite group and all other nucleophilic sites protected to provide regiocontrol of coupling reactions. The main nucleophilic sites on deoxynucleosides are the two sugar-based hydroxyl groups and the exocyclic amine groups on adenine, guanine, and cytosine nucleobases (Figure 2.1). To control the regiochemistry of coupling reactions, permanent protecting groups which are stable throughout the entire synthesis are placed on the exocyclic amine groups and on the phosphotriester backbone, whereas temporary protecting groups which can be removed after each coupling step are placed on one of the sugar-based hydroxyl groups. Stereochemistry is maintained throughout synthesis *via* the use of nucleosides in D-form derived from natural sources.<sup>6</sup>



**Figure 2.1** Main nucleophilic sites on nucleosides.

### 2.1.1 Synthesis of nucleoside phosphoramidite monomers

Synthesis of the nucleoside phosphoramidite monomers starts with the protection of the exocyclic amines with acyl protecting groups. The thymine nucleobase does not contain any exocyclic amine groups and thus does not need to be protected. The most common protecting group for adenine and cytosine nucleobases is the benzoyl protecting group, whereas the isobutyryl group is often used for the protection of guanine nucleobases (Figure 2.2).<sup>6</sup>

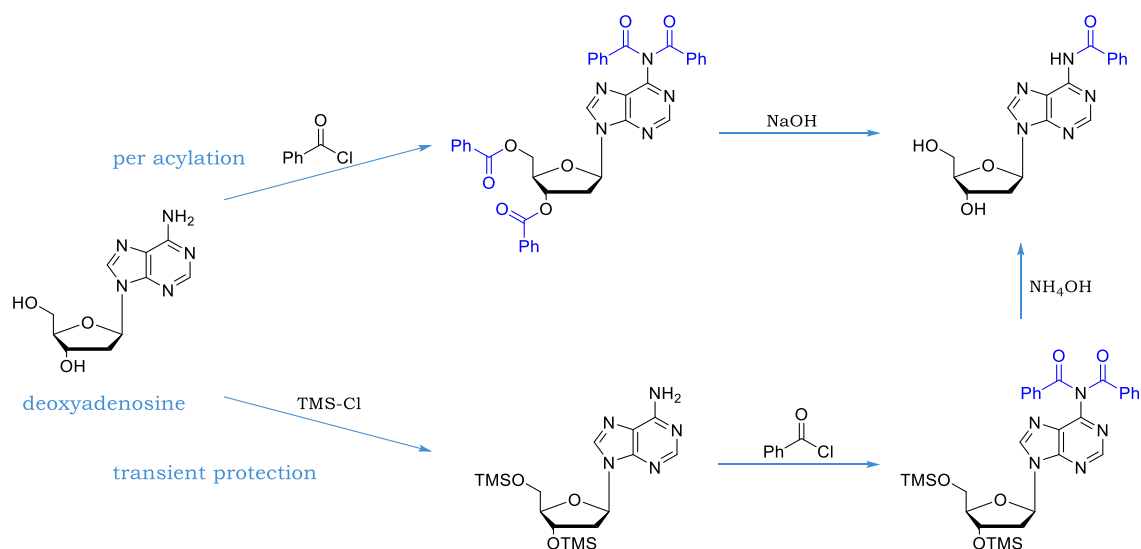


**Figure 2.2** Benzoyl and isobutyryl protecting groups used to protect exocyclic amine groups on deoxyadenosine, deoxycytidine, and deoxyguanosine nucleobases.

Acylation is performed *via* one of two procedures. The **per acylation** route uses excess acylating reagent to acylate both exocyclic amines and hydroxyl groups of the molecule. The esters formed are then selectively hydrolysed to recover the

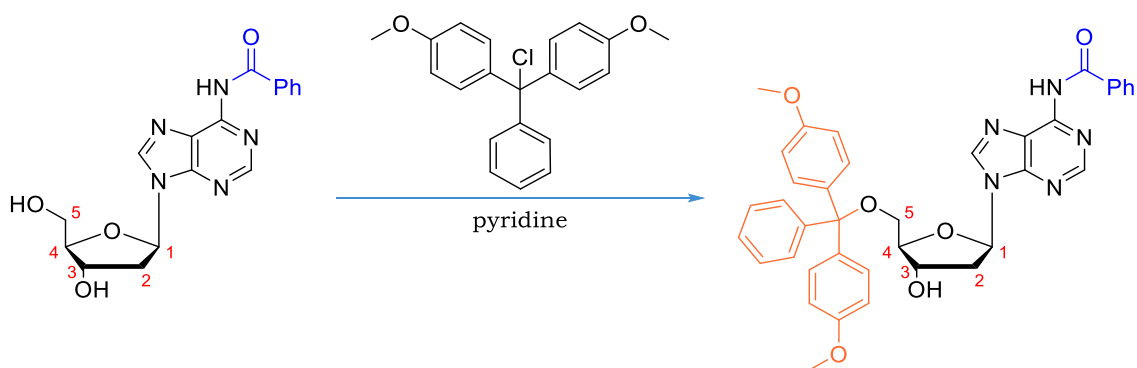


hydroxyl groups. The **transient protection** route selectively protects the hydroxyl groups using a silylation agent prior to acylation of the exocyclic amine. Desilylation then occurs to recover the hydroxyl groups (Scheme 2.2).<sup>6</sup>



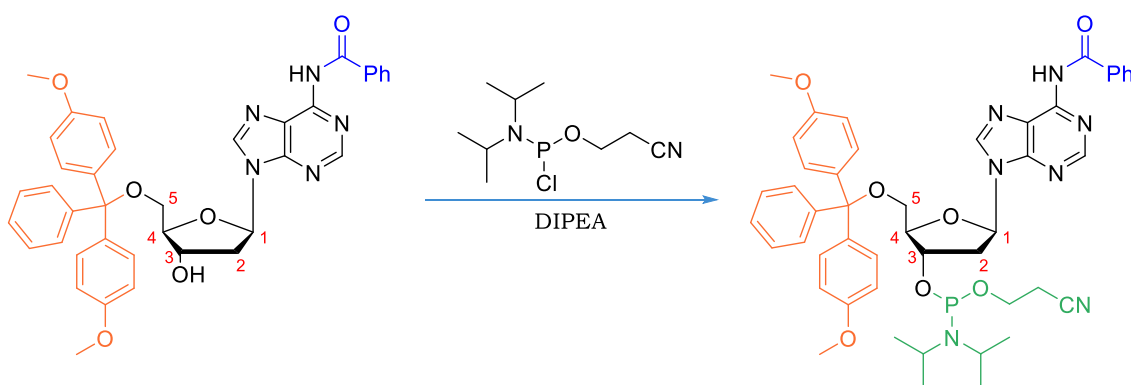
**Scheme 2.2** Different synthetic routes for the acylation of deoxyadenosine. Similar protocols are applied to the acylation of deoxycytidine and deoxyguanosine.

The phosphoramidite group can be attached to either hydroxyl group on the sugar. However, the 3' hydroxyl group is usually phosphitylated as the primary 5' hydroxyl group is a more effective nucleophile than the secondary 3' hydroxyl group. Prior to phosphitylation of the 3' hydroxyl group, the 5' hydroxyl group is first protected with a dimethoxytrityl (DMT) group (Scheme 2.3). The reaction is regioselective due to the steric bulk of the DMT group. DMT is chosen as a temporary protecting group as it is sensitive to mild acidic conditions, unlike the other protecting groups used. Thus, this group can be selectively removed for each coupling step. Another advantage of DMT is that upon deprotection, the liberated DMT cation is orange in colour and so can be used as a measure of nucleoside coupling.<sup>6</sup>



**Scheme 2.3** DMT reaction of acyl protected deoxyadenosine. Identical DMT reactions are applied to acyl protected deoxycytidine, acyl protected deoxyguanosine, and deoxythymidine.

The last step in the synthesis of the nucleoside phosphoramidite monomers is the phosphitylation of the 3' hydroxyl group using a phosphitylating reagent. The reagent often used is 2-cyanoethyl N,N-diisopropylchlorophosphoramidite (Scheme 2.4). The cyanoethyl group serves as a phosphate protecting group and is removed under basic conditions.<sup>6</sup>

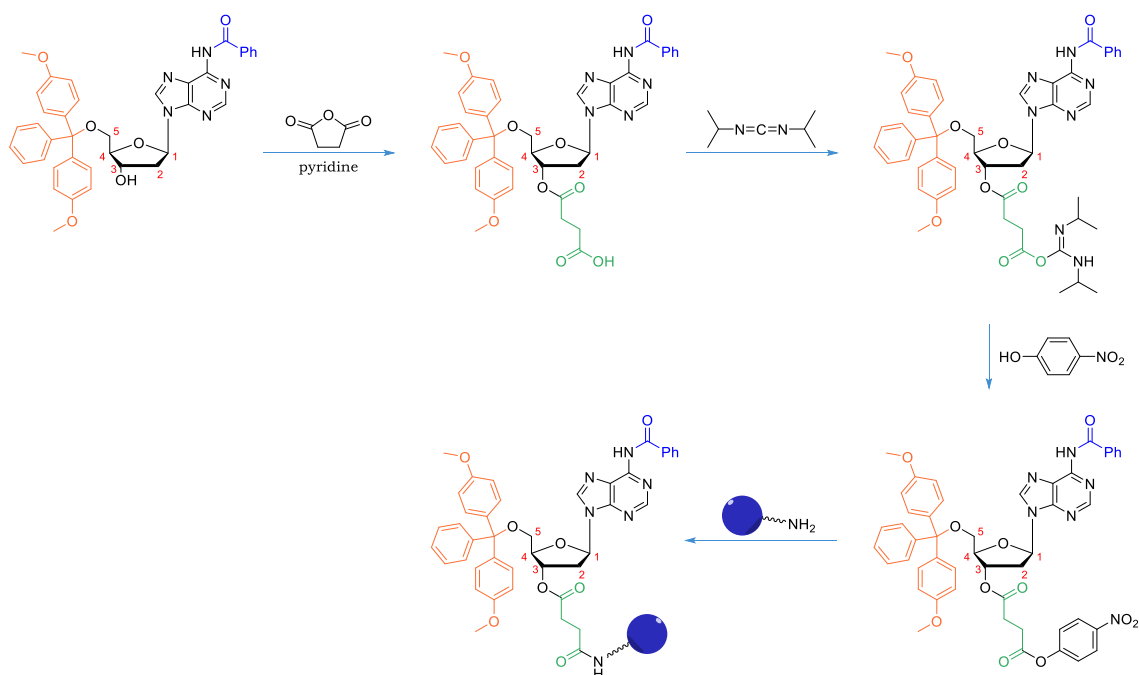


**Scheme 2.4** Phosphitylation of DMT and acyl protected deoxyadenosine. Identical conditions are applied to the other protected nucleosides.

### 2.1.2 Attachment of protected nucleosides to solid supports

The first nucleoside must be attached to the CPG resin before automated solid phase DNA synthesis can take place. A succinate linker is coupled to the 3'

hydroxyl position of a protected nucleoside. This linker is then activated and reacted with amine functionalised CPG resin to fix the protected nucleoside onto the solid support (Scheme 2.5).<sup>7</sup>



**Scheme 2.5** Attachment of protected deoxyadenosine to a solid support *via* a succinate linker. Identical conditions are applied to the other protected nucleosides.

### 2.1.3 DNA synthesis process

Automated DNA synthesis is performed in a  $3' \rightarrow 5'$  direction i.e. the chain is extended from the 3' hydroxyl group of the nucleoside. This direction is opposite to the biosynthesis of DNA. The complete synthesis cycle consists of 5 main steps, with acetonitrile washing of the resin occurring between steps (Scheme 2.6).

**Step 1:** The DMT group is removed from the attached nucleoside by addition of trichloroacetic acid (TCA) in DCM to generate the 5' hydroxyl group for nucleophilic attack. The absorbance intensity of the liberated

dimethoxytrityl cation at 495nm is monitored and used to determine the coupling efficiency of each nucleoside.

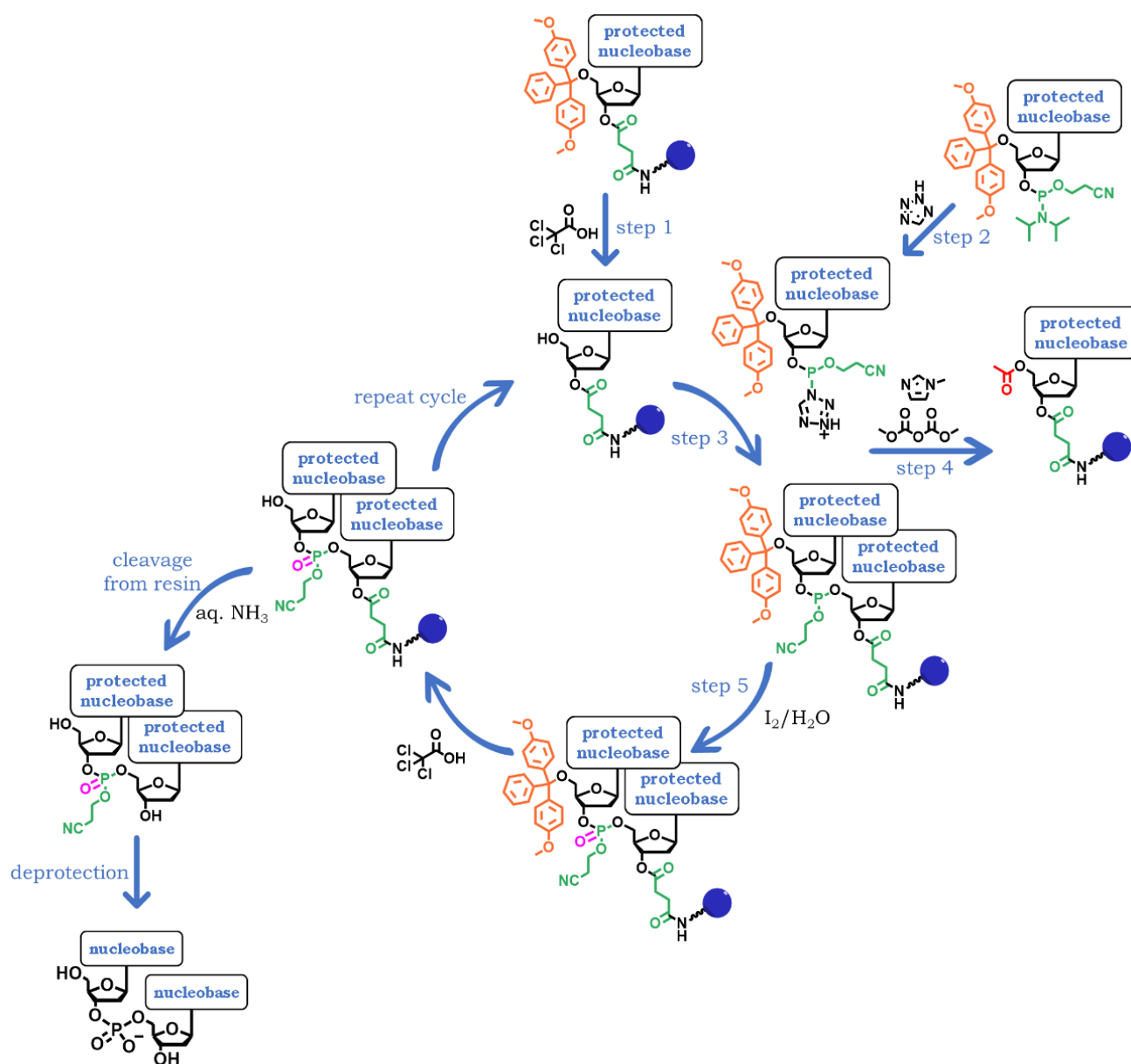
**Step 2:** The next nucleoside phosphoramidite monomer in the sequence is activated with tetrazole in acetonitrile.

**Step 3:** The activated nucleoside phosphoramidite is reacted with the exposed 5' hydroxyl group to form a phosphite triester.

**Step 4:** Unreacted hydroxyl groups are capped with acetic anhydride and N-methylimidazole to prevent them from reacting with subsequent nucleoside phosphoramidite monomers and causing failure sequences with missing internal residues.

**Step 5:** The formed phosphite triester is oxidised to a phosphate triester with iodine in pyridine and water. Pyridine is used to neutralise hydrogen iodide produced in the reaction.

The cycle is then repeated, adding one nucleoside at a time to produce the desired oligonucleotide. Once the final nucleoside has been added, the last DMT group is removed with TCA, and then concentrated aqueous ammonia is added to cleave the strands from the resin and remove cyanoethyl protecting groups. The ammonia solution is then heated to 60°C for 6 hours to remove acyl protecting groups from the nucleobases. The solvent is evaporated to leave the desired DNA sequence along with shortened capped strands and impurities from the basic deprotection stage. The crude is then purified to isolate the final product.<sup>7</sup>



**Scheme 2.6** Solid phase DNA synthesis.

## 2.2 Purification of oligonucleotides

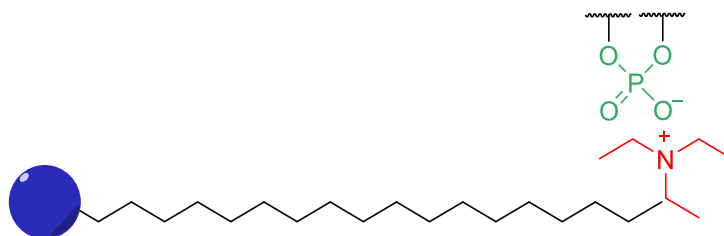
After solid phase DNA synthesis, the crude contains a mixture of truncated sequences, byproducts of the deprotection steps, and other impurities alongside desired DNA sequences. The two most frequently used techniques to isolate the desired strands are high performance liquid chromatography (HPLC) and polyacrylamide gel electrophoresis (PAGE). HPLC is suitable for the purification of strands of up to 50 bases with high sample recovery, whereas PAGE purification is better for longer sequences but is often low yielding.<sup>6</sup>

### 2.2.1 High performance liquid chromatography (HPLC)

HPLC separates mixtures by differential adsorption of each component onto an adsorbent stationary phase. Components are then independently desorbed from the stationary phase and eluted by altering solvent (mobile phase) conditions. A HPLC system usually comprises of an injector for sample introduction, an interchangeable column (stationary phase), a pump for delivery of solvents through the column, a solvent mixing system for gradient elutions, and a detector for analysis (UV/vis, fluorescent, refractive index), all connected by a computer. Two different HPLC methods can be applied to purify chemically synthesised DNA: reversed phase HPLC and anion exchange HPLC.<sup>8</sup>

**Reversed phase HPLC** resolves components in a mixture according to their hydrophobicity. The mixture is passed through a hydrophobic stationary phase (typically C18) and components are eluted with polar solvent (e.g. acetonitrile, methanol, water). Longer strands contain more hydrophobic nucleobases and are therefore retained more on the column and eluted later than short strands. Purification of oligonucleotides *via* this method utilises ion pairing whereby a

long-chained alkyl amine (e.g. triethylamine) is added to the mobile phase to indirectly increase DNA-stationary phase interactions and enhance resolution (Figure 2.3).<sup>9</sup>



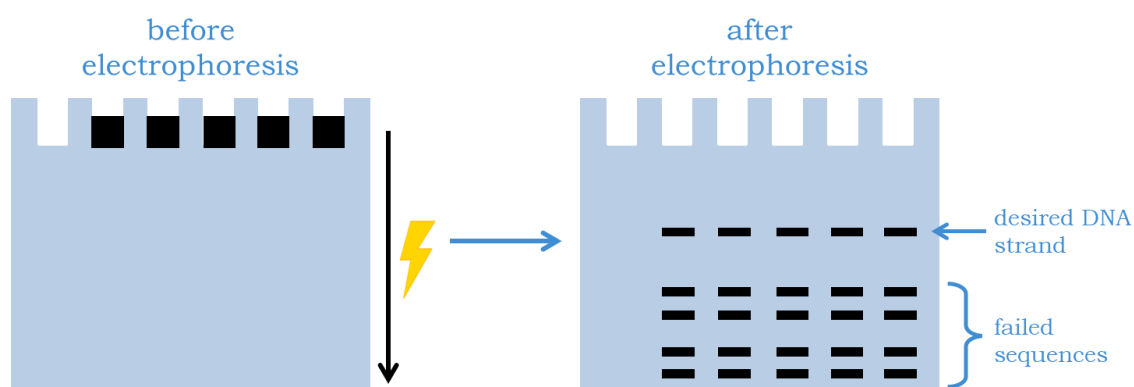
**Figure 2.3** Indirect adsorption of DNA to a C18 stationary phase *via* interaction of the phosphate backbone (shown in green) with an ion pair reagent (highlighted in red).

**Anion exchange HPLC** resolves components in a mixture according to their charge. A positively charged stationary phase is used for heightened DNA adsorption *via* the negatively charged phosphate backbone. DNA strands are then displaced according to charge by an aqueous mobile phase containing increasing concentrations of sodium salt, and eluted. Longer strands have a larger number of charged phosphodiester linkages and thus are eluted later as they require a higher concentration of sodium salt for displacement to occur.<sup>8</sup>

### 2.2.2 Polyacrylamide gel electrophoresis (PAGE)

PAGE separates the desired DNA product from failed sequences according to their electrophoretic mobility. This property is a function of the molecule's mass, size, conformation, and charge. The gels used are hydrogels formed from acrylamide and bisacrylamide crosslinkers, with acrylamide content generally ranging from 5-25%. A buffer (typically tris-borate-EDTA) is also incorporated to stabilise the pH of the gel. Gels can either be run natively where DNA interactions and conformations are conserved, or under denaturing conditions

where a denaturant such as urea is added to the gel to disrupt all non-covalent interactions. Wells are created in the gel where the sample is loaded, and an electrical potential gradient is applied across the gel to instigate the migration of DNA strands through the gel. The gel matrix provides a resistance to the movement of the strands, resulting in greater separation (Scheme 2.7). Shorter strands tend to have higher electrophoretic mobilities and thus travel quicker through the gel. The separate strands are stained and visualised as bands in the gel. The band corresponding to the desired strand is excised and the DNA extracted.<sup>10</sup>



**Scheme 2.7** PAGE purification of crude DNA mixture.



## 2.3 Characterisation of oligonucleotides by mass spectrometry

The purified single DNA strands are characterised prior to use in sensing experiments. If the sequence (i.e. the order of deoxyribonucleotides) is unknown, then the order is determined using the sequencing method originally developed by Fred Sanger.<sup>11</sup> However, the sequences of the oligonucleotides produced from automated solid phase DNA synthesis are already known, as the strands are synthesised in the order of the inputted sequence, and thus mass spectrometry is deemed an adequate and efficient method for characterisation of these oligonucleotides.

Mass spectrometry is a technique that analyses ions by their mass to charge ratios ( $m/z$ ). Mass spectrometers consist of an ionisation source, a mass analyser, and an ion detector, all controlled by a computer. The sample is vaporised and ionised by the ionisation source, the ions are then accelerated under vacuum through electric/magnetic fields into the mass analyser where their mass to charge ratios are determined. Mass spectra are recorded and compared to the calculated mass of the sample. Although destructive, very little material (1mg/ml) is required for analysis with this technique.<sup>12</sup>

### 2.3.1 Ionisation methods

The methods of ionisation routinely used in the analysis of oligonucleotides are either matrix laser desorption ionisation (MALDI) or electrospray ionisation (ESI). Both are soft ionisation techniques, whereby fragmentation of the individual molecules is significantly reduced. In MALDI, sample is added to a matrix solution (usually 3-hydroxypicolinic acid for oligos), the resulting

mixture is placed on a sample grid, and the grid is inserted into the instrument for analysis. The sample is vaporised by a strong laser pulse, resulting in ionisation of the oligo *via* protonation (i.e. the oligo gains a +1 charge).<sup>13</sup> With (negative mode) ESI, the sample is dissolved in a volatile solvent and introduced into the instrument as a fine spray in the presence of a strong electric field. The molecules are ionised *via* deprotonation and the droplets transported to a high vacuum region where solvent is removed. Multiple deprotonations within the same strand can occur resulting in a variety of species with multiple negative charges.<sup>14</sup>

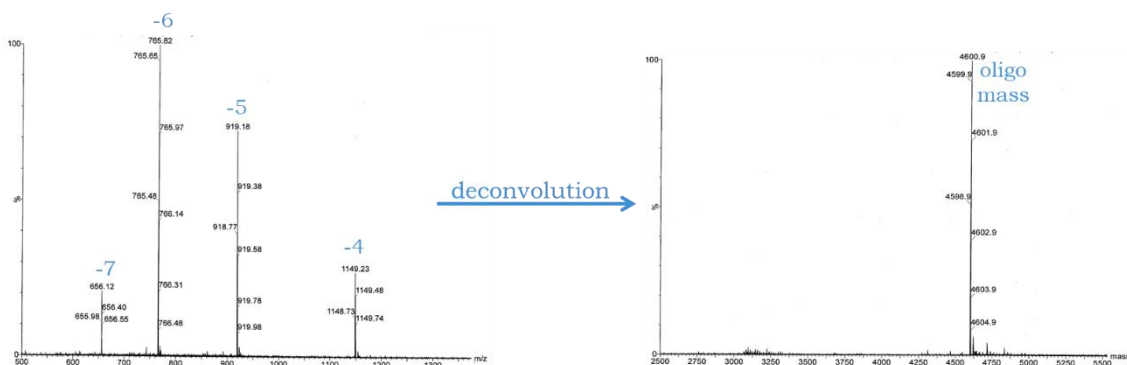
### 2.3.2 Mass analysers

Numerous mass analysers are available for mass determination of the ionised oligos. However, the mass analysers utilised in the analysis of DNA strands studied in this thesis were either time of flight (TOF), quadrupole, or a combination of the two. In TOF, the ions are accelerated by an electric field to matching kinetic energies (KE) and their time of flight is measured across a set distance. As the mass to charge ratio is directly proportional to the time of flight, masses can be calculated from the time taken for the ions to reach the detector, with small ions arriving faster than larger ions. With quadrupole mass analysers, a constant direct current and radio-frequency is applied between two pairs of two parallel electrodes. Ions are injected in-between the electrodes and the voltage is varied, creating an electric field which causes the ions to oscillate. Changing the radio frequency results in ejection of ions of a specific mass to charge ratio into an ion detector, and thus masses can be determined from the radio frequency applied.<sup>12</sup>

### 2.3.3 Interpretation of mass spectra

Oligo data produced from MALDI shows the mass to charge ratios of singly ionised species (i.e.  $M + H^+$ ). Signals corresponding to doubly charged species can occur but are rare. Determination of the molecular mass from the data requires subtraction of a proton mass from the signal given.

The raw data produced from ESI does not show the parent ion. Instead, many signals are generated corresponding to strands with different multiple negative charges (i.e.  $M - nH^+$ ). The mass of the parent ion is calculated by multiplying the measured signal by the charge of the species, and then adding the magnitude of the charges to that amount. The charge of the species (i.e. the number of protons removed) is determined from the spacing between a cluster of signals. It is calculated by taking the inverse of the spacing. For example, an oligo ion with a -2 charge would show a cluster of signals with 0.5  $m/z$  spacing in the ESI raw data. Oligo ESI traces shown in this thesis are processed results that have been treated with a deconvolution algorithm. The software applies the calculation to the raw data to give a spectrum showing the parent ion peak (Scheme 2.8).

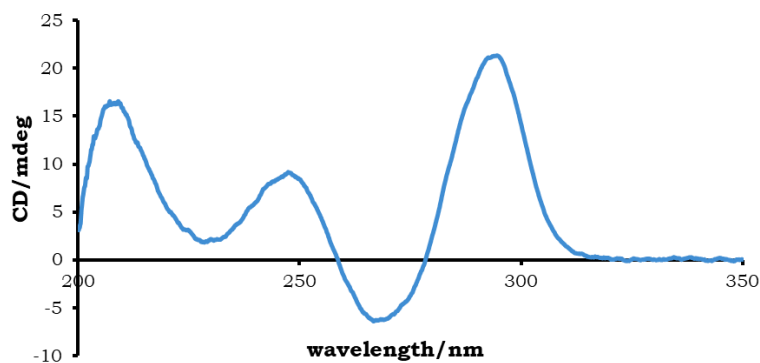


**Scheme 2.8** Deconvolution of ESI raw data.

Hail *et al.* compared MALDI and ESI for analysis of oligonucleotides and reported that results from both methods were accurate for short oligos, whereas MALDI was found to be ineffective for determining accurate masses of longer strands (>50 bases). Long strands fly poorly in the MALDI instrument setup, resulting in an insufficient number of ionised strands reaching the detector and subsequent low resolution spectra.<sup>15</sup>

## 2.4 Circular dichroism spectroscopy (CD)

Circular dichroism (CD) is observed when optically active molecules absorb left and right handed circularly polarised light differently. This slight difference in absorption of the two forms leads to an elliptical projection of the polarised light which is measured as ellipticity on a CD spectropolarimeter. The spectrum produced shows the ellipticity as a function of wavelength (Figure 2.4). Positive CD signals indicate a greater absorbance of left handed circularly polarised light over right handed circularly polarised light of a particular wavelength, whereas negative CD signals indicate the reverse.<sup>16</sup> Optically active molecules are chiral in nature and therefore this phenomenon is observed in oligonucleotides as they exhibit planar chirality and contain multiple chiral centres. CD spectroscopy is an efficient, highly sensitive, and non-destructive technique for studying the secondary structures of DNA in solution. Structures of both duplexed and single stranded DNA have unique CD spectral signatures in the UV range of 200-350nm, and interpretation of CD data is performed empirically.<sup>17</sup> DNA conformation is highly influenced by ligand binding, and sample environment (e.g. pH, temperature, solvent), and such changes are observed by shifts in the position and intensity of the CD signal of the solution.

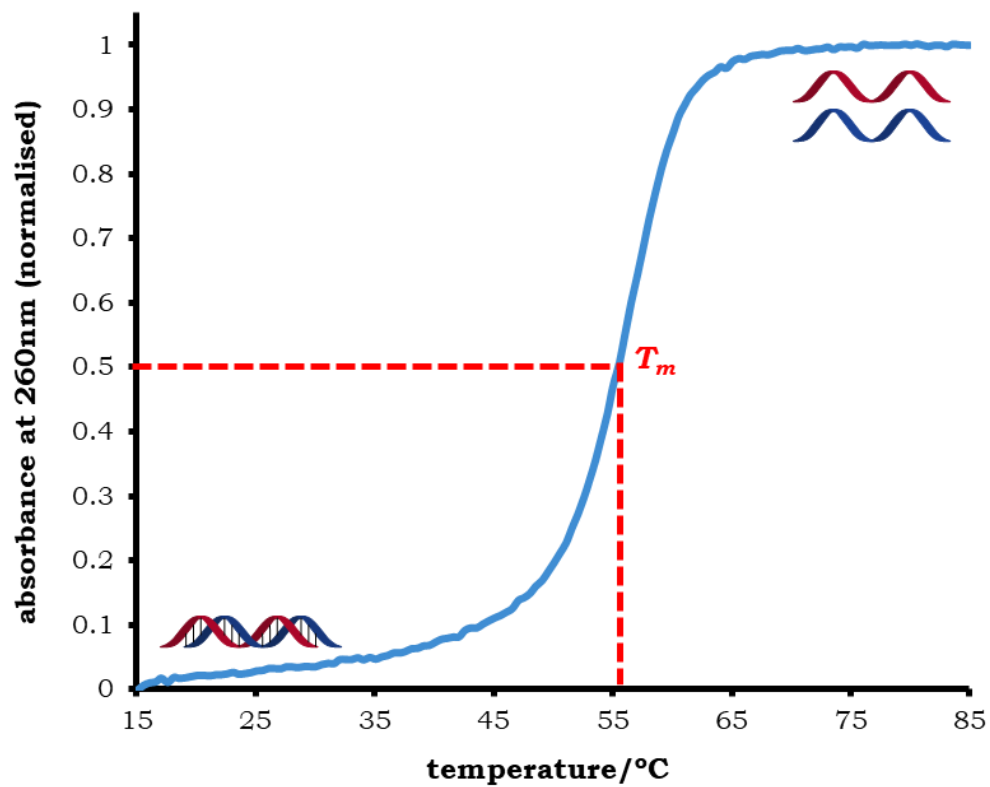


**Figure 2.4** CD spectrum of a solution containing single stranded DNA. CD peak at 295nm indicates a G-quadruplex conformation for the strands.

## 2.5 Thermal melting ( $T_m$ )

Thermal melting is used to probe oligonucleotide duplex formation and is based on the hyperchromic effect. The hyperchromic effect relates the UV absorbance of an oligonucleotide to the stacking of bases within the strand(s). The aromatic nature of the nucleobases leads the strands to have a high absorbance of UV light at 260nm. When an oligonucleotide is folded or duplexed, bases within the complex become stacked and hidden, causing a decrease in the overall amount of light that the molecule can absorb at 260nm. Upon denaturation (or melting/unravelling), bases become unstacked and thus the strand(s) can absorb more light. An increase in UV absorbance is called hyperchromicity, whereas a decrease in absorbance is called hypochromicity.<sup>18</sup>

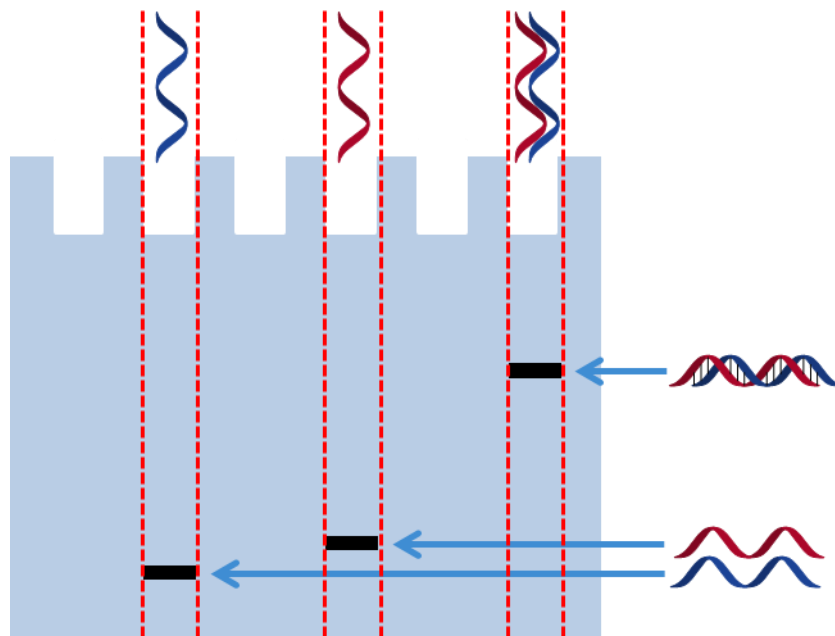
In thermal melting experiments, heat is used to denature the oligonucleotide. The duplexed sample is slowly heated to the point of full denaturation, and the absorbance of the solution at 260nm is measured at each temperature point. The sample is then slowly cooled to rehybridise the single strands back to the original duplex (thermal cooling). A melting curve is then plotted showing the UV absorbance at 260nm as a function of temperature (Figure 2.5). The midpoint of the curve is the point at which 50% of duplexes have denatured and is termed the melting temperature ( $T_m$ ) of the duplex. The  $T_m$  value gives an indication of the stability of the duplex, with more stable duplexes resulting in higher  $T_m$  values. Duplexes with more G-C content have higher  $T_m$  values, as more energy is required to break the 3 hydrogen bonds between the bases, compared to 2 hydrogen bonds with the A-T base pair.  $T_m$  values are used to probe the effect of strand modifications, intercalators and sample environment conditions (i.e. pH, salt concentration) on duplex stability.<sup>19</sup>



**Figure 2.5** DNA thermal melting curve. Determination of the  $T_m$  value is shown in red.

## 2.6 Gel electromobility shift assay (gel EMSA)

Gel EMSA is a technique used to study biomolecular interactions by showing how such interactions affect the electrophoretic mobility of the biomolecule of interest in a gel. Any complexation would add mass to the biomolecule of interest and reduce its mobility, resulting in a higher band in the gel (Figure 2.6).<sup>20</sup> Gel EMSA is run under native conditions and is used to confirm DNA duplex formation and DNA-protein binding.

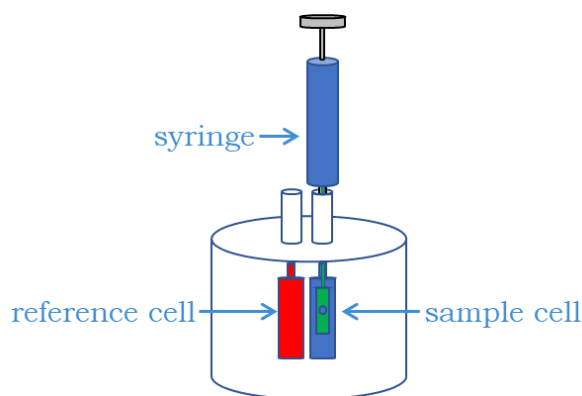


**Figure 2.6** Gel EMSA of DNA. Highest band shows duplex formation.



## 2.7 Isothermal titration calorimetry (ITC)

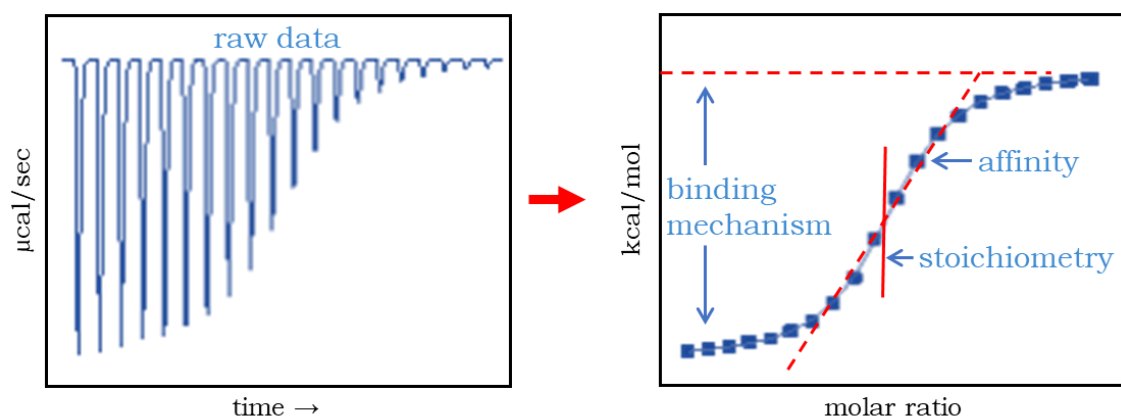
Isothermal titration calorimetry (ITC) is a label free technique used to study biomolecular interactions. It measures the heat generated or absorbed upon binding of a probe to a target, depending on whether the interaction is *exo* or *endothermic*.<sup>21</sup> A solution of target is loaded into a syringe and added dropwise to a solution of the probe in a sample cell (Figure 2.7). The heat of interaction is measured after each addition by calculating the difference in temperatures of the sample cell and a reference cell containing only solvent (usually water). The sample cell temperature is then readjusted back to baseline before the next addition.



**Figure 2.7** ITC setup. Target is loaded into the syringe and added dropwise to a solution of probe in the sample cell. The heat of interaction is measured after each addition, and the cell temperature readjusted back to baseline.

As probe binding sites become saturated with target, the intensity of the heat of interaction decreases and eventually plateaus to heats of dilution.<sup>22</sup> The heat of interaction is often referred to as the differential power (DP) and is a negative value for exothermic interactions and a positive value for endothermic interactions. The raw data shows changes in the differential power (in  $\mu\text{cal/sec}$ )

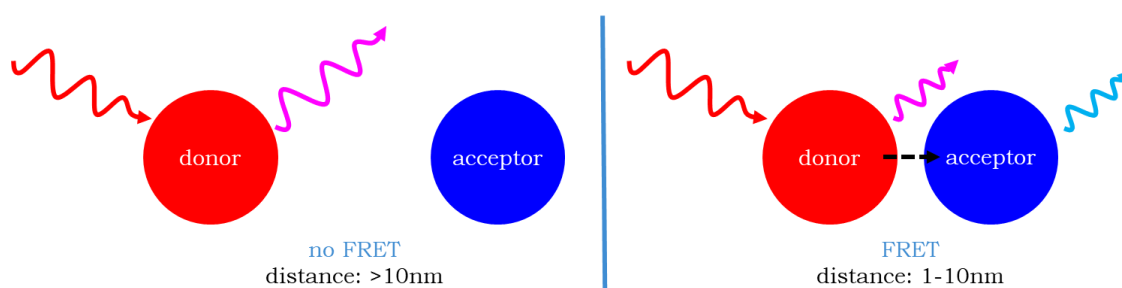
as a function of time (Figure 2.8). ITC software is used to convert the data into heat changes (in kcal/mol) as a function of the molar ratio. A suitable model is selected, and a binding curve fitted to determine the thermodynamics, stoichiometry and affinity of the interaction.<sup>23</sup>



**Figure 2.8** ITC raw and processed data used to determine the thermodynamics, stoichiometry, and affinity of exothermic probe-target interactions.

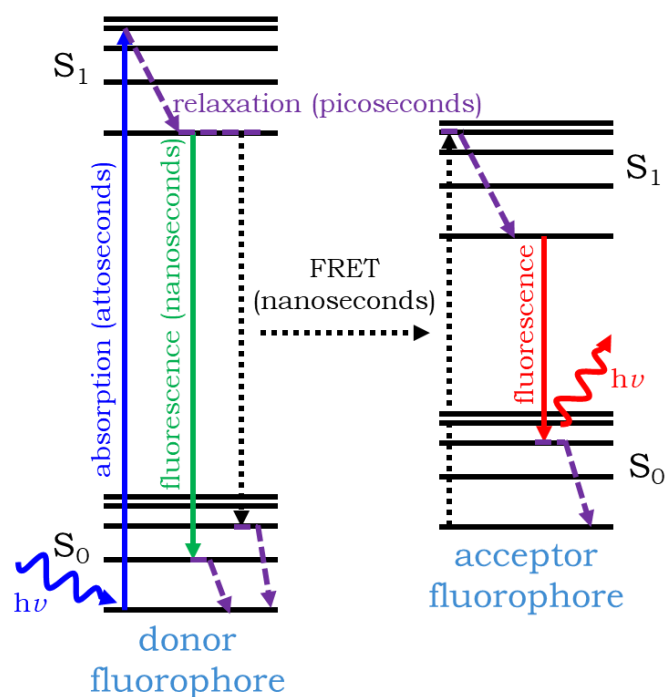
## 2.8 Förster resonance energy transfer (FRET)

FRET is a collision free, nonradiative transfer of energy from a donor chromophore in an excited state to an acceptor chromophore within range. When both chromophores are fluorescent, it results in reduced fluorescence emission of the donor and enhanced fluorescence emission of the acceptor. The principle is used to study interactions between fluorescent molecules by estimating interaction distances (Figure 2.9). Förster theory states that the efficiency of energy transfer ( $E$ ) is proportional to the inverse sixth power of the distance ( $r$ ) between the two chromophores ( $E \propto r^{-6}$ ).<sup>24</sup> Consequently, FRET is overly sensitive to minute changes in distance and can be used to monitor such changes.



**Figure 2.9** FRET between two fluorophores.

For FRET to occur, both fluorophores must be close in distance but not touching (1-10nm), be in the right orientation (for dipole-dipole interaction), the donor must have a high quantum yield, and the emission spectrum of the donor fluorophore must overlap with the absorption spectrum of the acceptor fluorophore.<sup>25</sup> The timescale of FRET events is represented in a Jablonski diagram (Figure 2.10).

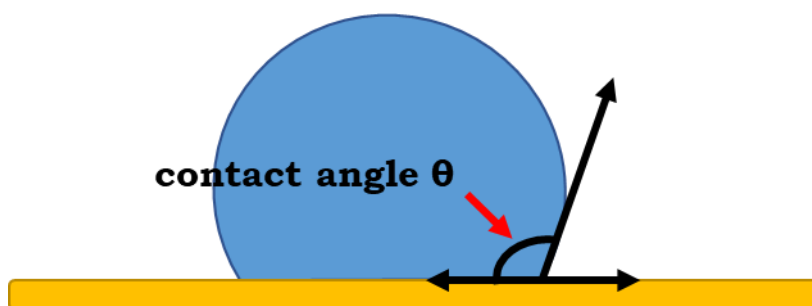


**Figure 2.10** Jablonski diagram showing timescales of FRET events. Solid arrows show radiative transitions whereas dashed arrows show nonradiative transitions.

Experimentally, a fluorimeter or fluorescence microscope is used to measure the change in fluorescence of samples. The FRET efficiency is determined by monitoring the fluorescence intensity or fluorescence lifetime of the donor fluorophore in the presence and absence of the acceptor fluorophore. Interaction distances are then calculated from the FRET efficiency. Oligonucleotides must first be labelled with two different fluorophores before FRET experiments can be undertaken. These labels can either be placed on the same strand to explore DNA conformations, or placed on different strands to study intermolecular interactions, such as duplex formation.<sup>26</sup>

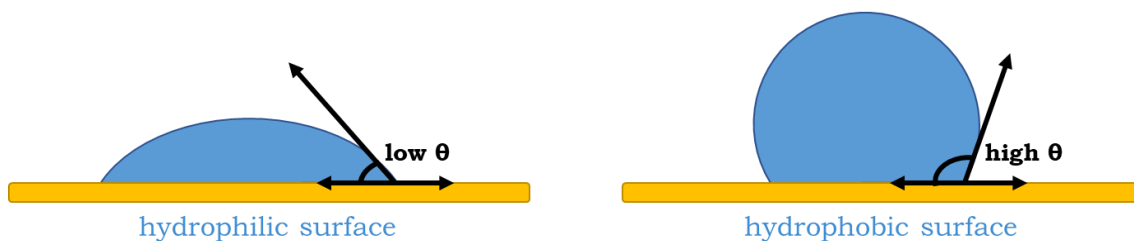
## 2.9 Contact angle

Contact angle is a surface analysis technique used to measure the wetting of a solid by a liquid. It is defined as the angle formed by a droplet at the three-phase boundary where a liquid, solid, and gas interact (Figure 2.11). Contact angles indicate the strength of interactions between liquids and solids and is used to identify the hydrophobicity/hydrophilicity of surfaces. Water is often employed as the liquid for contact angle analysis.<sup>27</sup>



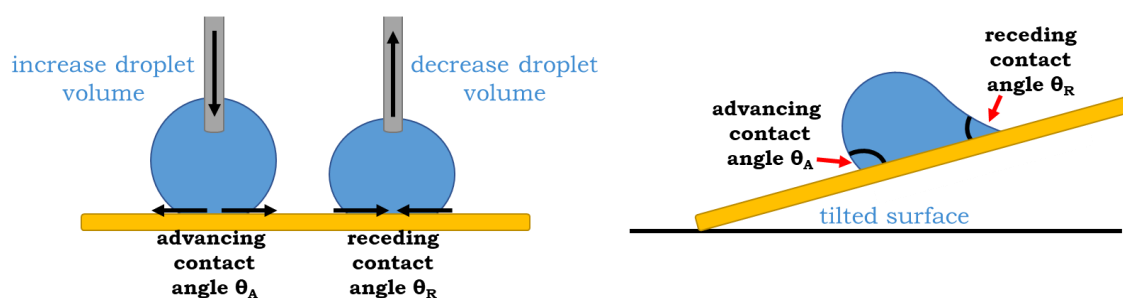
**Figure 2.11** Contact angle at the three-phase boundary where a liquid, solid, and gas interact.

Hydrophilic surfaces have strong interactions with aqueous solutions, resulting in low contact angles for the droplets. The reverse applies for hydrophobic surfaces (Figure 2.12). Contact angles are measured with a tensiometer, of which there are two types- force and optical. Force tensiometers measure the mass affecting a balance when a solid sample is brought into contact with a liquid whereas optical tensiometers record and automatically analyse the shape of a droplet on a surface over time.<sup>28</sup>



**Figure 2.12** Contact angle comparison of hydrophilic and hydrophobic surfaces with water.

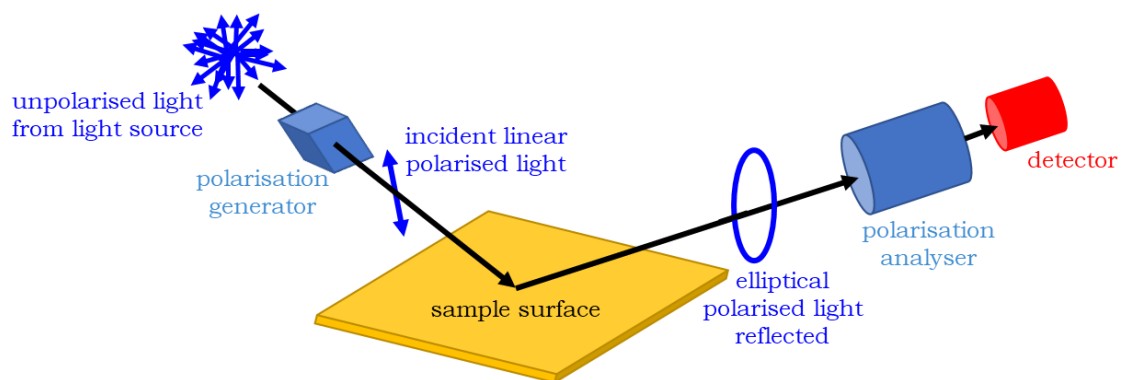
Contact angles can either be measured statically where the droplet is stationary, or dynamically where the droplet is moving. Dynamic contact angles are split into advancing and receding angles and give an indication of the roughness of a surface. They are determined by either altering the volume of the droplet on the surface or by tilting the surface (Figure 2.13). Advancing and receding contact angles give the maximum and minimum contact angle values of a surface, and the difference between the two is called the contact angle hysteresis. Rougher surfaces produce larger contact angle hysteresis values. When analysing surfaces covered with self-assembled monolayers (SAMs), the contact angle hysteresis can give an indication of the homogeneity of SAMs with a small hysteresis suggesting a well packed SAM.<sup>29</sup>



**Figure 2.13** Determination of dynamic contact angles by either changing the droplet volume (left) or by tilting the surface (right).

## 2.10 Ellipsometry

Ellipsometry is a non-invasive technique used to determine the optical properties of surfaces. It analyses the change in state of polarised light upon reflection at an oblique angle from a sample for the characterisation of surfaces, interfaces, and thin films. Ellipsometry measurements are performed by an ellipsometer which sets the polarisation state of the incident beam and detects the polarisation state of the reflected beam. Ellipsometers consist of a light source, polarisation generator, a polarisation analyser, and a detector (Figure 2.14). Light produced from the light source is polarised by the polarisation generator which allows light of a specific electric field orientation to pass. The axis of the generator is oriented to permit waves which are either parallel or perpendicular to the plane of incidence of the sample. The linearly polarised light reaches the sample where upon reflection becomes elliptically polarised and reflected into the polarisation analyser and detector.<sup>30</sup>



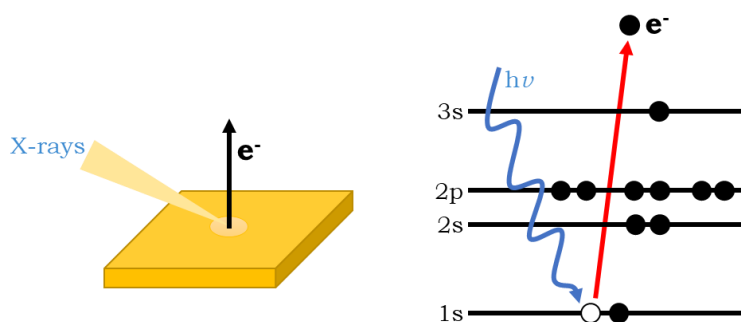
**Figure 2.14** Ellipsometer setup showing light source, polarisation generator, sample, polarisation analyser, and detector. Linearly polarised light becomes elliptically polarised upon reflection from the sample.

The detector converts the reflected light into an electronic signal and the information is compared to the inputted incident light to generate ellipsometry parameters. Computer software is used to construct models to describe the sample with calculated ellipsometry parameters. The experimental data is then fitted to the computed values using regression analysis. Film thicknesses and optical constants are then obtained from the calculations.<sup>31</sup>



## 2.11 X-ray photoelectron spectroscopy (XPS)

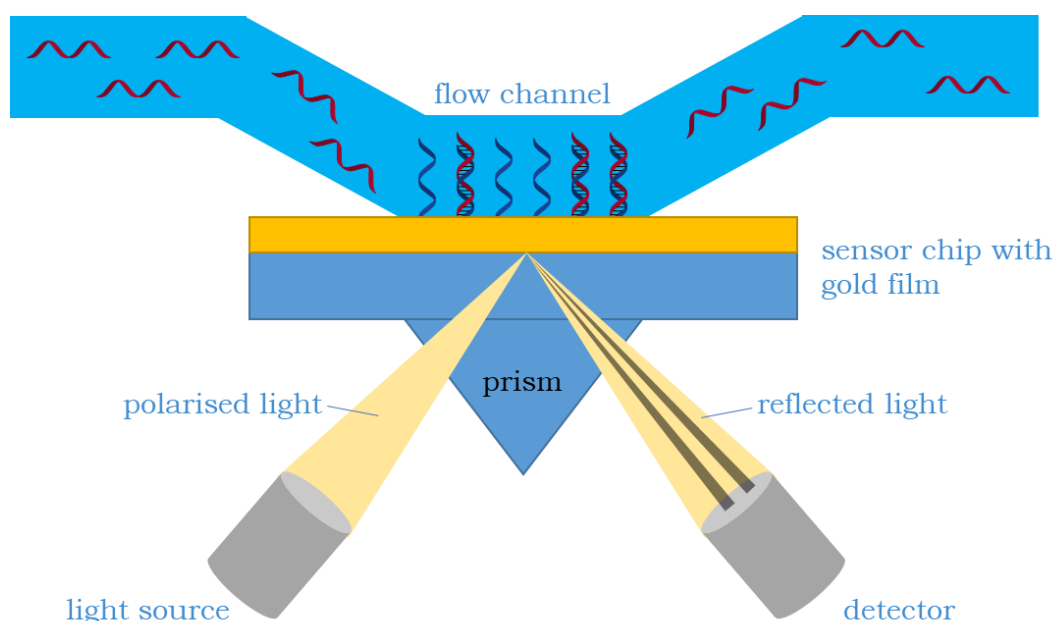
X-ray photoelectron spectroscopy (XPS) is a chemical analysis technique used to identify elemental compositions. The technique uses X-rays of a set energy to eject photoelectrons from samples and analyses the kinetic energies of the ejected photoelectrons to determine their elemental source. X-ray photoelectron spectrometers typically consist of a monochromatic X-ray source (fixed or tuneable), a vacuum chamber where the sample is placed, an electron energy analyser, and an electron detector, all controlled by a computer. When X-rays are applied to a specimen in a vacuum chamber, the rays have a probability of interacting with electrons in the material, causing photoelectrons to be ejected from the atoms or molecules of that sample (Figure 2.15). Every element has a defined electronic configuration and every electron in that element is bound by a certain amount of energy called the binding energy (BE). Binding energies are specific to each electron and vary between elements. With XPS, binding energies are related to the kinetic energies of the ejected photoelectrons and thus can be determined. XPS spectra show photoelectron intensity (in counts or counts per second) as a function of either kinetic energy or binding energy. Interpretation of XPS data is achieved empirically with peaks assigned using element line position tables or XPS analysis software.<sup>32</sup>



**Figure 2.15** Ejection of photoelectron from sample caused by X-ray absorption.

## 2.12 Surface plasmon resonance (SPR)

Surface plasmon resonance spectroscopy (SPR) is a mass based, label free technique used to study biomolecular interactions in real time. The technique monitors interactions between free analytes in solution and fixed probes on a metallic surface in a sensor chip, usually gold. The technique works by reflecting a light source through a prism, off the back of the sensor chip, and into a detector (Figure 2.16). At a certain incidence angle, known as the resonance angle, light is absorbed by electrons in the metallic film of the sensor chip, causing them to resonate. This effect is observed as a loss in intensity, or a dip, in the SPR reflection intensity spectrum. The resonating electrons are called surface plasmons and are sensitive to their surrounding environment. When a binding or unbinding event occurs, it affects the surface plasmons, causing a change in the size or location of the dip. The change in SPR response is used to convey information about interactions at the sensor surface.<sup>33</sup>



**Figure 2.16** SPR setup showing changes to the reflected light upon probe-target binding.

## 2.13 List of references

1. A. M. Michelson and A. R. Todd, *J. Chem. Soc.*, 1955, **0**, 2632-2638.
2. H. G. Khorana, *Some recent developments in the chemistry of phosphate esters of biological interest*, Wiley, New York, 1961.
3. R. L. Letsinger and K. K. Ogilvie, *J. Am. Chem. Soc.*, 1969, **91**, 3350-3355.
4. R. L. Letsinger, J. L. Finnan, G. A. Heavner and W. B. Lunsford, *J. Am. Chem. Soc.*, 1975, **97**, 3278-3279.
5. S. L. Beaucage and M. H. Caruthers, *Tetrahedron Lett.*, 1981, **22**, 1859-1862.
6. G. M. Blackburn and M. J. Gait, *Nucleic Acids in Chemistry and Biology*, Royal Society of Chemistry, Cambridge, 1990.
7. Solid-Phase Oligonucleotide Synthesis,  
<https://www.atdbio.com/content/17/Solid-phase-oligonucleotide-synthesis>, (accessed 05/09/2018).
8. Purification of Oligonucleotides,  
<https://www.atdbio.com/content/7/Purification-of-oligonucleotides>,  
(accessed 05/09/2018).
9. Method of Oligonucleotide Purification,  
<https://www.biosyn.com/tew/method-of-oligonucleotide-purification.aspx>, (accessed 05/09/2018).
10. Nucleic Acid Gel Electrophoresis- A Brief Overview and History,  
<https://www.thermofisher.com/uk/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/na-electrophoresis-education/na-separation-overview.html>,  
(accessed 05/09/2018).

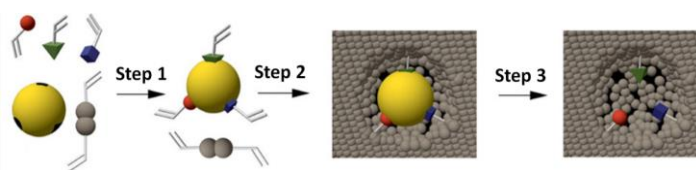
11. F. Sanger, S. Nicklen and A. R. Coulson, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, **74**, 5463-5467.
12. D. H. Williams and I. Fleming, *Spectroscopic Methods in Organic Chemistry*, McGraw-Hill, Berkshire, 1997.
13. J. T. Stults, *Curr Opin Struct Biol*, 1995, **5**, 691-698.
14. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse, *Science*, 1989, **246**, 64-71.
15. M. E. Hail, B. Elliott and K. Anderson, *Am Biotech Lab*, 2004, **22**, 12-14.
16. A. R. Urbach, *J. Chem. Educ.*, 2010, **87**, 891-893.
17. V. Michaela, K. Iva, B. Klára, R. Daniel and K. Jaroslav, *Chirality*, 2012, **24**, 691-698.
18. V. S. N., *Int. J. Quantum Chem*, 1979, **16**, 119-132.
19. J. L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515-537.
20. Gel Shift Assays- EMSA,  
<https://www.lifetechnologies.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/gel-shift-assays-emsa.html>, (accessed 05/09/2018).
21. M. M. Pierce, C. S. Raman and B. T. Nall, *Methods*, 1999, **19**, 213-221.
22. Isothermal Titration Calorimetry (ITC),  
<http://www.malvern.com/en/products/technology/isothermal-titration-calorimetry/>, (accessed 05/09/2018).
23. V. Linkuvienė, G. Krainer, W.-Y. Chen and D. Matulis, *Anal. Biochem.*, 2016, **515**, 61-64.
24. T. Förster, *Ann. Phys.*, 1948, **437**, 55-75.
25. D. Shrestha, A. Jenei, P. Nagy, G. Vereb and J. Szollosi, *Int J Mol Sci*, 2015, **16**, 6718-6756.

26. C.-L. Maria-Chantal and P. Michel, *Biochem. Educ.*, 1998, **26**, 320-323.
27. D. Y. Kwok and A. W. Neumann, *Adv. Colloid Interface Sci.*, 1999, **81**, 167-249.
28. Contact Angle, <https://www.biolinscientific.com/measurements/contact-angle>, (accessed 05/09/2018).
29. Y. Yuan and T. R. Lee, in *Surface Science Techniques*, eds. G. Bracco and B. Holst, Springer, Berlin, 2013, vol. 51, ch. Contact Angle and Wetting Properties, pp. 3-34.
30. N. J. Podraza and G. E. Jellison, in *Encyclopedia of Spectroscopy and Spectrometry*, eds. J. C. Lindon, G. E. Tranter and D. W. Koppenaal, Academic Press, Oxford, 3rd edn., 2017, ch. Ellipsometry, pp. 402-411.
31. M. Bass, *Handbook of Optics: Fundamentals, techniques, and design*, McGraw-Hill, New York, 1994.
32. D. Briggs and J. T. Grant, *Surface Analysis by Auger and X-ray Photoelectron Spectroscopy*, IM Publications, Chichester, 2003.
33. R. B. M. Schasfoort and A. J. Tudos, *Handbook of Surface Plasmon Resonance*, Royal Society of Chemistry, Cambridge, 2008.

# Chapter 3 Acrylamide modified aptamers for incorporation into molecularly imprinted polymers

## 3.1 Introduction

Molecularly imprinted polymers, or MIPs, are a class of synthetically derived recognition units, designed to mimic natural recognition elements such as antibodies in their ability to bind to target molecules with high selectivity and specificity.<sup>1</sup> These polymers are formed from a template derived process whereby an analyte of interest is incubated with receptor monomers with polymerisable functionalities. Radical initiators and cross linkers initiate polymerisation, forming a network around the analyte.<sup>2</sup> The analyte is then removed, leaving behind a binding pocket specific to the analyte, as shown in Scheme 3.1.<sup>3</sup> Target removal leaves multiple cavities in the polymer which can then be used as recognition elements for the detection, extraction, and/or purification of the imprinted substrate. This concept originated from work by Polyakov in 1931 whereby he investigated the adsorption properties of various solvent molecules onto silica.<sup>4</sup> Organic based MIPs were first introduced by Mosbach in the 1950s where polyacrylamide gels were employed for the entrapment of enzymes and cells.<sup>5</sup> Since then, the MIP field has grown substantially in incorporating numerous receptor monomers for a wide range of targets.

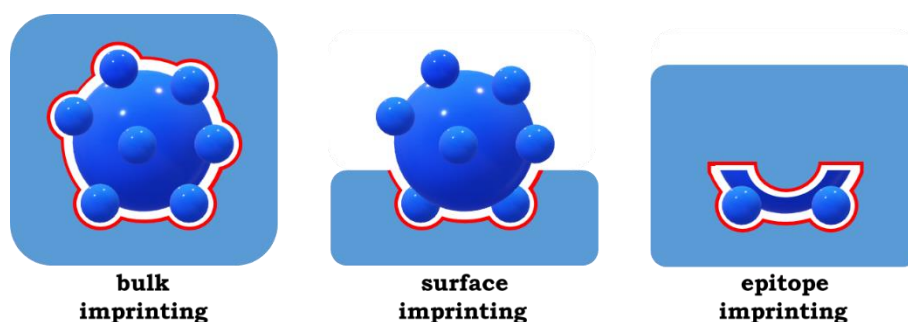


**Scheme 3.1** Molecular imprinting to form binding pockets which are complementary to target molecules.<sup>6</sup>

MIPs have several advantages over their biologically derived counterparts. They are robust, reusable, more stable to physical and chemical extremes, easier to synthesise, and integrate well with transducer elements for sensing prospects.<sup>7</sup> However, a major disadvantage of these polymers is their inability to sense analytes of high molecular weight (>1000Da). This is due to limitations in the pore size of the polymer resulting in restrictions to the diffusion of these molecules through the polymer network during the template removal process. If pore sizes are smaller than that of the target, it can lead to entrapment of large molecules within the bulk polymer. In addition, if the template is trapped, it further prevents target molecules reaching imprinted sites, leading to a loss in affinity. As a result, most MIPs are employed in the detection of targets of low molecular weights.<sup>8</sup>

Recently, techniques have been developed to overcome this limitation by increasing the surface area to volume ratio of the MIPs. One method is to grind the MIP to expose more imprinted sites, but this process can lead to the destruction of these sites.<sup>9</sup> Epitope imprinting entails the use of a specific portion of the large molecule as the template analyte, e.g. a short peptide chain of a protein is used to represent the whole molecule.<sup>10</sup> While this method avoids entrapment of the template in the bulk polymer, it does not overcome the diffusion problems faced when probing interactions with the larger target. Surface imprinting of large targets eliminates the problem of bulk polymer network diffusion as imprints are only formed at the surface of the polymer.<sup>11</sup> However, as only a portion of the template forms an interaction with the polymer during the imprinting process, the overall affinity is reduced compared to that

of bulk imprinting, as there are fewer interactions between the polymer and the analyte (Figure 3.1).



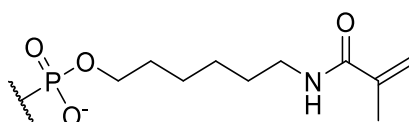
**Figure 3.1** Different methods used to form molecularly imprinted polymers of large targets.

Although surface imprinted MIPs show good affinities to large targets like proteins, a reduced selectivity is observed compared to their biological counterparts.<sup>8</sup> The bigger the target, the worse the selectivity. A possible solution to this issue is to use non classical monomers to enhance selectivities.<sup>12</sup> Alternatively, biologically derived elements could be introduced into MIPs by incorporating peptides and/or oligonucleotides into the polymer architectures. Recent studies by Turner and co-workers have introduced modified nucleotides,<sup>13</sup> short oligonucleotide sequences,<sup>14</sup> as well as DNA aptamers<sup>15</sup> into polymers for sensing different targets. The sequences were modified with polymerisable groups which, upon complexation to the analyte of interest, reacted with other monomers to form hybrid aptamer-MIP nanoparticles. In addition, it was shown that the nanoparticles were resistant to nuclease degradation. The Lui group created MIPs with DNA aptamer fragments as monomers. The adenosine aptamer was split into two halves, fluorescently labelled, and copolymerized into MIPs. They observed that the MIPs doubled the binding affinity of the aptamer to adenosine compared to a mixture of the free



aptamer fragments, with dissociation constants of 16 $\mu$ M for the MIP and 33 $\mu$ M for the free aptamer fragments.<sup>16</sup> Furthermore, Jolly *et al.* designed aptamer-MIP hybrids for sensing prostate specific antigen (PSA).<sup>17</sup> Higher PSA selectivities were described for the hybrid when compared to the aptamer alone, with limits of detection of 1pg/ml for the hybrid and 1 $\mu$ g/ml for the aptamer alone. The hybrid also gave low responses to the homologous protein human Kallikrein 2 and human serum albumin, demonstrating specificity towards PSA.

All examples of aptamer-MIP hybrids in the literature showed heightened interactions of small and large targets to MIPs when aptamer-based monomers were incorporated, yet none of the studies investigated the effect of anchorage of the aptamer monomer on the binding affinity of the hybrid. With the exception of Turner's work,<sup>15</sup> the aptamers mentioned<sup>16, 17</sup> were modified with an acrylamide group at the end of the sequence in the form of an Acrydite™ tag (Figure 3.2). Turner's work on cocaine aptamer MIPs briefly touched on the subject, commenting that multiple anchorage points of the DNA to the polymer network are advantageous for binding, although the results shown to verify this were limited.<sup>15</sup>

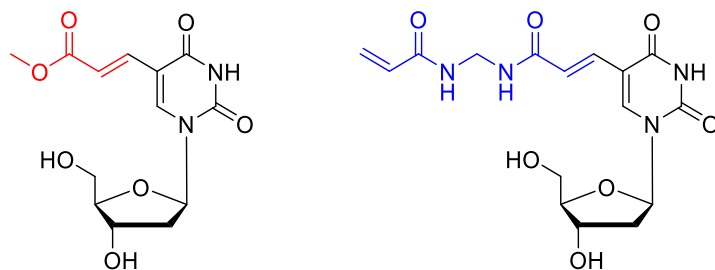


**Figure 3.2** Chemical structure of the Acrydite™ tag used to couple oligonucleotides with an acrylamide group.<sup>18</sup>

## 3.2 Aim

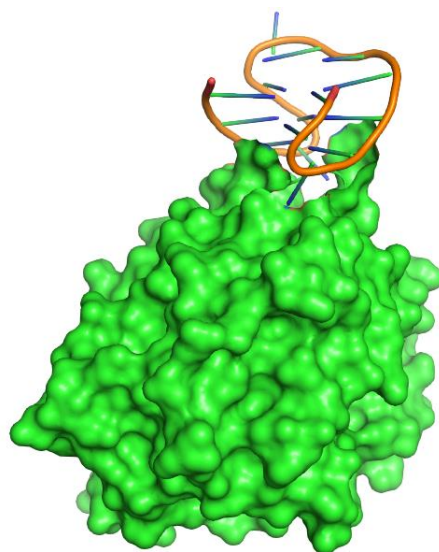
The aim of this chapter is to investigate how the flexibility of an aptamer incorporated into a molecularly imprinted polymer affects the binding capabilities of the aptamer-MIP hybrid. It is not known whether multiple anchorage points of the aptamer to the polymer would enhance or reduce interactions of the monomer to the target substrate. As aptameric sequences fold into specific 3D structures in order to bind to analytes with high specificity, it is anticipated that restricting the flexibility of this structure within the polymer would be beneficial to the binding capabilities of the hybrid. One key point to note is the placement of polymerisable groups within the aptamer's structure. If these groups are placed in regions where the target molecule associates with the aptamer, they could potentially destabilise (or stabilise) interactions between the aptamer and the template.

The first objective was to select an appropriate moiety for crosslinking. An acrylate modified thymidine base designed by the Turner group was chosen to insert multiple acrylate groups into an aptamer sequence. This would be compared with an acrylamide modified thymidine base for the incorporation of multiple acrylamide groups (Figure 3.3). The ability of the modified thymidine bases to polymerise would then be tested prior to insertion into DNA.



**Figure 3.3** Acrylate (left) and acrylamide (right) modified deoxythymidine molecules to be used in aptamer-MIP hybrids.

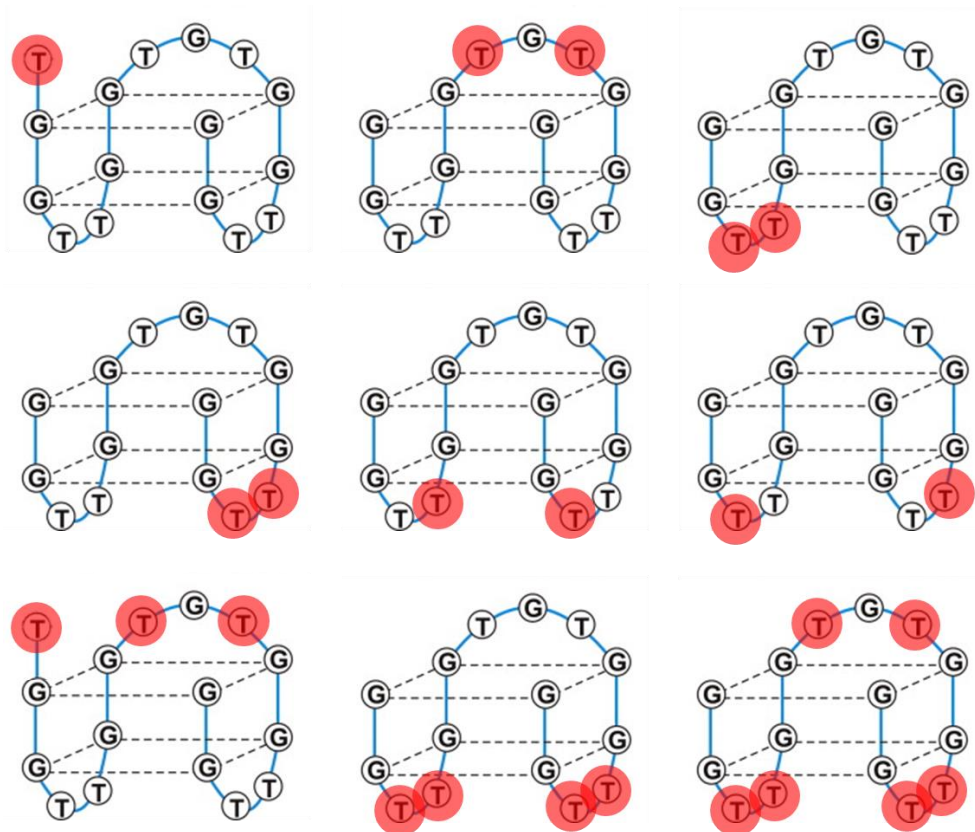
The second objective was to synthesise a suitable aptamer sequence modified with polymerisable groups in various positions. The 15 base thrombin binding aptamer sequence was chosen for modification as the structure of the aptamer and its association with the thrombin protein is widely studied due to its potential use in anticlotting therapeutics.<sup>19, 20</sup> Crystal structures of the aptamer-thrombin complex show the aptamer in an anti-parallel G-quadruplex conformation, with thrombin associating with the bottom T-T loops of the aptamer (Figure 3.4).<sup>21-23</sup> 40% of the thrombin binding aptamer sequence comprise of thymidine bases, giving plenty of scope to insert multiple polymerisable groups into the aptamer *via* this base.



**Figure 3.4** Aptamer-thrombin complex.

Modified thrombin binding aptamer sequences with single and multiple acrylate/acrylamide groups in various sites were to be constructed (Figure 3.5). Placement of the polymerisable groups in the two bottom T-T loops of strands were anticipated to affect aptamer-thrombin binding, whereas placement on other sites were expected to have less of an effect.

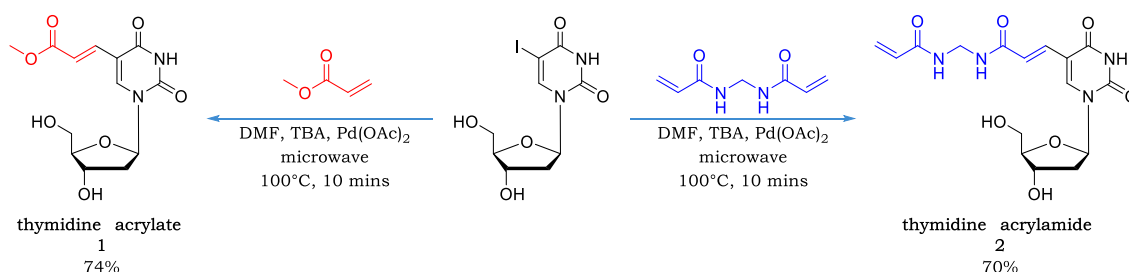
The third objective was to determine the ability of the modified strands to bind to thrombin/complementary DNA targets, with the final objective to react the best performing modified sequences with surface-based acrylamide polymers for biosensing with MIPs.



**Figure 3.5** Structures of the target modified thrombin binding aptamer strands. Location of residues with polymerisable groups is highlighted in red.

### 3.3 Synthesis of thymidine acrylate and thymidine acrylamide

Thymidine acrylate **1** was initially selected as a suitable base for aptamer-MIP studies as it has been previously used in the production of aptamer-MIP nanoparticles in the Turner group.<sup>15</sup> However, concerns over the reactivity of the double bond in this molecule due to sterics led to the design of thymidine acrylamide **2**, where the reacting alkene is placed in a terminal position. Both modified bases were synthesised from 5-iodo-2'-deoxyuridine using a microwave assisted palladium coupling procedure adapted from Ami and Fujimoto, as shown in Scheme 3.2.<sup>24</sup> The reaction was monitored by thin layer chromatography (TLC) which took approximately 10 minutes to complete. Purification was performed by recrystallization from chloroform.

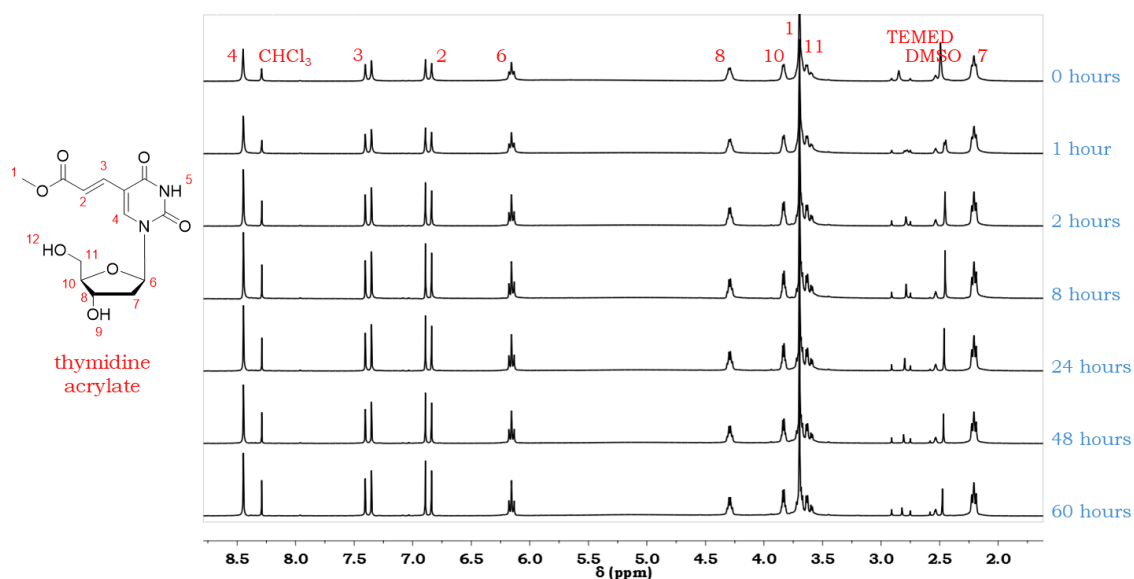


**Scheme 3.2** Synthesis of thymidine acrylate **1** and thymidine acrylamide **2**.

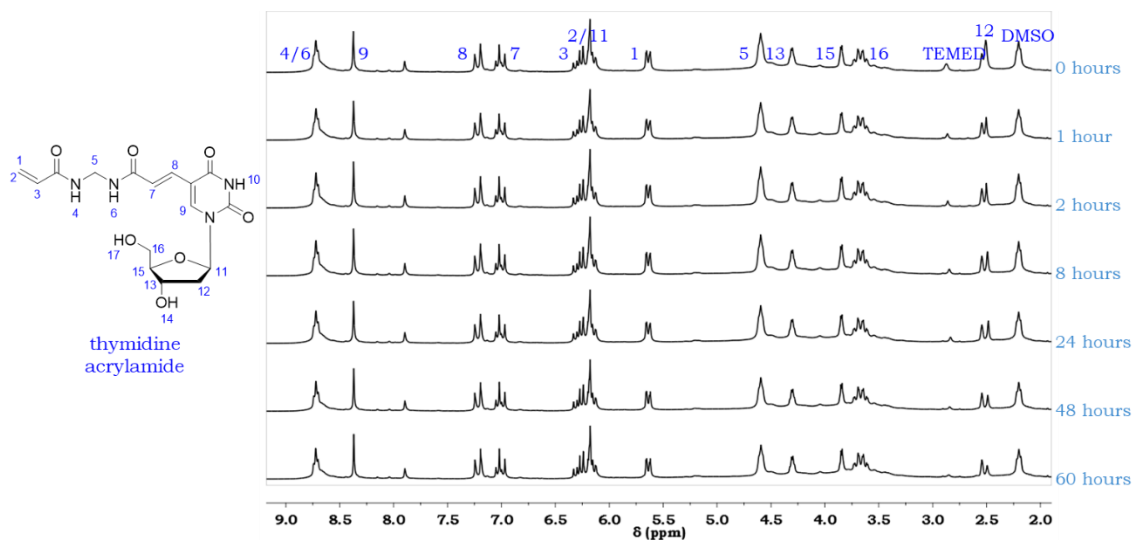
### 3.4 $^1\text{H}$ NMR homo- and co- polymerisation experiments with thymidine acrylate and thymidine acrylamide

#### 3.4.1 $^1\text{H}$ homopolymerisation experiments

$^1\text{H}$  NMR spectroscopy was utilised to monitor the homopolymerisation of the modified bases in deuterated DMSO, using ammonium persulphate (APS) and tetramethylethylenediamine (TEMED) as initiators in a sealed environment. Spectra were recorded after 0, 1, 2, 8, 24, 48, and 60 hours of reaction, as shown in Figure 3.6 and Figure 3.7. It was anticipated that resonances corresponding to the monomers would decrease as polymerisation proceeded. Moreover, new resonances associated with the polymer would appear and increase over time.



**Figure 3.6**  $^1\text{H}$  NMR spectra of the attempted homopolymerisation of thymidine acrylate **1** using APS/TEMED initiators in deuterated DMSO over 60 hours.



**Figure 3.7**  $^1\text{H}$  NMR spectra of the attempted homopolymerisation of thymidine acrylamide **2** using APS/TEMED initiators in deuterated DMSO over 60 hours.

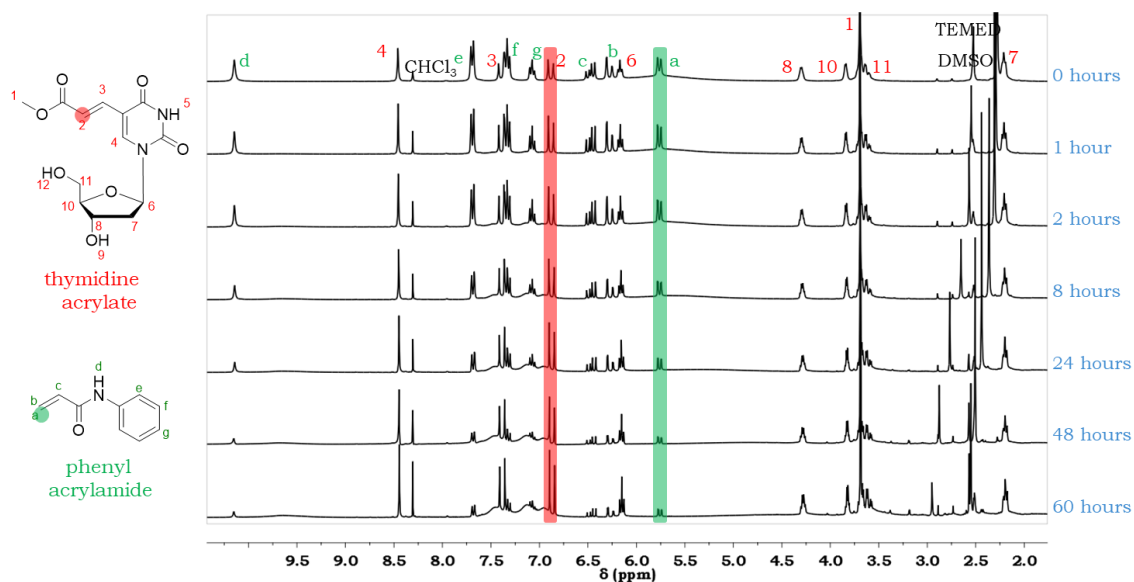
$^1\text{H}$  NMR results indicated no changes to either of the modified bases over 3 days, strongly suggesting that neither base could undergo homopolymerisation under the conditions given. This result was unexpected but advantageous as it indicated that aptamer sequences containing these bases would be unable to homopolymerise intra- or inter-molecularly in solution. The formation of the aptamer-MIP hybrid is dependent on the copolymerisation of reactive groups on the aptamer with acrylamide monomers in solution and on the surface, and thus any homopolymerisation between or within aptamer molecules would hinder this process.

### [3.4.2 \$^1\text{H}\$ copolymerisation experiments](#)

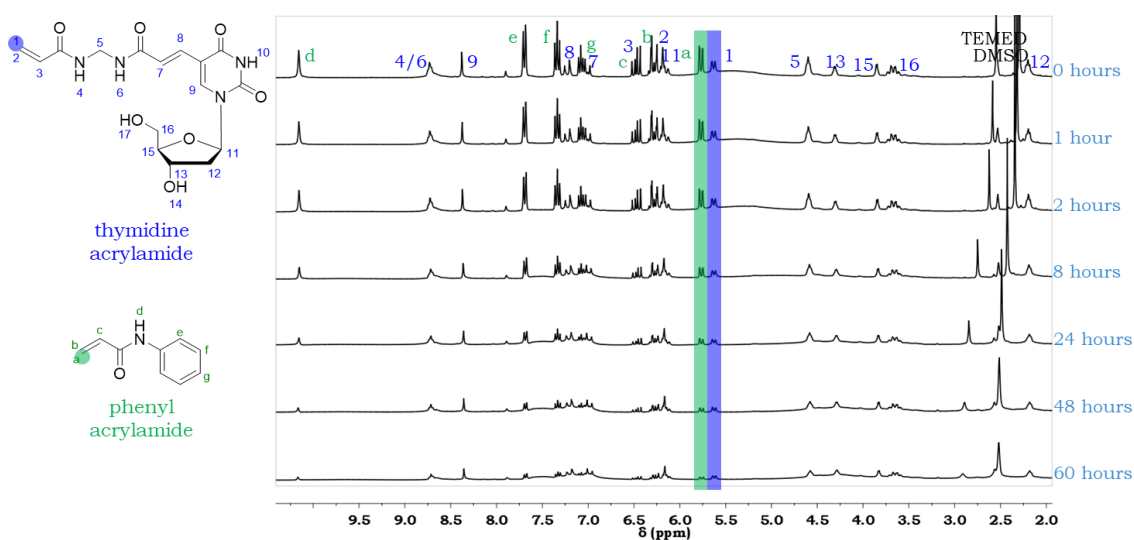
The same technique was next used to analyse the co-polymerisation of the modified bases in solution, as the ability of these bases to form copolymers is crucial to the formation of surface imprinted aptamer-MIPs. Identical conditions to those used for the homopolymerisation experiments were used with the



addition of an equal concentration of phenyl acrylamide as a co-monomer. Spectra were recorded after 0, 1, 2, 8, 24, 48, and 60 hours of reaction, as shown in Figure 3.8 and Figure 3.9.



**Figure 3.8**  $^1\text{H}$  NMR spectra of the attempted copolymerisation of thymidine acrylate **1** with phenyl acrylamide using APS/TEMED initiators in deuterated DMSO over 60 hours.



**Figure 3.9**  $^1\text{H}$  NMR spectra of copolymerisation of thymidine acrylamide **2** with phenyl acrylamide using APS/TEMED initiators in deuterated DMSO over 60 hours.

Over time, an insoluble product was observed at the bottom of the NMR tube in both copolymerisation experiments (Figure 3.10). The  $^1\text{H}$  NMR spectra for the copolymerisation of thymidine acrylate **1** with phenyl acrylamide showed a decrease in resonances associated with phenyl acrylamide but no change in the resonances matching the thymidine acrylate monomer. The resonances corresponding to alkene protons 2 and a in Figure 3.8 are highlighted to show this effect. New broad product peaks are observed suggesting that homopolymerisation of phenyl acrylamide is taking place instead of the expected copolymerisation reaction between the two monomers. However in contrast, the  $^1\text{H}$  NMR spectra for the copolymerisation of thymidine acrylamide **2** with phenyl acrylamide showed that all resonances for both monomers decreased over time. Interestingly, the resonances associated with phenyl acrylamide appeared to decrease at a faster rate, indicating faster reactions of this monomer than thymidine acrylamide. The signals related to alkene protons 1 and a in Figure 3.9 are highlighted to show this effect. No new product resonances are observed, presumably due to the insolubility of the copolymer product in the medium used for these experiments.

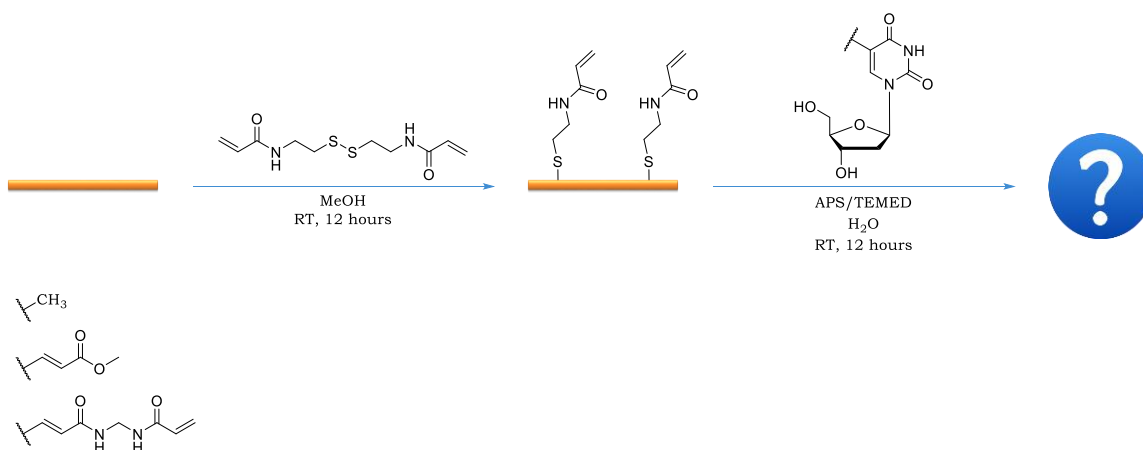


**Figure 3.10** Insoluble polymer products formed in deuterated DMSO during copolymerisation reactions.

In summary,  $^1\text{H}$  NMR analysis of the ability of the modified bases to polymerise in solution shows that thymidine acrylate **1** is unable to form homo- or co-polymers under the conditions tested, whereas thymidine acrylamide **2** can form copolymers.

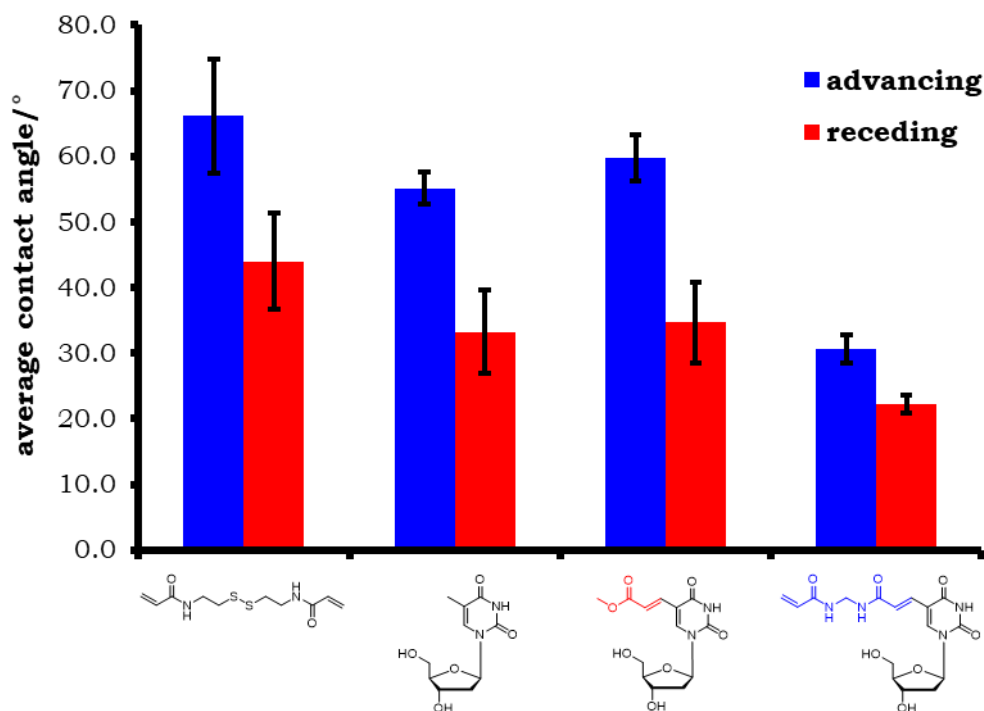
### 3.5 Contact angle experiments with thymidine acrylate and thymidine acrylamide

Contact angle analysis was used to examine the success of surface copolymerisation between the modified bases and an acrylamide monomer. Acrylamide SAMs on gold were formed by overnight incubation of gold chips in methanolic solutions of 1mM N,N'-bis(acryloyl)cystamine (BAC). The acrylamide SAMs were then reacted with 1 $\mu$ M unmodified or acrylate/acrylamide modified deoxythymidine in water at room temperature overnight, using APS and TEMED as initiators (Scheme 3.3).



**Scheme 3.3** Formation of acrylamide SAMs on gold and subsequent reaction with unmodified/modified deoxythymidine.

The dynamic contact angles of the reacted surfaces were then measured. It was predicted that the contact angles of the reacted surfaces should be lower than that of the acrylamide SAMs, as a result of the increased hydrophilicity of the surface now appended with the hydrophilic modified bases. Figure 3.11 shows the contact angles of all the series as an average of six experiments.



**Figure 3.11** Advancing and receding contact angles of acrylamide SAMs on gold before and after reaction with APS/TEMED initiators and modified or unmodified deoxythymidine.

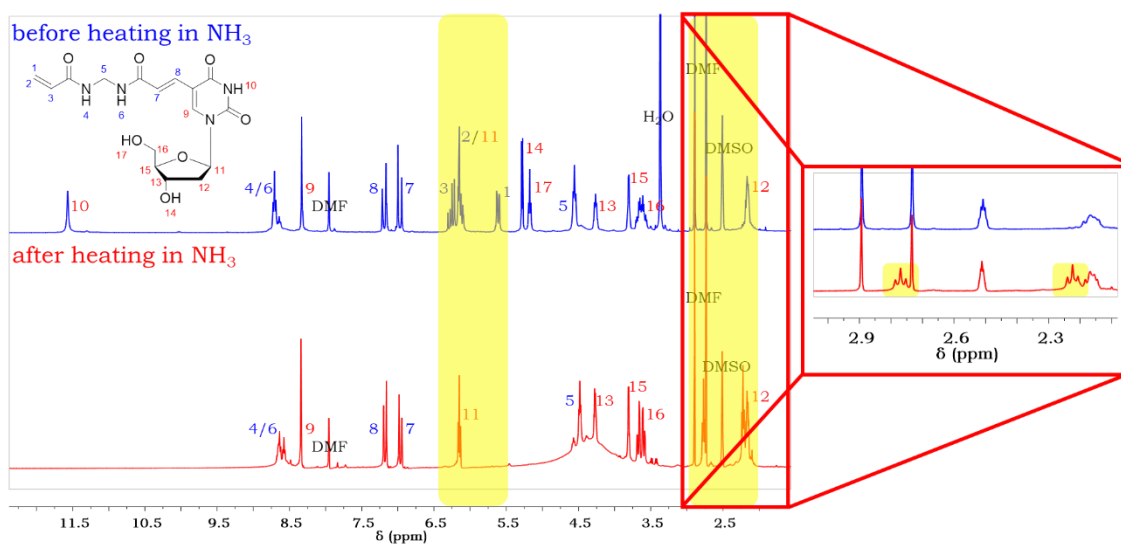
The contact angle data shows comparable results for all conditions with the exception of thymidine acrylamide reacted surfaces which were significantly lower, with advancing and receding angles of  $31^\circ \pm 2^\circ$  and  $22^\circ \pm 1^\circ$  compared to advancing and receding angles of  $66^\circ \pm 9^\circ$  and  $44^\circ \pm 7^\circ$  with acrylamide SAMs. This suggests that the thymidine acrylamide base had successfully coupled to the acrylamide SAMs, resulting in a more hydrophilic surface and lower contact angle. In agreement with  $^1\text{H}$  NMR spectroscopic analysis, thymidine acrylate does not appear to have reacted, as the contact angle data for this surface is similar to that of the acrylamide SAMs with advancing and receding angles of  $60^\circ \pm 4^\circ$  and  $35^\circ \pm 6^\circ$  respectively. However, thymidine acrylates have been demonstrated to incorporate into MIPs by Turner and co-workers using similar

initiators and monomers.<sup>15</sup> This would suggest that conditions used to conduct the polymerisation are highly sensitive to change, and thus require further development. It is noted that the alkenyl residue in the thymidine acrylate base is internal rather than terminal, which would hinder the monomer's ability to polymerise through steric effects.

To summarise, contact angle analysis demonstrated the ability of thymidine acrylamide to form copolymers with acrylamide SAMs on surfaces, whereas thymidine acrylate was unable to form copolymers under the conditions used. Due to poor results obtained with the thymidine acrylate system both in solution and on surfaces, it was decided to solely focus on the incorporation of the thymidine acrylamide base into DNA aptamers for integration into molecularly imprinted polymers.

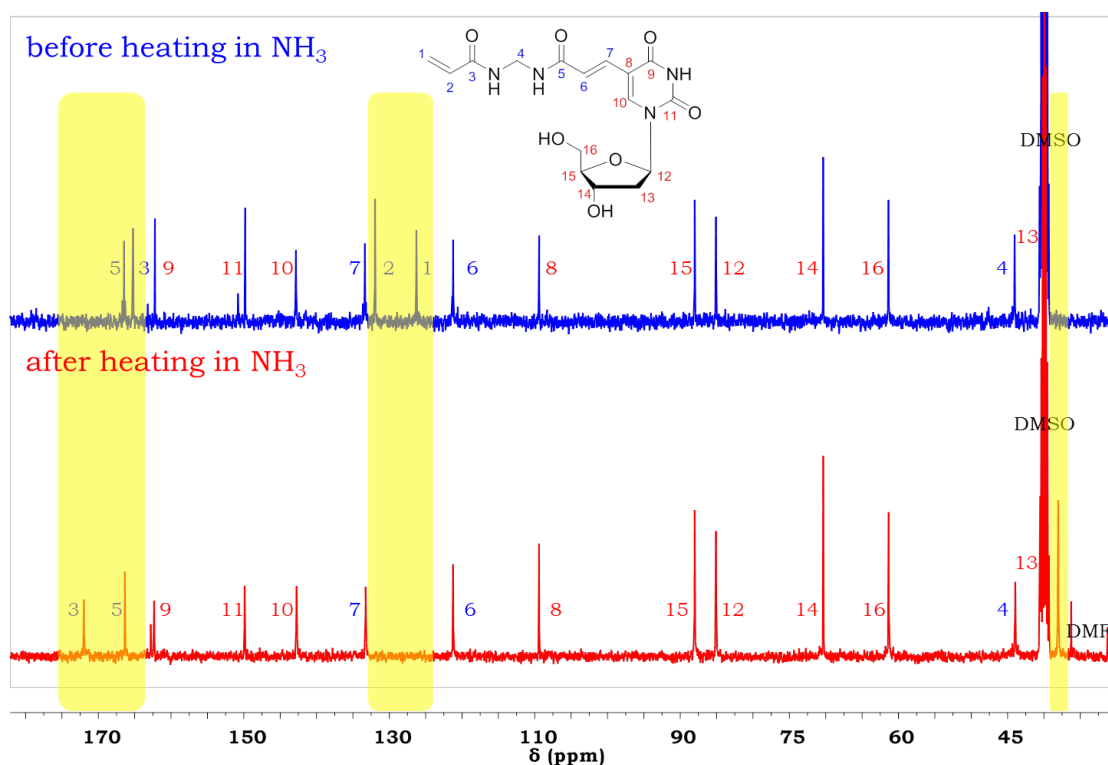
### 3.6 Testing the stability of thymidine acrylamide towards the oligonucleotide deprotection process

Automated oligonucleotide synthesis entails some relatively harsh chemical conditions, and any modification must be able to survive these conditions for incorporation to be successful. The harshest stage is nucleobase deprotection, where the DNA strands are heated to 60°C in aqueous ammonia for a minimum of six hours. There were concerns over whether the acrylamide group on thymidine acrylamide could survive this process, and so it was decided that this condition would be tested on the base prior to its insertion into DNA. A small amount of thymidine acrylamide was dissolved in aqueous ammonia, and the resulting solution was heated to 60°C for six hours. The solvent was removed, and the crude was analysed by  $^1\text{H}$  NMR spectroscopy (Figure 3.12).



**Figure 3.12**  $^1\text{H}$  NMR spectra of thymidine acrylamide in deuterated DMSO before and after heating in aqueous ammonia.

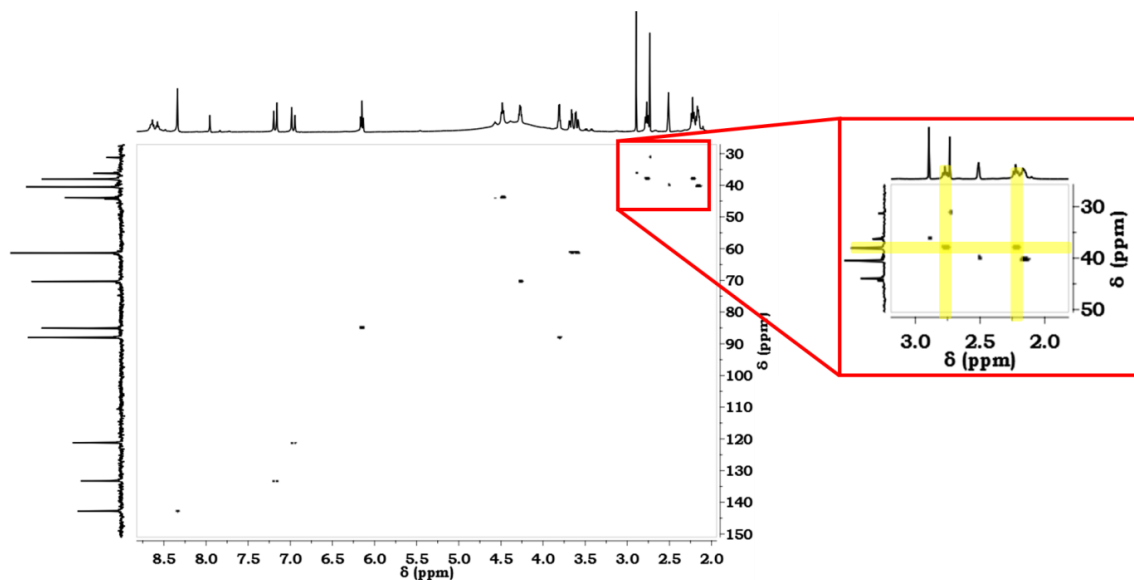
Figure 3.12 shows a notable change to the acrylamide moiety after heating. Signals corresponding to the terminal alkene protons have disappeared, and new upfield resonances have appeared, as shown in the inset. It was suspected that the amide groups on the molecule may have hydrolysed, and the crude was analysed further by  $^{13}\text{C}$  (Figure 3.13) and 2D NMR (Figure 3.14 and Figure 3.15) to test this.



**Figure 3.13**  $^{13}\text{C}$  NMR spectra of thymidine acrylamide in deuterated DMSO before and after heating in aqueous ammonia.

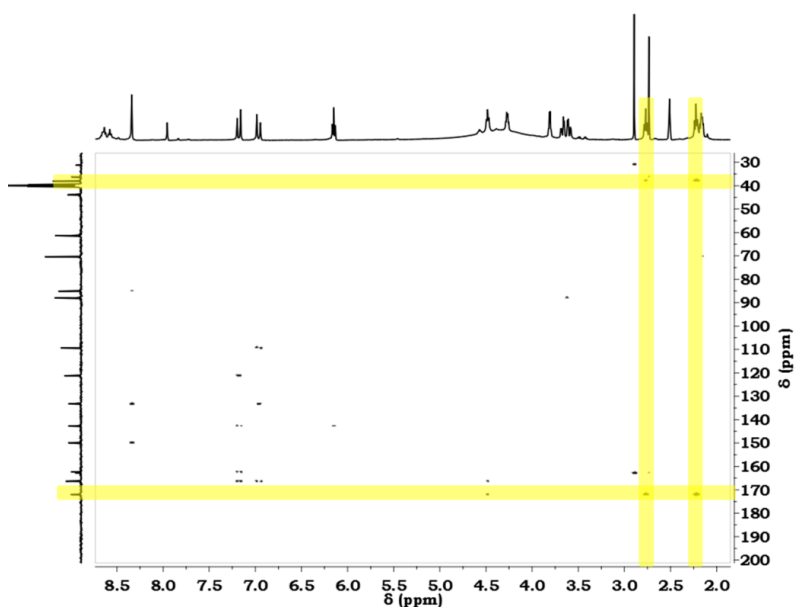
The bottom  $^{13}\text{C}$  NMR spectrum in Figure 3.13 showed the signal corresponding to the carbonyl carbon directly attached to the terminal alkene, indicating amide hydrolysis had not occurred, although this signal had shifted downfield from 165.2ppm to 172.0ppm. However, the resonances corresponding to the terminal alkene were absent, and there was a new peak at 38.0ppm.





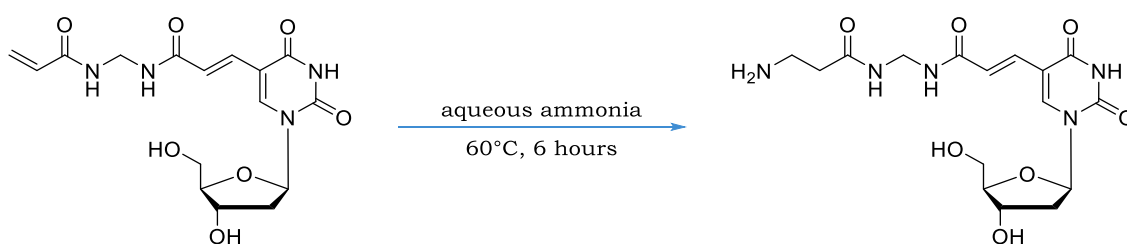
**Figure 3.14** HSQC NMR spectrum of thymidine acrylamide in deuterated DMSO after heating with aqueous ammonia.

The HSQC NMR spectrum in Figure 3.14 showed a direct connection between the two triplet signals at 2.77ppm and 2.22ppm, and the new carbon signal at 38.0ppm, indicating the formation of two methylene groups in the molecule.



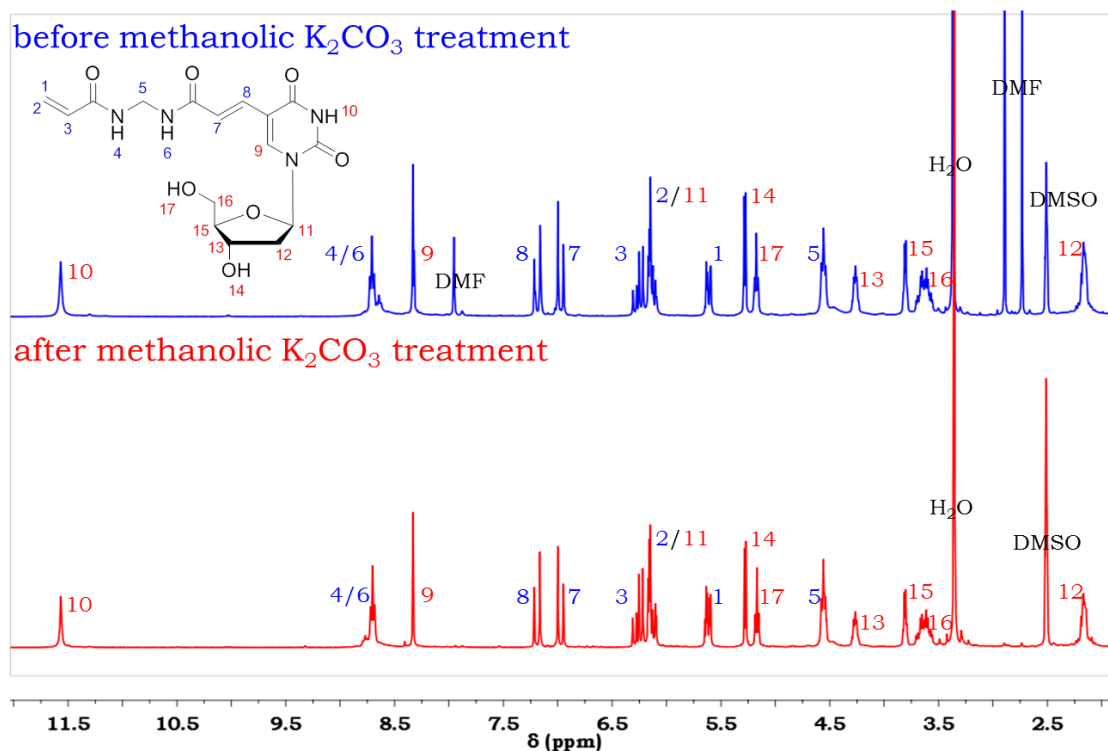
**Figure 3.15** HMBC NMR spectrum of thymidine acrylamide in deuterated DMSO after heating with aqueous ammonia.

HMBC NMR data (Figure 3.15) showed that the new methylene groups were close to the carbonyl carbon at 172.0ppm, implying reduction of the terminal alkene. The crude was further analysed by mass spectrometry and produced a mass 17 units higher than the starting material, strongly suggesting the addition of ammonia to the terminal double bond to form a primary amine (Scheme 3.4).



**Scheme 3.4** Reaction of thymidine acrylamide with hot aqueous ammonia.

These studies showed that thymidine acrylamide cannot withstand the standard oligonucleotide deprotection process, as the acrylamide group can react with ammonia to form a primary amine. Therefore, the molecule was tested with an alternative deprotection process which substitutes ammonia for methanolic potassium carbonate solution. Termed the ultramild deprotection process, alternative protecting groups are used on the nucleobases for deprotection under milder basic conditions. The process requires no heating of the strands and solutions are typically left at room temperature overnight for successful deprotection. Thymidine acrylamide was tested under ultramild deprotection conditions, and the crude analysed by  $^1\text{H}$  NMR spectroscopy, as shown in Figure 3.16.



**Figure 3.16**  $^1\text{H}$  NMR spectra of thymidine acrylamide in deuterated DMSO before and after treatment with potassium carbonate in methanol.

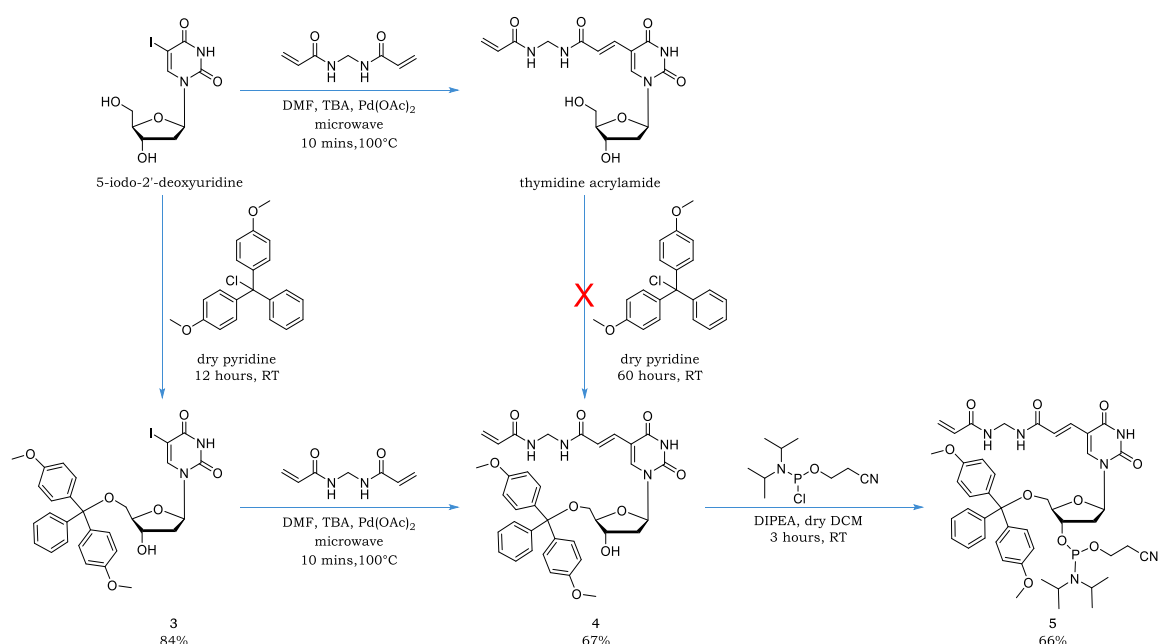
This time, the results showed no change to the structure of the molecule, with the terminal alkene still intact, and thus thymidine acrylamide can be incorporated into DNA under ultramild deprotection conditions.

To conclude, the acrylamide moiety cannot survive standard oligonucleotide deprotection conditions as the terminal alkene is susceptible to nucleophilic attack by ammonia. It was anticipated that alternative amines commonly used for nucleobase deprotection, e.g. methylamine, would experience the same result, although this theory was not tested. Fortunately, the terminal alkene can withstand ultramild deprotection conditions using methanolic potassium carbonate, and thus these conditions were to be applied when incorporating thymidine acrylamide into DNA.

## 3.7 Synthesis of acrylamide modified thrombin binding aptamer strands

### 3.7.1 Synthesis of thymidine acrylamide phosphoramidite

Thymidine acrylamide had to first undergo tritylation and phosphorylation to allow incorporation of the base into DNA sequences using automated synthesis. Unfortunately, tritylation of the 5' position of the modified base with a dimethoxytrityl group proved difficult. An alternative synthetic route was therefore devised in which tritylation was performed on 5-iodo-2'-deoxyuridine prior to the palladium catalysed addition of bisacrylamide. This route proved successful and the 3' alcohol of the modified base was phosphitylated to obtain the final phosphoramidite **5**, as shown in Scheme 3.5.

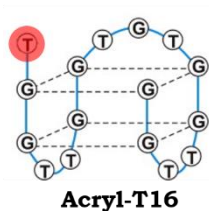


**Scheme 3.5** Synthesis of thymidine acrylamide derivatives.

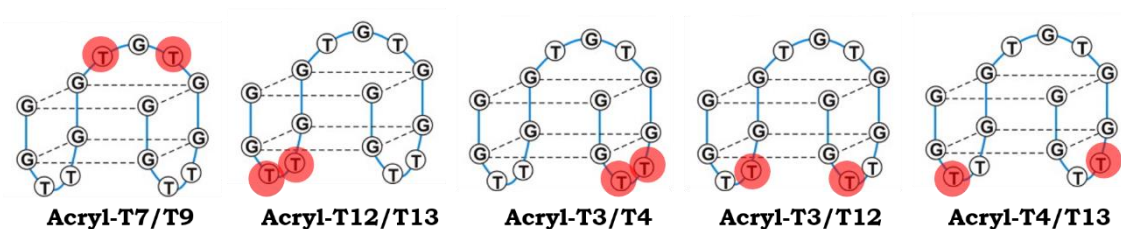
### 3.7.2 Synthesis of acrylamide modified aptamer strands

Thymidine acrylamide was successfully incorporated into DNA, generating high average stepwise yields of >90% for each coupling. Different acrylamide modified thrombin binding aptamer strands were produced under ultramild conditions, the structures of which are shown in Figure 3.17. The HPLC traces and mass spectra of these strands are provided in Appendix sections 6.6 and 6.7. As outlined in section 3.7.3, there were difficulties in purifying strands containing more than three acrylamide groups and so these were not studied further.

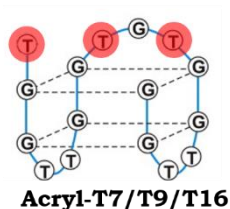
#### One acrylamide group



#### Two acrylamide groups



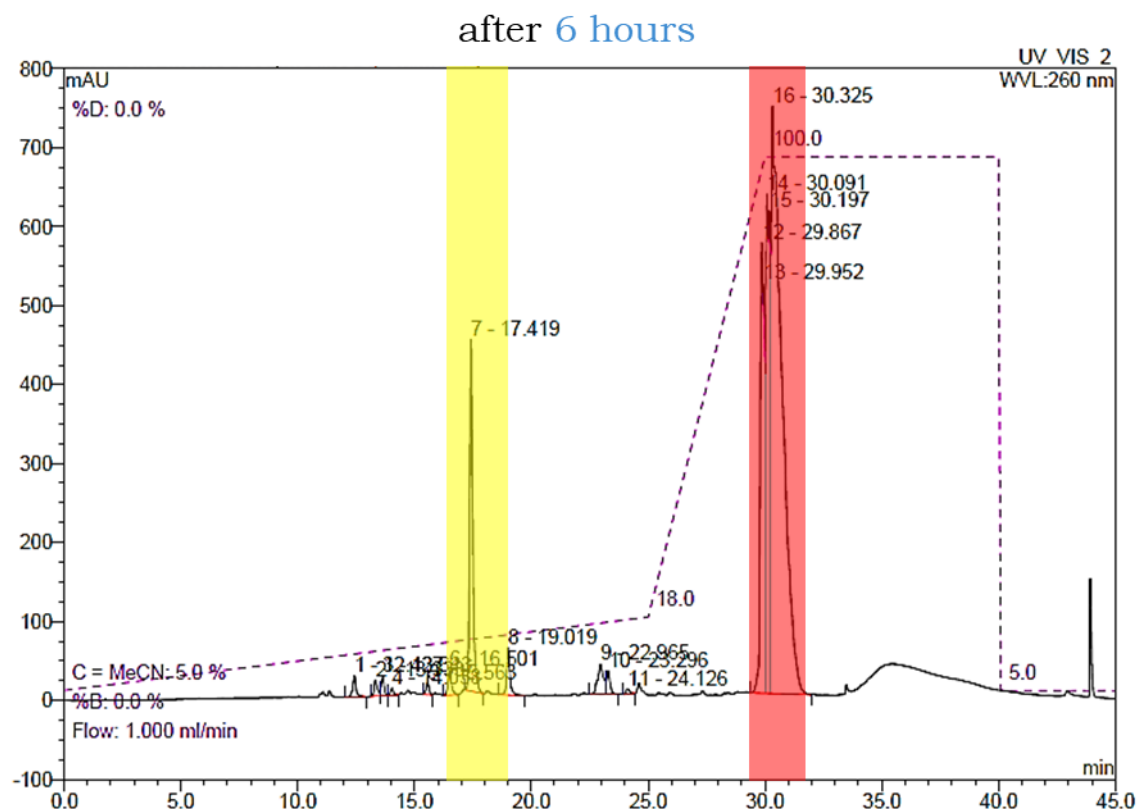
#### Three acrylamide groups

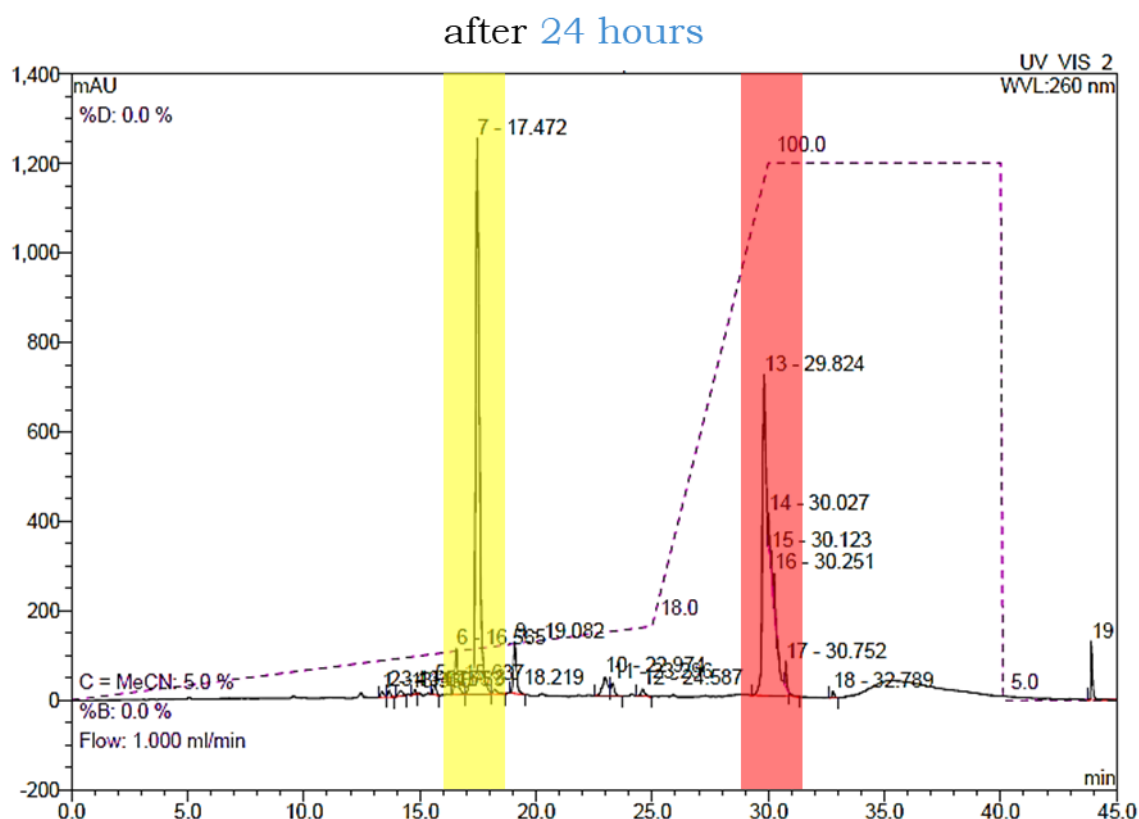


**Figure 3.17** G-quadruplex structures of successfully purified acrylamide modified thrombin binding aptamer strands. Highlighted T denotes incorporation of thymidine acrylamide.

### 3.7.3 Purification of acrylamide modified aptamer strands

The purification of the modified strands was hampered by the incomplete deprotection of the guanine bases. Protecting groups on guanine are inherently difficult to remove, and the step is considered to be the rate limiting step of the whole oligonucleotide synthesis process.<sup>25</sup> As 60% of the thrombin binding aptamer sequence consists of guanine bases, deprotection of these strands using ultramild conditions proved slower than anticipated. The introduction of multiple acrylamide groups into the strands appeared to slow this process even further, to a point where it became difficult to fully deprotect thrombin binding aptamer strands with more than three acrylamide groups. Figure 3.18 shows HPLC chromatograms of oligo **Acryl-T16** after ultramild deprotection.

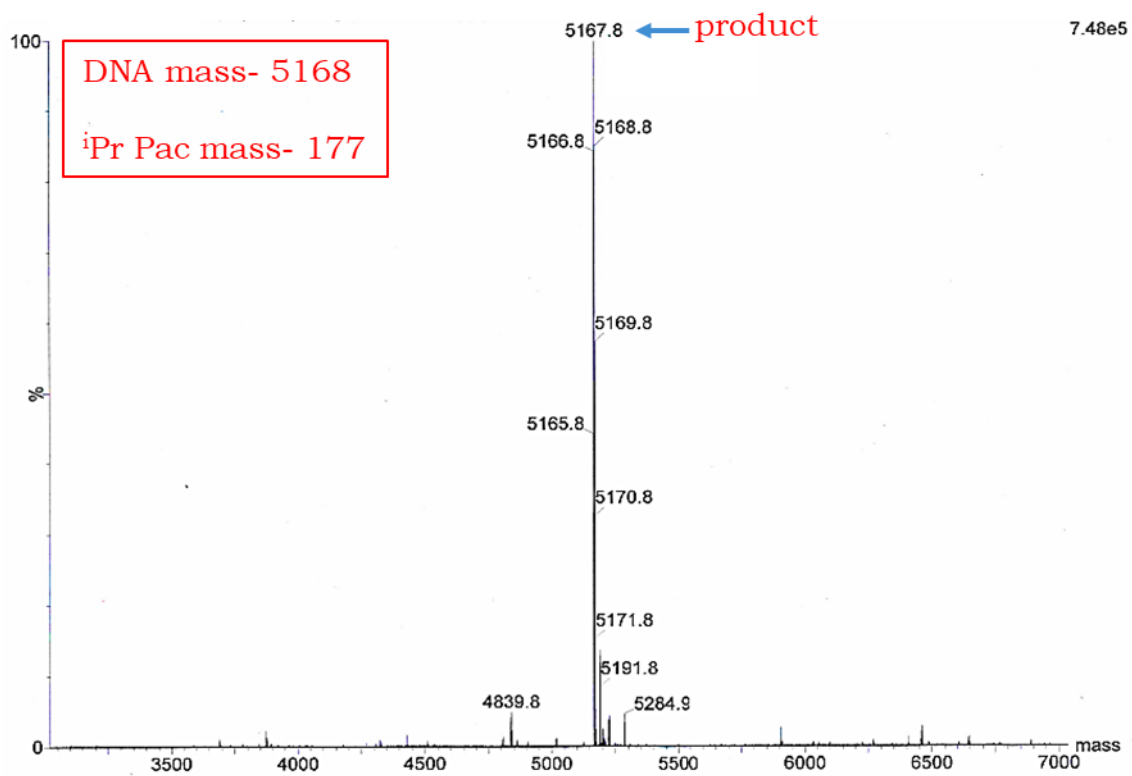
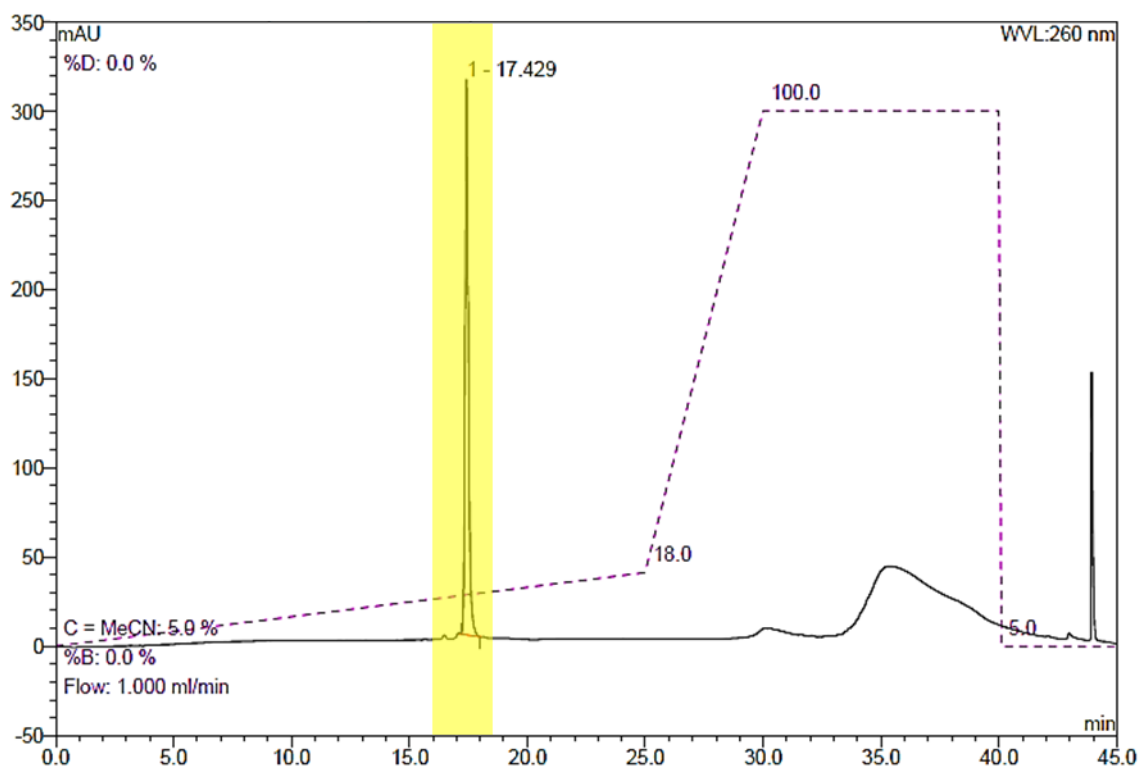




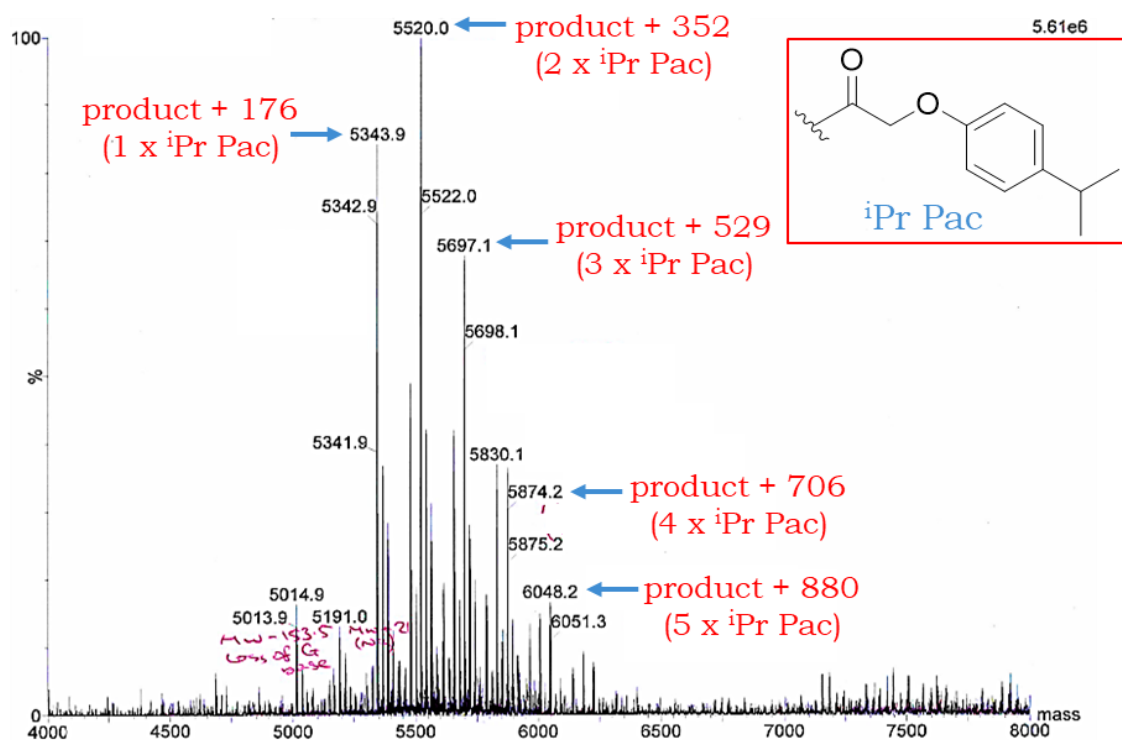
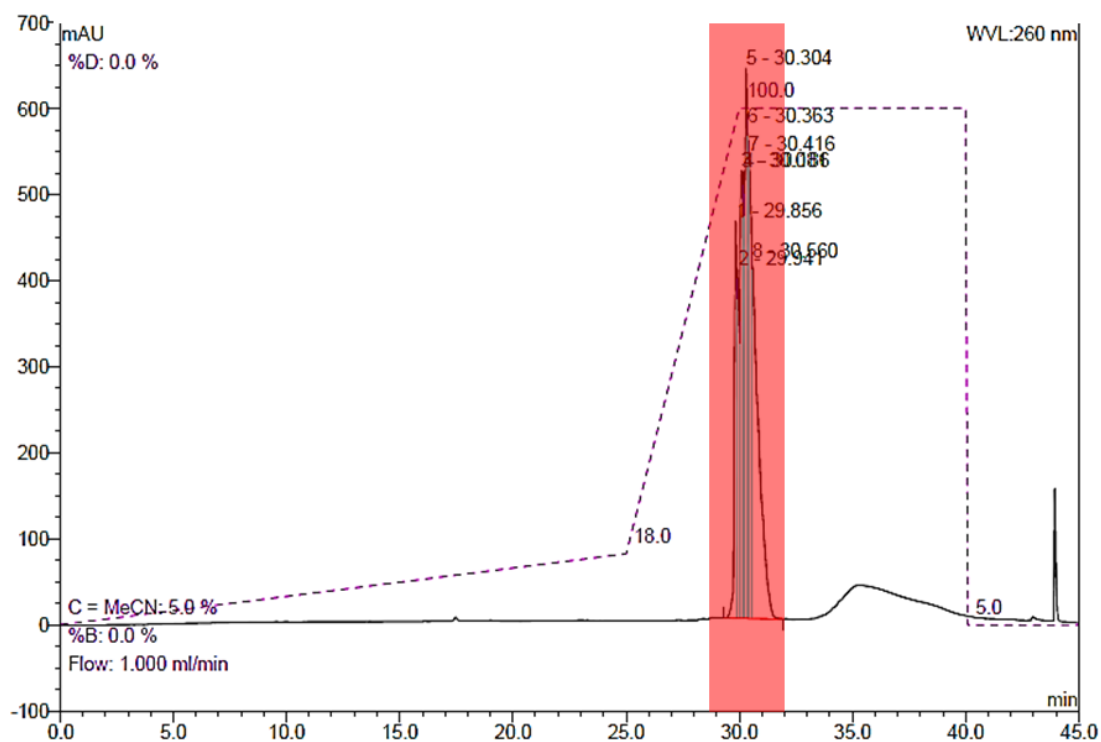
**Figure 3.18** HPLC chromatograms of oligo **Acryl-T16** after 6 and 24 hours of deprotection.

Fully deprotected strands are highlighted in yellow and partially protected strands are highlighted in red.

The crude solution produced two major peaks on HPLC. The peak at 17.5 minutes corresponds to the desired deprotected product, whereas the series of peaks around 30 minutes relates to various protected strands. Increasing the deprotection time from 6 to 24 hours led to an increase in the number of desired strands, but still incomplete deprotection. Isolation of the two peaks and analysis by mass spectrometry confirmed the presence of guanine protected strands. The mass spectrum of the peak at 30 minutes in Figure 3.19 shows mass differences of 177 units which match the mass of the *para*-isopropyl phenoxyacetyl group (*i*Pr Pac)- a group used to protect the amine moiety on the guanine base under ultramild conditions.

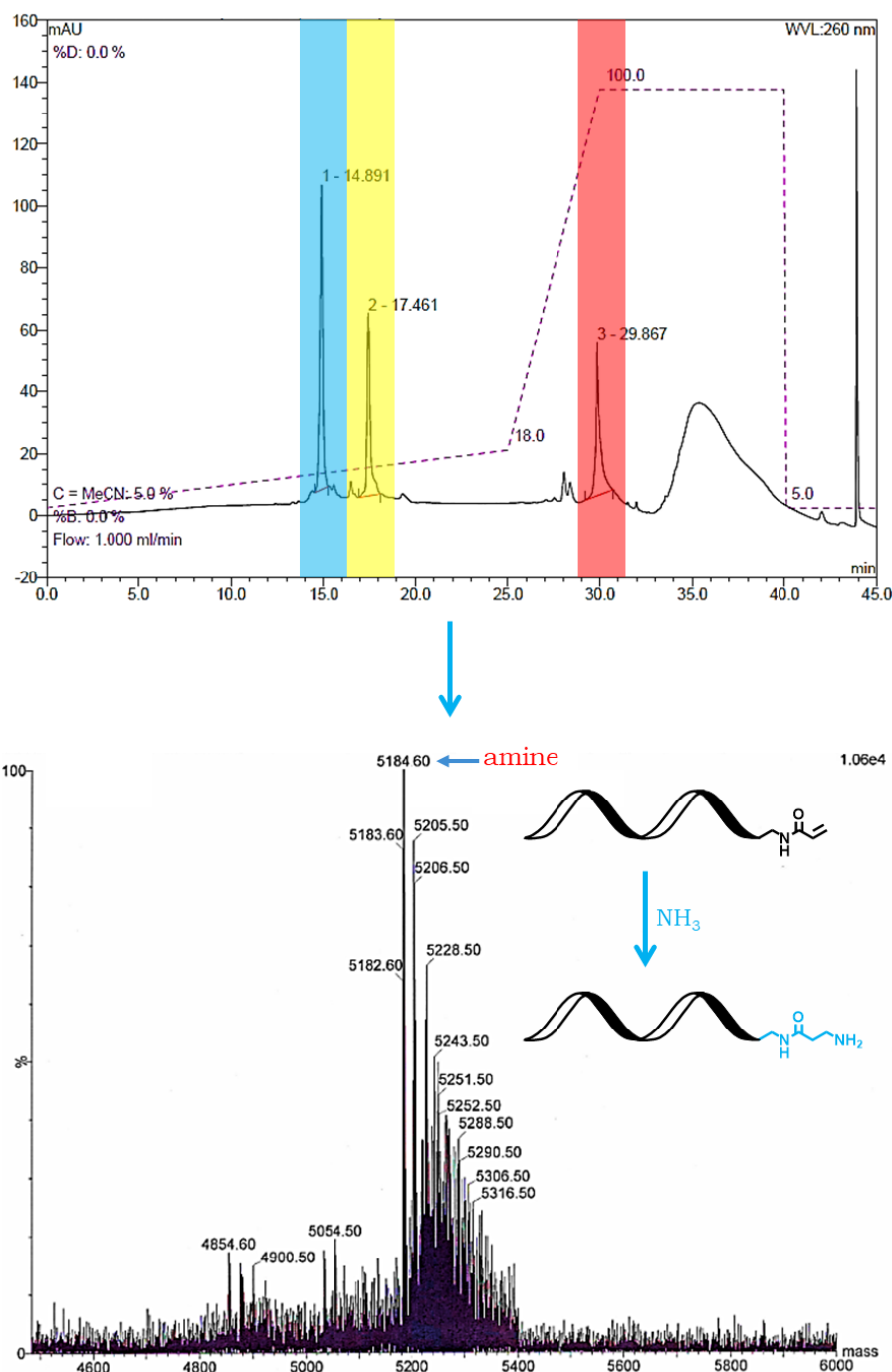






**Figure 3.19** Mass spectra of isolated peaks from the HPLC purification of the oligo **Acryl-T-16**.

The isolated protected strands were therefore incubated in aqueous ammonia overnight for further deprotection and the resulting crude was analysed by HPLC, as shown in Figure 3.20. This gave an additional peak at 14.9 minutes which was identified as the amine byproduct by mass spectrometry.

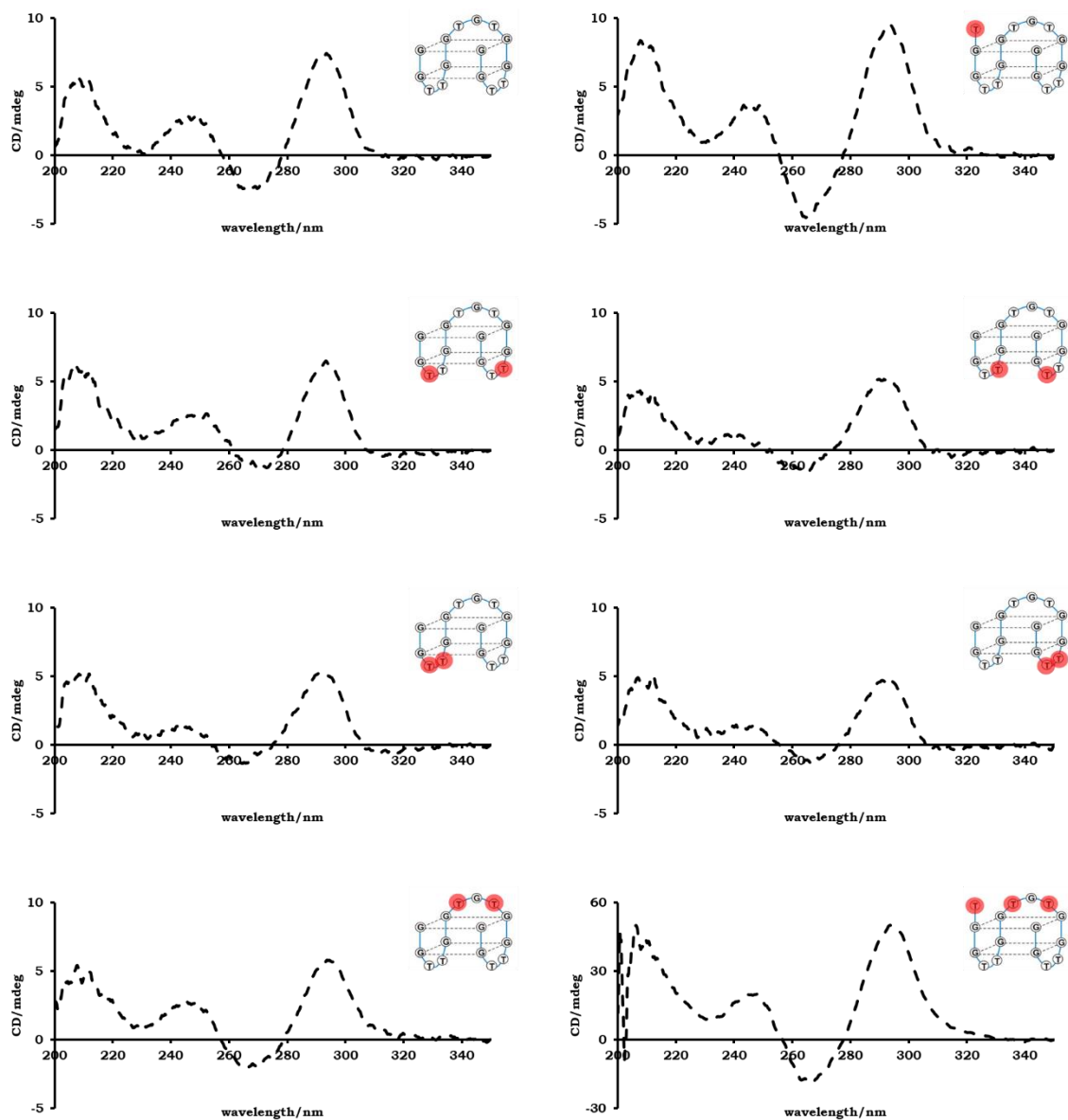


**Figure 3.20** HPLC chromatogram of the oligo **Acryl-T-16** after treatment with ammonia, and mass spectrum of the peak at 14.9 minutes.

Potassium is known to stabilise the formation of G-quadruplex structures in the thrombin binding aptamer sequence.<sup>26</sup> The folding of strands would hinder the deprotection of bases as the protecting groups are buried and hidden within the structure and thus inaccessible. Although the application of heat would denature the strands, allowing the protected sites to become accessible, this was not applied due to the low boiling point of methanol and the possibility of degrading the strands. It is currently not clear why the presence of multiple acrylamide groups in this particular sequence further inhibits the ultramild deprotection process, although it is presumed that this is due to steric effects that hinder reagent access.

### 3.8 Circular dichroism of unmodified and acrylamide modified thrombin binding aptamer strands

CD spectroscopy was used to probe the effect of acrylamide groups on the structure of the modified aptamer strands. In the presence of potassium, the thrombin binding aptamer has a characteristic CD maximum of around 295nm and a CD minimum of around 265nm, which is indicative of a G-quadruplex conformation. This conformation is essential for the aptamer to bind to its protein target.<sup>27</sup> It was anticipated that these signals would be observed for the modified sequences. Solutions of each strand in 1000-fold excess potassium chloride were prepared, annealed, and analysed by CD (Figure 3.21). The data showed no significant difference in peak positions of the CD signals of the modified strands compared to the unmodified strand. The strong peaks at ~295nm highlight this effect and indicate that all the modified strands can adopt a G-quadruplex conformation. However, the CD minima at ~265nm for strands containing acrylamide groups in the bottom T-T loops is not as pronounced as that for the unmodified strand, or for strands where acrylamide moieties are present in other sites. This reduction in signal indicates that positioning acrylamide groups in this region slightly weakens the G-quadruplex structure of the aptamer.



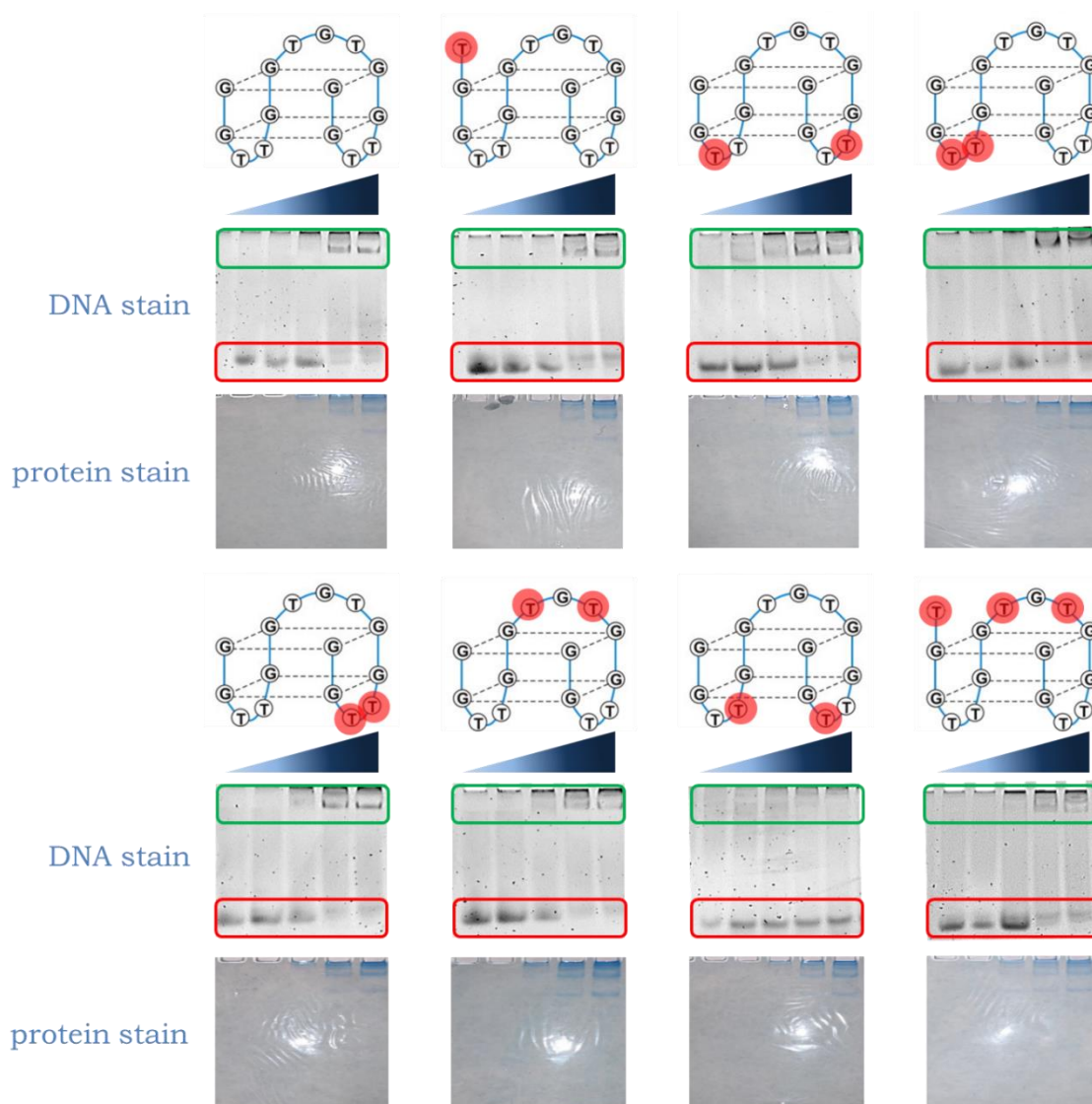
**Figure 3.21** CD spectra of unmodified and acrylamide modified thrombin binding aptamer strands.

### 3.9 Gel electromobility shift assay of unmodified and acrylamide modified thrombin binding aptamer strands with thrombin

To create successful aptamer-MIP hybrids, the acrylamide modified aptamer strands must first form a stable complex with the protein target. Gel electromobility shift assays were therefore performed to verify the presence of such complexes. Successful complexation of the aptamer and thrombin would result in a band on the polyacrylamide gel which is higher than that of the strand alone, as the complex would move more slowly down the gel when voltage is applied. Five solutions of 500nM unmodified or acrylamide modified aptamers containing Tris.HCl buffer and potassium chloride were made, each with increasing concentrations of thrombin to determine the binding ratio. Literature states a binding ratio of one aptamer strand for every thrombin molecule.<sup>28</sup> The solutions were incubated at room temperature for 1 hour prior to loading and the gels were run at 100V for 1 hour. They were then stained with Diamond™ nucleic acid dye and Coomassie blue dye to determine the positions of the aptamer, thrombin, and aptamer-thrombin complex (Figure 3.22).

The images show complete binding to thrombin for all but one of the seven modified stands at a ratio of 1:1 DNA: thrombin, matching the results observed with the unmodified aptamer. Only one strand exhibited no interaction of any kind to the thrombin protein: **Acryl-T3/T12** (bottom row, third gel). Acrylamide groups are positioned at T3 and T12 in this sequence and although CD data shows that the strand forms the G-quadruplex conformation necessary for thrombin binding, it appears that having acrylamide groups in these positions

is detrimental to thrombin binding. Oddly though, thrombin binding is observed with strands containing acrylamide groups in those sites in isolation, so the absence of binding in the T3/T12 strand appears to be synergistic.




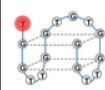
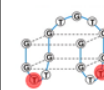

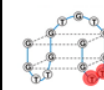
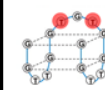
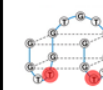
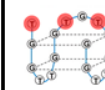
**Figure 3.22** Gel EMSA of unmodified and acrylamide modified aptamer (0.5 $\mu$ M) with increasing concentrations of thrombin (0 $\rightarrow$ 1 $\mu$ M). Positions of aptamer bands are highlighted in red and aptamer-thrombin complex bands are highlighted in green. Aptamer:thrombin ratios used in each lane are as follows: lane 1- 1:0, lane 2- 1:0.02, lane 3- 1:0.2, lane 4- 1:1, lane 5- 1:2.

### 3.10 Thermal melting experiments of unmodified and acrylamide modified thrombin binding aptamer strands with complementary sequence

Alongside the plan to use thrombin as a target for the aptamer-MIP hybrids, the idea of employing the complementary sequence as an alternative analyte was also explored. Thermal melting experiments were performed to observe whether the inclusion of acrylamide groups in the aptamer had any effect on the stability of the duplex formed. A more stable duplex would result in a higher  $T_m$  as more energy is required to disrupt interactions between the two strands. Denaturation and annealing of the duplexes were monitored by fluorescence of the intercalator dye SYBR Green™. SYBR Green™ binds specifically to the minor groove of double stranded DNA and fluoresces >1000 fold when bound.<sup>29</sup> The fluorescence of the dye is very weak in the presence of single stranded DNA, or on its own. Using these properties of the dye, a thermal melting profile for each duplex can be generated. Table 3.1 shows the  $T_m$  values of the unmodified and acrylamide modified stands duplexed with the 15 base complementary sequence for the thrombin binding aptamer. Samples were heated from 15°C → 90°C and cooled from 90°C → 20°C at a rate of 1°C/min, and  $T_m$  values were obtained from maxima of the negative first derivative of the melting curve (see Appendix section 6.10). The data shown is an average of three separate experiments.



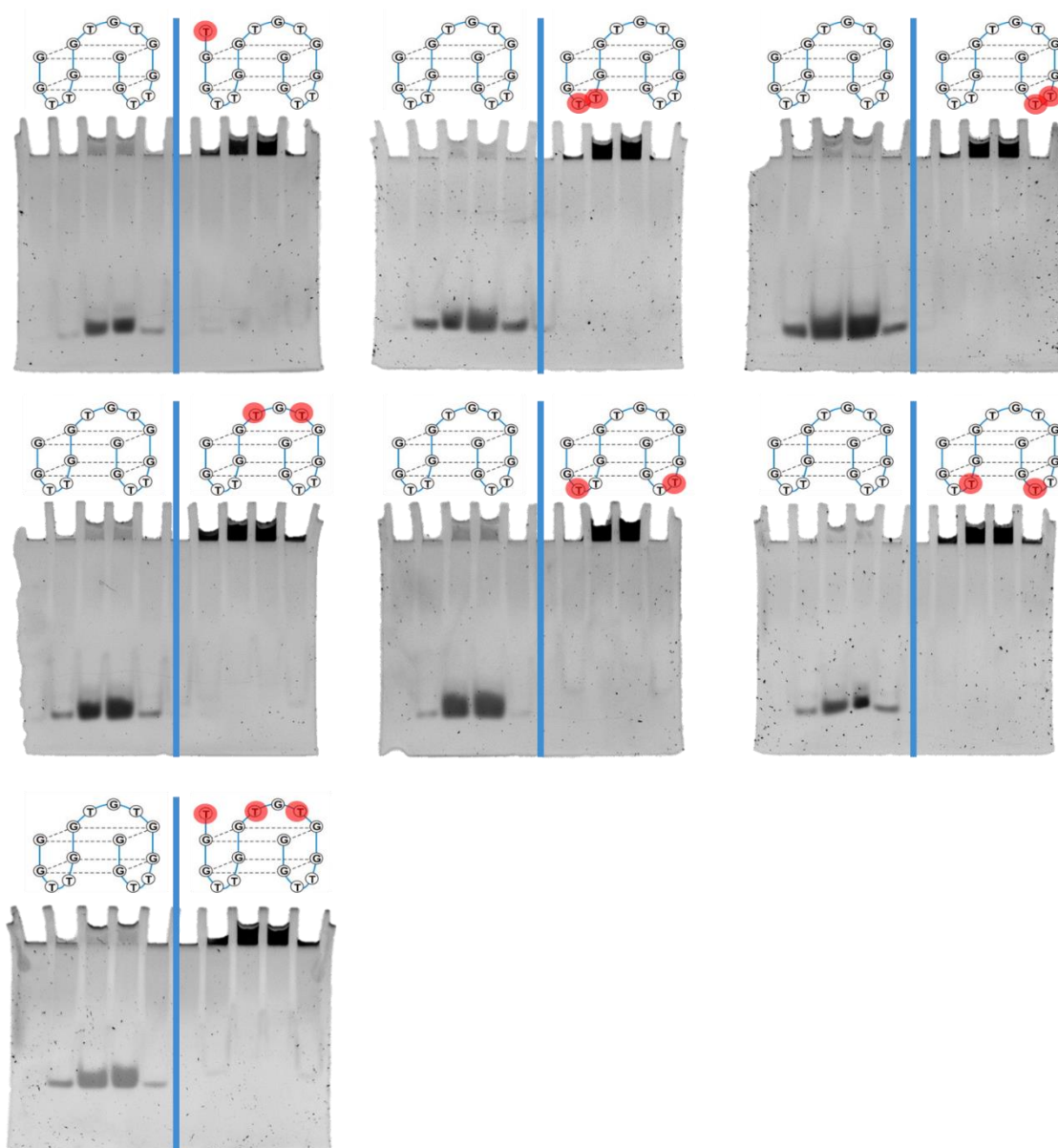
**Table 3.1**  $T_m$  values for duplexes formed with unmodified/acrylamide-modified aptamer strands.

								
average $T_m / ^\circ\text{C}$	62.7	65.8	62.8	62.7	63.3	64.2	63.7	66.7
std dev./ $^\circ\text{C}$	2.9	0.7	0.4	0.5	0.2	0.5	0.2	0.2

The average  $T_m$  values of the modified strands were very similar if not slightly higher than the unmodified aptamer. This suggests that the acrylamide groups enhance the stability of the duplex to some extent, possibly since the acrylamide moiety has sites for additional hydrogen bonding interactions. The 16-mer modified strands (**Acryl-T16** and **Acryl-T7/T9/T16**) produced  $T_m$  values which were up to 4°C higher than the 15-mer unmodified strand despite the use of a 15mer complementary strand for duplex formation, indicating greater stability with the slight overhang. It appears that the incorporation of acrylamide groups in the aptamer sequence is beneficial in its interactions with complementary sequences and thus could prove useful in designing hybrid DNA-MIPs for antisensing technologies.

### 3.11 DNA-acrylamide gel copolymerisation experiments

A polyacrylamide gel experiment was designed to test the ability of the acrylamide group(s) in the modified DNA strands to polymerise with other acrylamide monomers. A mixture of acrylamide, bisacrylamide, ammonium persulphate, TEMED, and DNA (unmodified or acrylamide modified) was loaded into wells of a polyacrylamide gel and allowed to set and chemically attach to the rest of the gel. An electrical current was then applied to the gel. It was anticipated that if the acrylamide group on the aptamer had reacted with the gel, the strand would be covalently bound and unable to move along with the current. Conversely, unmodified DNA strands would move with the current as they do not possess the acrylamide functionality to anchor the strand to the gel. Once run, the gels were stained and visualised to determine the position of the strands. As expected, Figure 3.23 shows that all acrylamide containing DNA strands remain at the wells of the gel, whereas unmodified versions have moved down the gel along with the current. This result confirms the ability of the incorporated acrylamide groups to react with other acrylamide monomers to anchor the DNA within the polymer network.



**Figure 3.23** Polyacrylamide gels of unmodified and acrylamide modified thrombin binding aptamer strands.

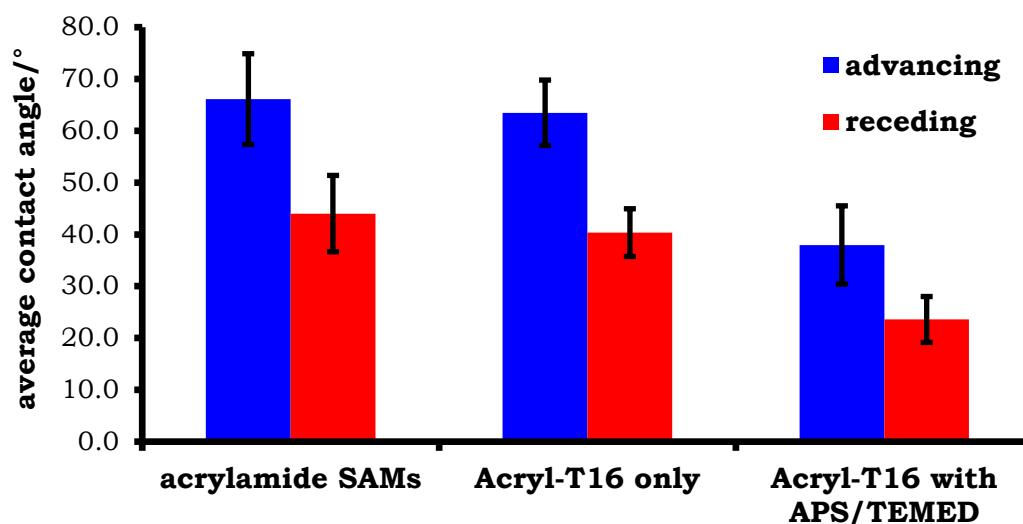
### 3.12 Surface coupling experiments with unmodified and acrylamide modified thrombin binding aptamer strands

The feasibility of acrylamide modified DNA to couple to acrylamide self-assembled monolayers on gold surfaces was studied by three techniques: contact angle, ellipsometry, and XPS. It was anticipated that successful coupling of the modified strands in the presence of APS and TEMED initiators would result in a change in the properties of the gold surface, and that no change would be observed with the unmodified sequence as it lacks the functionality to react with the surface bound acrylamide molecules. Acrylamide SAMs on gold chips were prepared from N,N'-bis(acryloyl)cystamine (BAC) and reacted with aqueous solutions of APS and TEMED with either unmodified or acrylamide modified DNA at room temperature overnight before analysis. Oligo **Acryl-T16** was the only modified strand tested in these experiments.

#### 3.12.1 Contact angle results

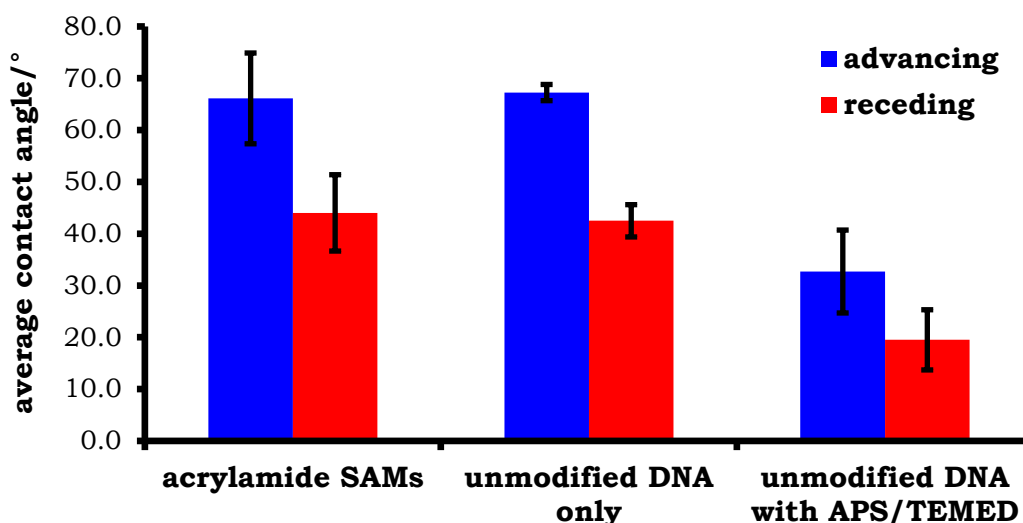
The contact angles of gold surfaces treated with only DNA, or both APS/TEMED initiators and DNA were tested and the results compared to that of the acrylamide SAMs. It was expected that a difference would only be observed with surfaces treated with both initiators and acrylamide modified DNA, as the attachment of DNA would increase the hydrophilicity of the surface. Initial contact angle results looked promising as it showed a significant reduction in the contact angle of the gold chips incubated with both initiators and acrylamide modified DNA, with advancing/receding angles of  $38^{\circ} \pm 8^{\circ}/24^{\circ} \pm 4^{\circ}$  compared

to advancing/receding angles of  $66^\circ \pm 9^\circ/44^\circ \pm 7^\circ$  with acrylamide SAMs, indicating successful coupling of the strand to the surface (Figure 3.24).



**Figure 3.24** Contact angle data for acrylamide SAMs on gold surfaces before and after treatment with either **Acryl-T16** or **Acryl-T16** with APS/TEMED.

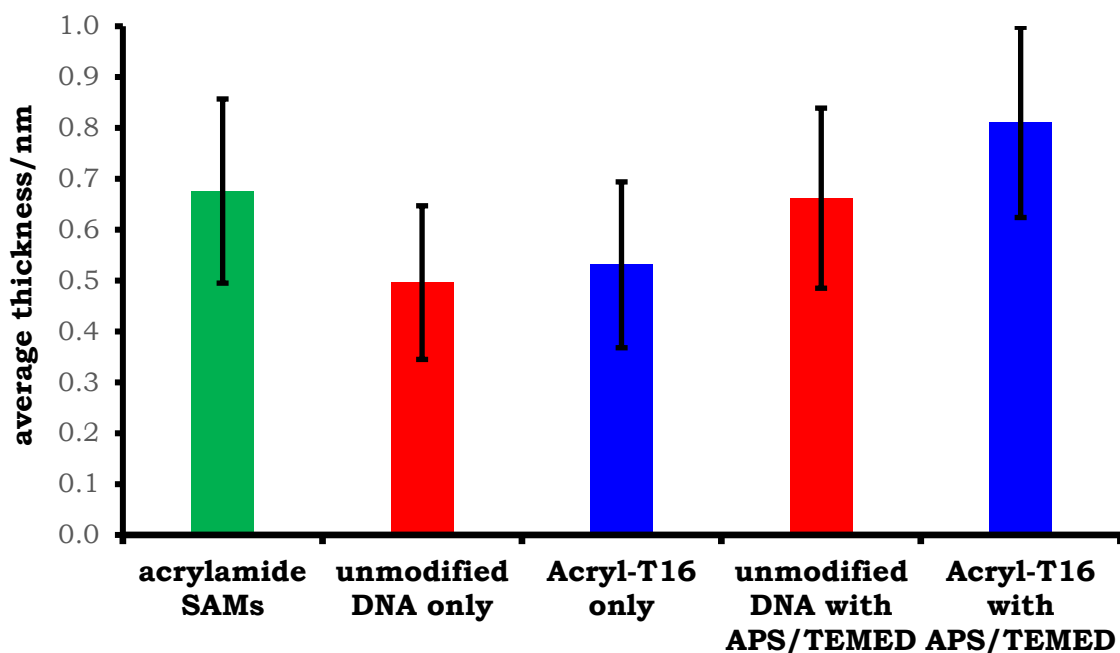
However, contact angle data for the unmodified strand with initiators also showed a significant decrease, with advancing/receding angles of  $33^\circ \pm 8^\circ/20^\circ \pm 6^\circ$  compared to advancing/receding angles of  $66^\circ \pm 9^\circ/44^\circ \pm 7^\circ$  with acrylamide SAMs, suggesting that acrylamide groups were not needed for the strand to react with the surface (Figure 3.25).



**Figure 3.25** Contact angle data for acrylamide SAMs on gold surfaces before and after treatment with either unmodified DNA or unmodified DNA with APS/TEMED.

### 3.12.2 Ellipsometry results

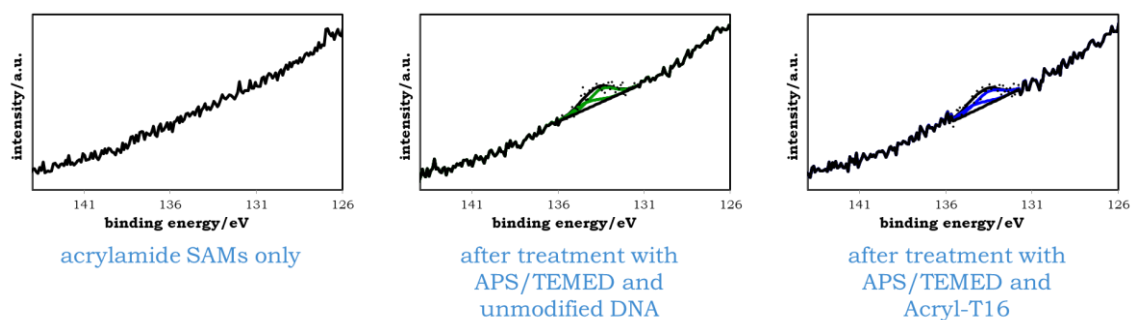
The thicknesses of self-assembled monolayers on gold surfaces treated with either DNA, or both initiators and DNA were next analysed by ellipsometry. The theoretical lengths of the acrylamide SAMs and the folded aptamer strands are 0.7nm and 3nm respectively, and thus a four-fivefold increase in thickness was expected with DNA reacted surfaces. Unfortunately, the ellipsometry results were not consistent with the contact angle results in that no significant increases for any of the reacted surfaces were observed (Figure 3.26). The increases in average thickness in the presence of initiators were not outside of standard deviations.



**Figure 3.26** Ellipsometry data for acrylamide SAMs on gold surfaces before and after treatment with unmodified/acrylamide modified DNA or unmodified/acrylamide modified DNA with APS/TEMED.

### 3.12.3 XPS results

The elemental composition of acrylamide SAMs on gold surfaces reacted with APS/TEMED initiators and either unmodified or acrylamide modified aptamer strands (**Acryl-T16**) was analysed by XPS and the results compared to that of unreacted SAMs. This technique would confirm the attachment of DNA to the surface by scanning for signals corresponding to phosphorus. The backbone of DNA consists of phosphate groups whereas the acrylamide SAMs do not possess any phosphorus containing moieties. Figure 3.27 shows the elemental scans in the phosphorus 2p range for the three different samples.



**Figure 3.27** Phosphorus 2p scans of acrylamide SAMs alone and treated with APS/TEMED initiators and unmodified or acrylamide modified (**Acryl-T16**) aptamer strands.

Very weak phosphorus signals at 133.4eV for P2p<sub>3/2</sub> and 134.3eV for P2p<sub>1/2</sub> were observed with DNA reacted samples whereas no phosphorus related signals were seen with the SAM only samples. In line with the contact angle data, the XPS scans indicated that both unmodified and acrylamide modified DNA had reacted with acrylamide SAMs.

XPS data was also used to estimate the thickness of the SAM before and after reaction with the DNA/initiator solutions by detecting how much the adsorbed film attenuates the gold substrate signal (Figure 3.28). Assuming that the surface is completely covered with an even film, the following equation can be applied to determine the gold substrate attenuation length and thus the thickness of the SAM:

$$d = -0.8 \times IMFP \times \ln\left(\frac{I_d}{I_0}\right)$$

Where  $d$  is the attenuation length in Å

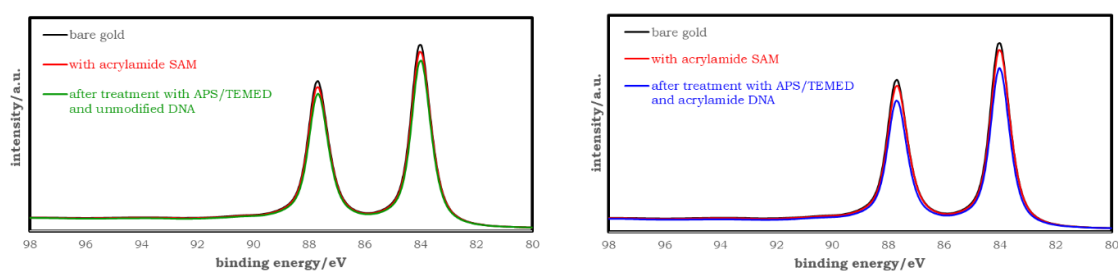
$IMFP$  is the electron inelastic mean free path (40Å for gold<sup>30</sup>)

$I_0$  is the gold 4f signal intensity for the bare substrate

$I_d$  is the gold 4f signal intensity for the SAM covered substrate



Applying the above equation to the XPS data resulted in a film thickness of 2.6Å (0.26nm) for the acrylamide SAMs alone, 4.0Å (0.4nm) for SAMs treated with APS/TEMED and unmodified DNA, and 5.1Å (0.5nm) for the SAMs treated with APS/TEMED and acrylamide modified DNA. The calculated thicknesses were lower than the theoretical thicknesses of the molecules and the thicknesses obtained by ellipsometry. However, the trend of slight increases in film thicknesses after treatment with DNA/initiator solutions is in agreement with the average thicknesses observed in ellipsometry data (Figure 3.26), although the changes in the ellipsometry data were not statistically significant.



**Figure 3.28** XPS data showing gold signal attenuation after acrylamide SAM formation and after treatment with unmodified DNA/initiator (left) or acrylamide modified DNA/initiator (right) solutions.

### 3.12.4 Summary

When exploring the coupling reaction of acrylamide modified DNA to acrylamide SAMs on gold surfaces, it was anticipated that the unmodified sequence would act as a negative control and not react with the surfaces in the presence of APS/TEMED initiators as it does not contain the acrylamide groups necessary for reaction with the SAMs. However, both contact angle and XPS techniques have shown noteworthy changes in the properties of the gold surfaces with an increase in hydrophilicity and signs of surface bound

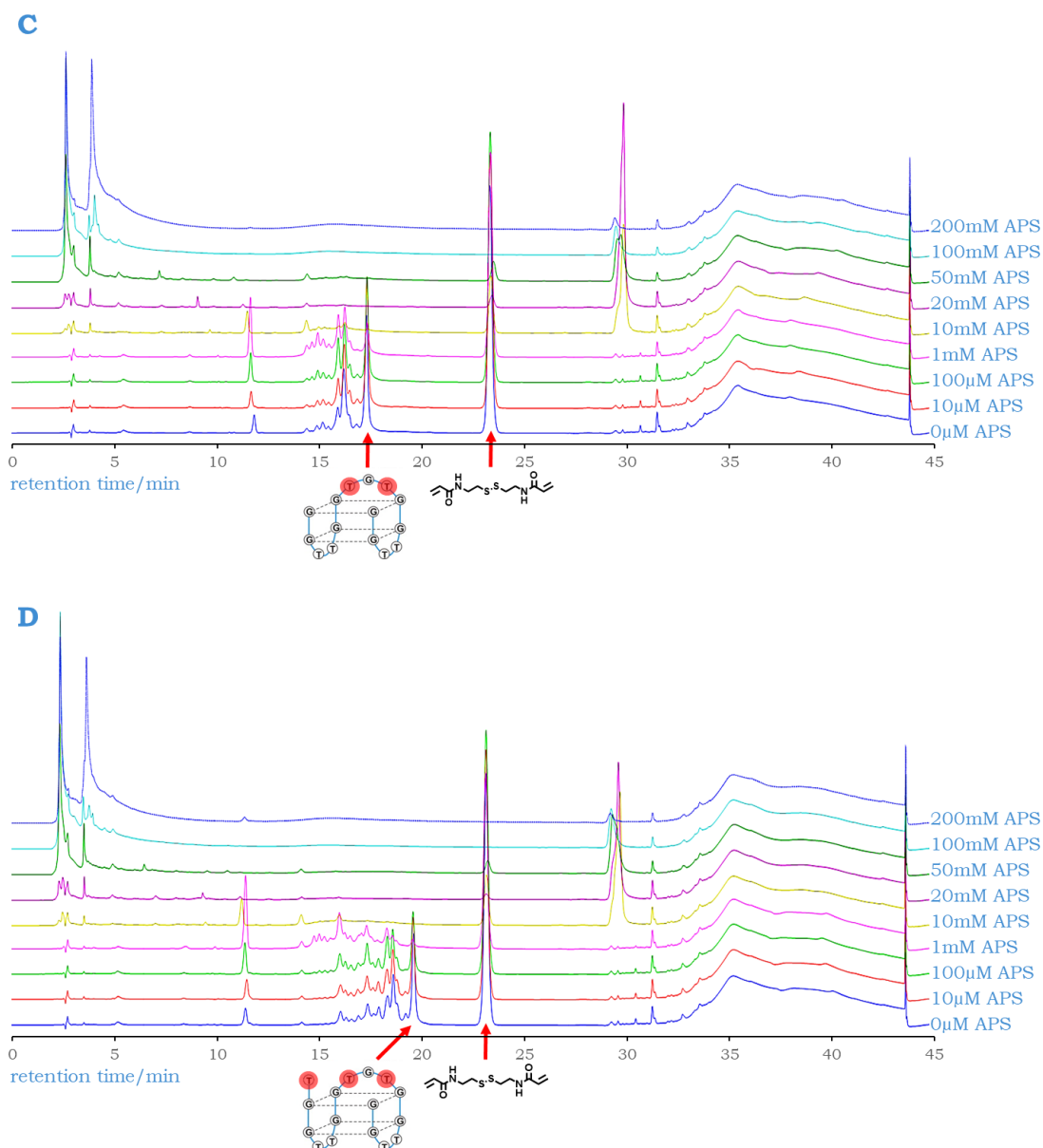
phosphorus after treatment with initiators and DNA. Ellipsometry and XPS data showed only a slight increase in the average thickness of the SAMs after DNA/initiator treatment, which would indicate partial addition of the strands to the surfaces. It is possible that the slight increases are down to coupling of fragments rather than whole strands of DNA. In these experiments, a high concentration of APS was used with respect to the concentration of DNA. In fact, ~17,000x more equivalents of APS was used w.r.t DNA and this vast excess appears to have degraded the DNA strands in solution prior to coupling on the surface. Although a similar ratio of DNA:APS is utilised in work published by the Turner group,<sup>15</sup> many other monomers at higher concentrations are present in the reaction mixture which would minimise the potential for DNA damage by ensuing radicals. This appears to also be the case with the gel work described in this chapter (section 3.11) on using APS/TEMED to fixate acrylamide DNA strands into acrylamide polymer gels. There, the presence of highly concentrated and more reactive acrylamide and bisacrylamide monomers in the mixture appears to have reduced the possibility of degradation of DNA by APS/TEMED radicals; the Diamond™ dye would not have stained anything had the DNA fragmented in those experiments.

### 3.13 Optimisation of APS concentration used in acrylamide DNA coupling experiments

#### 3.13.1 HPLC experiments

As high concentrations of APS resulted in apparent degradation of DNA in the surface coupling experiments, it was decided to optimise initiator concentrations to levels that would allow the radical coupling reaction to proceed without causing considerable damage to the DNA strands. HPLC was utilised to analyse solutions containing N,N'-bis(acryloyl)cystamine (BAC) (the compound used for acrylamide SAM formation), unmodified or acrylamide modified aptamer stands, and TEMED, with increasing concentrations of APS. BAC:DNA:TEMED ratios were identical to surface coupling experiments and kept constant. However, concentrations were increased by a factor of 10 to aid detection of the separate components. The following concentrations were used in solutions analysed by HPLC: 1mM BAC, 10 $\mu$ M DNA, 0.1 $\mu$ l/ml TEMED, and 0 $\rightarrow$ 200mM APS. As with the surface coupling experiments, samples were left to react overnight before analysis. Acrylamide strands **Acryl-T7/T9** and **Acryl-T7/T9/T16** were also tested alongside **Acryl-T16** to see whether stands with multiple acrylamide groups would require a different concentration of APS. Figure 3.29 shows a series of overlaid chromatograms with the different DNA strands.





**Figure 3.29** Overlaid HPLC spectra of BAC/DNA/TEMED solutions with increasing concentrations of APS. **A:** unmodified aptamer, **B:** Acryl-T16, **C:** Acryl-T7/T9, **D:** Acryl-T7/T9/T16.

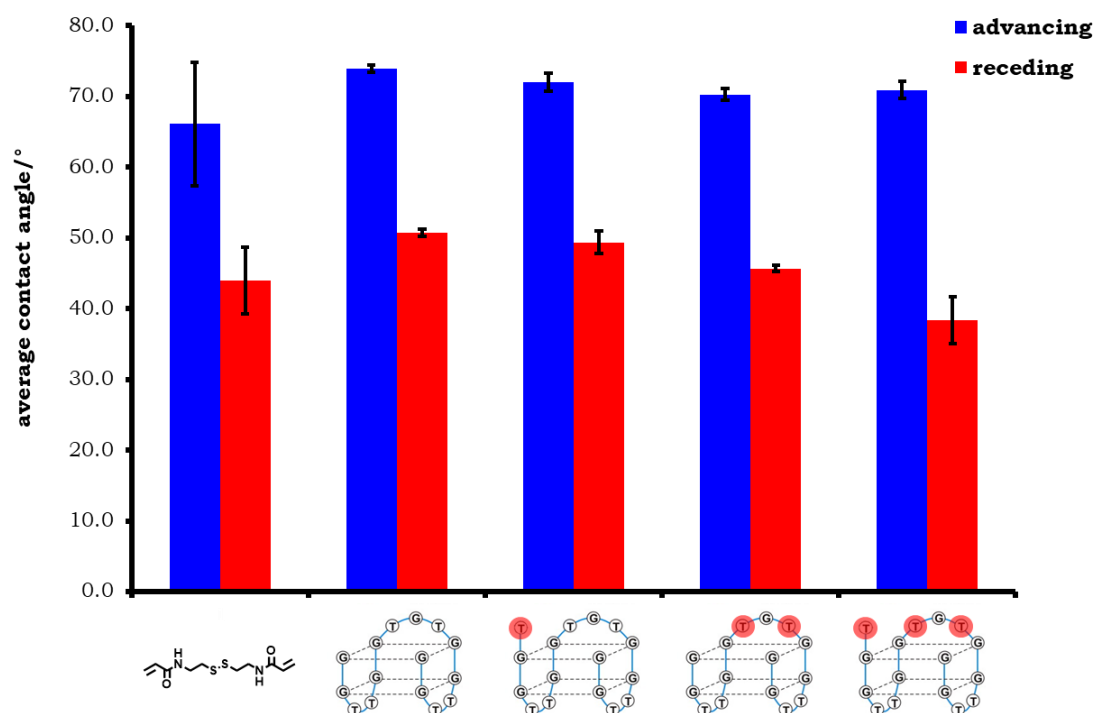
For the unmodified aptamer, peaks associated with BAC at 23 minutes and DNA at 14-19 minutes are present at low concentrations of APS whereas these peaks are not observed at concentrations above 20mM APS, due to reaction/degradation of the starting materials. Degradation products are observed at high concentrations of APS (>50mM) with a retention time of 2.5

minutes. The amounts of APS used in the surface coupling experiments are similar to the 200mM APS trace which would explain why SAM thicknesses were not as high as expected as it would appear that the degradation products of the strands have coupled to the surface instead. The spectra containing acrylamide strands (Figure 3.29 B, C & D) have additional signals around the main DNA peak. This effect is caused by TEMED as it is not observed in acrylamide DNA spectra with TEMED absent (see Appendix section 6.11). TEMED appears to react with the acrylamide groups on these strands as the additional signals are not seen with the unmodified strand and becomes more prominent as the number of acrylamide groups on each strand increases. It may be that as an amine (albeit a hindered one), TEMED adds to the double bond of the acrylamide group, forming numerous adducts with varying retention times. An additional peak at ~30 minutes is also observed with these samples when 10-100mM APS is used. This signal is not seen in BAC absent or DNA absent scans (see Appendix section 6.11) and can only be attributed to the formation of a BAC-DNA conjugate. Attempts to collect fractions at this retention time and characterise the samples by mass spectrometry proved unsuccessful as fractions were too weak in concentration for analysis. It was decided to revisit the surface coupling experiments and reduce the APS concentration almost 20-fold, working with a DNA:APS ratio of 1:1000 as the HPLC traces at this level appeared to form the BAC-DNA conjugate without causing substantial degradation to the DNA strand.

### [3.13.2 Contact angle experiments](#)

Contact angle experiments were conducted to ascertain whether the acrylamide DNA surface coupling reaction could proceed at lower

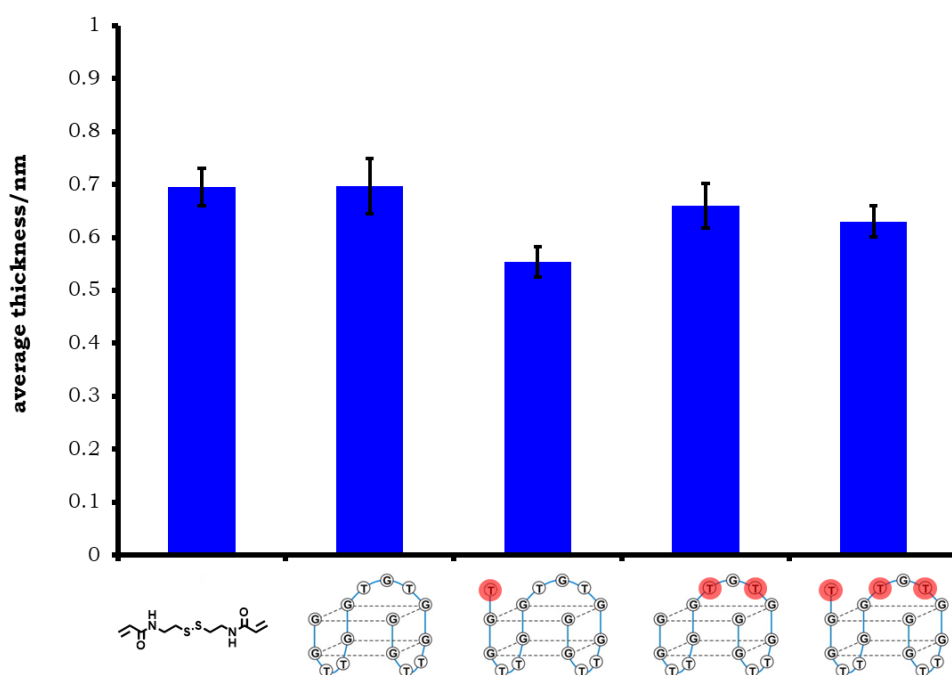
concentrations of APS. Gold chips with acrylamide SAMs were incubated in solutions containing TEMED, APS, and either unmodified or acrylamide modified DNA. Due to the volume of solution required for chip immersion and the availability of DNA stocks, DNA concentrations of 1  $\mu$ M were used instead of the 10  $\mu$ M concentrations applied in HPLC experiments. To keep the DNA:APS ratio the same, 1 mM of APS was used resulting in solutions which were 10x more diluted than those analysed in HPLC experiments. Unfortunately, the results show no significant change in the contact angles of the chips with acrylamide SAMs treated with any of the DNA/initiator solutions compared to that of just the SAM alone, indicating that the DNA coupling reaction was not successful (Figure 3.30).



**Figure 3.30** Contact angle data for acrylamide SAMs on gold surfaces and acrylamide SAMs treated with APS/TEMED and either unmodified or acrylamide modified DNA.

### 3.13.3 Ellipsometry experiments

The same samples prepared in contact angle experiments were then analysed by ellipsometry for changes in SAM thicknesses (Figure 3.31). Ellipsometry results displayed no increases in thicknesses of the acrylamide SAMs after treatment with DNA/low initiator solutions. This result is in agreement with the contact angle data and shows that the coupling reaction was unsuccessful under these conditions.

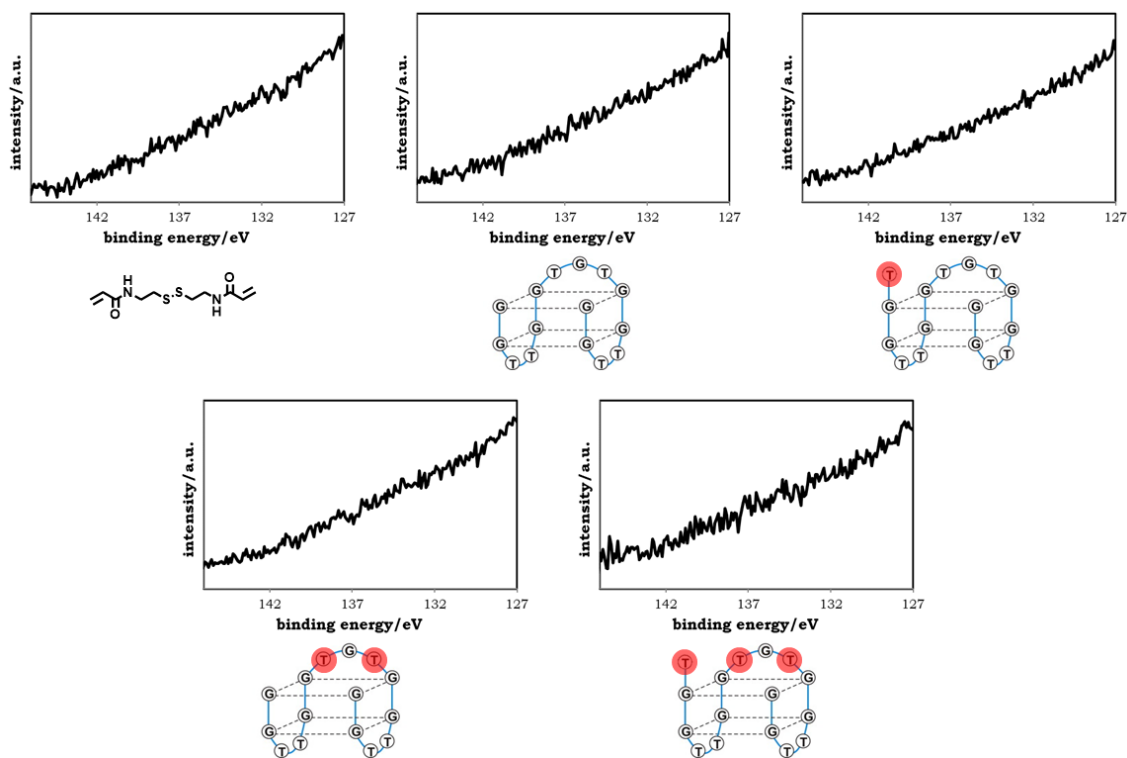


**Figure 3.31** Ellipsometry data for acrylamide SAMs on gold surfaces and acrylamide SAMs treated with APS/TEMED and either unmodified or acrylamide modified DNA.

### 3.13.4 XPS experiments

XPS was another technique used to confirm the unsuccessful coupling between DNA and acrylamide SAMs on surfaces. New samples prepared under the same low APS conditions as contact angle/ellipsometry experiments were analysed by XPS for the presence of phosphorus (Figure 3.32). None of the surfaces tested showed any traces of phosphorus.





**Figure 3.32** Phosphorus 2p scans of acrylamide SAMs alone and treated with APS/TEMED initiators and unmodified or acrylamide modified aptamer strands

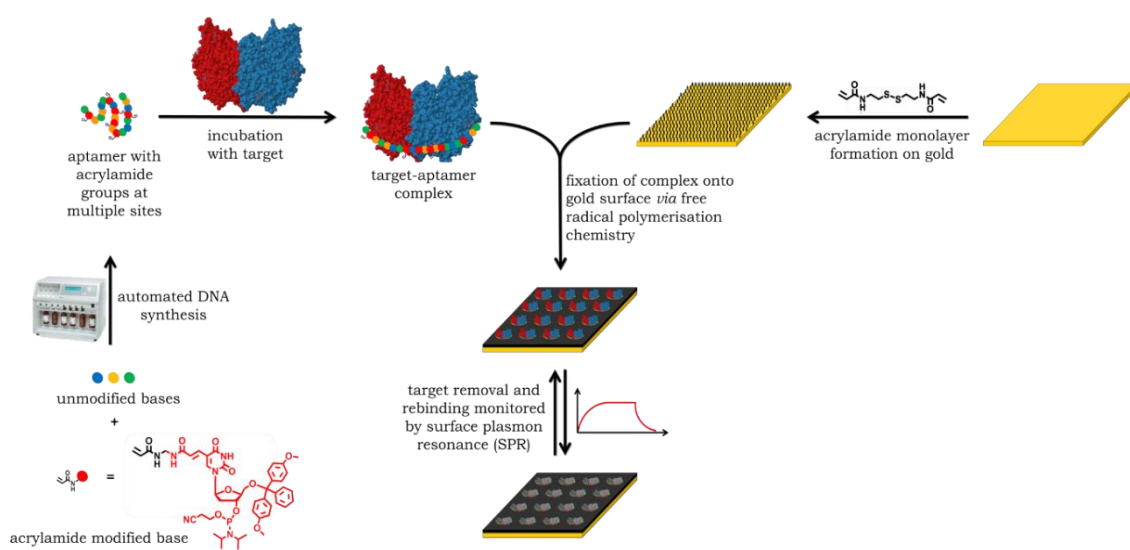
### 3.13.5 Summary

When exploring radical chemistry for coupling acrylamide modified DNA to acrylamide SAMs on gold surfaces, the use of high concentrations of APS with respect to DNA resulted in degradation of DNA and subsequent coupling of fragments to the surface. The APS concentration was optimised by investigating the effect of concentration on the coupling reaction in solution using the precursor BAC to represent acrylamide SAMs in solution. By studying the solutions with HPLC, a DNA:APS ratio of 1:1000 was found to cause the formation of a DNA:BAC adduct without degrading the DNA strands. This adduct was only observed with solutions containing acrylamide modified DNA. However, the TEMED catalyst was found to react with these strands in the absence of APS, with extra peaks observed in HPLC traces of strands with

additional acrylamide groups. When adopting the lower DNA:APS ratio to acrylamide SAMs on surfaces, the coupling reactions were found to be unsuccessful, as evidenced by contact angle, ellipsometry, and XPS. Two factors may have contributed to the different results observed between solution and surface coupling experiments. The first factor was the concentrations applied in both experiments. Although reacting ratios were identical, the solutions used in surface experiments were 10x more diluted than those used in HPLC experiments due to the volumes required and the availability of DNA stocks. This dilution may have resulted in inactivity in the surface experiments as solutions were too dilute for the reactions to proceed. Another factor is the concentration of the acrylamide SAMs. With surface experiments, the BAC molecules are first absorbed onto the surface prior to reaction, and so the true concentration of BAC used in surface experiments would be lower than that of HPLC experiments if all molecules had not absorbed onto the gold surface prior to the coupling reaction. A way to overcome these problems would be to increase the concentrations of DNA and APS used in surface experiments to those used in HPLC experiments, although DNA stocks would first have to be increased to proceed with this method. Another avenue to explore would be to increase the DNA:APS reacting ratio to 1:2000 as this condition also appears not to cause degradation of DNA. Reacting ratios of 1:3000 and 1:4000 were not investigated and it is a possibility that these ratios could be ideal for allowing the coupling reaction to proceed without significant degradation of DNA. However, due to time constraints these various options were not explored.

### 3.14 Conclusions and future work

To conclude, novel acrylamide modified thymidine residues were successfully incorporated into thrombin binding aptamer sequences at various positions. The modified strands were observed to bind to their desired targets (depending on the positioning of the acrylamide groups) and to react with other acrylamide monomers to be integrated into acrylamide polymers. Future work should continue probing the coupling of acrylamide modified DNA to acrylamide SAMs on gold surfaces. DNA:APS reacting ratios should be further explored and optimised for surface coupling reactions. There appears to be a fine balance between the amount of APS needed for the reaction to proceed and the amount required to cause substantial DNA damage. This balance can be shifted away from DNA damage by incorporating more reactive monomers and cross linkers into the reaction mixture. These additional molecules are required for MIP formation and thus proceeding to explore DNA surface coupling *via* the creation of the surface bound MIPs might prove more favourable than focusing efforts on solely coupling DNA to the surface bound acrylamide SAMs. Once optimised, SPR should then be utilised to study the effect of anchorage on the binding affinities of the aptamer incorporated surface bound MIPs to both thrombin and complementary DNA template molecules (Scheme 3.6). It is anticipated that different responses will be observed with the modified strands, with strands with multiple anchorage sites producing higher responses.



**Scheme 3.6** Proposed future plan for forming and testing surface-based aptamer-MIP hybrids.

### 3.15 List of references

1. G. Wulff, *Angew. Chem. Int. Ed. Engl.*, 1995, **34**, 1812-1832.
2. L. Ye and K. Haupt, *Anal. Bioanal. Chem.*, 2004, **378**, 1887-1897.
3. S. Manju, P. R. Hari and K. Sreenivasan, *Biosens. Bioelectron.*, 2010, **26**, 894-897.
4. M. Polyakov, *Zhurnal Fizieskoj Khimii/Akademiya SSSR*, 1931, **2**, 799-805.
5. K. Mosbach and R. Mosbach, *Acta Chem. Scand.*, 1966, **20**, 2807-2810.
6. G. Vlatakis, L. I. Andersson, R. Muller and K. Mosbach, *Nature*, 1993, **361**, 645-647.
7. G. Vasapollo, R. D. Sole, L. Mergola, M. R. Lazzoi, A. Scardino, S. Scorrano and G. Mele, *Int. J. Mol. Sci.*, 2011, **12**, 5908-5945.
8. D. R. Kryscio and N. A. Peppas, *Acta Biomater.*, 2012, **8**, 461-473.
9. S. Rimmer, *Chromatographia*, 1998, **47**, 470-474.
10. N. Hidekazu, H. Chin-Shiou and S. K. J., *Angew. Chem. Int. Ed.*, 2006, **45**, 2392-2396.
11. L. Chen, X. Wang, W. Lu, X. Wu and J. Li, *Chem. Soc. Rev.*, 2016, **45**, 2137-2211.
12. L. Chen, S. Xu and J. Li, *Chem. Soc. Rev.*, 2011, **40**, 2922-2942.
13. A. Poma, H. Brahmbhatt, J. K. Watts and N. W. Turner, *Macromolecules*, 2014, **47**, 6322-6330.
14. H. Brahmbhatt, A. Poma, H. M. Pendergraff, J. K. Watts and N. W. Turner, *Biomater. Sci.*, 2016, **4**, 281-287.
15. A. Poma, H. Brahmbhatt, H. M. Pendergraff, J. K. Watts and N. W. Turner, *Adv. Mater.*, 2015, **27**, 750-758.
16. Z. Zhang and J. Liu, *ACS Appl. Mater. Interfaces*, 2016, **8**, 6371-6378.

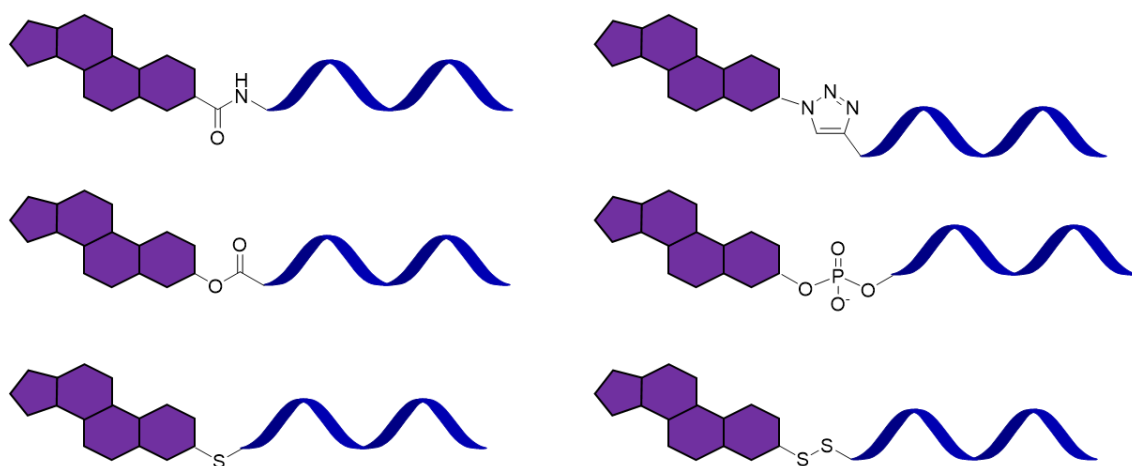
17. P. Jolly, V. Tamboli, R. L. Harniman, P. Estrela, C. J. Allender and J. L. Bowen, *Biosens. Bioelectron.*, 2016, **75**, 188-195.
18. 5' Acrydite,  
<https://www.idtdna.com/site/Catalog/Modifications/Product/1604>,  
(accessed 05/09/2018).
19. L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas and J. J. Toole, *Nature*, 1992, **355**, 564-566.
20. V. J. B. Ruigrok, M. Levisson, J. Hekelaar, H. Smidt, B. W. Dijkstra and J. van der Oost, *Int. J. Mol. Sci.*, 2012, **13**, 10537-10552.
21. K. Padmanabhan, K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler and A. Tulinsky, *J. Biol. Chem.*, 1993, **268**, 17651-17654.
22. K. Padmanabhan and A. Tulinsky, *Acta Crystallogr., Sect. D.*, 1996, **52**, 272-282.
23. I. Russo Krauss, A. Merlino, A. Randazzo, E. Novellino, L. Mazzarella and F. Sica, *Nucleic Acids Res.*, 2012, **40**, 8119-8128.
24. T. Ami and K. Fujimoto, *ChemBioChem*, 2008, **9**, 2071-2074.
25. E. A. Owens, N. Bruschi, J. G. Tawney and M. Henary, *Dyes Pigm.*, 2015, **113**, 27-37.
26. D. Sen and W. Gilbert, *Curr. Opin. Struct. Biol.*, 1991, **1**, 435-438.
27. L. Sun, H. Jin, X. Zhao, Z. Liu, Y. Guan, Z. Yang, L. Zhang and L. Zhang, *ChemMedChem*, 2014, **9**, 993-1001.
28. B. Deng, Y. Lin, C. Wang, F. Li, Z. Wang, H. Zhang, X.-F. Li and X. C. Le, *Anal. Chim. Acta*, 2014, **837**, 1-15.
29. A. I. Dragan, R. Pavlovic, J. B. McGivney, J. R. Casas-Finet, E. S. Bishop, R. J. Strouse, M. A. Schenerman and C. D. Geddes, *J Fluoresc*, 2012, **22**, 1189-1199.

30. NIST Standard Reference Database 71, <https://www.nist.gov/srd/nist-standard-reference-database-71>, (accessed 05/09/2018).

# Chapter 4 Lipid-aptamer conjugates for biosensing with liposomes

## 4.1 Introduction

Lipid-oligonucleotide conjugates (LOCs) are synthetically derived amphiphilic molecules which comprise of short strands of DNA or RNA connected to lipophilic moieties that are typically steroidal or hydrocarbon chain based. The lipid can be attached to the nucleobase, sugar, or phosphate backbone of the oligonucleotide, in single or multiple positions, but more often is tagged at the end of the strand.<sup>1</sup> LOCs are synthesised by either incorporating the lipid into solid phase oligonucleotide synthesis *via* phosphoramidite/H-phosphonate chemistries,<sup>2</sup> or by employing coupling strategies pre or post oligonucleotide synthesis to form amide,<sup>3</sup> ester,<sup>4</sup> thioether,<sup>5</sup> disulphide,<sup>6</sup> or triazole<sup>7</sup> linkages (Figure 4.1).



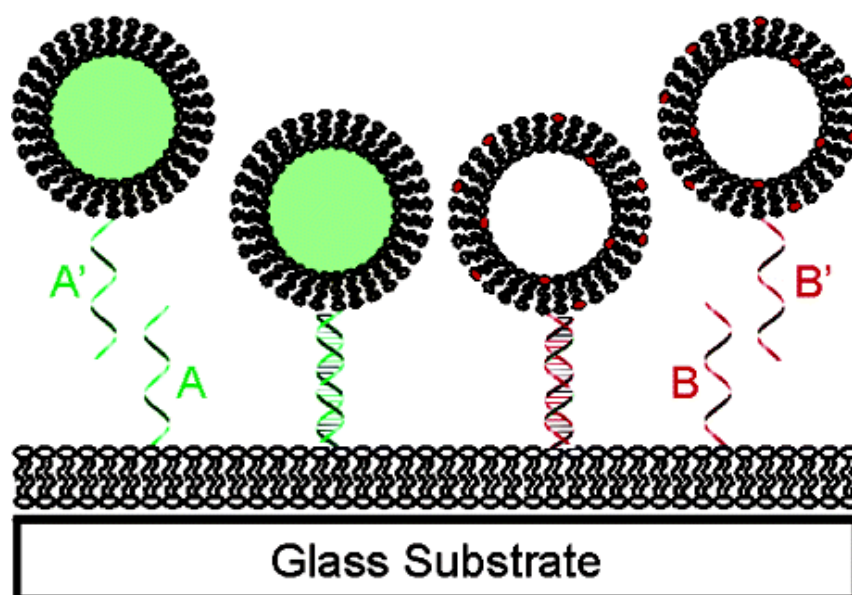
**Figure 4.1** Different coupling chemistries used to connect lipids to oligonucleotides.



LOCs have been designed to utilise the unique properties of both lipids and oligonucleotides. The lipid portion allows the LOC to interact with lipid membranes for anchorage and/or cell permeability, whereby the oligonucleotide acts as a recognition element for detecting complementary sequences, or other molecules if aptameric sequences are involved. The exclusive properties of LOCs make them appealing as building blocks for supramolecular applications. In sufficient concentrations, the LOC can aggregate and self-assemble to form micellar structures or liposomes.<sup>8, 9</sup> However, they are mainly used in low concentrations to functionalise lipid membranes with oligonucleotides for antisense applications.

There are numerous reports on liposome fusion and controlled liposome aggregation caused by duplex formation of DNA strands on different liposomes.<sup>10-13</sup> The most cited reports stem from the groups of Boxer and Höök who both independently used LOCs to mimic SNARE proteins by replicating their ability to fuse lipid membranes of different liposomes together.<sup>14-16</sup> For successful membrane fusion, the LOC must be able to bind to the complementary strand whilst anchored in the lipid membrane. This is highly dependent on the grafting density of the LOC in the membrane, with low coverages resulting in faster duplex formation.<sup>17</sup> LOCs anchored into the membranes of liposomes have been shown to bind to complementary LOCs in surface bound lipid bilayers to create 2D arrays of tethered mobile liposomes on supported lipid bilayers (Figure 4.2).<sup>18</sup>

In cell research, LOCs are used as an alternative method for introducing DNA/RNA into cells for gene silencing. The oligonucleotide alone cannot pass through the lipophilic cell membrane due to its polyanionic and polar nature. Earlier research utilised polycationic polymers for delivery of oligonucleotides into cells but these polymers were later found to be cytotoxic to varying degrees.<sup>19</sup> LOCs have been shown to be a non-toxic substitute for the cellular uptake of oligonucleotides.<sup>7</sup> The lipid moiety facilitates the internalisation of the LOC into the cell by insertion into the lipid membrane which is then taken into the cell by endocytosis for interaction with complementary mRNA to prevent protein translation. Wolfram and co-workers observed that cellular uptake of LOCs was enhanced when cholesteryl or long chain hydrocarbons were attached compared to short chain hydrocarbons and hypothesised the reason to be mediated internalisation by membrane bound lipoprotein receptors, as cholesterol is known to interact with serum lipoproteins.<sup>20</sup>

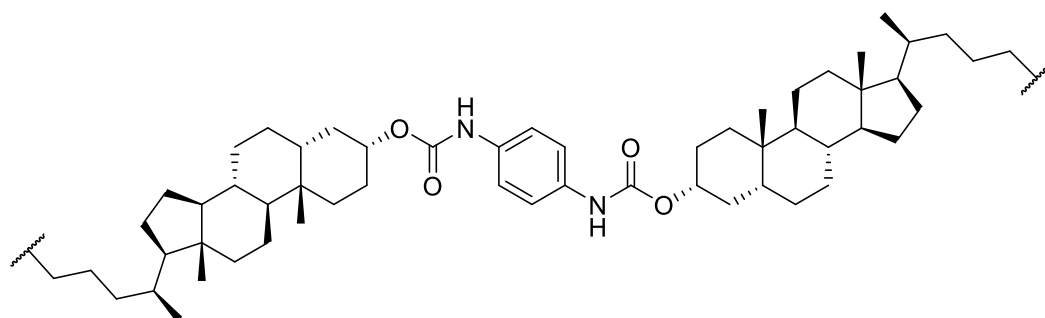


**Figure 4.2** 2D array of tethered mobile liposomes using duplexed DNA on a supported lipid bilayer.<sup>18</sup>

When designing LOCs for use in supramolecular assemblies, the LOC must be capable of spontaneous insertion with a stable anchorage. There is a fine balance between weakly lipophilic LOCs that may desorb from the lipid membrane and strongly lipophilic LOCs that may form stable self-assemblies. Cholesterol is often selected as the lipid constituent in LOCs over the lipid component of the membrane, as it is easier to synthesise (cholesterol phosphoramidites are commercially available) and membrane insertion is faster (happens in seconds as opposed to hours).<sup>21</sup> However, cholesterol-based anchoring of DNA is relatively weak. Höök and co-workers designed a novel LOC to strengthen this anchorage by featuring bivalent cholesterol coupling of duplexed DNA, with cholesterol units at the 3' end of one strand and the 5' end of the other.<sup>22</sup> They observed irreversible coupling of the bivalent cholesterol LOC to a supported phospholipid bilayer on a silicon dioxide surface. Whilst there are a few reports on cholesterol units positioned in-between nucleotides<sup>23</sup> and coupled to both ends of the same strand,<sup>24</sup> which dramatically alter the positioning of the cholesteryl-based oligonucleotide in lipid membranes, there are currently no reports on a bivalent cholesterol system being appended to one end of a single oligonucleotide strand.

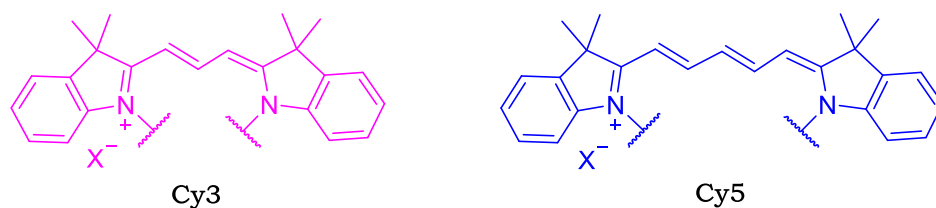
## 4.2 Aim

The aim of this chapter is to synthesise a cholesteryl-based DNA conjugate with two steroidal units tagged sequentially at the end of the strand for additional anchorage. When inserted into liposome membranes, this conjugate is expected to anchor across the whole lipid bilayer, providing extra stabilization to the supramolecular structure. Whilst there are a few reports on the insertion of conjugated steroidal units into lipid bilayers,<sup>25-28</sup> none so far have connected these molecules to oligonucleotides. The design of the lipid anchor couples two molecules of lithocholic acid to opposite ends of an aromatic ring (Figure 4.3). DNA is then added to one end of the anchor and a dye attached to the other end for fluorescence detection of the conjugate.



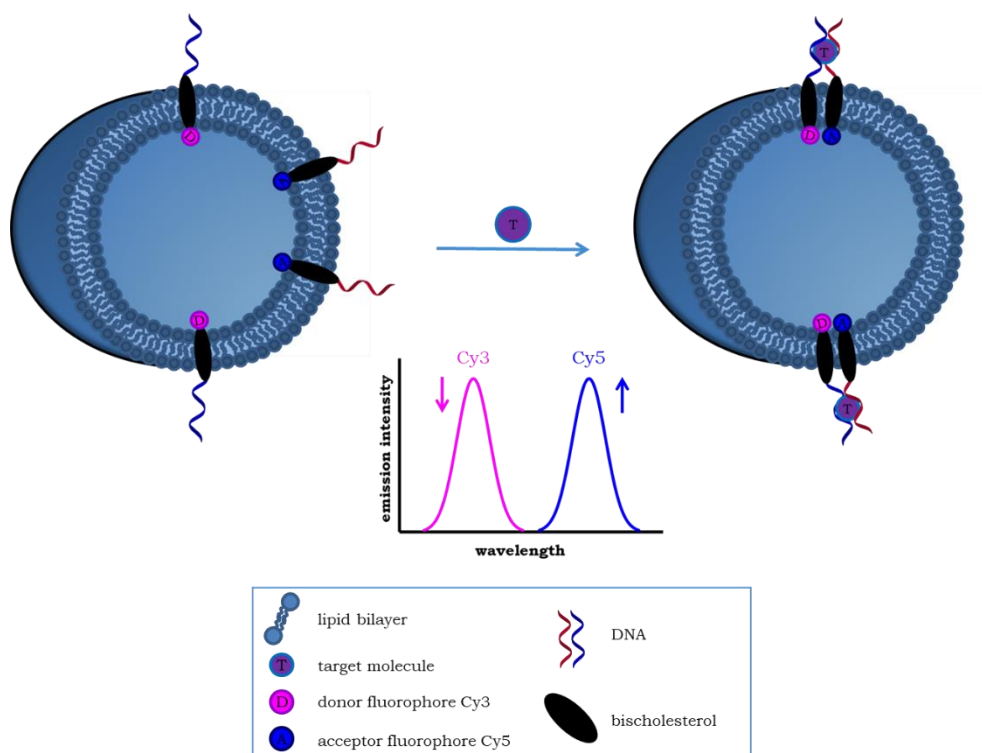
**Figure 4.3** Structure of lipid anchor.

Cyanine dyes Cy3 and Cy5 (Figure 4.4) were selected as their positive charge and hydrophobicity should encourage insertion of the conjugate in the correct orientation, since the liposome used will consist of a mixture of neutral and negative lipids.



**Figure 4.4** Structures of cyanine dyes Cy3 and Cy5.

One proposed function for the transmembrane molecules is to use the fluidic nature of the lipid bilayer to perform sandwich-based assays with the conjugates. Once anchored, the lateral movement of two different conjugates, with Cy3 on one and Cy5 on the other, can be monitored by FRET. If different DNA sequences that are known to bind to different sites of the same target are involved, then target recognition can drive the association of the two conjugates, resulting in a change in fluorescence inside the liposome (increased Cy5 emission and decreased Cy3 emission). In the absence of the target, the transmembrane probes will be kept apart and there will be no driving force for probe association as the DNA strands are repelled by their inherent negative charge (Figure 4.5). This concept is similar to the work of Dave and Liu who demonstrated the lateral movement of fluorescently tagged split aptamer LOCs in liposome membranes and subsequent assembly to bind to an adenosine target.<sup>29</sup>



**Figure 4.5** Proposed sandwich assay with transmembrane molecules.

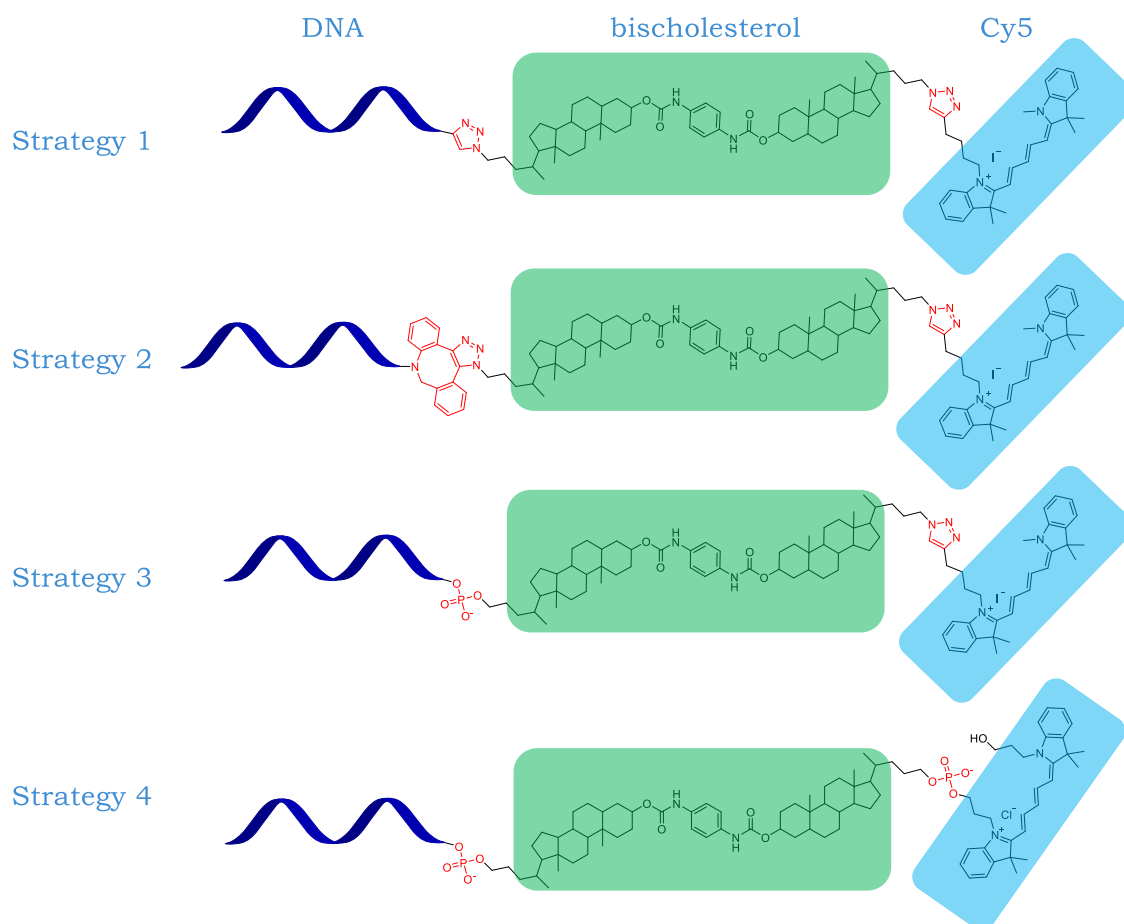
Four coupling strategies have been devised to construct the DNA-bischolesterol-dye conjugates:

1. Synthesise a bischolesterol anchor with azide groups on both ends. Use copper catalysed azide-alkyne cycloaddition to couple an alkyne modified cyanine dye to one end of the anchor and connect an alkyne modified DNA strand to the other end.
2. Synthesise a bischolesterol anchor with azide groups on both ends. Use copper catalysed azide-alkyne cycloaddition to couple an alkyne modified cyanine dye to one end of the anchor and copper free azide-alkyne cycloaddition to attach dibenzocyclooctyne (DBCO) modified DNA to the other end.

3. Synthesise a bischolesterol anchor with an azide group on one end and an alcohol on the other end. Use copper catalysed azide-alkyne cycloaddition to couple an alkyne modified cyanine dye to one end of the anchor. Convert the alcohol to a phosphoramidite group and attach the bischolesterol-dye conjugate to the 5' end of DNA using automated solid phase DNA synthesis.
4. Synthesise a bischolesterol anchor with a dimethoxytrityl (DMT) group on one end and a phosphoramidite group on the other end. Use automated solid phase DNA synthesis to couple the bischolesterol to (commercially available) cyanine dyes on solid supports. DNA is then attached *via* its 3' end to the other side of the bischolesterol anchor.

All four methods utilise either or both azide-alkyne cycloaddition and phosphoramidite chemistry to couple the bischolesterol anchor to DNA and the cyanine dye, resulting in triazole and phosphate linkages. Strategies 1-3 are most suitable for the desired application as they avoid the use of a negative phosphate linkage between the bischolesterol anchor and cyanine dye.

Figure 4.6 shows the structures of half of the transmembrane products formed from the different strategies.



**Figure 4.6** Structures of Cy5 derived transmembrane molecules from each coupling strategy. Bischolesterol anchor is highlighted in green and Cy5 dye is highlighted in blue. Triazole and phosphate linkages are highlighted in red.

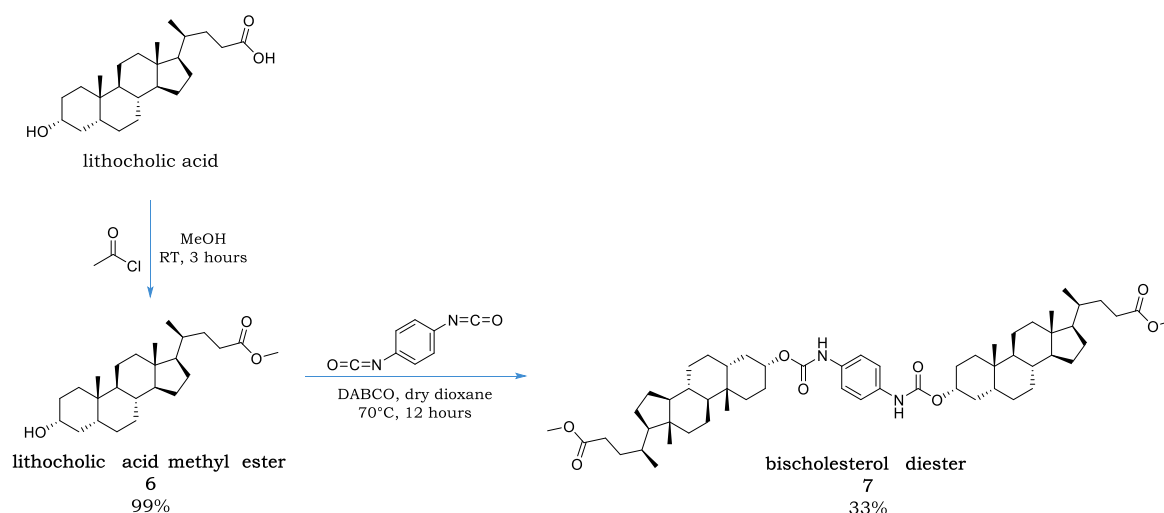
The first objective is to synthesise the bischolesterol derivatives necessary for all strategies. The second objective is to create the alkyne modified Cy3 and Cy5 dyes for use in strategies 1-3. The third objective is to couple the bischolesterol anchor to the two cyanine dyes as described in strategies 1-3. The fourth objective is to select and create alkyne and DBCO modified DNA strands for coupling strategies 1-2. The fifth objective is to complete the synthesis of the transmembrane molecules using the different synthetic strategies, and the final objective is to devise a suitable method for their purification.



## 4.3 Synthesis of bischolesterol derivatives

### 4.3.1 Synthesis of symmetrical bischolesterol derivatives

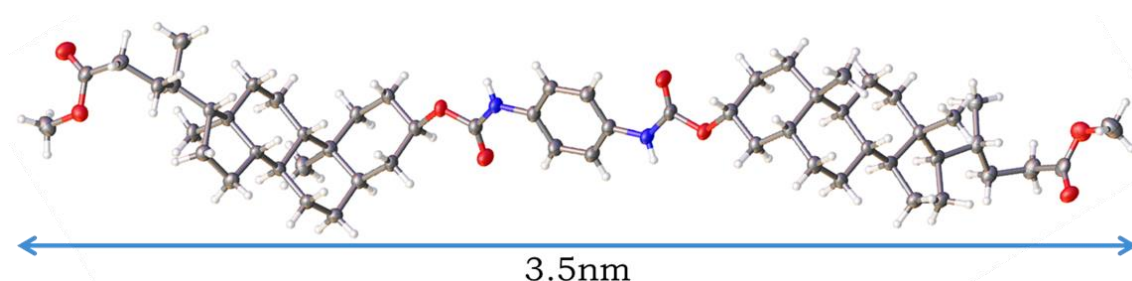
The central structure of the bischolesterol anchor was created by coupling two molecules of lithocholic acid to both sides of *para* phenylene diisocyanate using a procedure adapted from Kobuke,<sup>25</sup> to form carbamate linkages between the two steroid units and the aromatic ring. Prior to this coupling, lithocholic acid was esterified to prevent reaction of the acid with the isocyanate groups on the phenyl ring (Scheme 4.1).



**Scheme 4.1** Synthesis of the central structure of the bischolesterol anchor.

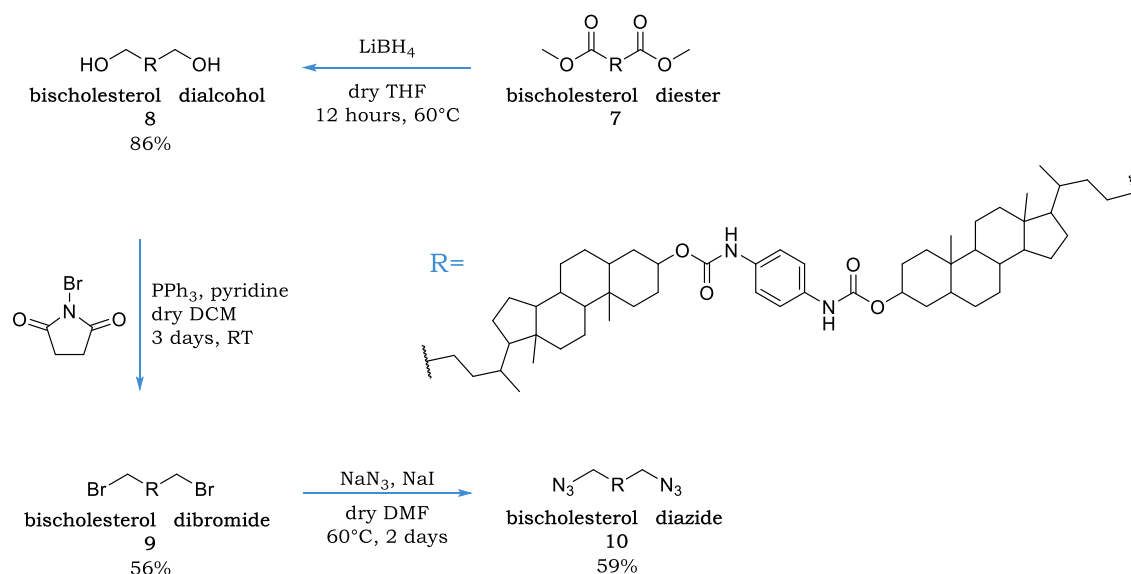
Bischolesterol diester **7** was found to precipitate during the course of the reaction and was inherently insoluble in all solvents tested (e.g. DCM, methanol, hexane, diethyl ether, water, DMSO, THF), making purification difficult. The product was soluble in hot chloroform at low concentrations but attempts to purify the compound by recrystallization from this solvent proved unsuccessful. Eventually the product was purified by flash chromatography. To avoid blockage

of the silica column caused by precipitation of the product, the crude was purified in small batches and loaded onto the column in hot chloroform. However, product loss during column purification resulted in low yields for this reaction. The pure product was a crystalline white solid allowing for a crystal structure of the compound to be determined (Figure 4.7). The length of the molecule was calculated to be 3.5nm and deemed appropriate for suitable anchorage across a whole lipid bilayer.



**Figure 4.7** Crystal structure of bischolesterol diester **7** with ellipsoids drawn at the 50% probability level. The structure contains a molecule of toluene which has been omitted for clarity (see Experimental section 5.14.6 for further details).

Strategies 1-2 require azide groups at both ends of the bischolesterol molecule for triazole coupling. The next synthetic steps involved a series of end functional group conversions (Scheme 4.2). Both ester groups were reduced to alcohols with lithium borohydride. This reducing agent was selected as it was strong enough to reduce the ester group without affecting the carbamate linkages. Sodium borohydride was deemed too weak a reducing agent, and lithium aluminium hydride too aggressive. The reaction proceeded to completion after heating to 60°C for a few hours to generate bischolesterol dialcohol **8**. The crude product was more soluble in a variety of solvents than its predecessor and was successfully purified by flash chromatography in good yield.

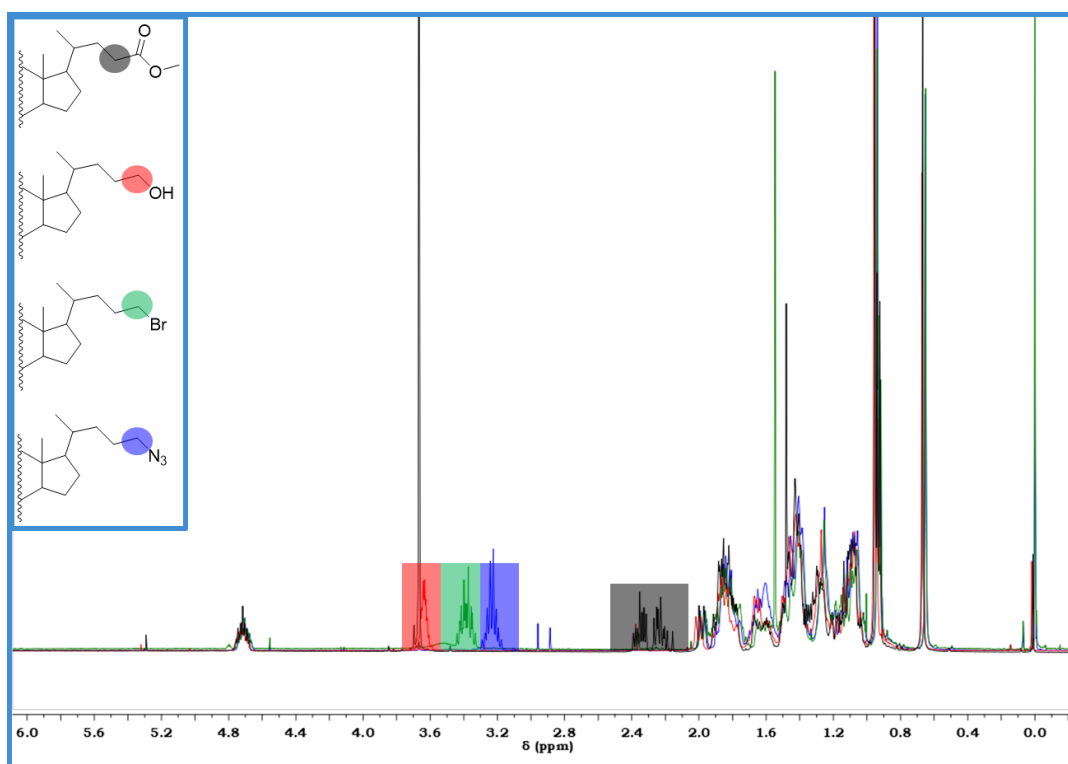


**Scheme 4.2** End functional group conversion of bischolesterol anchor.

Bischolesterol dialcohol **8** was converted to bischolesterol dibromide **9** by reaction with triphenylphosphine and N-bromosuccinimide. The crude was a brown/black sticky solid which proved difficult to work with when transferring to different vessels during work up and column purification, resulting in a loss of product. However, the product was successfully purified with acceptable yields. The last step entailed reaction of bischolesterol dibromide **9** with sodium azide to generate bischolesterol diazide **10**. Sodium iodide was added for halogen exchange between bromide and iodide ions (Finkelstein reaction),<sup>30</sup> leading to faster substitution to the final azide. Compounds **9** and **10** have identical retention factors ( $R_f$ ) and thus the reaction could not be monitored by TLC. The reaction was left heated at  $60^\circ\text{C}$  for 2 days to encourage completion. Product identification was determined by a mini workup and IR analysis of the crude, with the presence of a signal at  $2090\text{cm}^{-1}$  which is characteristic for organic azide groups. Reaction completion was verified by a mini workup and

NMR analysis of the reaction. The final product was isolated by precipitation from water and dried to give a white crystalline solid in suitable yields.

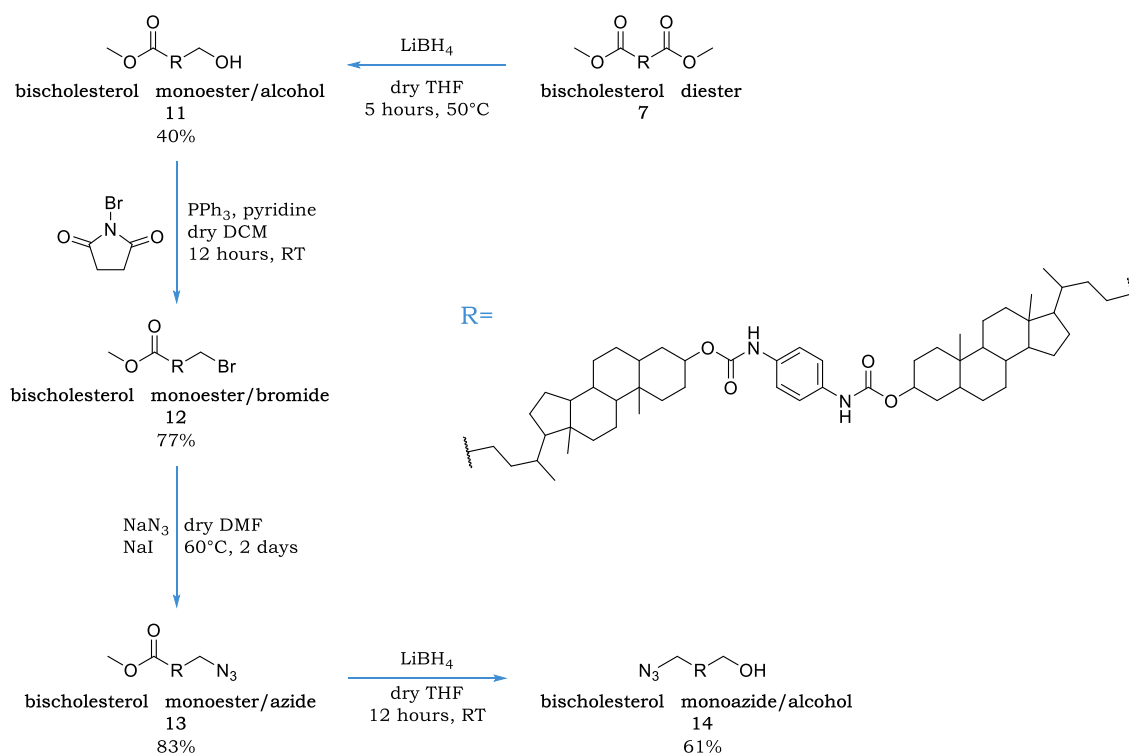
All symmetrical bischolesterol derivatives were characterised by NMR, mass spectrometry, and IR. The  $^1\text{H}$  NMR spectra of the compounds were almost identical except for signals corresponding to the methylene group directly attached to the functional group of interest (Figure 4.8). This allowed for the functional group conversions to be monitored by the shift in the environment of these particular protons. In addition to bischolesterol diester **7**, crystal structures were also obtained for bischolesterol dibromide **9** and bischolesterol diazide **10** (see Experimental section 5.14.6 for further details and Appendix section 6.5 for images).



**Figure 4.8** Overlay of  $^1\text{H}$  NMR spectra of bischolesterol derivatives in  $\text{CDCl}_3$ . Highlighted signals correspond to the methylene protons adjacent to the relevant functional group.

### 4.3.2 Synthesis of asymmetrical bischolesterol derivatives

Strategy 3 required desymmetrisation of the bischolesterol anchor prior to coupling with the cyanine dye. Functional group conversions were applied to one end of the molecule, and then to the other end (Scheme 4.3). Starting with bischolesterol diester **7**, one ester group was reduced to an alcohol with lithium borohydride to give bischolesterol monoester/alcohol **11**. As both ester groups have identical reactivities, selectively reducing one over the other proved difficult. To avoid over-reduction to the dialcohol **8**, the reducing agent was added in small batches over time, and reaction times were shortened. This resulted in a mixture of starting material, desired product, and over-reacted product. The crude was purified by column to isolate compound **11** in acceptable yields.



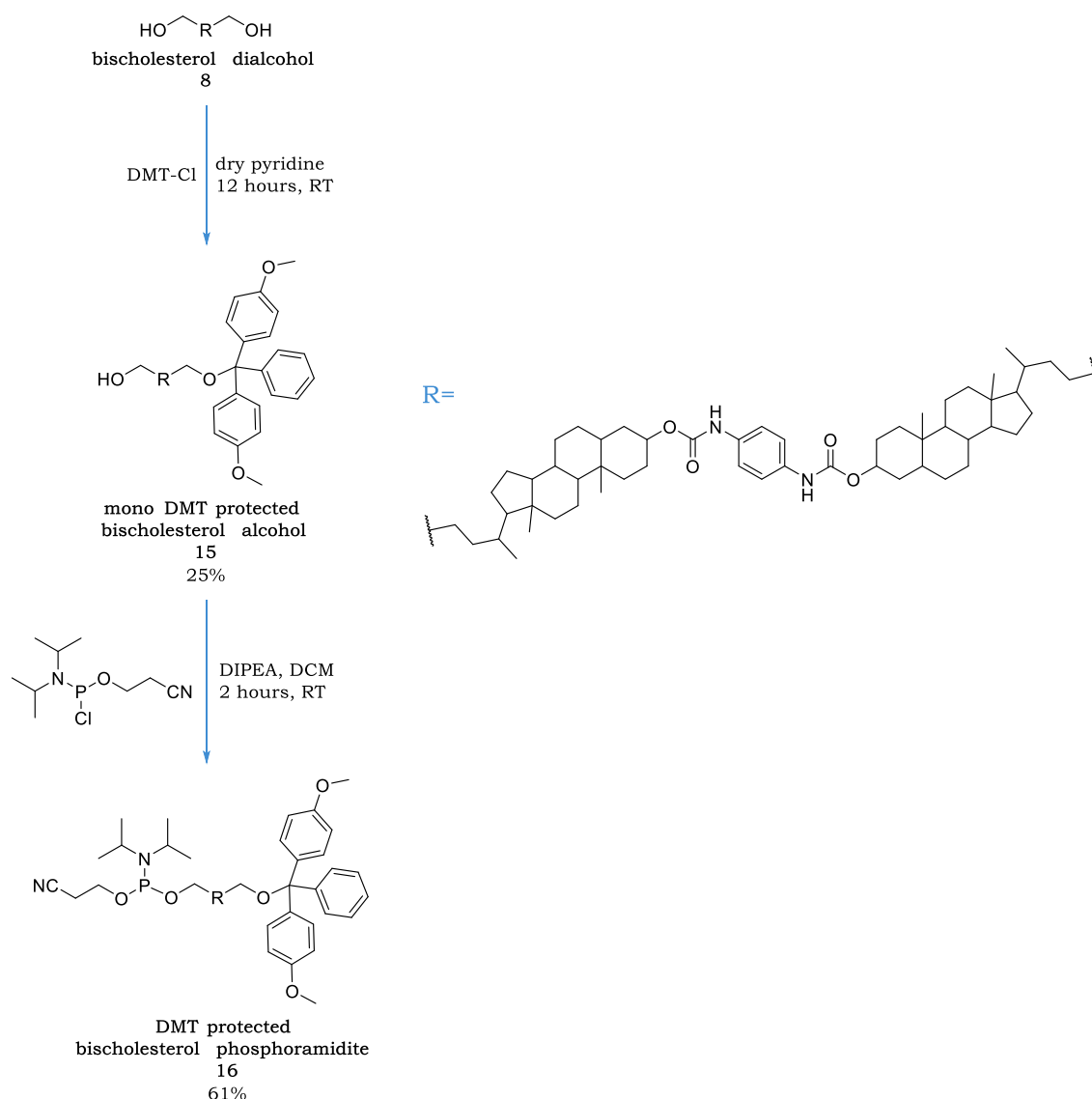
**Scheme 4.3** Synthesis of asymmetrical bischolesterol derivatives.

Bischolesterol monoester/alcohol **11** was then converted to bischolesterol monoester/bromide **12** with triphenylphosphine and N-bromosuccinimide. Compound **11** was only partially soluble in the DCM solvent used for the reaction and was thus added as a suspension, whereas dialcohol **8** had been fully soluble for the di-bromination reaction. Shorter reaction times were also applied to this reaction compared to the synthesis of the dibromide. Bischolesterol monoester/bromide **12** was reacted onwards with sodium azide to form bischolesterol monoester/azide **13** using similar conditions applied to the di-azidation reaction. As observed with the bischolesterol diazide **10** reaction, both starting material and product had the same  $R_f$  and thus reaction progression could not be determined by TLC. Once the azide group was formed on one end of the molecule, the last step was to reduce the ester group on the other end of the molecule to form bischolesterol monoazide/alcohol **14**. Lithium borohydride was employed as the reducing agent for this reaction. Minor reduction of the azide group was observed but the desired compound was found to be the major product and isolated by column chromatography in good yield.

#### 4.3.3 Synthesis of DMT protected bischolesterol phosphoramidite

Strategy 4 is different to the other strategies in that it does not utilise azide-alkyne cycloaddition reactions to couple either the DNA or the cyanine dye to the bischolesterol anchor. Instead, phosphoramidite techniques are applied to react both ends of the bischolesterol molecule. To prepare the bischolesterol anchor for use in automated solid phase DNA synthesis, one end must contain a DMT group for alcohol protection, and the other end a phosphoramidite group for phosphate conversion (Scheme 4.4). Bischolesterol dialcohol **8** was reacted with one equivalent of dimethoxytrityl chloride to couple DMT to one side of the

molecule. However, there was considerable over reaction to the di-DMT protected product, as both alcohol groups have matching reactivities. This resulted in a low yield of product. The remaining alcohol group was then reacted with the phosphitylating reagent 2-cyanoethyldiisopropylchlorophosphoramidite to produce the final product, DMT protected bischolesterol phosphoramidite **16**, in good yield.



**Scheme 4.4** Synthesis of DMT protected bischolesterol phosphoramidite **16** for incorporation into automated solid phase DNA synthesis.

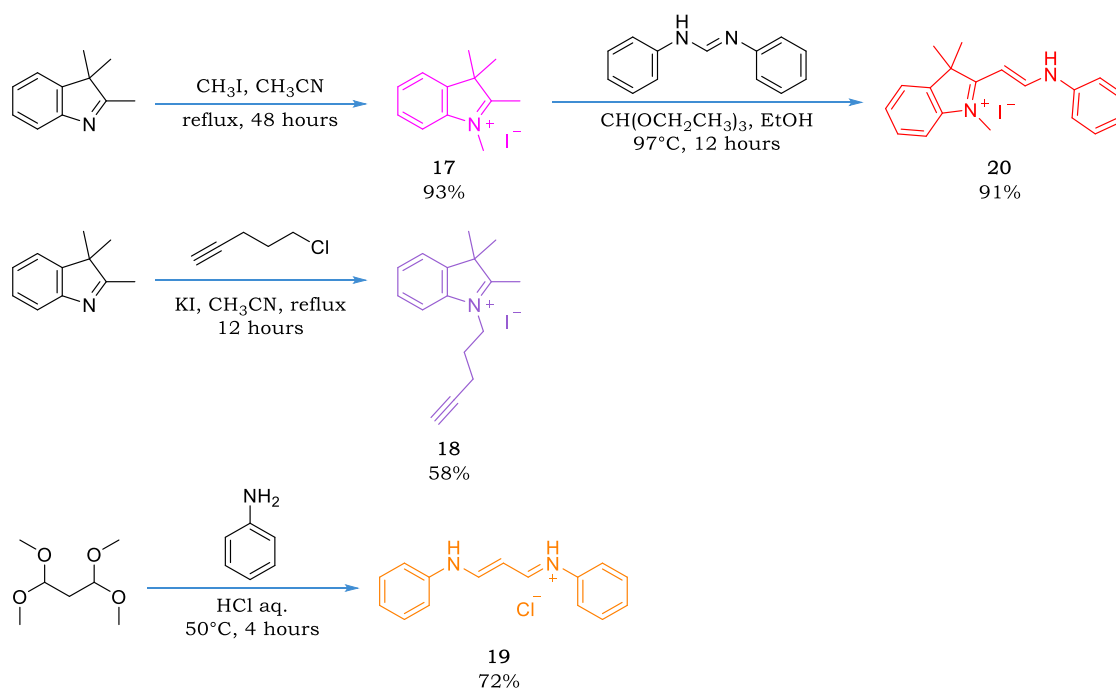
## 4.4 Synthesis of cyanine dyes

Although alkyne modified versions of the cyanine dyes Cy3 and Cy5 are commercially available, it was considered more cost effective to synthesise the dyes in-house.

### 4.4.1 Synthesis of cyanine dye precursors

Efforts towards the synthesis of cyanine dyes were conducted in an adapted literature procedure,<sup>31</sup> which first required the synthesis of two indolenine based precursors and a linker molecule (Scheme 4.5). Initially 2,3,3-trimethylindolenine was reacted with methyl iodide to produce 1,2,3,3-tetramethyl-3H-indolinium iodide **17** in high yield. Purification was performed by allowing the warm reaction mixture to cool slowly to room temperature, resulting in the formation of light purple crystals which were washed with cold diethyl ether to remove any excess starting materials. 2,3,3-trimethylindolenine was then reacted with 5-chloro-1-pentyne to yield the desired alkyne terminated indolenine derivative, 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide **18**. A crude <sup>1</sup>H NMR spectrum revealed the presence of a new triplet at 4.6ppm which was indicative of the formation of the new N-alkyl species. Initial attempts to purify the compound by column chromatography proved ineffective as the product was observed to degrade during purification. To simplify this method, the crude material was precipitated into a variety of solvents, with precipitation from diethyl ether yielding the pure product with acceptable yields.





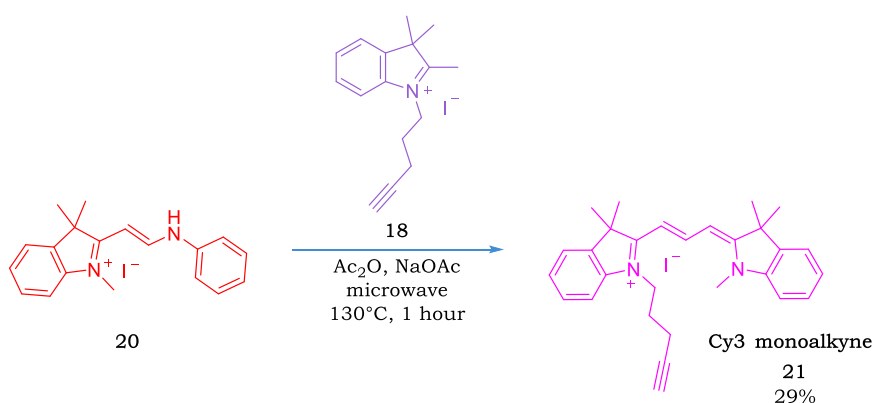
**Scheme 4.5** Synthesis of cyanine dye precursors.

Malondialdehyde bis(phenylimine) monohydrochloride **19** was synthesised by heating 1,1,3,3-tetramethoxypropane and aniline in 1M hydrochloric acid. The resulting solution was cooled and filtered to yield the product as a bright orange precipitate in good yield after washing with further portions of 1M HCl. 1,2,3,3-tetramethyl-3H-indolinium iodide **17** was reacted further with N,N'-diphenyl formamidine at reflux to yield the desired asymmetric cyanine dye precursor, 1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolinium iodide **20**. Although column chromatography proved effective, a simple precipitation in diethyl ether resulted in the desired product as an orange solid in a respectable yield. Analysis by  $^1\text{H}$  NMR spectroscopy (Appendix section 6.3.4) revealed doublets at 8.67ppm and 6.20ppm, corresponding to the double bond, whilst the singlet at 3.73ppm can be assigned to the N-methyl residue, demonstrating coupling of the two moieties.

#### 4.4.2 Synthesis of Cy3 monoalkyne

To create Cy3 monoalkyne **21**, coupling of 1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolium iodide **20** with 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide **18** was initially performed in pyridine and acetic anhydride at 50°C for 24 hours. The reaction was observed not to go to completion by TLC, despite increasing the reaction time and subsequently the temperature to 100°C. The resulting mixture was cooled and precipitated into cold diethyl ether in an attempt to purify the reaction, however TLC analysis showed three tightly packed bands which were resultantly inseparable by column chromatography.

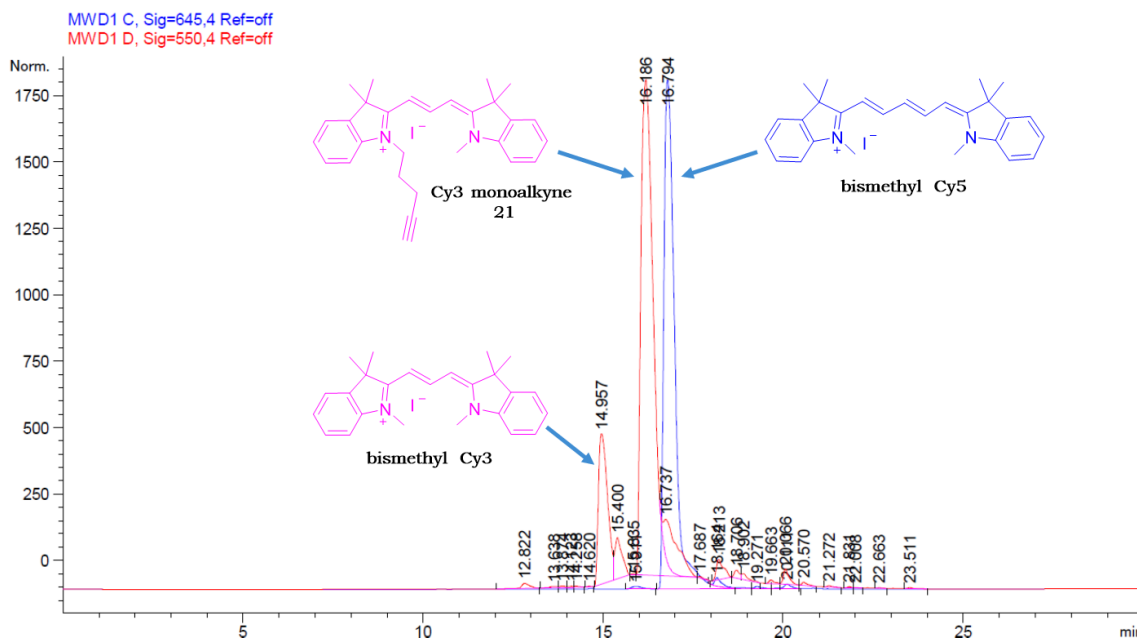
An alternative synthetic route<sup>32</sup> was investigated in which a microwave was utilised to conduct the coupling reaction (Scheme 4.6).



**Scheme 4.6** Synthesis of Cy3 monoalkyne **21**.

1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolium iodide **20** and 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide **18** were dissolved in acetic anhydride, sodium acetate was added, and the reaction was microwaved at 130°C under standard power and pressure. An irradiation time of 1 hour was

found to be sufficient for reaction completion. Longer irradiation resulted in over reaction and addition of iodide to the alkyne bond in the Cy3 monoalkyne product as evidenced by mass spectrometry with an increase in mass by 127 units, and  $^1\text{H}$  NMR analysis with the disappearance of the alkyne proton resonance at 2.14ppm. The reaction was cooled, and the crude product precipitated from cold diethyl ether and isolated as a purple solid. Initially, purification was performed by column chromatography, which isolated the desired dye from most coloured impurities. However, there were a few side products with retention closely similar to that of the desired product. This led to further purification of the dye with preparative HPLC using a C18 column and a water/methanol gradient eluent. With this method, the dye was loaded in low concentrations to prevent saturation of the column, which proved time consuming. Along with the final product, the impurities were isolated and identified by mass spectrometry to be symmetrical versions of Cy3 and Cy5 dyes (Figure 4.9).

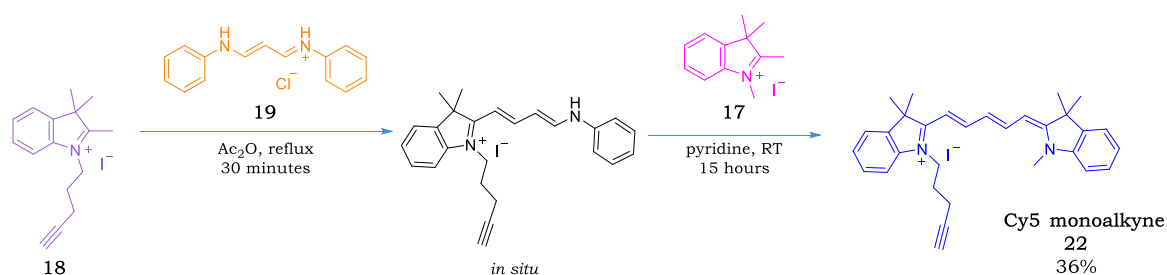


**Figure 4.9** Overlaid preparative HPLC chromatogram of the Cy3 monoalkyne **21** reaction after column purification showing detector signals at 550nm (red) and 645nm (blue).

The final yield for Cy3 monoalkyne **21** was fairly low due to formation of symmetrical versions of the dye (despite the use of an asymmetrical precursor), the low solubility of the crude in the eluent system used for HPLC, and close elution of Cy3 monoalkyne and the side product bismethyl Cy5, resulting in lost product.

#### 4.4.3 Synthesis of Cy5 monoalkyne

To synthesise Cy5 monoalkyne **22**, a suspension of 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide **18** and malondialdehyde bis(phenylimine) monohydrochloride **19** in acetic anhydride was refluxed to generate an asymmetric precursor *in situ*, which was then reacted onwards with 1,2,3,3-tetramethyl-3H-indolinium iodide **17** in pyridine to form the final product (Scheme 4.7).



**Scheme 4.7** Synthesis of Cy5 monoalkyne **22**.

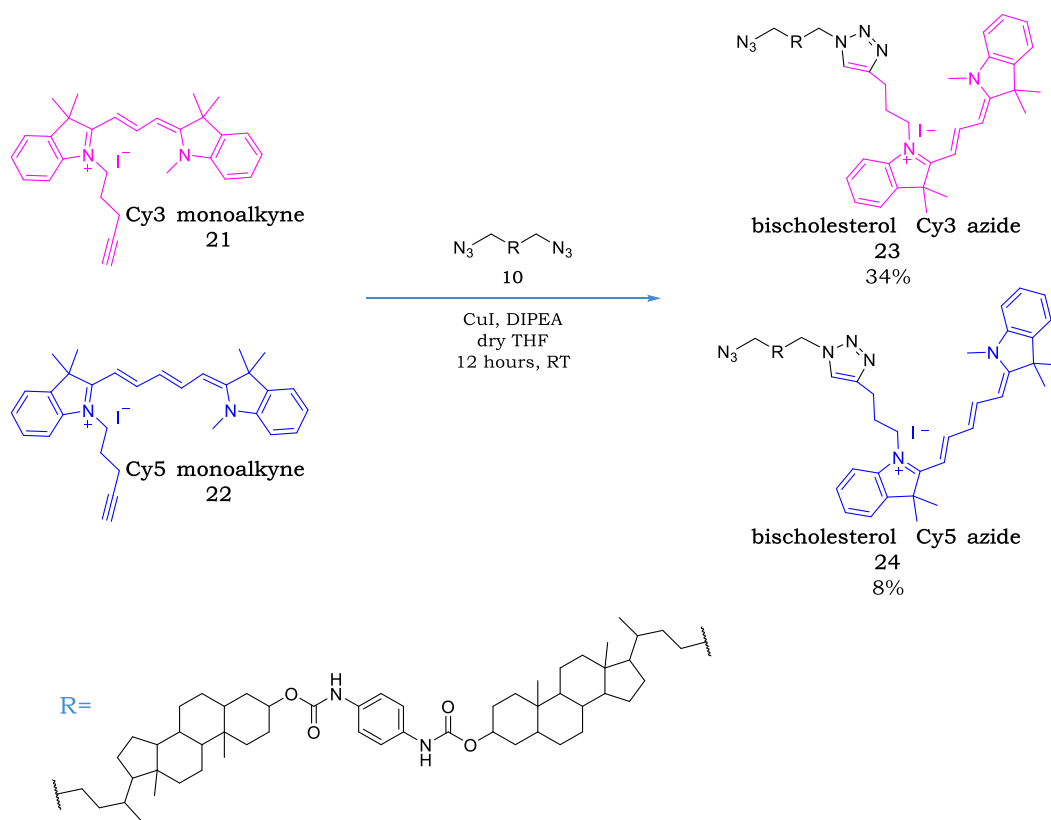
Initially, the asymmetric precursor was observed to react with residual compound **18** prior to the addition of **17** to form a bisalkyne version of Cy5. Symmetrical byproducts were detected even when all components were added as a one pot reaction. To minimise this side reaction, **18** was added gradually to the reaction, and acetic anhydride removed prior to the addition of **17** in pyridine. Precipitation of the reaction into cold diethyl ether as described in the literature<sup>31</sup> led to significant loss of product, and thus this step was avoided. The crude was first purified by column chromatography but then was further purified by preparative HPLC to separate the final compound from symmetrical byproducts (Figure 4.10). As with Cy3 monoalkyne **21**, the dye was purified in small batches which proved time consuming. Once again, the yield for Cy5 monoalkyne **22** was severely impacted by production of symmetrical forms of the dye, resulting in reduced amounts of product and difficult separation of **22**.



## 4.5 Synthesis of bischolesterol-dye conjugates

### 4.5.1 Synthesis of bischolesterol dye azides

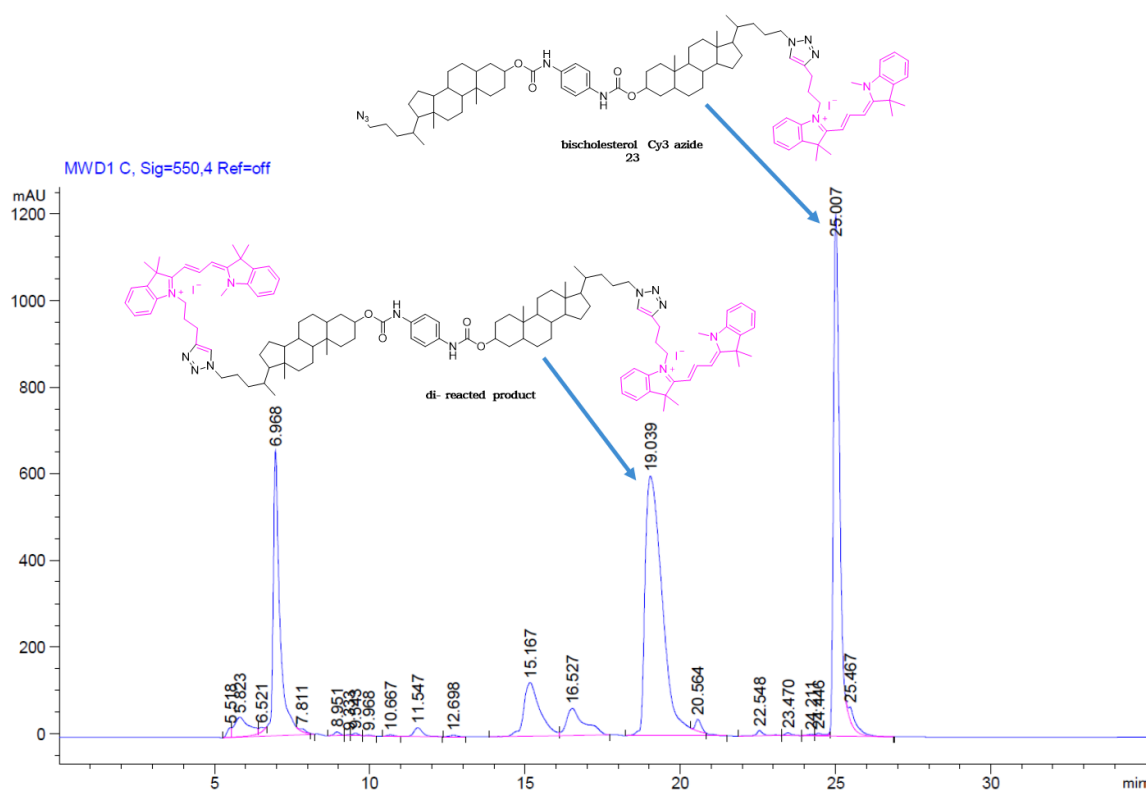
Strategies 1 and 2 require the copper catalysed azide-alkyne cycloaddition of cyanine monoalkyne dyes **21** and **22** to one end of bischolesterol diazide **10**, desymmetrising the molecule in the process (Scheme 4.8).



**Scheme 4.8** Synthesis of bischolesterol dye azides.

Bischolesterol Cy3 azide **23** was synthesised from a solution of Cy3 monoalkyne **21** and bischolesterol diazide **10** in dry THF. The reaction was thoroughly degassed prior to the addition of copper iodide to prevent oxidation of the catalyst and DIPEA was added as a base to further accelerate the cycloaddition reaction. The complete consumption of bischolesterol diazide **10**

was confirmed by TLC after 12 hours. Analysis of the reaction by mass spectrometry identified a di-reacted product alongside the desired mono-reacted product. The crude was purified in small batches by preparative HPLC using a C18 column and a gradient eluent of 90 → 100% methanol. The separation between mono- and di- Cy3 products was vast with a 6-minute difference in retention, making for easy purification by HPLC (Figure 4.11). Despite over reaction of the desired product, bischolesterol Cy3 azide **23** was isolated in a better than expected yield.

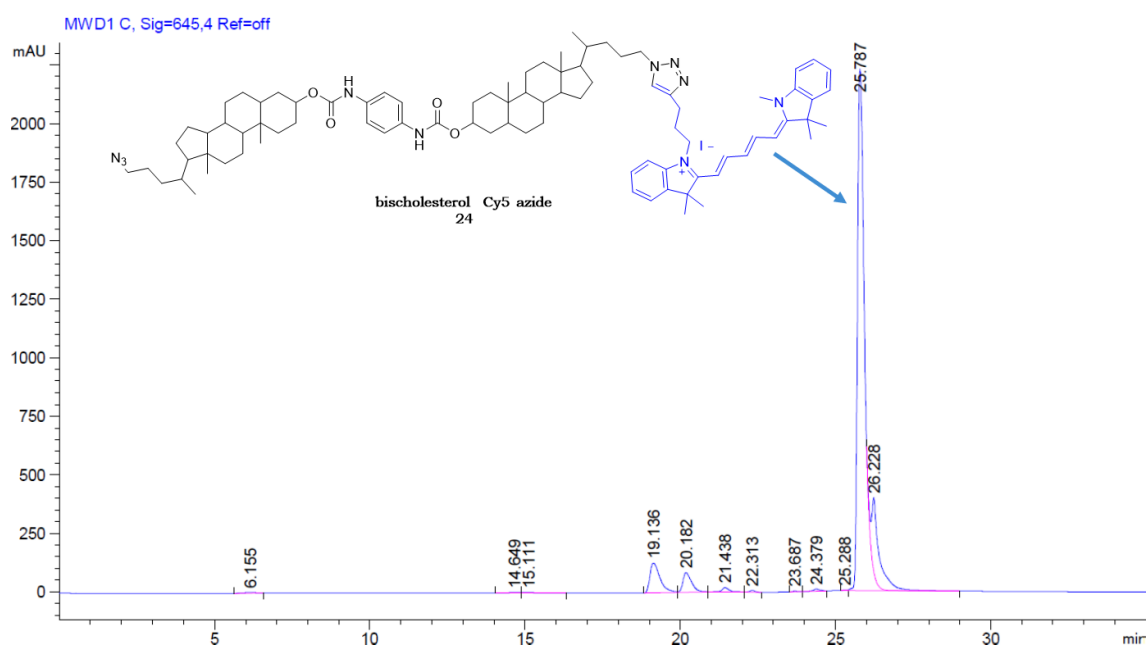


**Figure 4.11** Preparative HPLC chromatogram of bischolesterol Cy3 azide **23** reaction showing detector signal at 550nm.

Bischolesterol Cy5 azide **24** was produced from Cy5 monoalkyne **22** and bischolesterol diazide **10** in identical conditions to the synthesis of compound **23**. Analysis by TLC showed the reaction to be slower than that of the bischolesterol Cy3 conjugation reaction, with incomplete consumption of both



starting materials after 12 hours. The reaction was initially purified by column chromatography to remove unreacted starting materials, but the product was observed to degrade on the silica column. Product containing fractions were further purified by preparative HPLC using the same conditions as the Cy3 version (Figure 4.12). Compound **24** was successfully isolated but in poor yield mainly due to the inefficiency of the reaction and degradation of the product during column purification.

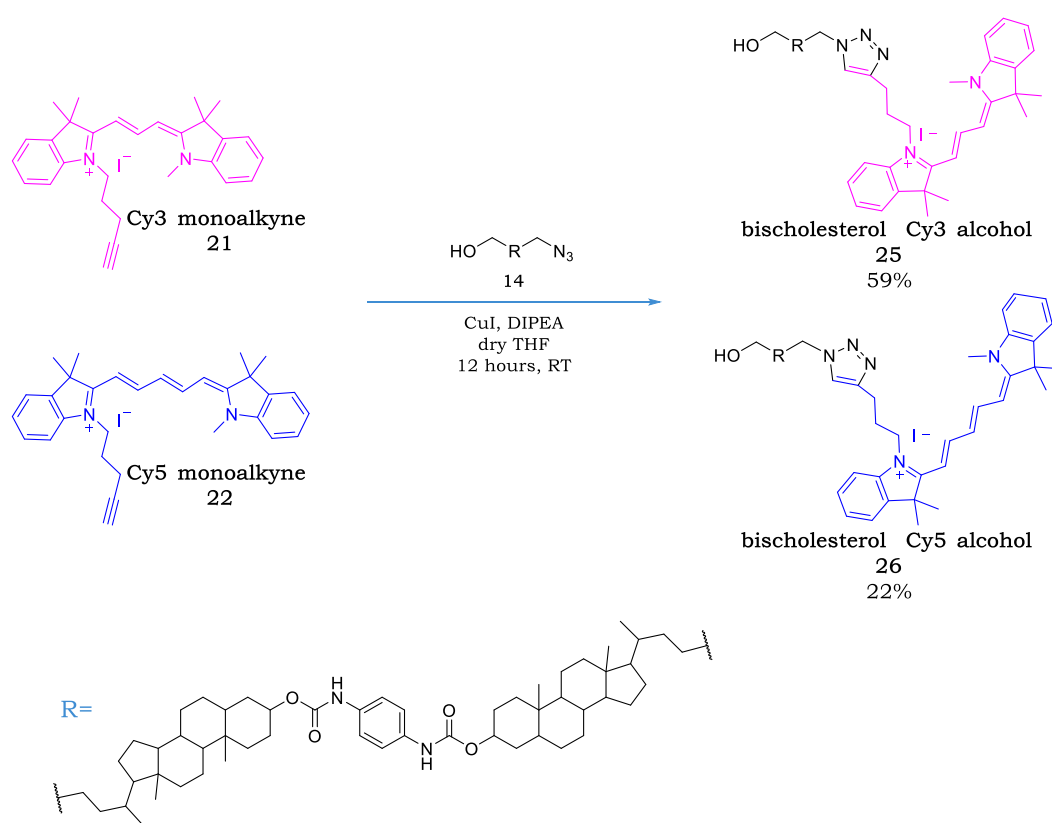


**Figure 4.12** Preparative HPLC chromatogram of bischolesterol Cy5 azide **24** reaction showing detector signal at 645nm.

#### 4.5.2 Synthesis of bischolesterol dye alcohols

The bischolesterol dye alcohol compounds required for strategy 3 were synthesised under identical conditions to their azide counterparts, using bischolesterol monoazide/alcohol **14** and their respective cyanine dye monoalkyne (**21** or **22**, Scheme 4.9). These reactions were more effective than that of the bischolesterol dye azides as there was only one azide site on the

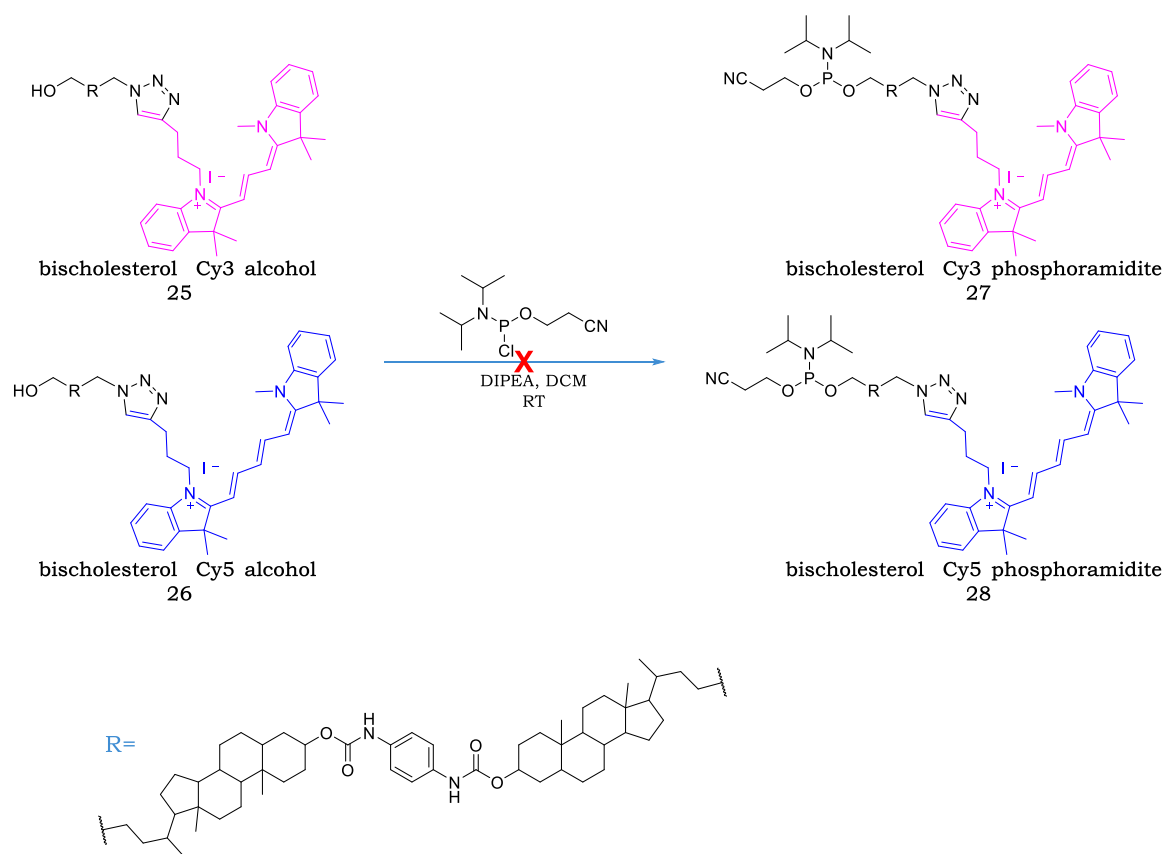
bischolesterol molecule for the dye to react with. As found with the diazide reaction, Cy3 coupling was quicker than that of Cy5 with full consumption of bischolesterol monoazide/alcohol **14** observed after 12 hours of reaction. In contrast to their azide counterparts, both bischolesterol Cy3 alcohol **25** and bischolesterol Cy5 alcohol **26** reactions were successfully purified by column chromatography using an eluent of 5% methanol in DCM. The bischolesterol-dye conjugates had similar retention times to the unreacted dyes, and yields could be further improved by the purification of minor mixed fractions by HPLC. However, this avenue was not explored due to time constraints.



**Scheme 4.9** Synthesis of bischolesterol dye alcohols.

#### 4.5.3 Synthesis of bischolesterol dye phosphoramidites

The penultimate synthetic step for strategy 3 was the conversion of the bischolesterol dye alcohols to bischolesterol dye phosphoramidites using the phosphitylating reagent 2-cyanoethyldiisopropylchlorophosphoramidite (Scheme 4.10). As the phosphitylating reagent is highly sensitive to water, significant efforts were made to ensure the exclusion of water from the reaction. Bischolesterol dye alcohols **25** and **26** were azeotroped with anhydrous DCM, the reaction was performed in anhydrous DCM, and the solution degassed prior to the addition of the phosphitylating reagent. Despite attempts to keep reactions dry, both phosphitylation reactions failed as evidenced by TLC and  $^{31}\text{P}$  NMR, which indicated successful coupling but an incorrect oxidation state for phosphorus. A lack of signal at ~148ppm and signals at 14.2ppm showed oxidation of phosphorus from P(III) to P(V). Copper is known to catalyse the oxidation of phosphorus,<sup>33</sup> and thus any residual copper from the previous cycloaddition step could be damaging to the phosphitylation reaction. Although it was thought that purification by column chromatography of the bischolesterol dye alcohols should have removed copper from these starting materials, these compounds were nonetheless washed with an ethylenediaminetetraacetic acid (EDTA) solution to extract any residual copper prior to the phosphitylation reaction. However, phosphitylation reactions were still unsuccessful even after this EDTA washing, meaning strategy 3 had to be abandoned. The failure to synthesise bischolesterol dye phosphoramidites **27** and **28** from bischolesterol dye alcohols **25** and **26** is currently unexplained.



**Scheme 4.10** Failed synthesis of the bischolesterol dye phosphoramidites.

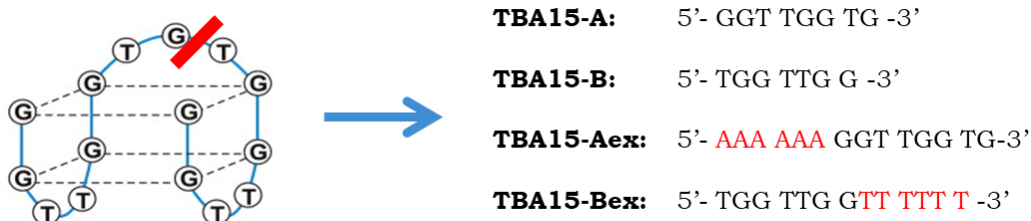
## 4.6 DNA selection and synthesis

### 4.6.1 DNA sequence selection

To coincide with research in the previous chapter, the 15 base thrombin binding aptamer (**TBA15**) was selected as a DNA sequence to append to the bischolesterol anchor. For use in LOC sandwich assays, the 29 base thrombin binding aptamer sequence (**TBA29**) was also selected as it is known to interact with a different site of the thrombin protein. Both complementary DNA and thrombin could be utilised as targets for this system. A non self-folding 15mer often adopted by the Tucker group for antisensing experiments (nina sequence), was also chosen for conjugation to bischolesterol.

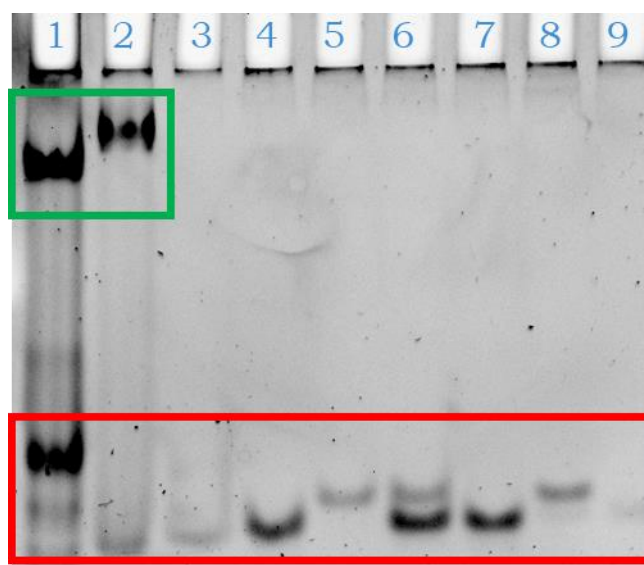
### 4.6.2 Splitting the thrombin binding aptamer sequence for sandwich-based sensing

The 15 base thrombin binding aptamer sequence was split into an 8mer and 7mer for use as an alternative system in the proposed LOC sandwich assays, as there are a handful of reports demonstrating thrombin binding with these split aptamer sequences.<sup>34-36</sup> The 8mer and 7mer strands were synthesised by automated DNA synthesis, along with extended versions, purified by reversed phase HPLC and characterised by mass spectrometry (Figure 4.13).



**Figure 4.13** Split thrombin binding aptamer sequences.

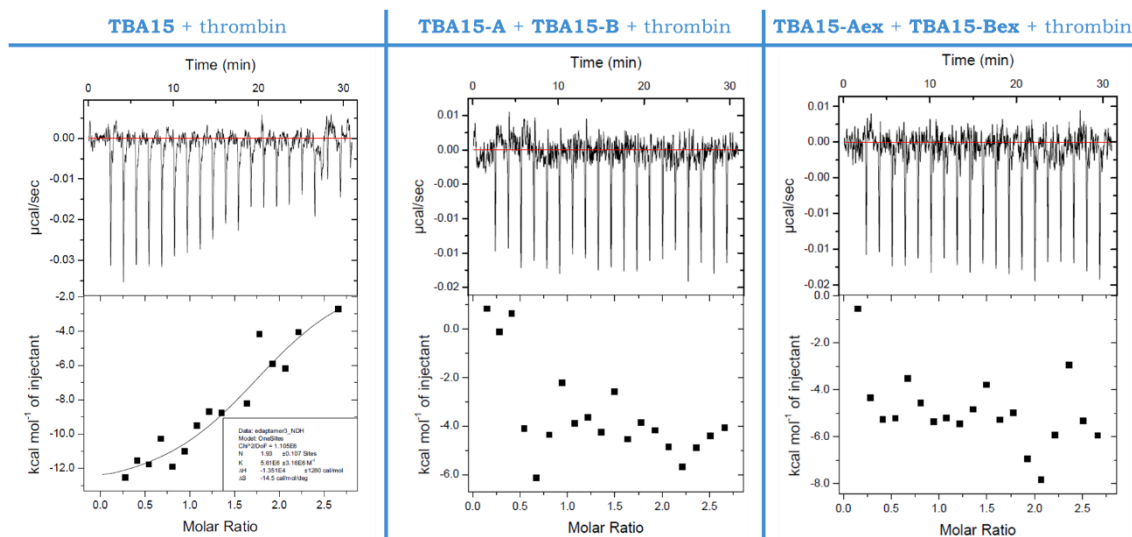
The binding of the split strands to the thrombin protein was first analysed by gel EMSA (Figure 4.14), using the full aptamer sequences **TBA15** and **TBA29** as positive controls. The Diamond™ dye used for staining DNA in gels could not stain the short strands **TBA15-A** and **TBA15-B** but was able to stain their extended counterparts. UV shadowing was thus employed to detect the shorter strands. Nonetheless, thrombin binding was not observed with any of the split aptamer combinations in contrast to the literature.



**Figure 4.14** DNA stain of gel EMSA of thrombin with full and split thrombin binding aptamer strands. DNA-thrombin bands are shown in the green box and DNA only bands are shown in the red box. **Lane 1:** **TBA29** + thrombin, **lane 2:** **TBA15** + thrombin, **lane 3:** **TBA15** only, **lane 4:** **TBA15-Aex** + thrombin, **lane 5:** **TBA15-Bex** + thrombin, **lane 6:** **TBA15-Aex** + **TBA15-Bex** + thrombin, **lane 7:** **TBA15-Aex** + **TBA15-B** + thrombin, **lane 8:** **TBA15-A** + **TBA15-Bex** + thrombin, **lane 9:** **TBA15-A** + **TBA15-B** + thrombin.

Isothermal titration calorimetry (ITC) was then employed to analyse the binding of the split strands to the thrombin protein using the full aptamer strand **TBA15** as a positive control. Thrombin (10μM) was loaded into the syringe and injected sequentially into a cell containing solutions of either the

full or two halves of the thrombin binding aptamer sequence (1 $\mu$ M), and the heats of interaction measured (Figure 4.15). Binding curves were obtained with the full aptamer whereas no binding was observed with the split aptamer sequences, in agreement with gel EMSA results.



**Figure 4.15** ITC binding data for thrombin with full and split thrombin binding aptamer strands (see experimental section 5.11 for further details).

As neither technique could verify binding of the split strands to the thrombin target, the plan to couple these strands to the bischolesterol anchor was abandoned.

#### 4.6.3 DNA synthesis for strategies 1-2

Automated DNA synthesis was utilised to create alkyne and DBCO modified DNA for use in strategies 1-2, the sequences of which are shown in Table 4.1. Alkyne and DBCO phosphoramidites were purchased from commercial suppliers and coupled to the 5' end of the selected sequences (Figure 4.16).





## 4.7 Synthesis and purification of transmembrane molecules

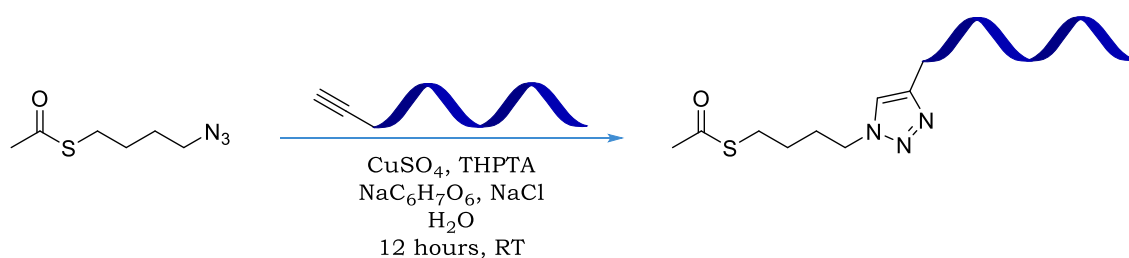
### 4.7.1 Strategy 1- coupling alkyne modified DNA to bischolesterol dye azide

The low amounts of DNA alkyne stocks required coupling reactions to be undertaken at low volumes and concentrations. Stock solutions of all components were made beforehand and aliquots taken from these stocks for the reactions. Literature suggests the use of excess copper for successful cycloaddition with DNA.<sup>37</sup> Due to the small scale and polar nature of DNA, reactions could not be monitored by normal TLC methods. Instead, reaction progression was tracked by mass spectrometry, HPLC, and/or analysis by polyacrylamide gels. Mass spectrometry proved to be a powerful technique in determining the success of the coupling reaction.

Initial attempts to couple the bischolesterol dye azides to alkyne modified DNA followed a literature procedure<sup>38</sup> which utilised copper sulphate, tris(carboxyethyl)phosphine (TCEP), and tris(benzyltriazolylmethyl)amine (TBTA), in a mixture of water, t-butanol, THF, and DMSO. TCEP reduces copper sulphate to generate the active copper(I) species required to catalyse the cycloaddition reaction, and TBTA is used as a stabilising ligand to prevent the oxidation and disproportionation of the copper(I) species.<sup>39</sup> Degassing of solutions was necessary to prevent oxidation of copper(I). The bischolesterol dye azides had limited solubility in this reaction at room temperature. When the mixtures were heated to 60°C, a white precipitate appeared which was attributed to degraded DNA. Copper is known to interact with and cause

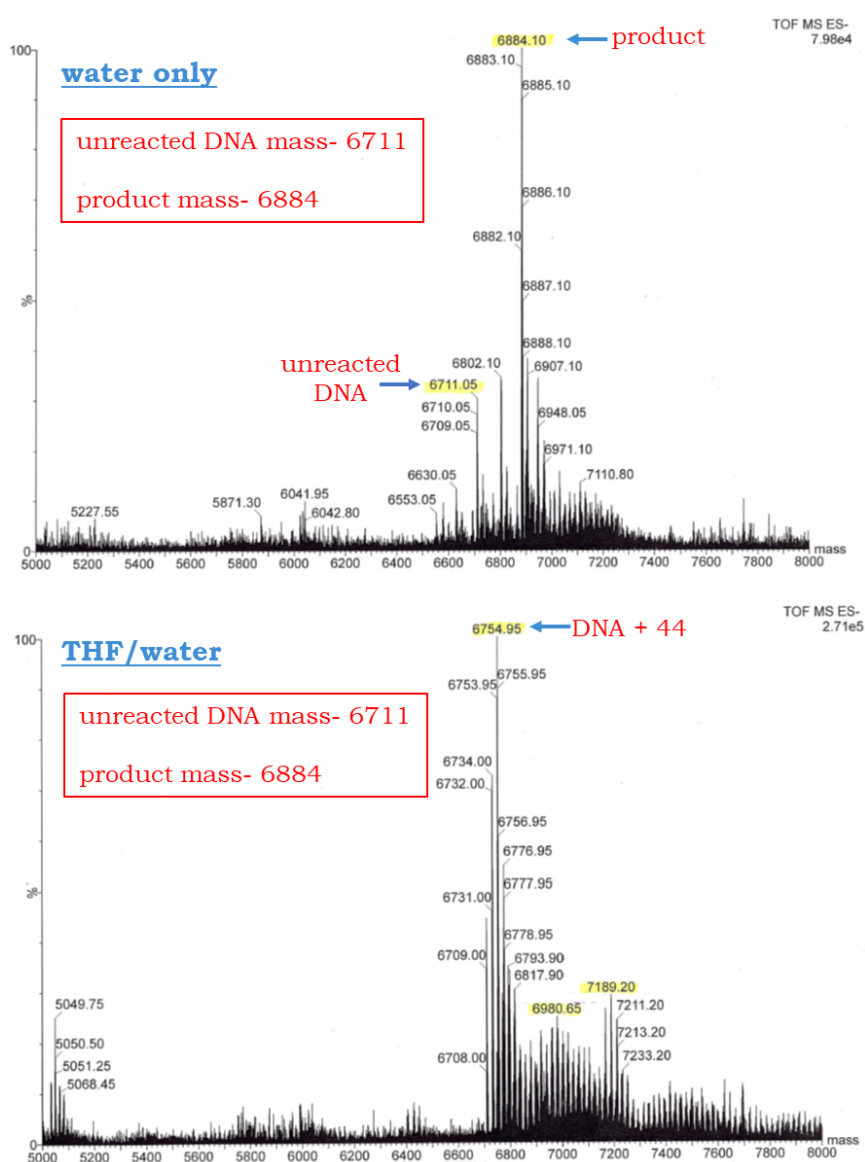
degradation of DNA if not properly chelated.<sup>40</sup> Mass spectrometry analysis of the reactions at both RT and 60°C indicated that this indeed was the case. TBTA:Cu ratios were then increased from 1:1 to 10:1, resulting in reduced but still significant DNA degradation, and addition of sodium chloride to the reaction reduced DNA degradation further. However, despite increasing TBTA:Cu ratios, replacing TBTA to the more water-soluble stabilising agent tris(hydroxypropyltriazolylmethyl)amine (THPTA), swapping TCEP for sodium ascorbate/ascorbic acid, increasing the volume of THF, and reducing/eliminating the volumes of water/DMSO/t-butanol (to encourage dissolution of the bischolesterol dye azides), the cycloaddition reactions with this procedure were ultimately unsuccessful.

There was a suspicion that copper sulphate was not a suitable copper source for the desired application. Alternative cycloaddition reactions were therefore performed with alkyne modified DNA and water soluble azide s-(4-azidobutyl)thioacetate with copper sulphate, THPTA, sodium ascorbate, and sodium chloride in water (Scheme 4.11).



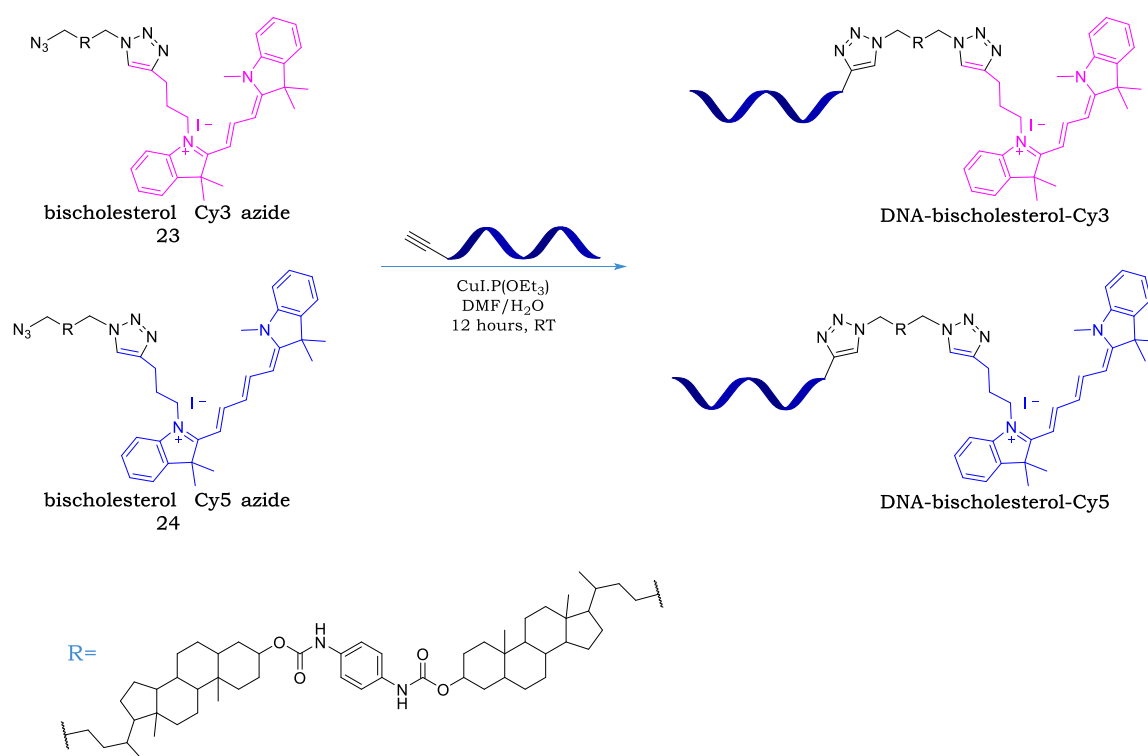
**Scheme 4.11** Copper catalysed cycloaddition of alkyne modified DNA and s-(4-azidobutyl)thioacetate.

A significant amount of product was observed after 12 hours, as evidenced by mass spectrometry. However, when repeated in 50% THF in water, no mass for the product was observed using this technique (Figure 4.17). This suggests that THF hinders the cycloaddition reaction with this catalyst. As the bischolesterol dye azides are insoluble in water it was decided that a copper catalyst which performs better in organic media would be explored.



**Figure 4.17** Mass spectra of alkyne modified DNA and s-(4-azidobutyl)thioacetate reaction in water showing the correct product mass (top), and in 1:1 water:THF showing incorrect masses (bottom).

The next attempts to connect alkyne modified DNA to bischolesterol dye azides were conducted using an alternative literature procedure which utilised copper(I) iodotriethylphosphite as a catalyst to couple DNA to organic polymers in DMF.<sup>39</sup> There was no need for reducing or stabilising agents with this catalyst as copper is already in the oxidation state required for catalysis and comes chelated to the triethylphosphite ligand. 1mM stock solutions of bischolesterol dye azides and CuI.P(OEt<sub>3</sub>) were prepared in degassed DMF, aliquots were taken, and 1ml reactions made with 10μM DNA alkyne, and DNA:bischolesterol dye azide:Cu ratios of 1:10:10 (Scheme 4.12). The solutions were then sealed and left for a minimum of 12 hours.

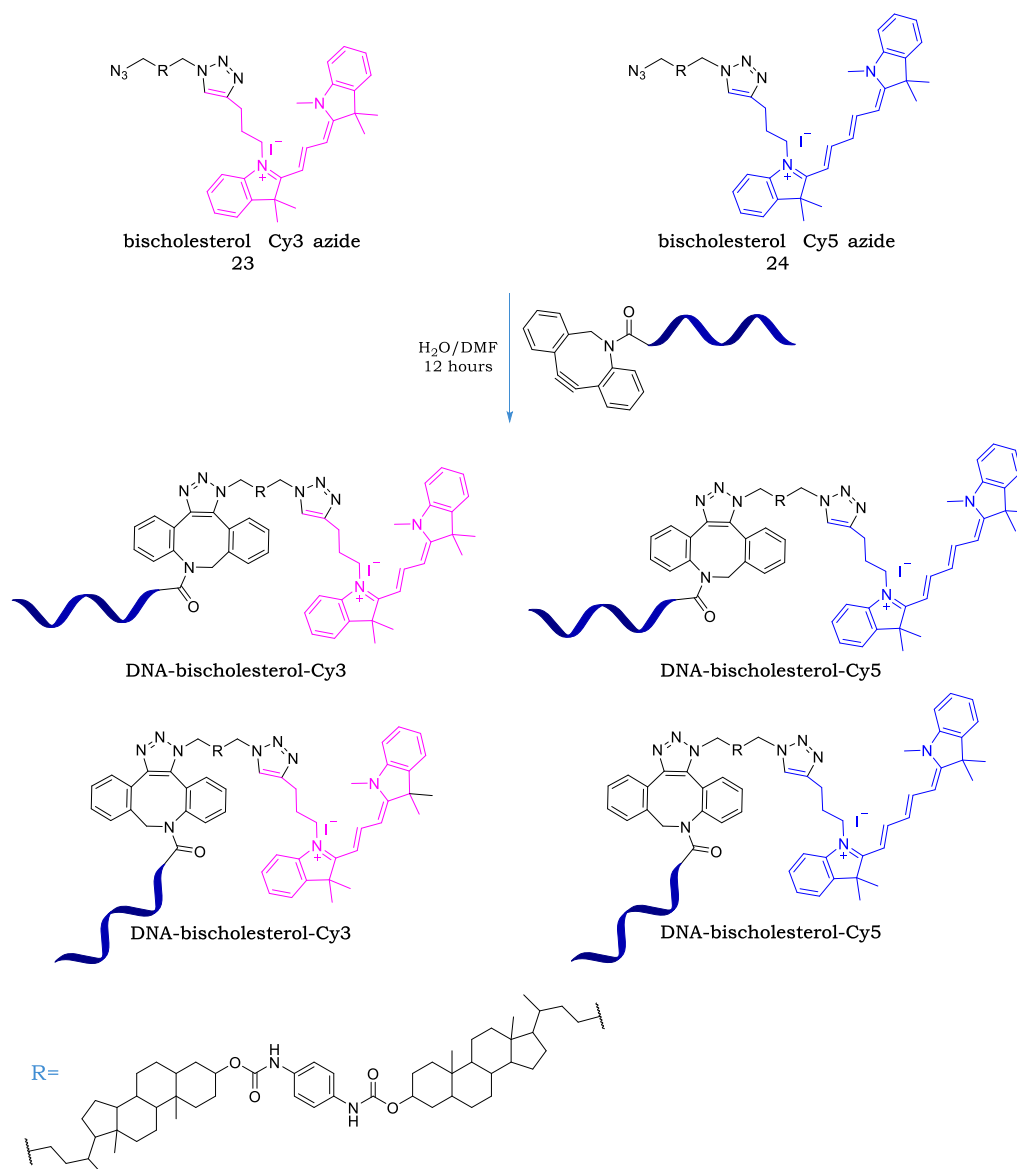


**Scheme 4.12** Copper catalysed cycloaddition of alkyne modified DNA and bischolesterol dye azides.

Analysis by mass spectrometry of the reactions showed product formation to varying degrees which was promising. However, further reactions fared worse with subsequent use of the catalyst as coupling reactions with newly purchased catalyst were successful whereas reactions with less fresh catalyst were not. It suggested that the commercially supplied  $\text{CuI.P(OEt}_3\text{)}$  catalyst was deteriorating, despite being stored and handled in an inert atmosphere (glove box).  $\text{CuI.P(OEt}_3\text{)}$  was therefore synthesised instead following a published procedure.<sup>41</sup> Reactions were now more successful, with no loss in activity of this reagent over time when stored under argon at 0°C. The conjugates TBA15-bischolesterol-Cy3, TBA29-bischolesterol-Cy3, nina-bischolesterol-Cy3, TBA15-bischolesterol-Cy5, and nina-bischolesterol-Cy5 were successfully synthesised *via* this method (Scheme 4.12).

#### 4.7.2 Strategy 2- coupling DBCO modified DNA to bischolesterol dye azide

The strained nature of the DBCO group allows the cycloaddition reaction to proceed quickly under ambient conditions in the absence of a copper catalyst. However, a major disadvantage of this method is that the reaction is unselective over the production of different regioisomers, whereas copper catalysed versions tend to only generate the 1,4-regioisomer.<sup>42</sup> Additionally, the DBCO-triazole group makes quite a bulky linkage between DNA and the bischolesterol anchor (Scheme 4.13), which might affect the orientation of the transmembrane molecule in the lipid bilayer.



**Scheme 4.13** Copper free cycloaddition of DBCO modified DNA and bischolesterol dye azides.

The conjugation of bischolesterol dye azides and DBCO modified DNA was initially performed in the fridge in solutions of 1:1 water:DMF with an excess of bischolesterol dye azide and left for a minimum of 12 hours. Analysis of solutions by mass spectrometry showed minor product formation. The next attempts to improve yields were to increase the temperature of the reactions to room temperature and 50°C. These increases were found to have no effect on the reaction yield with a minor amount of product still observed. Increasing the

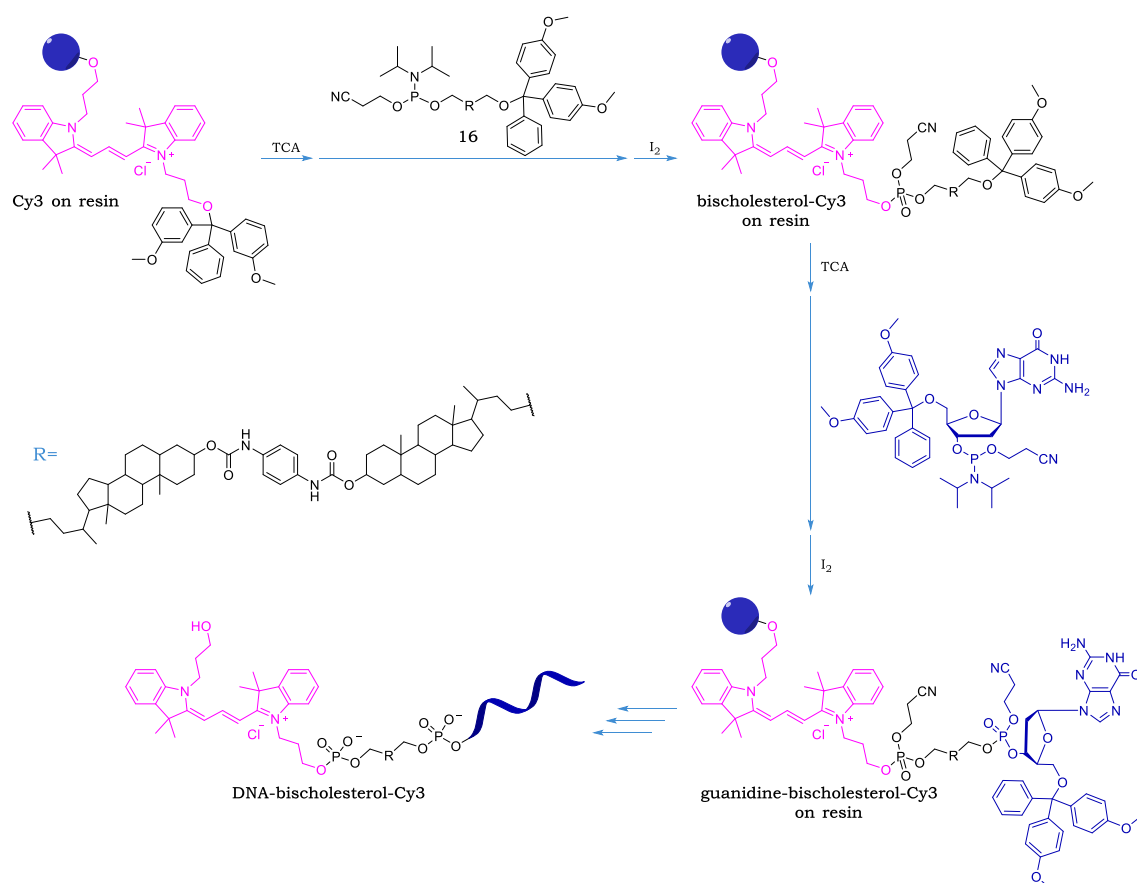
DMF content also appeared to have no effect, and a high content of water (>50%) resulted in precipitation of the bischolesterol dye azide. Final attempts used an (2x and 10x) excess of DNA instead which resulted in product formation, but the extent could not be determined by mass spectrometry. The conjugates TBA15-bischolesterol-Cy3, nina-bischolesterol-Cy3, TBA15-bischolesterol-Cy5, TBA29-bischolesterol-Cy5, and nina-bischolesterol-Cy5 were synthesised *via* this method, albeit in small amounts.

#### 4.7.3 Strategy 3- coupling bischolesterol dye phosphoramidite to DNA

Transmembrane molecules could not be produced following this strategy due to the failed synthesis of the bischolesterol dye phosphoramidites as discussed in section 4.5.3.

#### 4.7.4 Strategy 4- using the DNA synthesiser to couple components

DMT protected Cy3 and Cy5 molecules on solid supports were purchased from Link Technologies and placed on the DNA synthesiser for automated transmembrane molecule synthesis in the order of dye → bischolesterol → DNA *via* the 3' end. Commercially available Cy3 and Cy5 phosphoramidites could have been acquired for the synthesis of the conjugates in the reverse order (i.e. DNA → bischolesterol → dye) but were deemed too expensive for use.



**Scheme 4.14** Automated synthesis of DNA bischolesterol Cy3 conjugate.

Usually, phosphoramidites are dissolved in anhydrous acetonitrile for automated DNA synthesis but DMT protected bischolesterol phosphoramidite **16** was found to only be partially soluble in this eluent. Instead, **16** was dissolved in a solution of 1:1 anhydrous acetonitrile:DCM to a concentration of 0.1M and loaded onto the synthesiser alongside commercially supplied phosphoramidites and reagents required for automated synthesis. The conjugates were synthesised and deprotected under ultramild conditions. Trichloroacetic acid (TCA) was first introduced to the resin to deprotect the cyanine dyes. Compound **16** was then added along with tetrazole to activate coupling of the exposed alcohol on the cyanine dye with the phosphoramidite group of **16**. A coupling time of 10 minutes was used to ensure conjugation



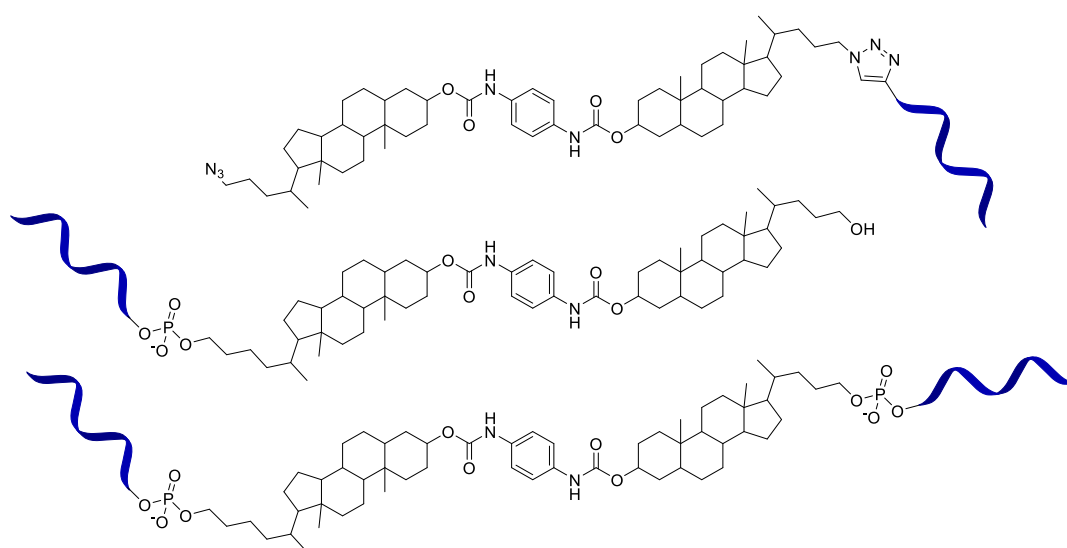
(typical coupling time for unmodified bases is 30 seconds). Any unreacted dye was capped with acetic anhydride and the phosphite linkage oxidised to phosphate with iodine. Trichloroacetic acid was then added to remove the DMT from the bischolesterol part of the now appended bischolesterol dye conjugate. The exposed alcohol was reacted with an activated nucleoside phosphoramidite, initiating the connection between the bischolesterol and the start of the DNA sequence. Additional nucleosides were followed to form the sequence as computed into the synthesiser program (Scheme 4.14).

As the removal of the DMT group is orange in colour, the coupling efficiency of bischolesterol to both dye and subsequent nucleoside can be monitored by the UV absorbance of the solution after detritylation. Coupling efficiencies were determined to be 100% between the cyanine dyes and bischolesterol, and 100% between bischolesterol and the first nucleoside, which looked promising. However, analysis of the crude solutions by mass spectrometry painted a different picture. The synthesiser was used to make 6 different transmembrane molecules: TBA15-bischolesterol-Cy3, TBA29-bischolesterol-Cy3, nina-bischolesterol-Cy3, TBA15-bischolesterol-Cy5, TBA29-bischolesterol-Cy5, and nina-bischolesterol-Cy5. Out of the 6 combinations, the desired product was only detected in three of the crude solutions (nina-bischolesterol-Cy3, nina-bischolesterol-Cy5, and TBA15 bischolesterol Cy5) and were present in low amounts despite the solutions being coloured. The mass of the major product in all solutions appeared to be that of the desired product minus 946 units, which was later determined to be the mass of the bischolesterol unit. The only explanation for this deletion is incomplete detritylation of the cyanine dyes prior to the addition of **16**. Had the dye been fully detritylated and failed to react with

**16**, then exposed alcohol groups on the dye should have reacted with acetic anhydride in the capping step of the process, effectively stopping them from reacting further. The high coupling efficiencies detected for the first coupling reactions could be attributed to DMT removal of the dyes and not the bischolesterol. It does seem odd however for the dyes to be detritylated on the second addition of TCA and not the first.

#### 4.7.5 Coupling bischolesterol to DNA

In addition to the synthesis of the transmembrane molecules *via* the described strategies, DNA-bischolesterol conjugates were also made. Alkyne modified DNA strands were reacted with bischolesterol diazide **10** *via* copper catalysed cycloaddition to form the conjugates TBA15-bischolesterol and nina-bischolesterol. DMT protected bischolesterol phosphoramidite **16** was also used to successfully insert the bischolesterol moiety into the middle and tag at the end of the nina sequence using automated DNA synthesis (Figure 4.18). Attaching bischolesterol to the end of the DNA strand using the synthesiser proved effective with coupling efficiencies of 92%.

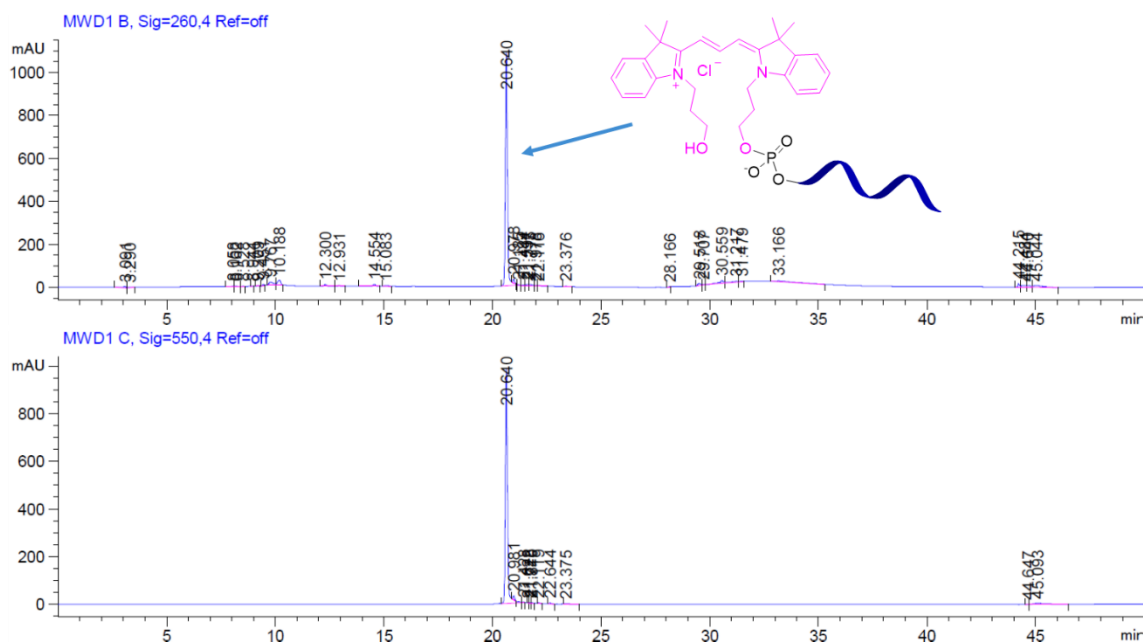


**Figure 4.18** DNA-bischolesterol conjugates synthesised.

#### 4.7.6 Purification of transmembrane molecules

Out of the four devised synthetic strategies, desired products were formed with strategies 1, 2 and 4. However, the products derived from strategies 2 and 4 were low yielding. DNA-bischolesterol-dye conjugates are inherently more lipophilic than their unmodified counterparts and therefore significantly longer retention times were anticipated with these molecules when utilising reversed phase HPLC for purification.

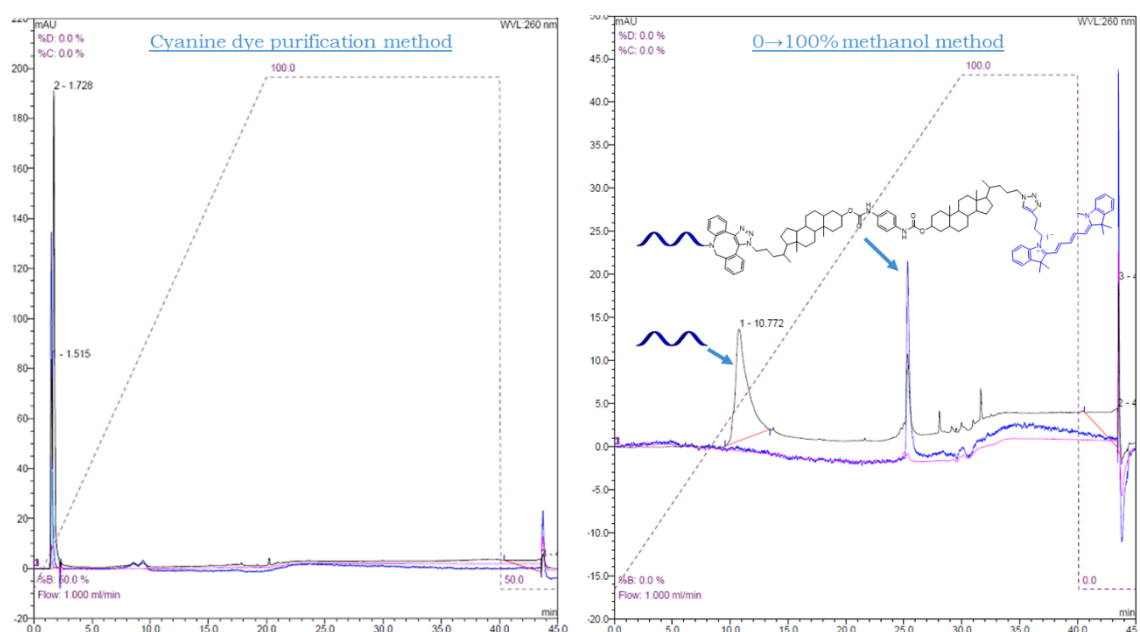
The first attempt at purification was carried out on transmembrane molecules made *via* strategy 4 on a Clarity Oligo semi prep column from Phenomenex using a 0.1M triethylammonium acetate (TEAA)/acetonitrile method adapted from the literature.<sup>38</sup> The samples were monitored at UV absorbances of 260nm for DNA, 550nm for Cy3 containing samples, and 645nm for Cy5 containing samples. Strong coinciding signals were observed at retention times of ~21mins (43% acetonitrile) for most samples, and fractions were manually collected at this timepoint. Analysis by mass spectrometry of the samples showed the peaks to correspond to DNA-dye conjugates (no bischolesterol, Figure 4.19). Dye absorbances were noticed in the acetonitrile flushes of successive blank and crude unmodified DNA samples. Small amounts of the product were collected, analysed by mass spectrometry, and found to contain masses matching desired DNA-bischolesterol-dye products. It appeared that the transmembrane molecules were retained on this particular column to an extent that they could only be removed under prolonged organic solvent conditions. Purification methods could not be further developed using strategy 4 derived transmembrane molecules as all samples had been used (and probably still remained) on the Clarity Oligo semi prep column.



**Figure 4.19** Crude HPLC chromatograms of strategy 4 derived nina-bischolesterol-Cy3 reaction on the Clarity oligo column at detector signals of 260nm (top) and 550nm (bottom).

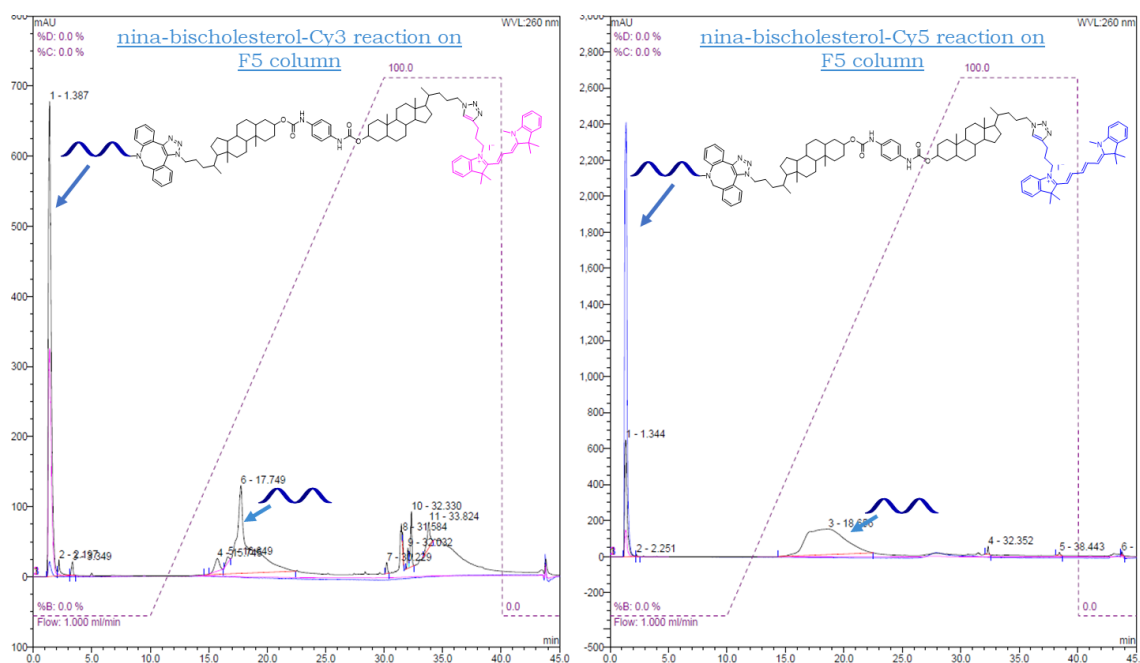
For transmembrane molecule reactions derived from strategies 1-2, samples were passed through a NAP-10 column from GE Healthcare to remove small molecules such as salts and unreacted bischolesterol dye azide prior to HPLC purification. The next set of samples to purify were DBCO derived DNA-bischolesterol-dye conjugates (strategy 2). Initial efforts repeated purification methods used on strategy 4 derived samples (Clarity oligo column, literature method). However, no dye signals were observed despite the solutions being coloured. The DBCO moiety makes these molecules even more hydrophobic than strategy 4 derived transmembrane molecules, leading to enhanced column retention. The method was then changed to one with a steeper acetonitrile gradient commonly used for DMT tagged oligonucleotides (DMT ON method). This method also failed to produce a dye signal from DBCO derived samples.

It was decided that the Clarity oligo column was unsuitable for the purification of transmembrane molecules and thus was replaced with a Kinetex C18 column from Phenomenex. This column had been previously used to purify cyanine dye and bischolesterol-dye conjugates using water/methanol methods and so appeared to be more appropriate for the desired application. Applying the cyanine dye purification method to the DBCO derived samples resulted in immediate elution of all components (Figure 4.20). Changing the method to run at a shallower gradient of 0 → 100% methanol over 30 minutes led to adequate separation from unreacted DNA with potential product peaks obtained at 25 minutes (~83% methanol). Unfortunately, there was significant carryover of this peak into subsequent blank injections, meaning prolonged retention of these conjugates on this column, as observed with the Clarity Oligo column.



**Figure 4.20** Different HPLC traces of the DBCO derived nina-bischolesterol-Cy5 reaction (strategy 2) on the Kinetex C18 column showing UV detector signals at 260nm (black), 550nm (pink) and 645nm (blue).

The next stationary phase to be utilised was an analytical F5 column from Phenomenex based on a pentafluorophenyl matrix. A 0 → 100% methanol gradient was applied, and the desired product was found to have no retention on the column, with complete elution observed under aqueous conditions (Figure 4.21). Fractions were collected at 1 and 18 minutes and were determined by mass spectrometry (negative mode ESI) to be the desired DNA-bischolesterol-dye conjugate and unreacted DNA respectively.



**Figure 4.21** HPLC traces of DBCO derived nina-bischolesterol-dye reactions (strategy 2) on the F5 column showing UV detector signals at 260nm (black), 550nm (pink) and 645nm (blue).

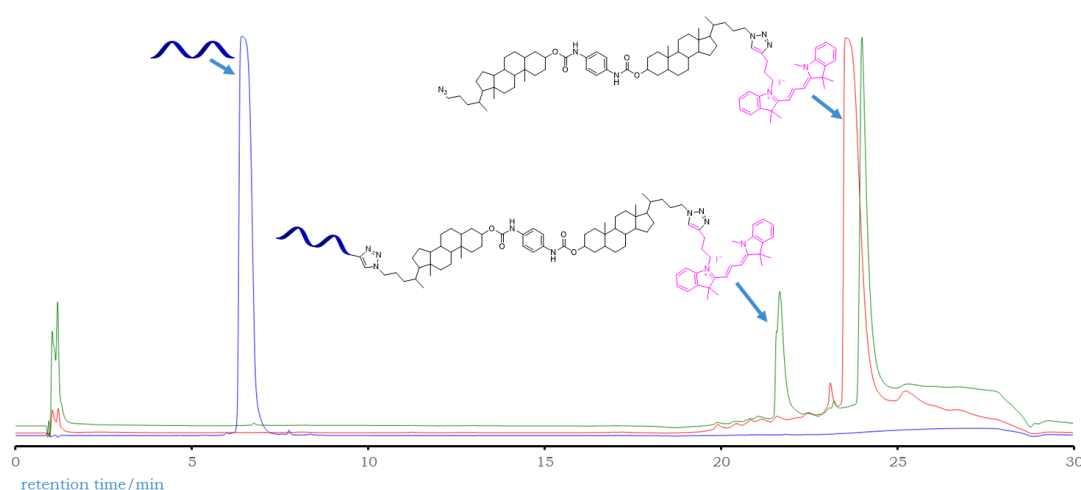
It had seemed that these coupling reactions could be successfully purified by reversed phase HPLC using the pentafluorophenyl column. However, analysis of pure fractions by positive mode ES identified masses corresponding to unreacted bischolesterol dye azides. The samples had been passed through a size exclusion (NAP-10) column prior to HPLC purification and it was assumed

that this process would remove any unreacted bischolesterol dye azide from the reaction. This now appeared to not be the case. Although the lack of retention of the transmembrane molecule to the F5 column allowed for separation of the product from any unreacted DNA, it seemed that the column was unable to separate the desired product from unreacted bischolesterol dye azide. On the other hand, this column proved more useful for purifying synthesiser derived 5' bischolesterol tagged DNA (Figure 4.18, middle), as all unreacted bischolesterol phosphoramidite had been washed from the resin prior to deprotection and HPLC purification.

Transmembrane molecule reactions derived from strategies 1 and 2 were then passed through an old C18 Ascentis column from Sigma Aldrich (>15 years), an old C18 Luna column from Phenomenex (>10 years), and a new C4 Jupiter column from Phenomenex in various eluent mixtures in efforts to reduce prolonged retention of the conjugates on these stationary phases. Unfortunately, significant carryover of product was observed and thus the columns were deemed unsuitable for the desired application.

One concern was that sample carryover may be attributed to product sticking to the injector needle of the instrument rather than the column. A quick experiment was performed to test this concern by changing columns after running a sample and then immediately running a blank. No signals were observed with the blank meaning the injector needle was not the cause of the product carryover.

An analytical reversed phase oligo column from Thermo Fisher was the final column to be utilised in the method development for the purification of the transmembrane molecules. Strategy 1 derived transmembrane molecules were purified using a 0.1M TEAA/acetonitrile method derived from the Thermo oligo RP column's handbook. Pleasingly, this column was able to isolate the desired product from both unreacted DNA and bischolesterol dye azide, with little to no carryover observed in subsequent samples (Figure 4.22). Further purification efforts should focus on this particular column to purify the conjugates at a semi preparative scale.



**Figure 4.22** Overlaid HPLC signals of starting materials and strategy 1 reaction with the Thermo Oligo RP column at a detector UV signal of 260nm. **Nina-alkyne** is shown in blue, bischolesterol Cy3 azide **23** is shown in red, and the nina-bischolesterol-Cy3 reaction (strategy 1) is shown in green.



## 4.8 Conclusions and future work

To conclude, out of the four devised strategies to construct the transmembrane molecules, the best approach appears to be *via* strategy 1 using a reversed phase oligo column from Thermo Fisher for purification. Only strategy 1 generated transmembrane molecules in sufficient yields and thus future efforts should focus on this strategy. Cycloaddition reactions could be attempted on solid supports on the DNA synthesiser prior to deprotection to allow for easier purification. An alternative synthetic strategy would be to revisit the construction of the bischolesterol-dye phosphoramidites (strategy 3) but change the process to circumvent the use of copper. One such proposal is to use bischolesterol dialcohol **8** and an alkyl halide modified cyanine dye to create an ether connection between the bischolesterol and dye rather than a triazole. Another synthetic strategy worth revisiting would be strategy 4 but performing the solid phase synthesis of the conjugate in the reverse order (DNA → bischolesterol → dye). It is believed that greater yields can be achieved by utilising automated DNA synthesis to construct the transmembrane molecules.

Future work should continue the purification of the transmembrane molecules on a preparative scale to isolate enough material to use in sensing experiments. Once sufficient amounts have been synthesised and isolated, the ability of the transmembrane molecules to insert into bilayers of liposomes in the preferred orientation should be studied using quenching experiments. If successful, then extravascular target-aptamer binding can be communicated inside vesicles by monitoring by FRET (as shown in Figure 4.5).

## 4.9 List of references

1. A. Patwa, A. Gissot, I. Bestel and P. Barthelemy, *Chem. Soc. Rev.*, 2011, **40**, 5844-5854.
2. R. L. Letsinger, G. R. Zhang, D. K. Sun, T. Ikeuchi and P. S. Sarin, *Proc Natl Acad Sci U S A*, 1989, **86**, 6553-6556.
3. S. M. Gryaznov and D. H. Lloyd, *Nucleic Acids Res.*, 1993, **21**, 5909-5915.
4. M. Acedo, G. Tarrason, J. Piulats, M. Mann, M. Wilm and R. Eritja, *Bioorg. Med. Chem. Lett.*, 1995, **5**, 1577-1580.
5. A. Guzaev and M. Manoharan, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 3671-3676.
6. J. M. Tomkins, K. J. Barnes, A. J. Blacker, W. J. Watkins and C. Abell, *Tetrahedron Lett.*, 1997, **38**, 691-694.
7. G. Godeau, C. Staedel and P. Barthélémy, *J. Med. Chem.*, 2008, **51**, 4374-4376.
8. C. Gosse, A. Boutorine, I. Aujard, M. Chami, A. Kononov, E. Cogné-Laage, J. F. Allemand, J. Li and L. Jullien, *J. Phys. Chem. B*, 2004, **108**, 6485-6497.
9. M. P. Thompson, M.-P. Chien, T.-H. Ku, A. M. Rush and N. C. Gianneschi, *Nano Lett.*, 2010, **10**, 2690-2693.
10. O. Ries, P. M. G. Löffler, A. Rabe, J. J. Malavan and S. Vogel, *Org. Biomol. Chem.*, 2017, **15**, 8936-8945.
11. M. Zhuojun, Y. Jian, L. Qing, d. V. J. Willem, G. Agnieszka, R. P. Alberto, C. B. J., K. Alexander and H. Andreas, *Chem. Eur. J.*, 2017, **23**, 9391-9396.

12. O. Ries, P. M. G. Löffler and S. Vogel, *Org. Biomol. Chem.*, 2015, **13**, 9673-9680.
13. L. P. M. G., R. Oliver, R. Alexander, O. A. H., T. R. P., K. Jørgen and V. Stefan, *Angew. Chem.*, 2017, **129**, 13410-13414.
14. G. Stengel, L. Simonsson, R. A. Campbell and F. Höök, *J. Phys. Chem. B*, 2008, **112**, 8264-8274.
15. B. van Lengerich, Robert J. Rawle, Poul M. Bendix and Steven G. Boxer, *Biophys. J.*, 2013, **105**, 409-419.
16. Y. H. Chan, B. van Lengerich and S. G. Boxer, *Biointerphases*, 2008, **3**, 17-21.
17. M. Banchelli, F. Betti, D. Berti, G. Caminati, F. B. Bombelli, T. Brown, L. M. Wilhelmsson, B. Nordén and P. Baglioni, *J. Phys. Chem. B*, 2008, **112**, 10942-10952.
18. C. Yoshina-Ishii and S. G. Boxer, *J. Am. Chem. Soc.*, 2003, **125**, 3696-3697.
19. H. Lv, S. Zhang, B. Wang, S. Cui and J. Yan, *J. Controlled Release*, 2006, **114**, 100-109.
20. C. Wolfrum, S. Shi, K. N. Jayaprakash, M. Jayaraman, G. Wang, R. K. Pandey, K. G. Rajeev, T. Nakayama, K. Charrise, E. M. Ndungo, T. Zimmermann, V. Kotliansky, M. Manoharan and M. Stoffel, *Nat. Biotechnol.*, 2007, **25**, 1149-1157.
21. M. Schade, D. Berti, D. Huster, A. Herrmann and A. Arbuzova, *Adv. Colloid Interface Sci.*, 2014, **208**, 235-251.
22. I. Pfeiffer and F. Höök, *J. Am. Chem. Soc.*, 2004, **126**, 10224-10225.
23. M. Banchelli, F. Gambinossi, A. Durand, G. Caminati, T. Brown, D. Berti and P. Baglioni, *J. Phys. Chem. B*, 2010, **114**, 7348-7358.

24. G. Zhang, F. Farooqui, O. Kinstler and R. L. Letsinger, *Tetrahedron Lett.*, 1996, **37**, 6243-6246.
25. Y. Kobuke and T. Nagatani, *J. Org. Chem.*, 2001, **66**, 5094-5101.
26. C. Goto, M. Yamamura, A. Satake and Y. Kobuke, *J. Am. Chem. Soc.*, 2001, **123**, 12152-12159.
27. S. Kumar, P. Bhargava, V. Sreekanth and A. Bajaj, *J. Colloid Interface Sci.*, 2015, **448**, 398-406.
28. L. Ma, M. Melegari, M. Colombini and J. T. Davis, *J. Am. Chem. Soc.*, 2008, **130**, 2938-2939.
29. N. Dave and J. Liu, *Chem. Commun.*, 2012, **48**, 3718-3720.
30. H. Finkelstein, *Ber. Dtsch. Chem. Ges.*, 1910, **43**, 1528-1532.
31. M. Gerowska, L. Hall, J. Richardson, M. Shelbourne and T. Brown, *Tetrahedron*, 2012, **68**, 857-864.
32. E. A. Owens, N. Bruschi, J. G. Tawney and M. Henary, *Dyes Pigm.*, 2015, **113**, 27-37.
33. S. E. Allen, R. R. Walvoord, R. Padilla-Salinas and M. C. Kozlowski, *Chem. Rev.*, 2013, **113**, 6234-6458.
34. J. Chen, J. Zhang, J. Li, H.-H. Yang, F. Fu and G. Chen, *Biosens. Bioelectron.*, 2010, **25**, 996-1000.
35. L. Xingfen, Y. Yonghong, H. Xiaoxiao, F. Xiaomiao, S. Shao, H. Yanqin, F. Quli, W. Lianhui and H. Wei, *Chin. J. Chem.*, 2015, **33**, 981-986.
36. X. Liu, L. Shi, X. Hua, Y. Huang, S. Su, Q. Fan, L. Wang and W. Huang, *ACS Appl. Mater. Interfaces*, 2014, **6**, 3406-3412.
37. S. I. Presolski, V. P. Hong and M. G. Finn, *Curr. Protoc. Chem. Biol.*, 2011, **3**, 153-162.
38. S. S. Pujari and F. Seela, *J. Org. Chem.*, 2012, **77**, 4460-4465.

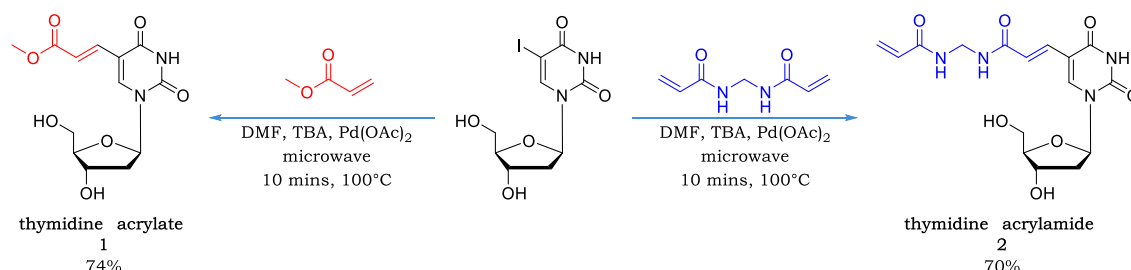
39. T. R. Chan, R. Hilgraf, K. B. Sharpless and V. V. Fokin, *Org. Lett.*, 2004, **6**, 2853-2855.
40. A. H. El-Sagheer and T. Brown, *Chem. Soc. Rev.*, 2010, **39**, 1388-1405.
41. F. E. Ziegler, K. W. Fowler, W. B. Rodgers and R. T. Wester, in *Ambient-Temperature Ullman Reaction: 4,5,4',5'-Tetramethoxy-1,1'-Biphenyl-2,2'-Dicarboxaldehyde in Organic Syntheses*, John Wiley & Sons, New Jersey, 2003.
42. J. Zhu, J. Hu and R. E. Marchant, in *Biomimetic Biomaterials*, Woodhead Publishing, Cambridge, 2013, ch. 9, pp. 238-275.

## Chapter 5 Experimental

### 5.1 Materials and methods

Reagents and solvents were purchased from commercial suppliers and used without further purification, unless otherwise stated. Column chromatography was carried out using open columns packed with Merck grade 60 silica gel topped with 0.5cm of sand. TLC analysis was performed on Merck silica gel 60 silica sheets.  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{31}\text{P}$  NMR spectra were obtained on Bruker AVIII300 or AVIII400 spectrometers. Chemical shifts ( $\delta$ ) are given in ppm and are relative to the residual solvent peak. Data was acquired using Bruker Topspin v3.2 and analysed with MestReNova v12.0.2-20910. Electrospray mass (ESI-MS) spectra were measured by either Waters micromass LCT electrospray time-of-flight (ES-TOF), Waters Xevo G2-XS, or Synapt G2S mass spectrometers. IR spectroscopy was performed on dry samples with a Varian 660-IR FT-IR spectrometer. A Stuart SMP10 melting point apparatus was used to determine melting points of compounds in open glass capillary tubes. Milli-Q water purified with a Millipore Elix-Gradient A10 system (resistivity  $> 18\mu\Omega\cdot\text{cm}$ , TOC  $\leq 5\text{ppb}$ , Millipore, France) was used for DNA sample preparation and in DNA containing and surface experiments. Bovine thrombin from Helena Biosciences was used in gel EMSA and ITC experiments. Polycrystalline gold substrates from George Albert PVD were utilised for surface coupling experiments and consisted of a 100nm gold layer deposited onto a silicon wafer, precoated with titanium as the adhesion layer.

## 5.2 Synthesis of thymidine acrylate and thymidine acrylamide



**Scheme 5.1** Synthesis of thymidine acrylate **1** and thymidine acrylamide **2**.

*Adapted from Ami and Fujimoto.<sup>1</sup>*

5-iodo-2'-deoxyuridine (1.00g, 2.82mmol) and palladium acetate (60mg, 0.28mmol) were suspended in DMF in a 10ml microwave tube equipped with a small magnetic stirring bar. Tributylamine (0.67ml, 2.82mmol) and the corresponding reagent, either methyl acrylate (0.38ml, 4.24mmol) or N,N'-methylenebisacrylamide (1.09g, 7.06mmol), were then added. The suspension was stirred and degassed with argon for 10 minutes. The tube was then tightly sealed with an aluminium/Teflon crimp top, and the mixture irradiated in a microwave for 10 minutes at 100°C. After the irradiation period, the reaction vessel was rapidly cooled to 50°C by gas jet cooling before opening. Reaction completion was verified by TLC. The reaction was then filtered and evaporated under vacuum. The crude was purified by recrystallization from chloroform.

**Thymidine acrylate** was synthesised in accordance to the general procedure. Product was obtained as a white solid (652mg, 74%). <sup>1</sup>H NMR (400MHz, CD<sub>3</sub>CN) δ 9.13 (s, 1H), 8.29 (s, 1H), 7.33 (d, J= 15.7Hz, 1H), 6.85 (d, J= 15.8Hz, 1H), 6.15 (t, J= 6.3Hz, 1H), 4.38-4.34 (m, 1H), 3.89-3.87 (m, 1H), 3.81-3.76 (m, 1H),

3.73-3.67 (m, 4H), 3.36 (d,  $J = 4.4\text{Hz}$ , 1H), 3.29 (t,  $J = 5.1\text{Hz}$ , 1H), 2.30-2.17 (m, 2H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CD}_3\text{CN}$ )  $\delta$  168.5 (C=O), 162.4 (C=O), 150.2 (C=O), 144.5 (HC=C), 138.2 (HC=C), 118.1 (HC=C), 109.7 (C), 88.6 ( $\text{CH}_2$ ), 86.6 (CH), 71.1 (CH), 62.0 ( $\text{CH}_2$ ), 52.0 ( $\text{CH}_3$ ), 41.4 ( $\text{CH}_2$ ). MS (ES +ve) ( $m/z$ ): 335.2  $[\text{M}+\text{Na}]^+$

Thymidine acrylamide was synthesised in accordance to the general procedure, yielding an off white solid (748mg, 70%).  $^1\text{H}$  NMR (400MHz,  $(\text{CD}_3)_2\text{SO}$ )  $\delta$  11.56 (s, 1H), 8.69 (t,  $J = 5.9\text{Hz}$ , 2H), 8.32 (s, 1H), 7.18 (d,  $J = 15.6\text{Hz}$ , 1H), 6.96 (d,  $J = 15.6\text{Hz}$ , 1H), 6.30-6.21 (m, 1H), 6.16-6.09 (m, 2H), 5.64-5.58 (m, 1H), 5.27 (d,  $J = 4.3\text{Hz}$ , 1H), 5.16 (t,  $J = 5.0\text{Hz}$ , 1H), 4.57-4.52 (m, 2H), 4.29-4.23 (m, 1H), 3.82-3.78 (m, 1H), 3.70-3.55 (m, 1H), 2.18-2.13 (m, 2H).  $^{13}\text{C}$  NMR (100MHz,  $(\text{CD}_3)_2\text{SO}$ )  $\delta$  166.0 (C=O), 164.8 (C=O), 161.8 (C=O), 149.4 (C=O), 142.4 (HC=C), 132.9 (HC=C), 131.5 (HC=C), 125.8 ( $\text{H}_2\text{C}=\text{C}$ ), 120.8 (HC=C), 108.9 (C), 87.6 (CH), 84.6 (CH), 69.9 (CH), 60.9 ( $\text{CH}_2$ ), 43.6 ( $\text{CH}_2$ ), 40.0 ( $\text{CH}_2$ ). HRMS (ES +ve) ( $m/z$ ):  $[\text{M}+\text{Na}]^+$  calcd for  $\text{C}_{16}\text{H}_{20}\text{N}_4\text{O}_7\text{Na}$ , 403.1230; found 403.1229



### 5.3 $^1\text{H}$ NMR homo- and co- polymerisation experiments with thymidine acrylate and thymidine acrylamide

#### 5.3.1 Homopolymerisation experiments

Two 1ml deuterated DMSO solutions were prepared for timed  $^1\text{H}$  NMR analysis; one for each base. Each solution contained 0.25M of modified base, 0.09M APS, 0.67M TEMED, and 0.25M TMS. The solutions were transferred to NMR tubes fitted with Teflon sealed screw caps. The tubes were then sonicated and  $^1\text{H}$  NMR measurements taken after 0, 1, 2, 8, 24, 48, and 60 hours of sonication.

#### 5.3.2 Copolymerisation experiments

Two 1ml deuterated DMSO solutions were prepared for timed  $^1\text{H}$  NMR analysis; one for each base. Each solution contained 0.25M of modified base, 0.25M phenyl acrylamide, 0.09M APS, 0.67M TEMED, and 0.25M TMS. The solutions were transferred to NMR tubes fitted with Teflon sealed screw caps. The tubes were then sonicated and  $^1\text{H}$  NMR measurements taken after 0, 1, 2, 8, 24, 48, and 60 hours of sonication.

## 5.4 Contact angle experiments with thymidine acrylate and thymidine acrylamide

### 5.4.1 Fabrication of acrylamide SAMs

Polycrystalline gold substrates were cut into 1cm x 1cm chips and used for surface experiments. To clean the surfaces, the 1cm x 1cm gold chips were immersed in acidic piranha solution (7:3 H<sub>2</sub>SO<sub>4</sub>:H<sub>2</sub>O<sub>2</sub>) at room temperature for 10 minutes (**caution:** *piranha reacts violently with all organic compounds and should be handled with care*). The chips were then immediately rinsed with Milli-Q water (50ml) followed by methanol (50ml), before submerging into methanolic solutions of N,N bis(acryloyl)cystamine (1mM, 1ml). The chip containing solutions were then shaken for 24 hours.

### 5.4.2 Copolymerisation of acrylamide SAMs with modified bases

A number of 3ml Milli-Q water solutions were prepared, one for each chip. Each solution contained 10% v/v DMF, 1mM modified base, 17mM APS, and 1% v/v TEMED. Control solutions substituting the modified base for an unmodified base (thymidine) were also prepared. After 24 hours, gold chips were removed from the N,N-bis(acryloyl)cystamine solutions, thoroughly rinsed with methanol (50ml) followed by water (50ml), and immediately transferred to the different 3ml solutions. The solutions were then left to shake for 24 hours. Afterwards, the chips were removed from the solutions, rinsed with Milli-Q water (100ml), and dried with argon.

### 5.4.3 Contact angle measurements

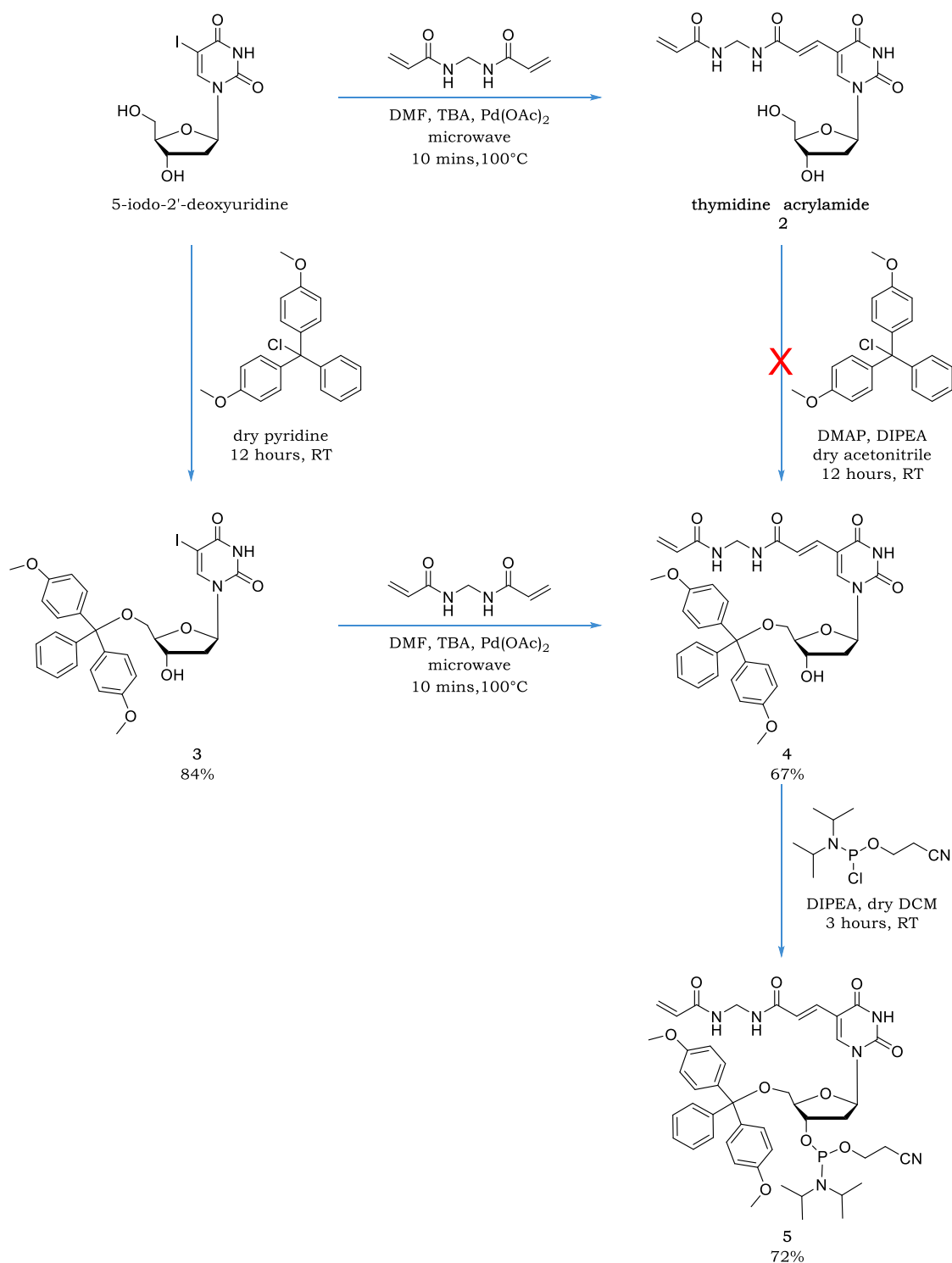
Contact angle measurements were conducted on an Attension Theta contact angle meter from Biolin Scientific. The dynamic contact angles were recorded as an automated micro-syringe was used to add a drop of Milli-Q water to the gold chip to give advancing contact angle data, and the same automated micro-syringe was used to remove liquid from the drop to give receding contact angle data. The drop was shown as a live video image with an acquisition rate of 32 frames per second. OneAttension software was used to analyse the contact angle of the drop at the three-phase intersection. The averages and standard deviations of contact angles were determined from six different measurements made for each condition.

## 5.5 Testing the stability of thymidine acrylamide towards the oligonucleotide deprotection process

**For standard deprotection conditions:** Thymidine acrylamide (10mg) was added to a 1ml solution of aqueous ammonia (30%). The resulting suspension was sealed, shaken, and heated to 60°C for 6 hours. The solvent was then evaporated, and the crude dissolved in deuterated DMSO and analysed by  $^1\text{H}$ ,  $^{13}\text{C}$ , and 2D NMR. The solid crude was also analysed by mass spectrometry.

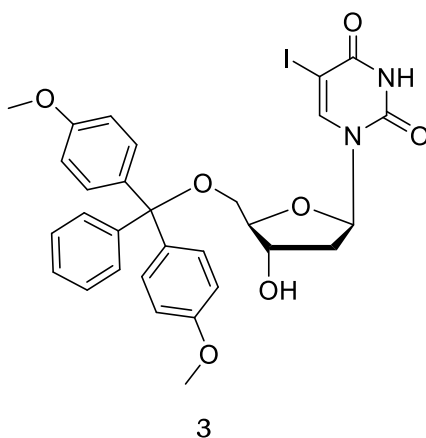
**For ultramild deprotection conditions:** Thymidine acrylamide (10mg) was added to a 1ml freshly prepared solution of potassium carbonate in methanol (0.05M). The resulting suspension was sealed, shaken, and left standing at room temperature for 12 hours. The solvent was then evaporated, and the crude dissolved in deuterated DMSO, and analysed by  $^1\text{H}$  NMR.

## 5.6 Synthesis of thymidine acrylamide derivatives



**Scheme 5.2** Synthesis of thymidine acrylamide derivatives.

### 5.6.1 Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine



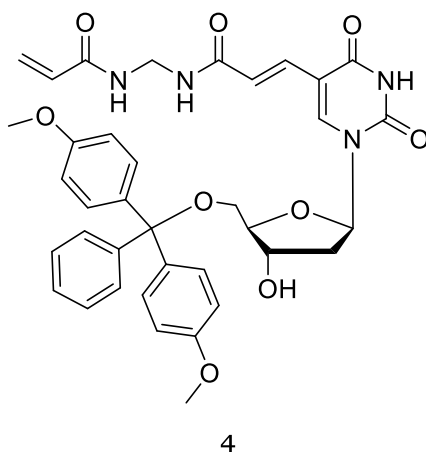
*Adapted from Vasseur<sup>2</sup>*

A solution of 5-iodo-2'-deoxyuridine (1.5g, 4.2mmol) in dry pyridine (25ml) was stirred in a Schlenk tube under argon. 4,4'-dimethoxytrityl chloride (1.9g, 5.5mmol) was added, and the reaction was left stirring under argon overnight. The solution was neutralised with saturated sodium bicarbonate solution (100ml) and the product was extracted with DCM (3 x 50ml). The organic phases were collected, dried over magnesium sulphate, and evaporated under vacuum (co-evaporating with 3 x 100ml toluene to remove pyridine). The crude product was purified by flash chromatography with an eluent of 9:1 DCM:7N ammonia in methanol. The appropriate fractions were combined and evaporated to give a white fluffy solid (2.34g, 84%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 9.58 (s, 1H), 8.15 (s, 1H), 7.42-7.40 (m, 2H), 7.34-7.25 (m, 6H), 7.23-7.19 (m, 1H), 6.84 (d, J = 8.9Hz, 4H), 6.34-6.31 (m, 1H), 4.56-4.54 (m, 1H), 4.13-4.11 (m, 1H), 3.77 (s, 6H), 3.41-3.33 (m, 2H), 2.55-2.49 (m, 1H), 2.32-2.25 (m, 1H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 160.5 (C=O), 158.7 (Ar-C), 150.4 (C=O), 144.5 (C=CH), 144.5 (Ar-C), 135.6 (Ar-C), 135.5 (Ar-C), 130.2 (Ar-CH), 128.2 (Ar-CH), 127.1 (Ar-CH), 113.5 (Ar-CH), 87.1 (C), 86.8 (CH), 85.9 (CH),

72.6 (CH), 69.0 (C=Cl), 63.7 (CH<sub>2</sub>), 55.4 (OCH<sub>3</sub>), 41.5 (CH<sub>2</sub>). TOF-MS (ES +ve) (*m/z*): 679.1 [M+Na]<sup>+</sup>.

#### 5.6.2 Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine



*Adapted from Ami and Fujimoto<sup>1</sup>*

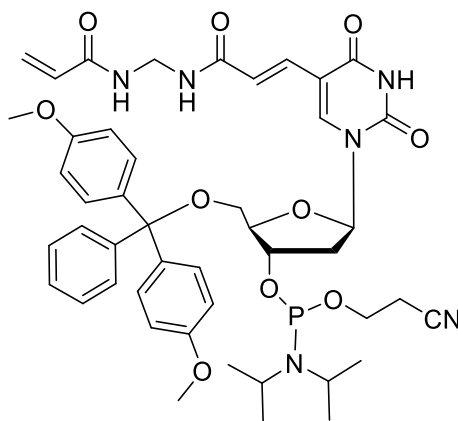
5'-O-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine **3** (1.01g, 1.54mmol) and palladium acetate (35mg, 0.16mmol) were suspended in DMF (3ml) in a 10ml microwave tube equipped with a small magnetic stirring bar. Tributylamine (0.37ml, 1.56mmol) and N,N'-methylenebis(acrylamide) (475mg, 3.08mmol) were then added. The suspension was stirred and degassed with argon for 10 minutes. The tube was tightly sealed with an aluminium/Teflon crimp top, and the mixture irradiated in a microwave for 10 minutes at 100°C. After the irradiation period, the reaction vessel was rapidly cooled to 50°C by gas jet cooling before opening. The reaction was then filtered through Celite®, rinsing with DCM (50ml), and the filtrate washed with water (50ml). The organic phase was dried over magnesium sulphate and evaporated under vacuum. The crude product was purified by flash chromatography with an eluent of 9:1 DCM:7N

ammonia in methanol. The appropriate fractions were combined and evaporated to give an off white solid (700mg, 67%).

$^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  10.87 (s, 1H), 8.14 (s, 1H), 7.86 (s, 1H), 7.39-7.37 (m, 3H), 7.29-7.25 (m, 7H), 7.20-7.14 (m, 2H), 6.84 (d,  $J$ = 8.7Hz, 4H), 6.36-6.28 (m, 2H), 6.14-6.08 (m, 1H), 5.66 (d,  $J$ = 11.1Hz, 1H), 4.75 (t,  $J$ = 6.6Hz, 2H), 4.46 (s, 1H), 4.24-4.21 (m, 1H), 3.74 (s, 6H), 3.44-3.31 (m, 2H), 2.70-2.66 (m, 1H), 2.23-2.17 (m, 1H), 1.81 (s, 1H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  168.1 (C=O), 166.8 (C=O), 162.2 (C=O), 158.8 (Ar-C), 149.6 (C=O), 144.4 (Ar-C), 142.2 (HC=C), 135.7 (Ar-C), 135.5 (Ar-C), 133.9 (HC=C), 130.3 (HC=C), 130.2 (Ar-CH), 130.0 (Ar-CH), 128.3 ( $\text{H}_2\text{C}=\text{C}$ ), 128.2 (Ar-CH), 127.2 (Ar-CH), 121.9 (HC=C), 113.5 (Ar-CH), 110.1 (C), 87.2 (C), 87.0 (CH), 86.8 (CH), 72.8 (CH), 64.0 ( $\text{CH}_2$ ), 55.4 ( $\text{OCH}_3$ ), 44.5 ( $\text{CH}_2$ ), 41.6 ( $\text{CH}_2$ ). HRMS (ES +ve) ( $m/z$ ):  $[\text{M}+\text{H}]^+$  calcd for  $\text{C}_{37}\text{H}_{39}\text{N}_4\text{O}_9$ , 683.2717; found 683.2715. IR neat ( $\text{cm}^{-1}$ ): 3483 (m, OH), 2921 (s, CH, aliphatic), 2851 (m, CH, aliphatic), 1608 (m, C=O, amide), 1508 (s, C=C, aromatic).



5.6.3 Synthesis of 5'-O-(4,4'-dimethoxytrityl)-3'-O-[2-cyano ethoxy-(N,N'-diisopropylamino)-phosphino]-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine



5

5'-O-(4,4'-dimethoxytrityl)-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine **4** (771mg, 1.13mmol) was dried *via* azeotroping with dry DCM (3 x 10ml). The solid was then redissolved in dry DCM (15ml), and the solution stirred under argon. DIPEA (0.49ml, 2.84mmol) was added *via* syringe to the solution. Then 2-cyanoethyl N,N-diisopropylchlorophosphoramidite (0.3ml, 1.34mmol) was cautiously added to the mixture. The solution was left to stir under argon until reaction completion was confirmed by TLC (~3 hours). The reaction was diluted with degassed DCM (50ml), and the solution was washed with degassed saturated sodium bicarbonate solution (2 x 50ml). The organic phases were collected, dried over sodium sulphate, and evaporated under vacuum. The crude product was purified by flash chromatography with an eluent of 9:1 DCM:7N ammonia in methanol. The appropriate fractions were combined and evaporated to give an off white solid (717mg, 72%).

$^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  8.65-8.59 (m, 1H), 7.79 (d,  $J$  = 21.1Hz, 1H), 7.43-7.40 (m, 2H), 7.33-7.27 (m, 8H), 7.23-7.18 (m, 1H), 7.07 (dd,  $J$  = 15.3, 2.7Hz, 1H), 6.87-6.83 (m, 4H), 6.44 (dd,  $J$  = 17.0, 1.1Hz, 1H), 6.28-6.23 (m, 1H), 6.16-6.09 (m, 1H), 5.69 (dd,  $J$  = 10.4, 1.0Hz, 1H), 4.73 (t,  $J$  = 6.0Hz, 2H), 4.56-4.49 (m, 1H), 4.28-4.22 (m, 1H), 3.87-3.52 (m, 11H), 3.38-3.29 (m, 2H), 2.72-2.59 (m, 2H), 2.45 (t,  $J$  = 6.4Hz, 1H), 2.28-2.17 (m, 1H), 1.18-1.08 (m, 12H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  167.8 (C=O), 166.6 (C=O), 162.8 (C=O), 158.6 (Ar-C), 149.1 (C=O), 144.3 (Ar-C), 142.1 (HC=C), 135.6 (Ar-C), 135.4 (Ar-C), 133.5 (HC=C), 130.1 (HC=C), 130.0 (Ar-CH), 128.4 ( $\text{H}_2\text{C}=\text{C}$ ), 128.1 (Ar-CH), 128.0 (Ar-CH), 127.1 (Ar-CH), 122.0 (HC=C), 117.5/117.4 (CN), 113.4 (Ar-CH), 109.9/109.8 (C), 86.8 (C), 86.3-86.0 (CH), 85.7 (CH), 74.0-73.5 (CH), 63.4/63.3 ( $\text{CH}_2$ ), 58.4/58.2 ( $\text{CH}_2$ ), 55.2 ( $\text{OCH}_3$ ), 44.6 ( $\text{CH}_2$ ), 43.4-43.2 (CH), 40.4 ( $\text{CH}_2$ ), 24.6 ( $\text{CH}_3$ ), 20.4-20.2 ( $\text{CH}_2$ ).  $^{31}\text{P}$  NMR (121MHz,  $\text{CDCl}_3$ )  $\delta$  149.1, 148.6. HRMS (TOF-ES +ve) ( $m/z$ ):  $[\text{M}]^+$  calcd for  $\text{C}_{46}\text{H}_{56}\text{N}_6\text{O}_{10}\text{P}$ , 883.3796; found 883.3793.

## 5.7 Synthesis and purification of oligonucleotides

All oligonucleotides were synthesised using solid phase synthesis on an Applied Biosystems ABI 394 DNA/RNA synthesiser using commercially supplied DNA synthesis grade solvents and reagents.

### 5.7.1 Standard synthesis of unmodified and alkyne modified strands

Standard phosphoramidites of Bz-dA, iBu-dG, Ac-dC, and dT from Link Technologies, and 5'-hexynyl phosphoramidite from Glen Research were used for synthesis. The phosphoramidites were dissolved in anhydrous acetonitrile to 0.1M prior to synthesis. Strands were synthesised at a 1 $\mu$ mol scale on SynBase™ CPG 1000/110 solid supports from Link Technologies. The resins have average pore sizes of 1000Å, nominal particle sizes of 110 $\mu$ m, and nucleoside loadings of 25-40 $\mu$ mol/g. Phosphoramidites were activated with 5-ethylthio-1H-tetrazole (0.5M) in acetonitrile prior to coupling. Coupling times of 25 seconds and 10 minutes were used for the nucleosides and hexynyl group respectively. Acetic anhydride and methylimidazole were added to cap unreacted material, and then iodine (0.1M) in THF/pyridine/water (78:20:2) was added to oxidise the phosphotriester formed. Upon sequence completion, oligonucleotides were treated with aqueous ammonia (30%) for 1 hour to cleave the strands from the resin. Protecting groups on the strands were removed by heating in aqueous ammonia (30%) at 60°C for 6 hours. The solvent was then removed on a Thermo Scientific speed vac in preparation for purification.

### 5.7.2 Ultramild synthesis of acrylamide, DBCO, and bischolesterol modified strands, and strategy 4 derived transmembrane molecules

Standard phosphoramidites of Pac-dA, iPr-Pac-dG, Ac-dC, dT from Link Technologies, DMT protected Cy3 and Cy5 cyanine dyes on CPG supports from Link Technologies, 5'-DBCO-TEG phosphoramidite from Glen Research, and phosphoramidite compounds **5** and **16** were used for ultramild synthesis. The phosphoramidites were dissolved in anhydrous acetonitrile to 0.1M prior to synthesis bar from **16** which was dissolved to the same concentration in 1:1 anhydrous acetonitrile:DCM. Strands were synthesised at a 1µmol scale on SynBase™ CPG 1000/110 solid supports from Link Technologies. Phosphoramidites were activated with 5-ethylthio-1H-tetrazole (0.25M) in acetonitrile prior to coupling. Coupling times of 25 seconds was used for the nucleoside phosphoramidites, and 10 minutes for phosphoramidites **5**, **16**, and DBCO-TEG. Then, phenoxyacetic anhydride and methylimidazole were added to cap unreacted material, and iodine (0.02M) in THF/pyridine/water (7:2:1) was added to oxidise the phosphotriester formed. Upon sequence completion, the resins were placed in freshly prepared 1ml solutions of potassium carbonate (0.05M) in methanol and left overnight to cleave strands from the resin and remove protecting groups. The solutions were neutralised with acetic acid (6µl) and the solvent was removed on a Thermo Scientific speed vac. The dried powders were redissolved in 1ml Milli-Q water and desalted with a NAP-10 column from GE Healthcare to remove residual resin and potassium carbonate. The solutions were then concentrated to 1ml and stored in the freezer for purification.

### 5.7.3 Purification of unmodified, acrylamide modified, and alkyne modified strands

Semi preparative HPLC purification was performed on an Agilent Technologies 1260 Infinity system using a Phenomenex Clarity 5µm Oligo-RP LC 250x10mm column. 1ml of sample was injected with a run time of 45 minutes for each sample, at a flow rate of 3ml/min. The column was heated to 60°C prior to sample injection for modified and unmodified thrombin binding aptamer sequences. The UV/vis absorbance of each run was monitored at 260nm. The solvent gradients used are listed in the table below:

**Table 5.1** Solvent gradients used in the HPLC purification of oligonucleotides.

Time/mins	0.1M TEAA in HPLC water/%	acetonitrile/%
0	95	5
30	82	18
30.1	0	100
40	0	100
40.1	95	5
45	95	5

Collected fractions were evaporated to dryness, diluted to 1ml in Milli-Q water, and desalted using a NAP-10 column (GE Healthcare), whilst eluting to 1.5ml. Purity of oligonucleotides was determined by analytical HPLC using a Phenomenex Clarity 5µm Oligo RP LC 250x4.6mm column on either a Shimadzu HPLC or an Agilent Technologies 1260 Infinity system. The column was heated to 60°C prior to sample injection for modified and unmodified thrombin binding aptamer strands. 20µl of sample was injected with a run time of 45 minutes for each sample, at a flow rate of 1ml/min. Solvent gradients used were identical to semi preparative HPLC (Table 5.1). The UV/vis absorbance of each run was monitored at 260nm.

#### 5.7.4 Purification of DCBO modified strands

DBCO modified strands were purified and analysed with the same columns and systems as previous strands, with the exception of a different solvent gradient being applied to account for the enhanced lipophilicity of the strands (Table 5.2).

**Table 5.2** Solvent gradients used in the HPLC purification of DBCO modified oligonucleotides.

Time/mins	0.1M TEAA in HPLC water/%	acetonitrile/%
0	95	5
15	76	24
20	57	43
25	57	43
25.1	0	100
35	0	100
35.1	95	5
41	95	5

#### 5.7.5 Oligonucleotide characterisation

Samples showing >95% purity by analytical HPLC were deemed sufficiently pure for use in experiments. Samples showing <95% purity were repurified by semi preparative HPLC. The characterisation of pure oligonucleotide samples was performed by negative mode electrospray mass spectrometry. Sample concentrations were determined by optical density at 260nm using a BioSpec-nano micro-volume UV-Vis spectrophotometer (nanodrop) from Shimadzu and the Beer Lambert law, with extinction coefficients obtained from Integrated DNA Technologies' OligoAnalyzer.

## 5.8 Polyacrylamide gel experiments

Experiments were performed on 12% native polyacrylamide gels with 1x TBE buffer and 10mM potassium chloride, using 1x TBE buffer with 10mM potassium chloride as a running buffer. Gels were run on Bio-rad Mini-PROTEAN® gel kits with a Bio-rad PowerPac (highest voltage: 5000V/500mA/400W). After electrophoresis, gels were stained with Diamond™ nucleic acid dye and visualised under UV with an AlphaImager HP gel imager from Alpha Innotech. For thrombin binding experiments, gels were further stained with Imperial protein stain from Thermo Fisher Scientific.

### 5.8.1 DNA-acrylamide gel copolymerisation experiments

100µl aqueous samples containing 1µM DNA, 12% acrylamide, 1x TBE buffer, and 10mM potassium chloride were prepared. 10% APS in water (1µl) and TEMED (1µl) were added to samples, and 40µl of each sample was immediately loaded into two adjacent wells of a gel (20µl per well). A three well gap was left between unmodified and acrylamide modified DNA as the solutions were found to react into neighbouring wells. The solution containing wells were then left to set (typically 5 minutes). Once set, 10µl of 12% polyacrylamide gel solution was added to the same wells and left to set. The gels were then rinsed with HPLC water and run at 100V for 1 hour.

### 5.8.2 Gel electromobility shift assay of unmodified and acrylamide modified thrombin binding aptamer strands with thrombin

50µl aqueous solutions of DNA (0.5µM) and thrombin (0µM, 0.01µM, 0.1µM, 0.5µM, or 1µM) in KCl (10mM), tris.HCl buffer (10mM, pH 7.5), and glycerol (3%)

were prepared. 10 $\mu$ l of each sample was loaded into wells and gels were run at 100V for 1 hour.

### 5.8.3 Gel electromobility shift assay of full and split thrombin binding aptamer strands with thrombin

50 $\mu$ l aqueous solutions of DNA (1 $\mu$ M) and thrombin (1 $\mu$ M) in KCl (10mM), tris.HCl buffer (10mM, pH 7.5), and glycerol (3%) were prepared. 10 $\mu$ l of each sample was loaded into wells and gels were run at 100V for 1 hour.



## 5.9 Circular dichroism of unmodified and acrylamide modified thrombin binding aptamer strands

500µl aqueous solutions of DNA (5µM) in KCl (10mM) and tris.HCl buffer (10mM, pH 7.5) were prepared. Prior to CD analysis, the samples were heated to 95°C for 5 minutes and cooled slowly to room temperature. CD spectra of samples were recorded on a Jasco J-810 spectropolarimeter, scanning from 350-200nm at a rate of 100nm/min. 3 accumulations were performed for each sample and the data produced an average of the three scans. A baseline correction was manually performed on each sample by subtracting the blank and offsetting results at 350nm.

### 5.10 Thermal melting experiments of unmodified and acrylamide modified thrombin binding aptamer strands with complementary sequence

100µl aqueous solutions of aptamer (10µM) and complementary DNA (10µM) in NaCl (100mM) and sodium phosphate buffer (10mM, pH 7) were prepared. 10µl of each sample was withdrawn, added to 10µl of SsoAdvanced™ universal SYBR® green supermix from Bio-rad, and the thermal melting of the resulting solutions was performed on a M550 double beam scanning UV/visible spectrophotometer. Samples were heated from 15°C → 90°C and cooled from 90°C → 20°C at a rate of 1°C/min, and values were obtained from the maxima of the negative first derivative of the melting curve.

### 5.11 ITC experiments on the binding of full and split thrombin binding aptamer strands with thrombin

ITC experiments were performed at the University of Cambridge on a Malvern MicroCal PEAQ-ITC at 25°C. 300µl of DNA (1µM) in binding buffer (10mM tris HCl, pH 7.5 and 10mM KCl) was used in the cell for each experiment. 60µl of thrombin (10µM) in binding buffer was placed in the syringe for each experiment.

ITC injection conditions are as follows:

- reference power: 5µcal s<sup>-1</sup>
- stir speed: 750rpm
- initial delay: 60 seconds
- initial injection volume: 0.4µl
- initial injection duration: 0.8 seconds
- injection volume: 2µl
- injection duration: 4 seconds
- injection spacing: 120 seconds
- number of injections: 20

## 5.12 Surface coupling experiments with unmodified and acrylamide modified thrombin binding aptamer strands

See 5.4.1 for the fabrication of acrylamide SAMs on gold chips.

### 5.12.1 Reaction of acrylamide SAMs with unmodified/acrylamide modified aptamer strands at high APS concentrations

A number of 2ml Milli-Q water solutions were prepared, one for each chip. Each solution contained 1 $\mu$ M of either unmodified or acrylamide modified DNA, 17mM APS, and 0.1% v/v TEMED. Control solutions excluding either or both DNA and APS/TEMED were also prepared. After 24 hours, gold chips were removed from the N,N-bis(acryloyl)cystamine solutions, thoroughly rinsed with methanol (50ml) followed by water (50ml), and immediately transferred to the different 2ml solutions. The solutions were then left to shake for 24 hours. Afterwards, the chips were removed from the solutions, rinsed with Milli-Q water (100ml), and dried with argon.

### 5.12.2 Reaction of acrylamide SAMs with unmodified/acrylamide modified aptamer strands at low APS concentrations

Conditions were identical to 5.12.1 with the exception of 1mM APS being used instead of 17mM.

### 5.12.3 Contact angle measurements

See 5.4.3 for details on contact angle measurements. The averages and standard deviations of contact angles were determined from six different measurements made for each condition.

#### [5.12.4 Ellipsometry measurements](#)

For high APS samples, ellipsometry measurements were conducted on a UVISSEL spectroscopic ellipsometer from Horiba and data was processed with DeltaPsi 2 software. For low APS samples, ellipsometry measurements were conducted on an Alpha-SE ellipsometer from J.A. Woollam and data was processed with CompleteEASE software. All samples were measured at an angle of incidence of 70° and film thickness calculations were based on a three-phase ambient/SAM/Au model in which the SAM was assumed to be isotropic and assigned a refractive index of 1.5. Averages and standard deviations were determined from six different measurements made for each condition.

#### [5.12.5 XPS measurements](#)

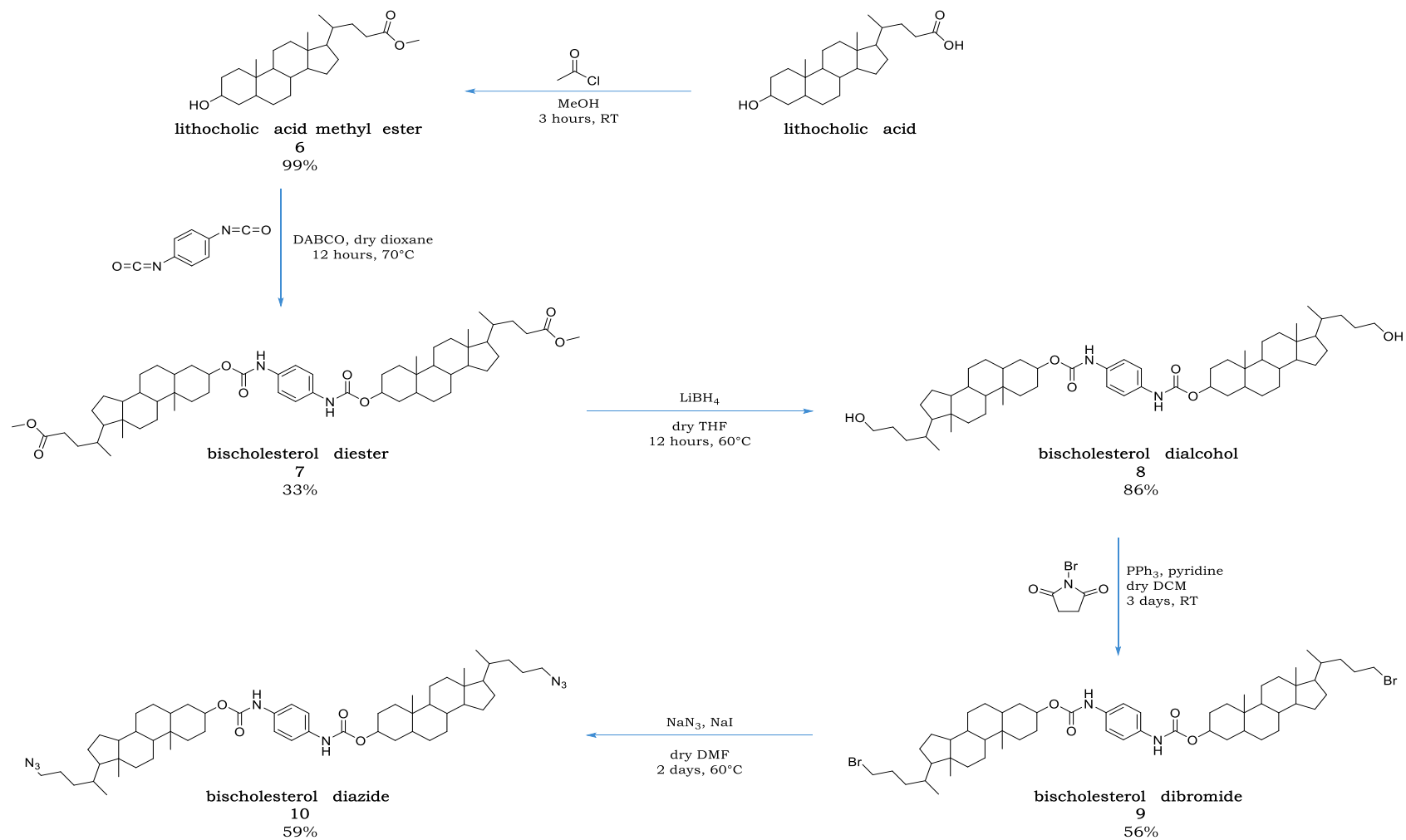
XPS measurements were carried out through the provision of the EPSRC X-ray photoelectron spectroscopy service for the academic community through the 'Access to research equipment initiative' using Thermo Scientific K-Alpha XPS systems.

For high APS samples, XPS spectra were obtained at the University of Newcastle (NEXUS), whereas for low APS samples, XPS spectra were obtained at the University of Cardiff (HarwellXPS). XPS experiments were performed using a monochromatic Al K $\alpha$  X-ray source (1486.7eV) at a take-off angle of 90° to the surface plane. High-resolution scans of Au (4f), C (1s), O (1s), N (1s), S (2p) and P (2p) were recorded using pass energies of 40eV at a step size of 0.1eV. Fitting of XPS peaks was performed using CasaXPS software using the following sensitivity factors: Au (4f) 17.10, C (1s) 1.00, O (1s) 2.93, N (1s) 1.80, S (2p) 1.68, P (2p) 1.19.

### 5.13 Optimisation of APS concentration used in acrylamide DNA coupling experiments by HPLC analysis

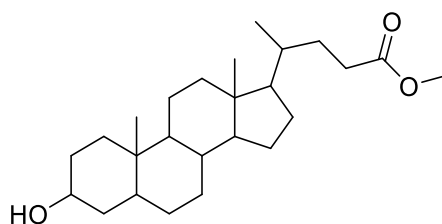
100µl aqueous solutions of unmodified/acrylamide modified DNA (10µM), BAC (10mM), TEMED (1% v/v), and APS (0µM, 10µM, 100µM, 1mM, 10mM, 20mM, 50mM, 100mM, or 200mM) were prepared. The solutions were shaken overnight prior to HPLC analysis. Samples were analysed by analytical HPLC using a Phenomenex Clarity 5µm Oligo RP LC 250x4.6mm column on an Agilent Technologies 1260 Infinity system. The column was heated to 60°C prior to sample injection. 70µl of sample was injected with a run time of 45 minutes for each sample, at a flow rate of 1ml/min. The solvent gradients used can be found in Table 5.1.

## 5.14 Synthesis of symmetrical bischolesterol derivatives



**Scheme 5.3** Synthesis of symmetrical bischolesterol derivatives

#### 5.14.1 Synthesis of lithocholic acid methyl ester



6

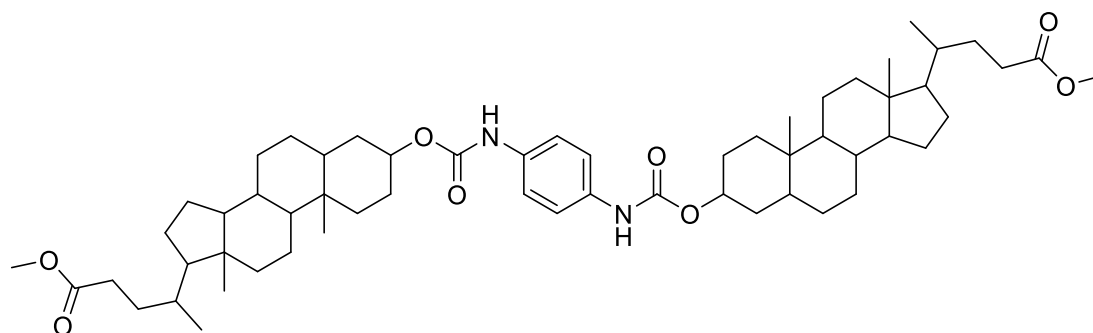
*Adapted from Ma and Davis<sup>3</sup>*

Acetyl chloride (0.95ml, 13.3mmol) was added to a suspension of lithocholic acid (10g, 26.6mmol) in MeOH (100ml). The reaction mixture was stirred under argon for 3 hours at RT. Water (100ml) was then added to the reaction mixture and the resulting white precipitate was collected by filtration and dried *in vacuo*. The crude product was crystallized from acetonitrile to give a white solid (10.3g, 99%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 3.66 (s, 3H), 3.63-3.59 (m, 1H), 2.39-0.89 (m, 38H), 0.63 (s, 3H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 174.9 (C=O), 72.0 (CH), 56.6 (CH), 56.1 (CH), 51.6 (OCH<sub>3</sub>), 42.9 (C), 42.2 (CH), 40.6 (CH), 40.3 (CH<sub>2</sub>), 36.6 (CH<sub>2</sub>), 36.0 (CH), 35.5 (CH), 35.5 (CH<sub>2</sub>), 34.7 (C), 31.2 (CH<sub>2</sub>), 31.1 (CH<sub>2</sub>), 30.7 (CH<sub>2</sub>), 28.3 (CH<sub>2</sub>), 27.3 (CH<sub>2</sub>), 26.6 (CH<sub>2</sub>), 24.3 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 21.0 (CH<sub>2</sub>), 18.4 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). MS (APCI +ve) (*m/z*): 373.6 [M-OH]<sup>+</sup>.



#### 5.14.2 Synthesis of bischolesterol diester



7

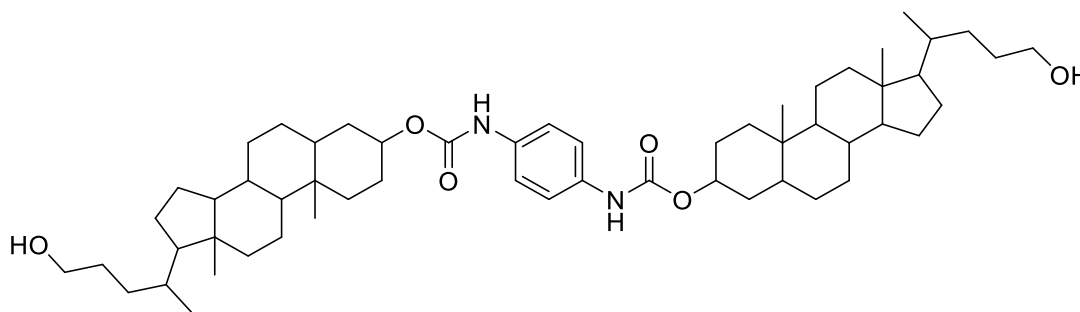
*Adapted from Kobuke<sup>4</sup>*

Phenylene diisocyanate (205mg, 1.28mmol) and 1,4-diazabicyclo[2.2.2]octane (DABCO, 30mg, 0.26mmol) were added to a solution of lithocholic acid methyl ether **6** (1g, 2.56mmol) in dry dioxane (20ml) under argon at room temperature. The solution was heated to 70°C and left to stir under argon at this temperature overnight. The reaction was then cooled to RT, and the solvent evaporated under reduced pressure. Aqueous 1N HCl was then added to the white solid and the resulting suspension was extracted with DCM until the organic phase was clear. The organic phases were combined, dried over magnesium sulphate, and evaporated under vacuum. The crude product was washed with methanol, dissolved in a small volume of hot chloroform, and purified by flash chromatography with an eluent of 3:1 diethyl ether:hexane. The appropriate fractions were combined and evaporated to afford a white solid (394mg, 33%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 7.30 (s, 4H), 6.41 (s, 2H), 4.76-4.68 (m, 2H), 3.67 (s, 6H), 2.39-0.92 (m, 75H), 0.67 (s, 6H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 174.8 (C=O), 153.6 (C=O), 134.0 (Ar-C), 119.9 (Ar-CH), 75.7 (CH), 56.9 (CH), 56.4 (CH), 51.5 (OCH<sub>3</sub>), 43.1 (C), 42.3 (CH), 40.9 (CH), 40.5 (CH<sub>2</sub>), 36.1 (CH), 35.6 (CH), 35.4 (CH<sub>2</sub>), 34.9 (C), 32.9 (CH<sub>2</sub>), 31.4 (CH<sub>2</sub>), 31.3 (CH<sub>2</sub>), 28.4 (CH<sub>2</sub>), 27.3 (CH<sub>2</sub>),

27.2 (CH<sub>2</sub>), 26.6 (CH<sub>2</sub>), 24.4 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 21.1 (CH<sub>2</sub>), 18.5 (CH<sub>3</sub>), 12.3 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): [M+Na]<sup>+</sup> calcd for C<sub>58</sub>H<sub>88</sub>N<sub>2</sub>O<sub>8</sub>Na, 963.6438; found 963.6440. IR neat (cm<sup>-1</sup>): 3361 (m, NH, carbamate), 2938 (s, CH, aliphatic), 2864 (s, CH, aliphatic), 1718 (s, C=O, ester/carbamate), 1548 (s, C=C, aromatic), 1522 (s, C=C, aromatic). Melting point: 184-185°C.

#### 5.14.3 Synthesis of bischolesterol dialcohol



8

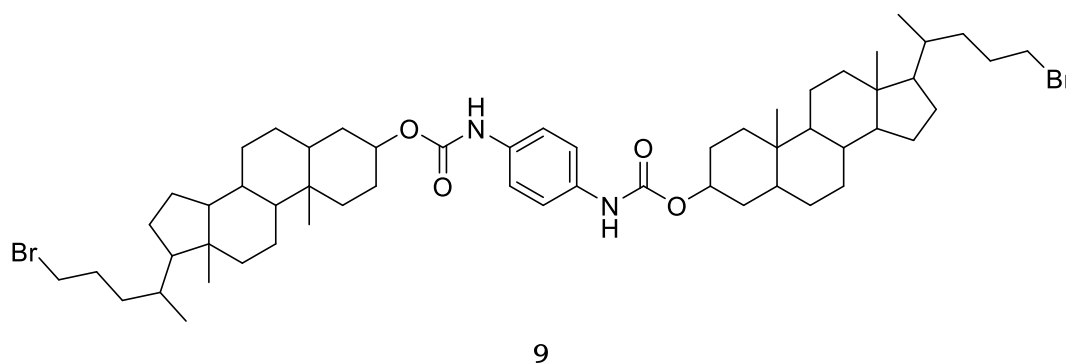
*Adapted from Kobuke<sup>5</sup>*

Bischolesterol diester **7** (558mg, 0.59mmol) was suspended in dry THF (25ml) under argon and heated to 50°C. LiBH<sub>4</sub> (50mg, 2.30mmol) was then added and the reaction heated to 60°C and left to stir at this temperature for 4 hours. The reaction was cooled to RT and left stirring under argon overnight. 1N aqueous HCl was added dropwise until effervescence had ceased, and the reaction mixture was diluted with water, extracted with chloroform, and dried over MgSO<sub>4</sub>. After evaporation of the solution, the residue was purified by column chromatography (4% methanol in chloroform) to yield the final product (450mg, 86%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 7.30 (s, 4H), 6.53 (s, 2H), 4.74-4.66 (m, 2H), 3.63-3.58 (m, 4H), 1.99-0.65 (m, 77H), 0.65 (s, 6H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ

153.5 (C=O), 133.8 (Ar-C), 119.6 (Ar-CH), 75.5 (CH), 63.8 (CH<sub>2</sub>), 56.7 (CH), 56.4 (CH), 42.9 (C), 42.1 (CH), 40.6 (CH), 40.3 (CH<sub>2</sub>), 36.0 (CH), 35.7 (CH), 35.2 (CH<sub>2</sub>), 34.7 (C), 32.7 (CH<sub>2</sub>), 32.0 (CH<sub>2</sub>), 29.5 (CH<sub>2</sub>), 28.5 (CH<sub>2</sub>), 27.2 (CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 26.5 (CH<sub>2</sub>), 24.4 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 21.0 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): [M+Na]<sup>+</sup> calcd for C<sub>56</sub>H<sub>88</sub>N<sub>2</sub>O<sub>6</sub>Na, 907.6540; found 907.6541. IR neat (cm<sup>-1</sup>): 3424 (w, OH), 2938 (s, CH, aliphatic), 2865 (s, CH, aliphatic), 1720 (s, C=O, carbamate), 1624 (s, C=C, aromatic). Melting point: 203°C.

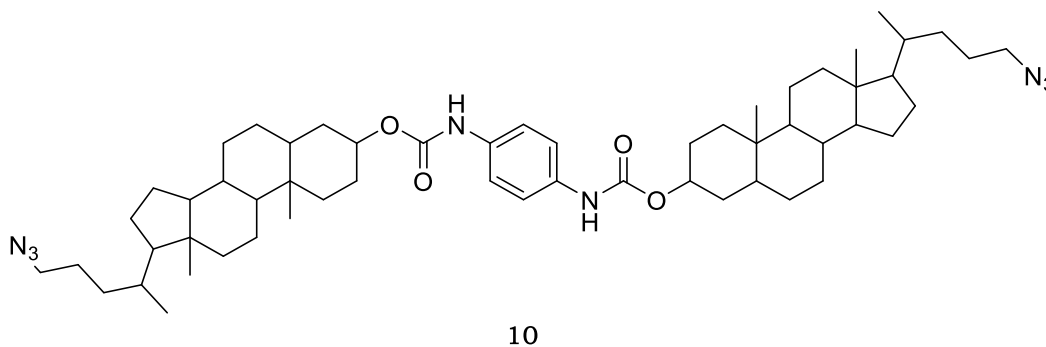
#### [5.14.4 Synthesis of bischolesterol dibromide](#)



Under an inert atmosphere, pyridine (0.2ml, 2.58mmol) was added to a solution of triphenylphosphine (1.58g, 6.01mmol) and N-bromosuccinimide (1.15g, 6.44mmol) in dry DCM (20ml). A solution of bischolesterol dialcohol **8** (950mg, 1.07mmol) in dry DCM (50ml) was then added, and the resulting reaction was left stirring under argon at RT for 3 days. The solvent was removed, co-evaporating with toluene (3 x 100ml), to give a black/brown residue which was purified by column chromatography (1:1 toluene:ethyl acetate) to yield the final product as an off white solid (603mg, 56%).

$^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  7.30 (s, 4H), 6.44 (s, 2H), 4.73-4.67 (m, 2H), 3.44-3.33 (m, 4H), 1.99-0.88 (m, 70H), 0.65 (s, 6H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  153.5 (C=O), 133.8 (Ar-C), 119.6 (Ar-CH), 75.5 (CH), 56.7 (CH), 56.2 (CH), 42.9 (C), 42.1 (CH), 40.6 (CH), 40.3 ( $\text{CH}_2$ ), 36.0 (CH), 35.4 (CH), 35.2 ( $\text{CH}_2$ ), 34.7 ( $\text{CH}_2$ ), 34.7 (C), 32.7 ( $\text{CH}_2$ ), 29.8 ( $\text{CH}_2$ ), 28.4 ( $\text{CH}_2$ ), 27.2 ( $\text{CH}_2$ ), 27.1 ( $\text{CH}_2$ ), 26.5 ( $\text{CH}_2$ ), 24.3 ( $\text{CH}_2$ ), 23.5 ( $\text{CH}_3$ ), 21.0 ( $\text{CH}_2$ ), 18.8 ( $\text{CH}_3$ ), 12.2 ( $\text{CH}_3$ ). HRMS (TOF-ESI +ve) ( $m/z$ ):  $[\text{M}+\text{Na}]^+$  calcd for  $\text{C}_{56}\text{H}_{88}\text{N}_2\text{O}_4\text{Na}^{79}\text{Br}^{81}\text{Br}$ , 1033.4832; found 1033.4838. IR neat ( $\text{cm}^{-1}$ ): 3321 (w, NH, carbamate), 2936 (s, CH, aliphatic), 2865 (s, CH, aliphatic), 1726 (m, C=O, carbamate), 1704 (m, C=O, carbamate), 1524 (m, C=C, aromatic). Melting point: 133-134°C.

#### [5.14.5 Synthesis of bischolesterol diazide](#)



*Adapted from Sharma and Gilmer<sup>6</sup>*

To a stirred solution of bischolesterol dibromide **9** (560mg, 0.55mmol) in anhydrous DMF (16ml) was added sodium azide (216mg, 3.3mmol) and sodium iodide (17mg, 0.11mmol). The reaction was heated to 60°C and left to stir at this temperature under argon for 2 days. Water (50ml) was added to precipitate the final product, which was isolated by filtration and washed with water to give an off white solid (290mg, 59%).

$^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  7.30 (s, 4H), 6.48 (s, 2H), 4.74-4.66 (m, 2H), 3.29-3.17 (m, 4H), 1.99-0.92 (m, 80H), 0.65 (s, 6H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  153.5 (C=O), 133.8 (Ar-C), 119.6 (Ar-CH), 75.6 (CH), 56.7 (CH), 56.2 (CH), 52.1 (CH<sub>2</sub>), 42.9 (C), 42.1 (CH), 40.6 (CH), 40.3 (CH<sub>2</sub>), 36.0 (CH), 35.6 (CH), 35.2 (CH<sub>2</sub>), 34.7 (C), 33.1 (CH<sub>2</sub>), 32.7 (CH<sub>2</sub>), 28.4 (CH<sub>2</sub>), 27.2 (CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 26.5 (CH<sub>2</sub>), 25.7 (CH<sub>2</sub>), 24.3 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 21.0 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) ( $m/z$ ):  $[\text{M}+\text{Na}]^+$  calcd for  $\text{C}_{56}\text{H}_{86}\text{N}_8\text{O}_4\text{Na}$ , 957.6670; found 957.6671. IR neat ( $\text{cm}^{-1}$ ): 3322 (w, NH, carbamate), 2930 (s, CH, aliphatic), 2865 (s, CH, aliphatic), 2092 (m, N=N=N, azide), 1710 (m, C=O, carbamate), 1523 (s, C=C, aromatic). Melting point: 196-198°C.

#### 5.14.6 Crystal structure determination of bischolesterol diester, bischolesterol dibromide, and bischolesterol diazide

##### Crystal data for bischolesterol diester (compound **7**)

$C_{65}H_{96}N_2O_8$  ( $M=1033.43\text{g/mol}$ ): triclinic, space group  $P1$  (no. 1),  $a = 7.8291(3)\text{\AA}$ ,  $b = 10.2983(4)\text{\AA}$ ,  $c = 18.1014(7)\text{\AA}$ ,  $\alpha = 97.459(3)^\circ$ ,  $\beta = 92.901(3)^\circ$ ,  $\gamma = 92.677(3)^\circ$ ,  $V = 1443.16(9)\text{\AA}^3$ ,  $Z = 1$ ,  $T = 99.95(18)\text{K}$ ,  $\mu(\text{CuK}\alpha) = 0.602\text{mm}^{-1}$ ,  $D_{\text{calc}} = 1.189\text{g/cm}^3$ , 23539 reflections measured ( $8.672^\circ \leq 2\theta \leq 144.194^\circ$ ), 10702 unique ( $R_{\text{int}} = 0.0332$ ,  $R_{\text{sigma}} = 0.0384$ ) which were used in all calculations. The final  $R_1$  was 0.0562 ( $I > 2\sigma(I)$ ) and  $wR_2$  was 0.1690 (all data).

##### Crystal data for bischolesterol dibromide (compound **9**)

$C_{62}H_{100}Br_2N_4O_6$  ( $M=1157.27\text{g/mol}$ ): monoclinic, space group  $P2_1$  (no. 4),  $a = 8.68918(6)\text{\AA}$ ,  $b = 59.9830(5)\text{\AA}$ ,  $c = 17.67491(16)\text{\AA}$ ,  $\beta = 97.1524(7)^\circ$ ,  $V = 9140.54(13)\text{\AA}^3$ ,  $Z = 6$ ,  $T = 100.01(10)\text{K}$ ,  $\mu(\text{CuK}\alpha) = 2.084\text{mm}^{-1}$ ,  $D_{\text{calc}} = 1.261\text{g/cm}^3$ , 82168 reflections measured ( $5.25^\circ \leq 2\theta \leq 136.502^\circ$ ), 32056 unique ( $R_{\text{int}} = 0.0734$ ,  $R_{\text{sigma}} = 0.0718$ ) which were used in all calculations. The final  $R_1$  was 0.0764 ( $I > 2\sigma(I)$ ) and  $wR_2$  was 0.2006 (all data).

##### Crystal data for bischolesterol diazide (compound **10**)

$C_{62}H_{100}N_{10}O_6$  ( $M=1081.51\text{g/mol}$ ): orthorhombic, space group  $P2_12_12_1$  (no. 19),  $a = 8.85680(10)\text{\AA}$ ,  $b = 20.1304(3)\text{\AA}$ ,  $c = 34.4364(5)\text{\AA}$ ,  $V = 6139.70(15)\text{\AA}^3$ ,  $Z = 4$ ,  $T = 100.00(10)\text{K}$ ,  $\mu(\text{CuK}\alpha) = 0.599\text{mm}^{-1}$ ,  $D_{\text{calc}} = 1.170\text{g/cm}^3$ , 40102 reflections measured ( $8.786^\circ \leq 2\theta \leq 147.748^\circ$ ), 12232 unique ( $R_{\text{int}} = 0.0329$ ,  $R_{\text{sigma}} = 0.0252$ ) which were used in all calculations. The final  $R_1$  was 0.0942 ( $I > 2\sigma(I)$ ) and  $wR_2$  was 0.2751 (all data).

The datasets were measured on an Agilent SuperNova diffractometer using an Atlas detector. The data collections were driven and processed, and absorption corrections was applied using CrysAlisPro. The structures of compounds **7** and **9** were solved using ShelXS<sup>7</sup> while the structure of compound **10** was solved using ShelXT.<sup>8</sup> All three structures were refined by a full-matrix least-squares procedure on  $F^2$  in ShelXL.<sup>8</sup> All non-hydrogen atoms were refined with anisotropic displacement parameters. In compound **7** the hydrogen atoms bonded to N(1) and N(2) were located in the electron density and their positions refined freely, with their  $U(\text{iso})$  values based on the  $U(\text{eq})$  values of the parent atoms. All remaining hydrogen atoms in compound **7** and all hydrogen atoms in compounds **9** and **10** were fixed as riding models and the isotropic thermal parameters ( $U_{\text{iso}}$  based on the  $U_{\text{eq}}$  of the parent atom. Figures and reports were produced using OLEX2.<sup>9</sup>

All three structures occupy chiral space groups. For compounds **7** and **10** it was not possible to determine the absolute structure from the diffraction data. However, for compound **9** the Flack parameter was determined as -0.005(10) which means that the absolute structure can confidently be assigned from the diffraction data.

The structure of compound **7** contains a molecule of toluene, (C(101)-C(107) and C(11')-C(17')) which is disordered over two positions at a refined percentage occupancy ratio of 61.7(11):38.3(11) respectively.

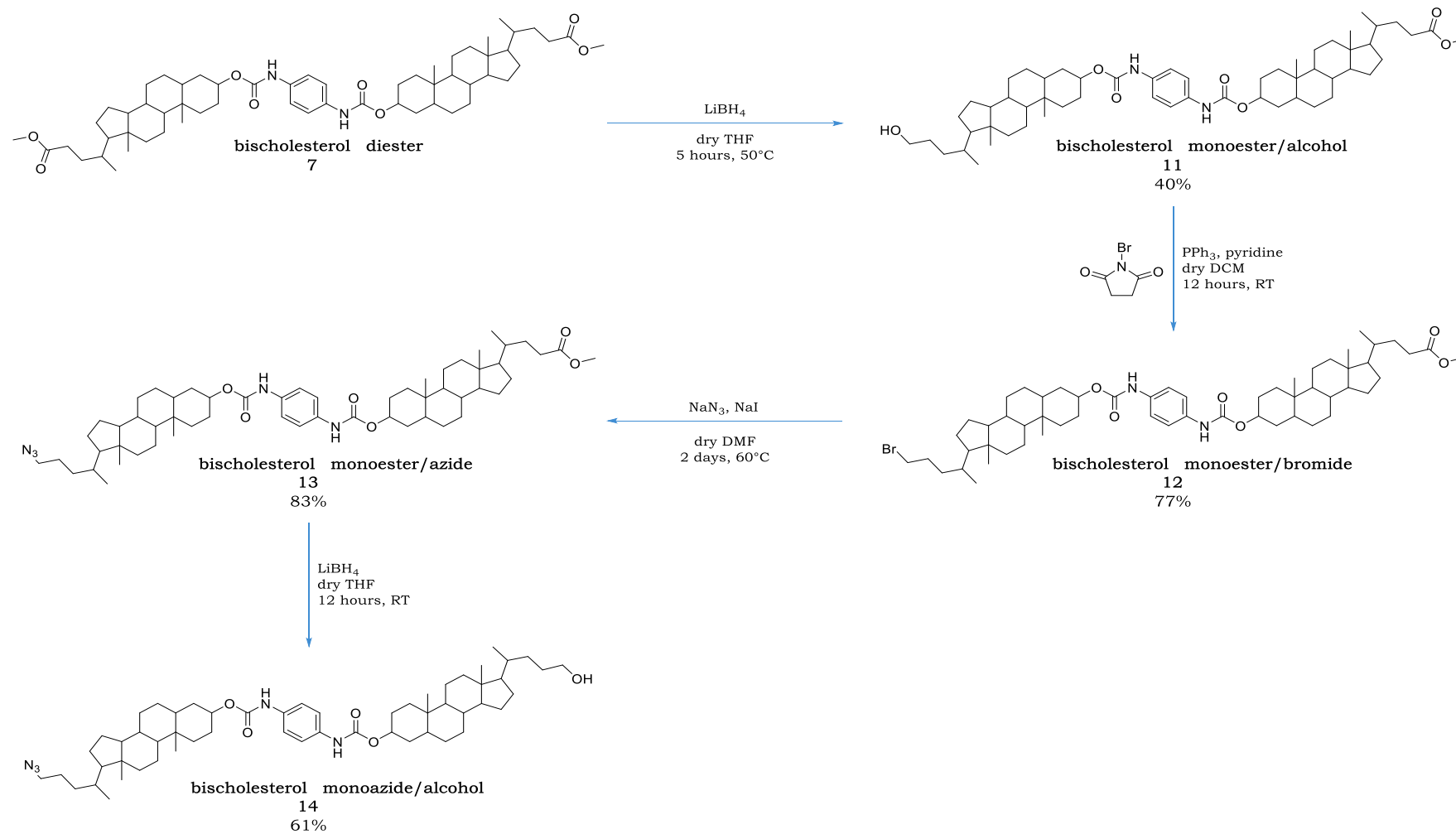
The structure of compound **9** contains three crystallographically-independent molecules and six molecules of DMF. One of the bromine atoms belonging to molecule 2 (C(101)-C(156), N(101)-N(102), O(101)-O(104), Br(11)-Br(12)) is

disordered over two positions labelled Br(12) / Br(2') at a refined occupancy ratio 89.7(5):10.3(5) respectively. In molecule 3 (C(201)-C(256), N(201)-N(202), O(201)-O(204), Br(21)-Br(22)) the C(255), C(256), Br(22) / C(55'), C(56'), Br(3)' group is disordered over two positions at a refined occupancy ratio of 67.7(6):32.3(6). One of the six molecules of DMF that are present, (C(801)-C(803), N(801), O(801) / C(81')-C(83'), N(81'), O(81')) is disordered over two positions with a refined percentage occupancy ratio of 63.7(17):36.3(17).

The structure of compound **10** contains two molecules of DMF, (C(101)-C(103), N(101), O(101) / C(11')-C(13'), N(11'), O(11')) and (C(201)-C(203), N(201), O(201) / C(21')-C(23'), N(21'), O(21')) which are both disordered over two positions with a refined percentage occupancy ratios of 71.6(6):28.4(6) and 54.1(5): 45.9(5) respectively.

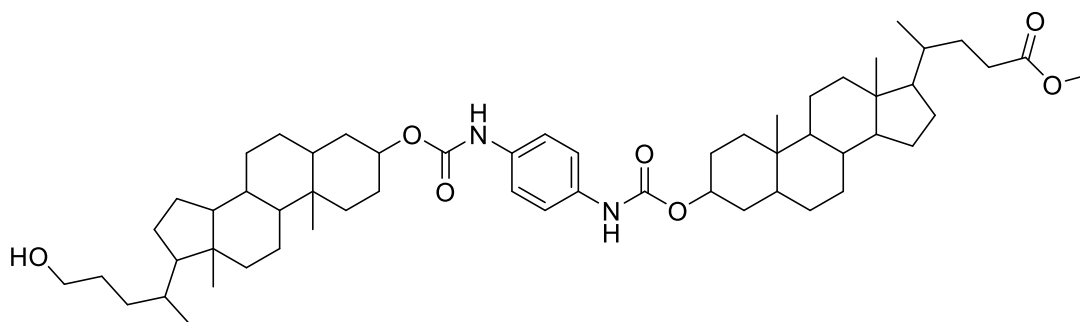


## 5.15 Synthesis of asymmetrical bischolesterol derivatives



**Scheme 5.4** Synthesis of asymmetrical bischolesterol derivatives.

#### 5.15.1 Synthesis of bischolesterol monoester/alcohol



11

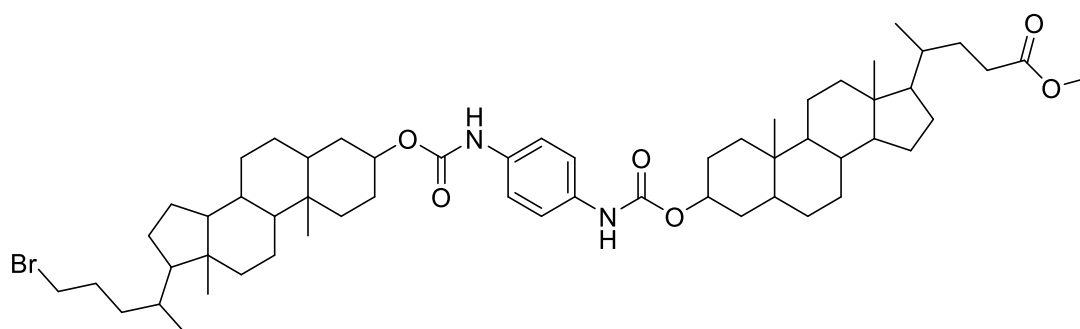
*Adapted from Kobuke<sup>5</sup>*

Bischolesterol diester **7** (1g, 1.06mmol) was suspended in dry THF (50ml) under argon. The suspension was stirred at RT for 4 hours and then heated to 50°C for 1 hour. During this time, LiBH<sub>4</sub> (13mg, 0.55mmol) was added sequentially (~4mg every 1.5 hours) to the solution. The reaction was then cooled to RT. 1N aqueous HCl was added dropwise until effervescence had ceased, and the reaction mixture was diluted with water, extracted with chloroform, and dried over MgSO<sub>4</sub>. After evaporation of the solution, the residue was purified by column chromatography (2:1 toluene: ethyl acetate) to yield the final product as a white solid (385mg, 40%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 7.30 (s, 4H), 6.49 (s, 2H), 4.74-4.66 (m, 2H), 3.66 (s, 3H), 3.61 (s, 2H), 2.39-2.17 (m, 2H), 2.00-0.85 (m, 73H), 0.65 (s, 6H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 174.9 (C=O), 153.5 (C=O), 133.8 (Ar-C), 119.6 (Ar-CH), 75.6 (CH), 63.8 (CH<sub>2</sub>), 56.7 (CH), 56.4 (CH), 56.2 (CH), 51.6 (OCH<sub>3</sub>), 42.9 (C), 42.1 (CH), 40.7 (CH), 40.3 (CH<sub>2</sub>), 36.0 (CH), 35.7 (CH), 35.5 (CH), 35.2 (CH<sub>2</sub>), 34.7 (C), 32.7 (CH<sub>2</sub>), 32.0 (CH<sub>2</sub>), 31.2 (CH<sub>2</sub>), 31.2 (CH<sub>2</sub>), 29.6 (CH<sub>2</sub>), 28.5 (CH<sub>2</sub>), 27.2 (CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 26.5 (CH<sub>2</sub>), 24.3 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 21.0 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>), 18.4 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): [M+Na]<sup>+</sup> calcd for

C<sub>57</sub>H<sub>88</sub>N<sub>2</sub>O<sub>7</sub>Na, 935.6489; found 935.6484. IR neat (cm<sup>-1</sup>): 3439 (w, OH), 3329 (w, NH, carbamate), 2932 (s, CH, aliphatic), 2865 (s, CH, aliphatic), 1702 (s, C=O, ester/carbamate), 1522 (s, C=C, aromatic). Melting point: 195-196°C.

#### 5.15.2 Synthesis of bischolesterol monoester/bromide



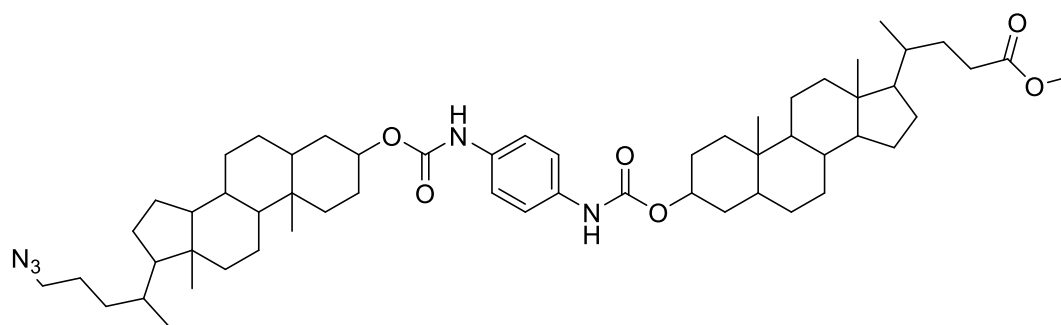
12

Under an inert atmosphere, pyridine (0.05ml, 0.66mmol) was added to a solution of triphenylphosphine (402mg, 1.53mmol) and N-bromosuccinimide (292mg, 1.64mmol) in dry DCM (10ml). A suspension of bischolesterol monoester/alcohol **11** (250mg, 0.27mmol) in dry DCM (30ml) was then added, and the resulting reaction was left stirring under argon at RT overnight. The solvent was removed, co-evaporating with toluene (3 x 50ml), to give a black/brown residue. The crude product was redissolved in a small volume of chloroform and purified by flash chromatography with an eluent of 2:1 diethyl ether:hexane to yield the final product as an off white solid (206mg, 77%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 7.30 (s, 4H), 6.46 (s, 2H), 4.73-4.67 (m, 2H), 3.66 (s, 3H), 3.42-3.35 (m, 2H), 2.39-2.19 (m, 2H), 1.98-0.88 (m, 77H), 0.65 (s, 6H).  
<sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 174.9 (C=O), 153.5 (C=O), 133.8 (Ar-C), 119.6 (Ar-CH), 75.6 (CH), 56.7 (CH), 56.2 (CH), 56.2 (CH), 51.6 (OCH<sub>3</sub>), 42.9 (C), 42.1 (CH), 40.7 (CH), 40.3 (CH<sub>2</sub>), 36.0 (CH), 35.5 (CH), 35.4 (CH), 35.2 (CH<sub>2</sub>), 34.7 (CH<sub>2</sub>),

34.7 (C), 34.7 (CH<sub>2</sub>), 32.7 (CH<sub>2</sub>), 31.2 (CH<sub>2</sub>), 31.2 (CH<sub>2</sub>), 29.8 (CH<sub>2</sub>), 28.4 (CH<sub>2</sub>), 27.2 (CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 26.5 (CH<sub>2</sub>), 24.3 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 21.0 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>), 18.4 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): [M+Na]<sup>+</sup> calcd for C<sub>57</sub>H<sub>87</sub>N<sub>2</sub>O<sub>6</sub>Na<sup>79</sup>Br, 997.5645; found 997.5649. IR neat (cm<sup>-1</sup>): 3340 (w, NH, carbamate), 2937 (s, CH, aliphatic), 2865 (s, CH, aliphatic), 1724 (s, C=O, carbamate), 1523 (s, C=C, aromatic).

### 5.15.3 Synthesis of bischolesterol monoester/azide



13

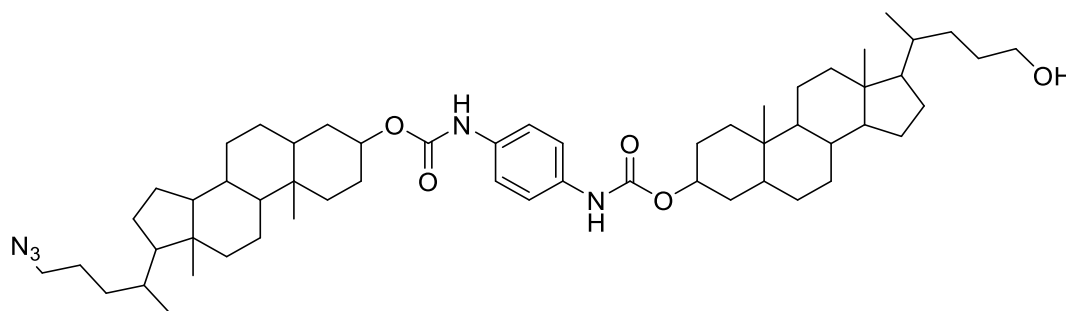
*Adapted from Sharma and Gilmer<sup>6</sup>*

To a stirred solution of bischolesterol monoester/bromide **12** (345mg, 0.36mmol) in anhydrous DMF (10ml) was added sodium azide (69mg, 1.07mmol) and sodium iodide (5.3mg, 0.04mmol). The reaction was heated to 60°C and left to stir at this temperature under argon for 2 days. Water (100ml) was added to precipitate the final product, which was isolated by filtration. The product was dissolved in chloroform, and the solution was washed with water and dried over magnesium sulphate. Chloroform was removed to yield the product as a white solid (271mg, 83%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 7.30 (s, 4H), 6.45 (s, 2H), 4.73-4.67 (m, 2H), 3.66 (s, 3H), 3.27-3.19 (m, 2H), 2.39-2.19 (m, 2H), 1.98-0.83 (m, 94H), 0.65 (s, 6H).

$^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  174.9 (C=O), 153.5 (C=O), 133.8 (Ar-C), 119.6 (Ar-CH), 75.6 (CH), 56.7 (CH), 56.2 (CH), 56.1 (CH), 52.1 ( $\text{CH}_2$ ), 51.6 ( $\text{OCH}_3$ ), 42.9 (C), 42.1 (CH), 40.6 (CH), 40.3 ( $\text{CH}_2$ ), 36.0 (CH), 35.6 (CH), 35.5 (CH), 35.2 ( $\text{CH}_2$ ), 34.7 (C), 33.1 ( $\text{CH}_2$ ), 32.7 ( $\text{CH}_2$ ), 31.2 ( $\text{CH}_2$ ), 31.2 ( $\text{CH}_2$ ), 28.4 ( $\text{CH}_2$ ), 27.2 ( $\text{CH}_2$ ), 27.1 ( $\text{CH}_2$ ), 26.5 ( $\text{CH}_2$ ), 25.7 ( $\text{CH}_2$ ), 24.3 ( $\text{CH}_2$ ), 23.5 ( $\text{CH}_3$ ), 21.0 ( $\text{CH}_2$ ), 18.8 ( $\text{CH}_3$ ), 18.4 ( $\text{CH}_3$ ), 12.2 ( $\text{CH}_3$ ). HRMS (TOF-ESI +ve) ( $m/z$ ):  $[\text{M}]^+$  calcd for  $\text{C}_{57}\text{H}_{88}\text{N}_5\text{O}_6$ , 938.6735; found 938.6739. IR neat ( $\text{cm}^{-1}$ ): 3333 (w, NH, carbamate), 2937 (s, CH, aliphatic), 2865 (s, CH, aliphatic), 2093 (m,  $\text{N}=\text{N}=\text{N}$ , azide), 1725 (s, C=O, carbamate), 1523 (s, C=C, aromatic). Melting point: 185-187°C.

#### [5.15.4 Synthesis of bischolesterol monoazide/alcohol](#)



14

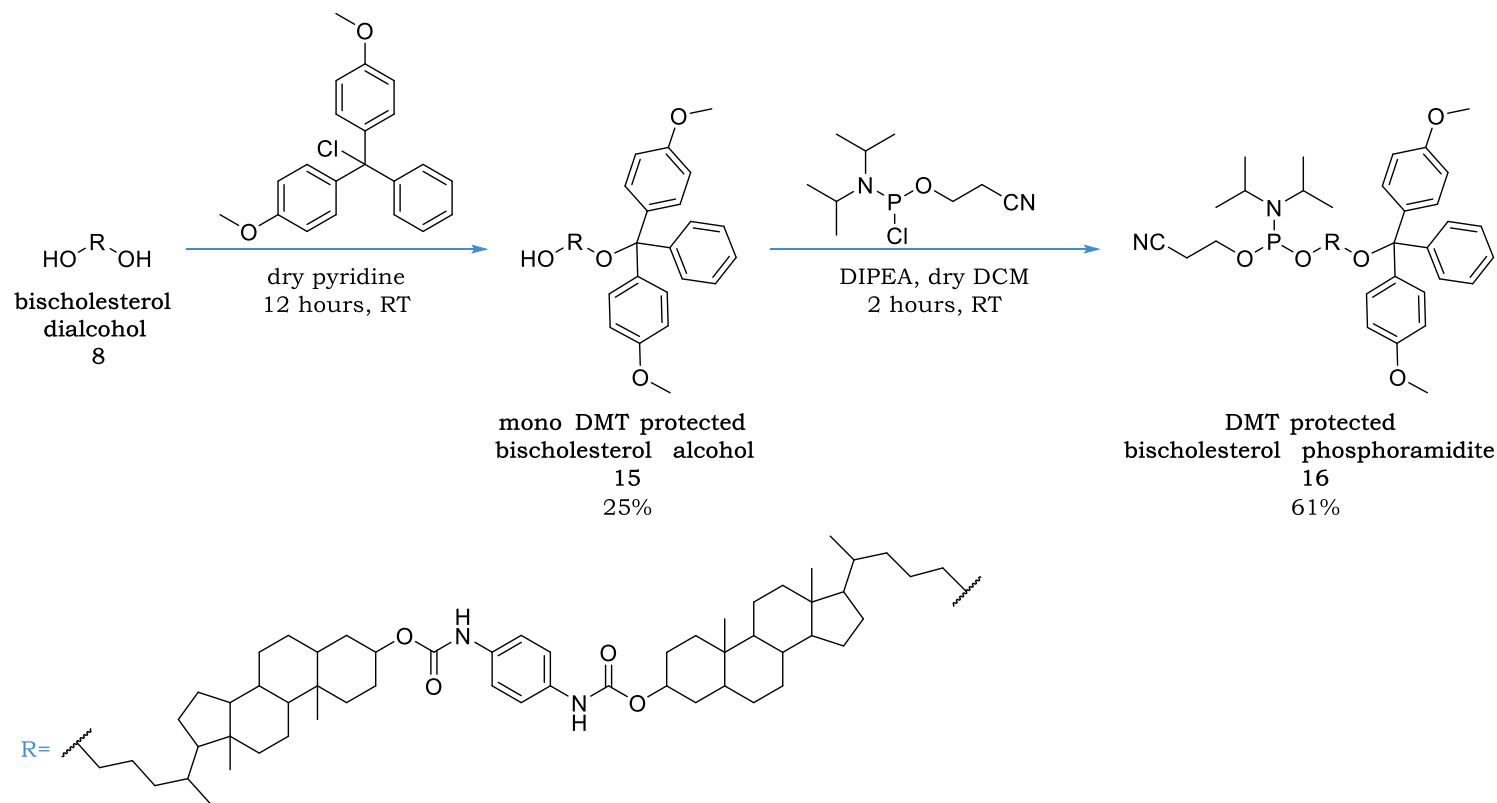
*Adapted from Kobuke<sup>5</sup>*

Bischolesterol monoester/azide **13** (170mg, 0.18mmol) was dissolved in dry THF (10ml) under argon. 0.2ml of  $\text{LiBH}_4$  solution (2M in THF) was added and the reaction was left stirring under argon overnight. Inspection by TLC showed incomplete conversion and thus a further 0.1ml of  $\text{LiBH}_4$  solution was added and the reaction heated to 50°C for 1 hour, after which the reaction was shown to be complete. 5M aqueous HCl was added dropwise to the reaction until effervescence had ceased, and the mixture was diluted with water, extracted

with DCM, and dried over MgSO<sub>4</sub>. After evaporation of the solution, the residue was purified by flash chromatography using an eluent of 3:1 ether:hexane to yield the final product as a white solid (130mg, 61%).

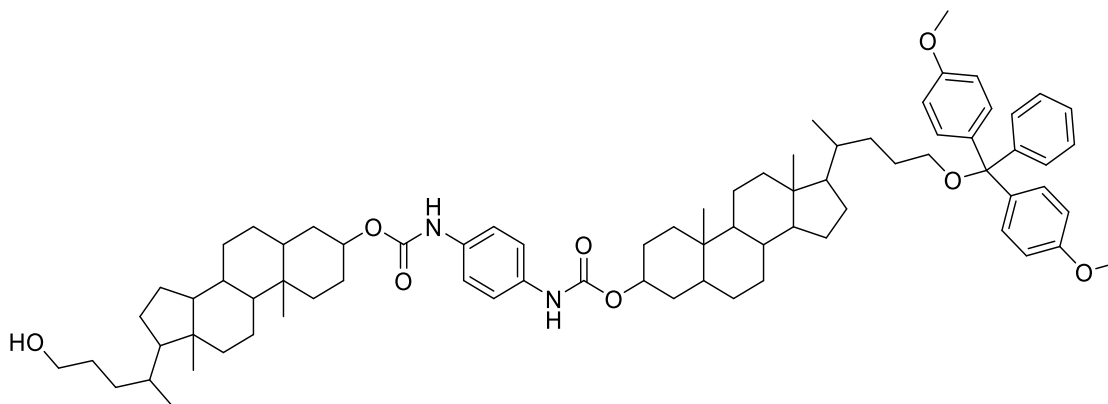
<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 7.30 (s, 4H), 6.52 (s, 2H), 4.74-4.66 (m, 2H), 3.66-3.57 (m, 2H), 3.29-3.17 (m, 2H), 1.99-0.88 (m, 71H), 0.65 (s, 6H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 153.5 (C=O), 133.8 (Ar-C), 119.7 (Ar-CH), 75.6 (CH), 63.7 (CH<sub>2</sub>), 56.7 (CH), 56.4 (CH), 56.2 (CH), 52.1 (CH<sub>2</sub>), 42.9 (C), 42.1 (CH), 40.6 (CH), 40.3 (CH<sub>2</sub>), 36.0 (CH), 35.7 (CH), 35.6 (CH), 35.2 (CH<sub>2</sub>), 34.7 (C), 33.1 (CH<sub>2</sub>), 32.7 (CH<sub>2</sub>), 32.0 (CH<sub>2</sub>), 29.5 (CH<sub>2</sub>), 28.4 (CH<sub>2</sub>), 27.2 (CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 26.5 (CH<sub>2</sub>), 25.7 (CH<sub>2</sub>), 24.3 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 21.0 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>), 18.8 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): M<sup>+</sup> calcd for C<sub>56</sub>H<sub>88</sub>N<sub>5</sub>O<sub>5</sub>, 910.6785; found 910.6793. IR neat (cm<sup>-1</sup>): 3440 (w, OH), 3324 (w, NH, carbamate), 2932 (s, CH, aliphatic), 2865 (s, CH, aliphatic), 2093 (m, N=N=N, azide), 1702 (s, C=O, carbamate), 1522 (s, C=C, aromatic).

# 5.16 Synthesis of DMT protected bischolesterol phosphoramidite



**Scheme 5.5** Synthesis of DMT protected bischolesterol phosphoramidite.

#### 5.16.1 Synthesis of mono DMT protected bischolesterol alcohol



15

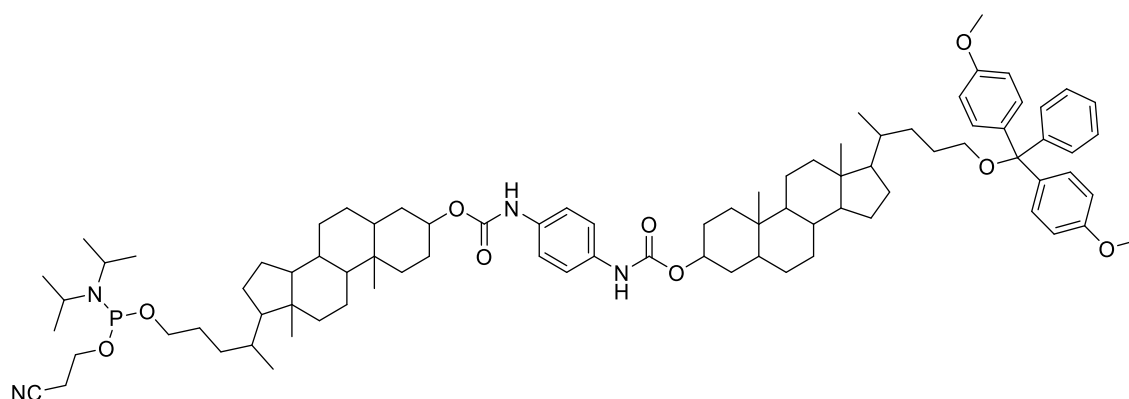
Under an inert atmosphere, dimethoxytrityl chloride (202mg, 0.60mmol) was added to a solution of bischolesterol dialcohol **8** (480mg, 0.54mmol) in dry pyridine (20ml) and the resulting yellow solution was left to stir overnight. The reaction was quenched with saturated sodium bicarbonate solution (100ml) and the product extracted with chloroform (3 x 50ml). The organics were combined and dried over sodium sulphate. The solvent was then removed, coevaporating with toluene (2 x 50ml) to remove pyridine, resulting in a crude yellow solid. The crude was redissolved in a small volume of chloroform, purified by flash chromatography with silica pretreated with triethylamine, and a gradient eluent of 3:1 → 4:1 diethyl ether:hexane. The appropriate fractions were combined, and solvent removed to give the product as a white solid (162mg, 25%).

$^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  7.45-7.42 (m, 2H), 7.34-7.26 (m, 10H), 7.22-7.17 (m, 1H), 6.84-6.80 (m, 4H), 6.50 (s, 2H), 4.74-4.66 (m, 2H), 3.79 (s, 6H), 3.61 (s, 2H), 3.05-2.95 (m, 2H), 2.00-0.86 (m, 78H), 0.65 (s, 3H), 0.63 (s, 3H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  158.4 (Ar-C), 153.5 (C=O), 145.6 (Ar-C), 136.9 (Ar-C), 133.8 (Ar-C), 130.2 (Ar-CH), 128.4 (Ar-CH), 127.8 (Ar-CH), 126.7 (Ar-CH), 119.6 (Ar-



CH), 113.1 (Ar-CH), 85.8 (C), 75.6 (CH), 64.2 (CH<sub>2</sub>), 63.8 (CH<sub>2</sub>), 56.7 (CH), 56.4 (CH), 55.3 (OCH<sub>3</sub>), 42.9 (C), 42.1 (CH), 40.7 (CH), 40.3 (CH<sub>2</sub>), 36.0 (CH), 35.7 (CH), 35.2 (CH<sub>2</sub>), 34.7 (C), 32.7 (CH<sub>2</sub>), 32.5 (CH<sub>2</sub>), 32.0 (CH<sub>2</sub>), 31.7 (CH<sub>2</sub>), 29.6 (CH<sub>2</sub>), 28.4 (CH<sub>2</sub>), 27.2 (CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 26.5 (CH<sub>2</sub>), 24.4 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 21.0 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): [M+Na]<sup>+</sup> calcd for C<sub>77</sub>H<sub>106</sub>N<sub>2</sub>O<sub>8</sub>Na, 1209.7847; found 1209.7844. IR neat (cm<sup>-1</sup>): 3324 (w, NH, carbamate), 2932 (s, CH, aliphatic), 2865 (s, CH, aliphatic), 1703 (m, C=O, carbamate), 1521 (s, C=C, aromatic), 1509 (s, C=C, aromatic).

#### 5.16.2 Synthesis of DMT protected bischolesterol phosphoramidite



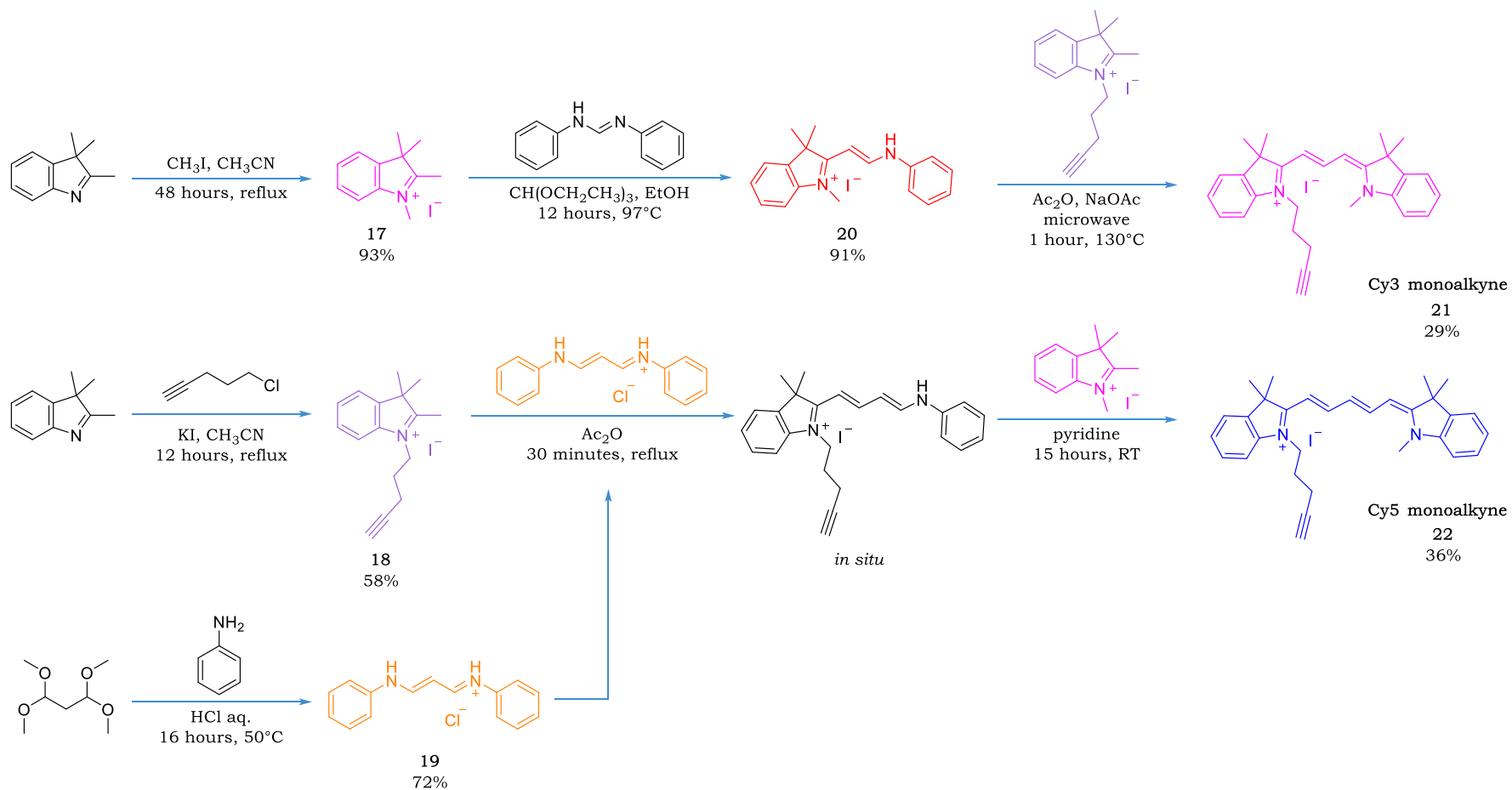
16

Mono DMT protected bischolesterol alcohol **15** (152mg, 0.13mmol) was azeotroped with dry DCM (3 x 5ml) and then redissolved in dry DCM (5ml) for reaction. The solution was degassed for five minutes before and during the addition of DIPEA (60μl, 0.32mmol) and 2-cyanoethyl diisopropyl chlorophosphoramidite (30μl, 0.15mmol), the latter of which was added dropwise to the reaction. The clear reaction was left stirring under argon until completion which was typically 2 hours. Once complete, the solution was diluted with degassed DCM (10ml) and washed with degassed sodium bicarbonate solution

(30ml). The organics were isolated, dried over sodium sulphate, and solvent removed to leave a clear oil. The oil was purified by flash chromatography using an eluent of 3:1 degassed diethyl ether: hexane and 1% triethylamine to give the final product as a white solid (180mg, 61%).

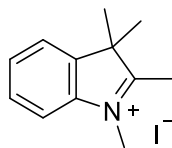
$^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  7.45-7.43 (m, 2H), 7.34-7.25 (m, 10H), 7.19 (t,  $J=7.2$  Hz, 1H), 6.83-6.80 (m, 4H), 6.67 (s, 2H), 4.72-4.67 (m, 2H), 3.89-3.53 (m, 12H), 3.06-2.96 (m, 2H), 2.70-2.44 (m, 7H), 1.99-0.86 (m, 99H), 0.65 (s, 3H), 0.63 (s, 3H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  158.3 (Ar-C), 153.5 (C=O), 145.6 (Ar-C), 136.9 (Ar-C), 133.7 (Ar-C), 130.1 (Ar-CH), 128.3 (Ar-CH), 127.7 (Ar-CH), 126.6 (Ar-CH), 119.6 (Ar-CH), 117.7 (CN), 113.0 (Ar-CH), 85.7 (C), 75.4 (CH), 64.3 ( $\text{CH}_2$ ), 64.1 ( $\text{CH}_2$ ), 58.4 ( $\text{CH}_2$ ), 56.6 (CH), 56.3 (CH), 55.2 ( $\text{OCH}_3$ ), 43.1 (CH), 42.8 (C), 42.0 (CH), 40.6 (CH), 40.2 ( $\text{CH}_2$ ), 35.9 (CH), 35.6 (CH), 35.5 (CH), 35.1 ( $\text{CH}_2$ ), 34.7 (C), 32.7 ( $\text{CH}_2$ ), 32.4 ( $\text{CH}_2$ ), 32.1 ( $\text{CH}_2$ ), 29.8 ( $\text{CH}_2$ ), 28.3 ( $\text{CH}_2$ ), 27.9 ( $\text{CH}_2$ ), 27.1 ( $\text{CH}_2$ ), 27.0 ( $\text{CH}_2$ ), 26.4 ( $\text{CH}_2$ ), 24.7 ( $\text{CH}_3$ ), 24.3 ( $\text{CH}_2$ ), 23.4 ( $\text{CH}_3$ ), 20.9 ( $\text{CH}_2$ ), 20.4 ( $\text{CH}_2$ ), 18.7 ( $\text{CH}_3$ ), 12.1 ( $\text{CH}_3$ ).  $^{31}\text{P}$  NMR (121MHz,  $\text{CDCl}_3$ )  $\delta$  147.2, 147.1, 139.0.

## 5.17 Synthesis of cyanine dyes



**Scheme 5.6** Synthesis of cyanine dyes.

#### 5.17.1 Synthesis of 1,2,3,3-tetramethyl-3H-indolinium iodide



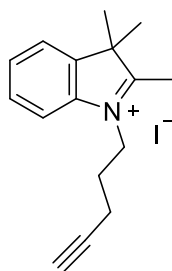
17

*Adapted from Brown<sup>10</sup>*

Methyl iodide (2.1ml, 34mmol) was added to a solution of 2,3,3-trimethylindolenine (5ml, 31mmol) in acetonitrile (25ml). The reaction mixture was refluxed for 48 hours and then cooled to room temperature. The precipitate formed was collected by filtration, washed with diethyl ether, and dried in vacuo over potassium hydroxide pellets to afford **17** as pink crystals (8.8g, 93%).

<sup>1</sup>H NMR (400MHz, (CD<sub>3</sub>)<sub>2</sub>SO) δ 7.92-7.90 (m, 1H), 7.84-7.83 (m, 1H), 7.65-7.59 (m, 2H), 3.97 (s, 3H), 2.77 (s, 3H), 1.53 (s, 6H). <sup>13</sup>C NMR (100MHz, (CD<sub>3</sub>)<sub>2</sub>SO) δ 196.0 (C), 142.1 (Ar-C), 141.6 (Ar-C), 129.3 (Ar-CH), 128.8 (Ar-CH), 123.3 (Ar-CH), 115.2 (Ar-CH), 54.0 (C), 34.7 (NCH<sub>3</sub>), 21.7 (CH<sub>3</sub>), 14.1 (CH<sub>3</sub>). MS (TOF-ESI +ve) (*m/z*): 174.1 [M-I]<sup>+</sup>.

### 5.17.2 Synthesis of 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide



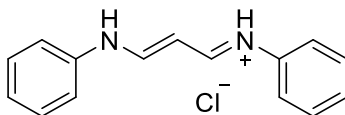
18

*Adapted from Brown<sup>10</sup>*

5-chloro-1-pentyne (5.0ml, 31.40mmol) was added to a suspension of potassium iodide (11.5g, 69.08mmol) in acetonitrile (50ml), and the yellow suspension was stirred at 50°C. After 10 minutes, 2,3,3-trimethylindolenine (5.05ml, 31.40mmol) was added dropwise and the reaction mixture was refluxed overnight and then cooled to room temperature. The inorganic salt was filtered off, washed with DCM, and the filtrate was evaporated to dryness. The residue was purified by column chromatography (0 → 10% MeOH in DCM) and the pure product was obtained as a purple solid (6.41g, 58%).

<sup>1</sup>H NMR (400MHz, CD<sub>3</sub>OD) δ 7.93-7.89 (m, 1H), 7.81-7.76 (m, 1H), 7.68-7.64 (m, 2H), 4.66-4.62 (m, 2H), 3.32-3.30 (m, 4H), 2.52-2.45 (m, 3H), 2.23-2.16 (m, 2H), 1.62 (s, 6H). <sup>13</sup>C NMR (100MHz, CD<sub>3</sub>OD) δ 198.5 (C=N), 143.4 (Ar-C), 142.6 (Ar-C), 131.2 (Ar-CH), 130.5 (Ar-CH), 124.7 (Ar-CH), 116.5 (Ar-CH), 83.1 (C≡CH), 72.0 (HC≡C), 56.0 (C), 48.4 (NCH<sub>2</sub>), 27.5 (CH<sub>2</sub>), 22.8 (CH<sub>3</sub>), 16.5 (CH<sub>2</sub>). MS (TOF-ESI +ve) (*m/z*): 352.1 [M-I]<sup>+</sup>.

### 5.17.3 Synthesis of malondialdehyde bis(phenylimine) monohydrochloride



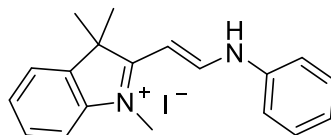
19

*Adapted from Brown<sup>10</sup>*

An aqueous solution of hydrochloric acid (110ml, 1M), and 1,1,3,3-tetramethoxypropane (9.9ml, 0.06mol) was heated to 50°C. A solution of aniline (11ml, 0.12mol) in aqueous hydrochloric acid (185ml, 1M) was added, and the resulting reaction was stirred at 50°C for 16 hours. The suspension was cooled to room temperature and the orange precipitate afforded was isolated by filtration, washed with hydrochloric acid (1M), and dried under vacuum to yield the final product as an orange powder (11g, 72%).

<sup>1</sup>H NMR (400MHz, CD<sub>3</sub>OD) δ 8.72 (d, J= 11.6Hz, 2H), 7.48-7.44 (m, 4H), 7.40-7.37 (m, 4H), 7.30-7.25 (m, 2H), 6.28 (t, J= 11.6Hz, 1H). <sup>13</sup>C NMR (100MHz, CD<sub>3</sub>OD) δ 159.8 (HC=C), 139.8 (Ar-C), 131.1 (Ar-CH), 127.7 (Ar-CH), 118.8 (Ar-CH), 99.4 (HC=C). MS (TOF-ESI +ve) (*m/z*): 223.1 [M-Cl]<sup>+</sup>.

#### 5.17.4 Synthesis of 1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolium iodide



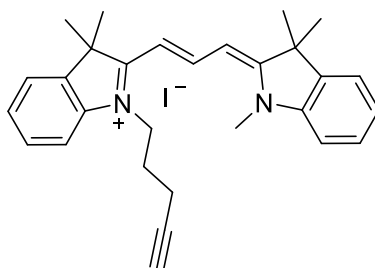
20

*Adapted from Brown<sup>10</sup>*

A solution of 1,2,3,3-tetramethyl-3H-indolinium iodide **17** (500mg, 1.7mmol) and N,N'-diphenyl formamidine (360mg, 1.8mmol) in ethanol (25ml) was stirred for 10 minutes. Triethyl orthoformate (270mg, 1.8mmol) was added and the resulting solution was stirred at reflux for 18 hours and then slowly cooled to room temperature. The precipitate was removed by filtration and washed in cold diethyl ether to yield 1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolium iodide **20** as orange crystals (610mg, 91%).

<sup>1</sup>H NMR (400MHz, CD<sub>3</sub>OD) δ 8.67 (d, J= 12.6Hz, 1H), 7.58 (d, J= 7.5Hz, 1H), 7.53-7.42 (m, 6H), 7.39-7.29 (m, 2H), 6.20 (d, J =12.6Hz, 1H), 3.73 (s, 3H), 1.75 (s, 6H). <sup>13</sup>C NMR (100MHz, CD<sub>3</sub>OD) δ 180.5 (C=N), 153.1 (HC=C), 143.7 (Ar-C), 142.2 (Ar-C), 140.0 (Ar-C), 131.2 (Ar-CH), 129.9 (Ar-CH), 127.6 (Ar-CH), 127.3 (Ar-CH), 123.4 (Ar-CH), 119.0 (Ar-CH), 112.7 (Ar-CH), 91.9 (HC=C), 51.3 (C), 32.4 (NCH<sub>3</sub>), 28.5 (CH<sub>3</sub>). MS (TOF-ESI +ve) (*m/z*): 277.3 [M-I]<sup>+</sup>.

#### 5.17.5 Synthesis of Cy3 monoalkyne



21

*Adapted from Henary<sup>11</sup>*

To a 20ml microwave tube was added 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide **18** (360mg, 1.02mmol), 1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolinium iodide **20** (330mg, 0.82mmol), and acetic anhydride (12ml). The solution was degassed for 5 minutes before and 5 minutes after the addition of sodium acetate (150mg, 1.83mmol). The tube was tightly sealed with an aluminium/Teflon crimp top, and the mixture irradiated in a microwave for 1 hour at 130°C. After the irradiation period, the reaction vessel was cooled to room temperature and added to cold ether (800ml). The resulting precipitate was isolated by suction filtration and washed with cold ether. The solid was then partially dissolved in DCM, the filtrate was isolated, and solvent removed. The purple crude was purified by flash chromatography using a gradient eluent of 0 → 50% ethyl acetate in DCM followed by 10% methanol in DCM. Product containing fractions were combined and the solvent was evaporated to leave a pink solid which was further purified by preparative HPLC on a Kinetex C18 prep column using the method described in the following table:



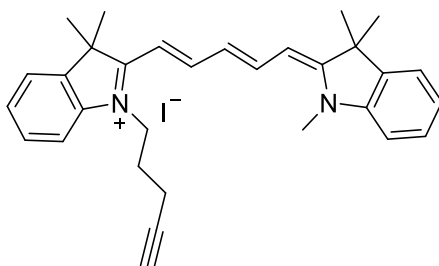
**Table 5.3** Solvent gradients used in the HPLC purification of cyanine dyes.

Time /mins	Water + 0.1% TFA/%	Methanol + 0.1% TFA/%
0	50	50
20	0	100
25	0	100
25.1	50	50
30	50	50

The purification was monitored at 550nm and the desired product was found to elute at ~16 minutes. The product was collected, and the solvent removed by rotary evaporation and freeze drying to give the final product as a pink solid (129mg, 29%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 8.39 (s, 1H), 7.42-7.36 (m, 4H), 7.29-7.23 (m, 4H), 7.15 (d, J= 7.8Hz, 1H), 6.59-6.52 (m, 2H), 4.25 (s, 2H), 3.69 (s, 3H), 2.41 (s, 2H), 2.14 (s, 1H), 2.06 (s, 2H), 1.72 (s, 12H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 175.0 (C), 174.1 (C), 150.8 (HC=C), 142.7 (Ar-C), 142.2 (Ar-C), 140.6 (Ar-C), 140.6 (Ar-C), 129.1 (Ar-CH), 125.8 (Ar-CH), 125.6 (Ar-CH), 122.3 (Ar-CH), 122.2 (Ar-CH), 111.1 (Ar-CH), 111.0 (Ar-CH), 104.2 (HC=C), 103.7 (HC=C), 83.1 (C≡CH), 70.2 (HC≡C), 49.3 (C), 49.2 (C), 43.2 (CH<sub>2</sub>), 31.6 (NCH<sub>3</sub>), 28.3 (CH<sub>3</sub>), 28.1 (CH<sub>3</sub>), 26.0 (CH<sub>2</sub>), 16.0 (CH<sub>2</sub>). MS (TOF-ESI +ve) (*m/z*): 409.3 [M-I]<sup>+</sup>.

#### 5.17.6 Synthesis of Cy5 monoalkyne



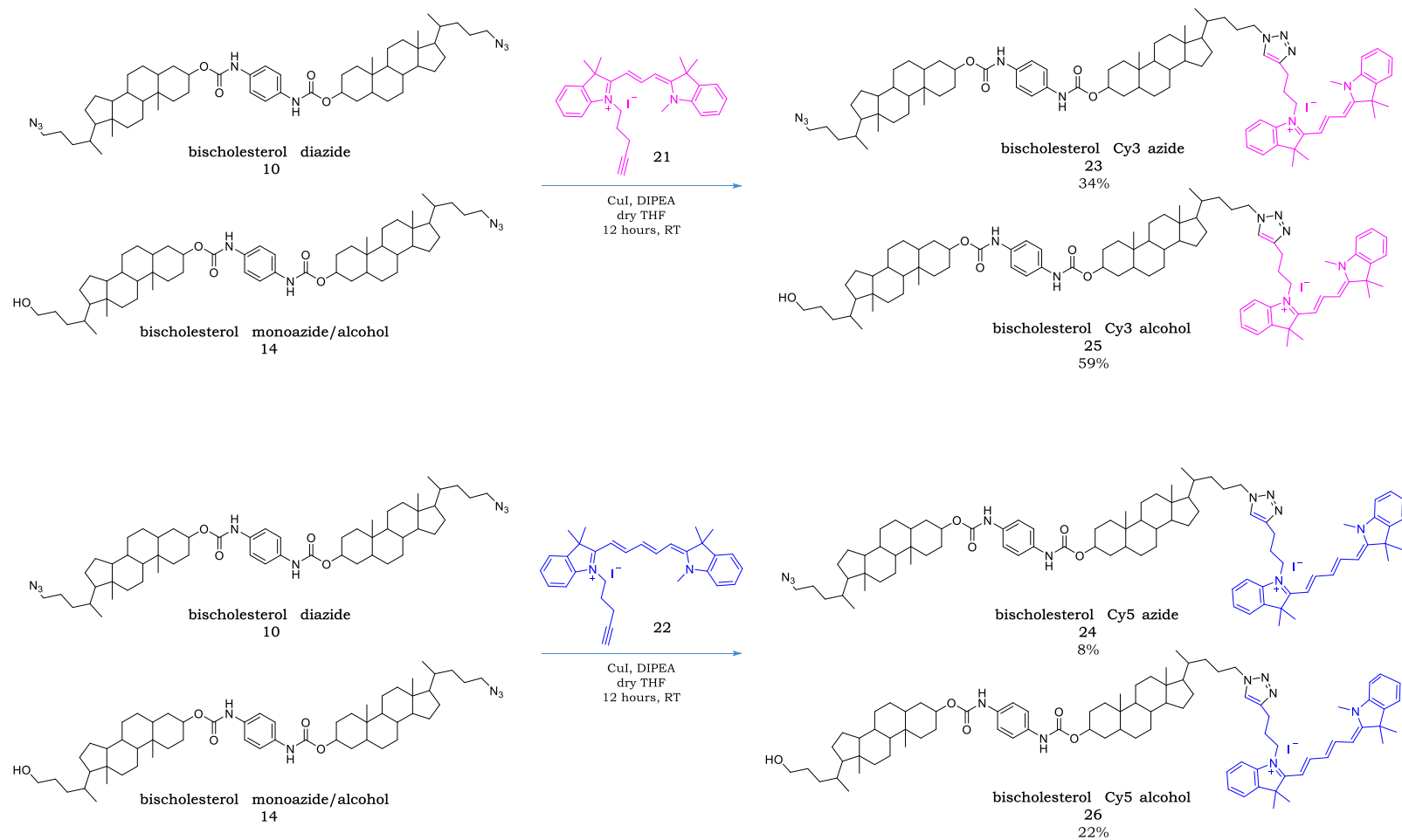
22

*Adapted from Brown<sup>10</sup>*

Under an inert atmosphere, a suspension of 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide **18** (100mg, 0.28mmol) and malondialdehyde bis(phenylimine) monohydrochloride **19** (475mg, 1.84mmol) in acetic anhydride (5ml) was heated to reflux for 30 minutes. During this time, a further 300mg (0.85mmol) of 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide **18** was added sequentially to the reaction in 100mg batches every 5 minutes. After 30 minutes, acetic anhydride was removed, the resulting green solid was redissolved in a solution of 1,2,3,3-tetramethyl-3H-indolinium iodide **17** (852mg, 2.82mmol) in dry pyridine (6ml), and the reaction was left stirring under argon overnight. Solvent was removed and the blue crude was purified by flash chromatography with a gradient eluent of 0 → 5% methanol in DCM. Product containing fractions were combined, and the solvent was evaporated to leave a blue solid which was further purified by preparative HPLC on a Kinetex C18 prep column using the method described in Table 5.3. The purification was monitored at 645nm and the desired product was found to elute at ~16 minutes. The product was collected, and the solvent removed by rotary evaporation and freeze drying to give the final product as a blue solid (229mg, 36%).

$^1\text{H}$  NMR (400MHz,  $\text{CD}_3\text{CN}$ )  $\delta$  8.13-8.05 (m, 2H), 7.48 (t,  $J$ = 7.0Hz, 2H), 7.44-7.37 (m, 2H), 7.29-7.22 (m, 4H), 6.52 (t,  $J$ = 12.5Hz, 1H), 6.23 (dd,  $J$ = 13.8, 12.0Hz, 2H), 4.08 (t,  $J$ = 7.5Hz, 2H), 3.55 (s, 3H), 2.37-2.32 (m, 3H), 2.02-1.95 (m, 3H), 1.68 (s, 12H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CD}_3\text{CN}$ )  $\delta$  175.5 (C), 173.9(C), 155.1 (HC=C), 154.6 (HC=C), 143.9 (Ar-C), 143.3 (Ar-C), 142.3 (Ar-C), 142.2 (Ar-C), 129.4 (Ar-CH), 129.4 (Ar-CH), 126.1 (Ar-CH), 125.7 (Ar-CH), 125.6 (HC=C), 123.2 (Ar-CH), 123.1 (Ar-CH), 111.9 (Ar-CH), 111.5 (Ar-CH), 104.4 (HC=C), 103.5 (HC=C), 84.0 ( $\text{C}\equiv\text{CH}$ ), 71.1 ( $\text{HC}\equiv\text{C}$ ), 50.2 (C), 50.0 (C), 43.6 ( $\text{CH}_2$ ), 32.0 ( $\text{NCH}_3$ ), 27.7 ( $\text{CH}_3$ ), 27.5 ( $\text{CH}_3$ ), 26.7 ( $\text{CH}_2$ ), 16.3 ( $\text{CH}_2$ ). MS (TOF-ESI +ve) ( $m/z$ ): 435.3  $[\text{M-I}]^+$ .

## 5.18 Synthesis of bischolesterol dye conjugates



**Scheme 5.7** Synthesis of bischolesterol dye conjugates.

#### 5.18.1 General procedure

Equimolar equivalents of the bischolesterol azide derivative and the corresponding cyanine dye were dissolved in dry THF in an inert atmosphere. DIPEA and copper iodide were then added, and the reaction was left stirring at room temperature for 12 hours. The reaction was quenched with water, and the product extracted with DCM. Organics were combined and dried over magnesium sulphate which was subsequently removed by filtration along with undissolved copper iodide. The solvent was removed to leave crude product which was then purified by either flash chromatography, HPLC or both techniques.

#### 5.18.2 HPLC purification of bischolesterol dye conjugates

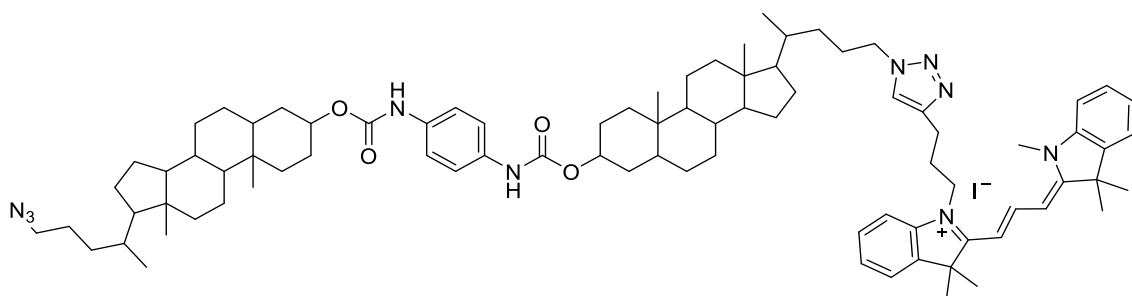
HPLC purification of the conjugates was performed on a Kinetex C18 prep column from Phenomenex, following the method described in the table below:

**Table 5.4** Solvent gradients used in the HPLC purification of bischolesterol dye conjugates.

Time/mins	Water + 0.1% TFA/%	Methanol + 0.1% TFA/%
0	10	90
10	10	90
20	0	100
30	0	100
30.1	10	90
35	10	90

HPLC purification of Cy3 containing products were monitored at 260nm and 550nm whereas Cy5 containing products were monitored at 260nm and 645nm. Desired products eluted at ~25 minutes.

### 5.18.3 Synthesis of bischolesterol Cy3 azide



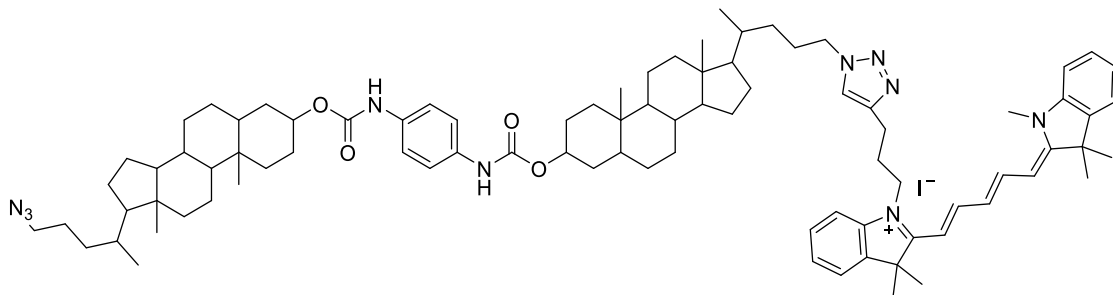
23

Bischolesterol Cy3 azide **23** was synthesised in accordance to the general procedure, using bischolesterol diazide **10** (39mg, 0.04mmol), Cy3 monoalkyne **21** (22.6mg, 0.04mmol), DIPEA (7 $\mu$ l, 0.004mmol), and copper iodide (4mg, 0.02mmol) in dry THF (5ml). The crude was purified by the HPLC purification procedure described in section 5.18.2 to yield the final product as a pink solid (21mg, 34%).

$^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  8.38 (t,  $J$ = 13.2Hz, 1H), 7.89 (s, 1H), 7.43-7.24 (m, 12H), 7.12 (d,  $J$ = 7.9Hz, 1H), 6.58-6.51 (m, 4H), 4.73-4.66 (m, 2H), 4.37-4.26 (m, 2H), 4.17 (s, 2H), 3.66 (s, 3H), 3.29-3.17 (m, 2H), 3.04 (s, 2H), 2.26 (s, 2H), 1.99-0.83 (m, 84H), 0.65 (s, 3H), 0.61 (s, 3H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  174.5 (C), 174.2 (C), 153.5 (C=O), 150.7 (HC=C), 142.8 (Ar-C), 142.0 (Ar-C), 140.5 (Ar-C), 140.5 (Ar-C), 133.8 (Ar-C), 129.4 (Ar-CH), 129.1 (Ar-CH), 125.8 (Ar-CH), 125.6 (Ar-CH), 122.2 (HC=C), 122.2 (Ar-CH), 119.6 (Ar-CH), 111.7 (Ar-CH), 110.9 (Ar-CH), 104.0 (HC=C), 103.9 (HC=C), 75.6 (CH), 56.7 (CH), 56.6 (CH), 56.2 (CH), 55.9 (CH), 52.1 ( $\text{CH}_2$ ), 51.7 ( $\text{CH}_2$ ), 49.3 (C), 49.1 (C), 44.1 ( $\text{CH}_2$ ), 42.9 (C), 42.1 (CH), 40.6 (CH), 40.3 ( $\text{CH}_2$ ), 36.0 (CH), 35.6 (CH), 35.4 (CH), 35.2 ( $\text{CH}_2$ ), 34.7 (C), 33.1 ( $\text{CH}_2$ ), 32.7 ( $\text{CH}_2$ ), 31.4 ( $\text{NCH}_3$ ), 29.8 ( $\text{CH}_2$ ), 28.4 ( $\text{CH}_2$ ), 28.2 ( $\text{CH}_3$ ), 27.2 ( $\text{CH}_2$ ), 27.1 ( $\text{CH}_2$ ), 26.8 ( $\text{CH}_2$ ), 26.5 ( $\text{CH}_2$ ), 25.7 ( $\text{CH}_2$ ), 24.3 ( $\text{CH}_2$ ), 23.5 ( $\text{CH}_3$ ),

22.8 (CH<sub>2</sub>), 21.0 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>), 18.6 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): [M-I]<sup>+</sup> calcd for C<sub>85</sub>H<sub>119</sub>N<sub>10</sub>O<sub>4</sub>, 1344.9416; found 1344.9415.

#### 5.18.4 Synthesis of bischolesterol Cy5 azide



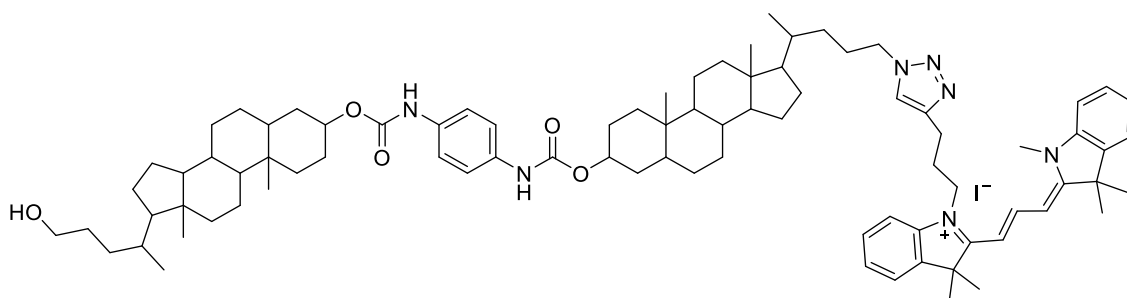
24

Bischolesterol Cy5 azide **24** was synthesised in accordance to the general procedure, using bischolesterol diazide **10** (50mg, 0.05mmol), Cy5 monoalkyne **22** (30mg, 0.05mmol), DIPEA (9μl, 0.005mmol), and copper iodide (5mg, 0.025mmol) in dry THF (5ml). The crude was purified by flash chromatography with an eluent system of 95:5 DCM:MeOH and further purified by the HPLC purification procedure described in section 5.18.2 to yield the final product as a blue solid (13mg, 8%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 7.84-7.77 (m, 3H), 7.39-7.21 (m, 15H), 7.07 (d, J= 7.9Hz, 1H), 6.67 (s, 1H), 6.54 (s, 1H), 6.48 (s, 1H), 6.35 (d, J= 13.3Hz, 1H), 6.19 (d, J= 13.0Hz, 1H), 4.72-4.67 (m, 2H), 4.32 (s, 2H), 4.14 (s, 2H), 3.58 (s, 3H), 3.28-3.18 (m, 2H), 3.00 (s, 2H), 2.24 (s, 2H), 1.99-0.83 (m, 86H), 0.65 (s, 3H), 0.62 (s, 3H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 173.3 (C), 172.9 (C), 153.6 (HC=C), 153.6 (C=O), 152.9 (HC=C), 145.1 (C), 142.9 (Ar-C), 142.1 (Ar-C), 141.1 (Ar-C), 140.8 (Ar-C), 133.9 (Ar-C), 129.1 (Ar-CH), 128.8 (Ar-CH), 126.5 (HC=C), 125.6 (Ar-CH), 125.2 (Ar-CH), 122.2 (HC=C/Ar-CH), 119.7 (Ar-CH), 111.3 (Ar-CH),

110.4 (Ar-CH), 104.5 (HC=C), 103.7 (HC=C), 75.6 (CH), 56.7 (CH), 56.6 (CH), 56.3 (CH), 55.9 (CH), 52.1 (CH<sub>2</sub>), 51.4 (CH<sub>2</sub>), 49.5 (C), 49.2 (C), 44.0 (CH<sub>2</sub>), 42.9 (C), 42.1 (CH), 40.7 (CH), 40.3 (CH<sub>2</sub>), 36.0 (CH), 35.6 (CH), 35.4 (CH), 35.2 (CH<sub>2</sub>), 34.8 (C), 33.1 (CH<sub>2</sub>), 32.7 (CH<sub>2</sub>), 31.3 (NCH<sub>3</sub>), 29.8 (CH<sub>2</sub>), 28.4 (CH<sub>2</sub>), 28.1 (CH<sub>3</sub>), 27.2 (CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 26.9 (CH<sub>2</sub>), 26.5 (CH<sub>2</sub>), 25.7 (CH<sub>2</sub>), 24.3 (CH<sub>2</sub>), 23.5 (CH<sub>3</sub>), 22.4 (CH<sub>2</sub>), 21.0 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>), 18.7 (CH<sub>3</sub>), 12.2 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): [M-I]<sup>+</sup> calcd for C<sub>87</sub>H<sub>121</sub>N<sub>10</sub>O<sub>4</sub>, 1370.9572; found 1370.9568.

#### 5.18.5 Synthesis of bischolesterol Cy3 alcohol



25

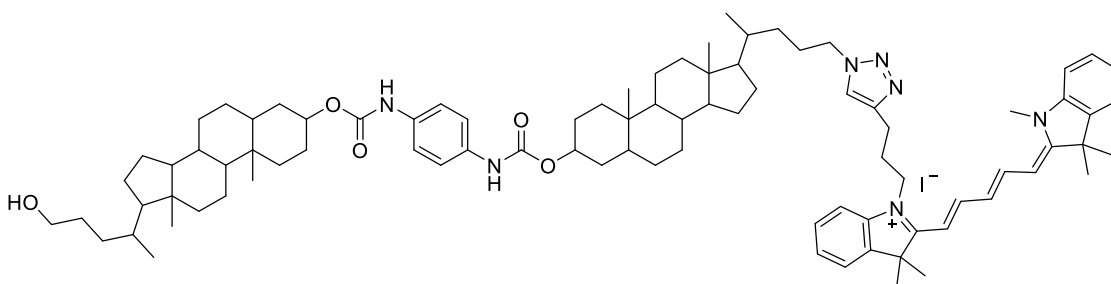
Bischolesterol Cy3 alcohol **25** was synthesised in accordance to the general procedure, using bischolesterol monoazide/alcohol **14** (132mg, 0.15mmol), Cy3 monoalkyne **21** (78mg, 0.15mmol), DIPEA (25μl, 0.015mmol), and copper iodide (138mg, 0.7mmol) in dry THF (15ml). The crude was purified by flash chromatography with an eluent system of 95:5 DCM:MeOH to give the final product as a pink solid (115mg, 59%).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ 8.43 (t, J= 13.4Hz, 1H), 7.89 (s, 1H), 7.41-7.12 (m, 13H), 7.13 (d, J= 7.9Hz, 1H), 6.73 (s, 2H), 4.72-4.65 (m, 2H), 4.35-4.20 (m, 4H), 3.80 (s, 1H), 3.67-3.57 (m, 2H), 3.23-3.18 (m, 2H), 2.34-2.27 (m, 2H), 1.99-0.89 (m, 89H), 0.64 (s, 3H), 0.62 (s, 3H). <sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>) δ 174.1 (C), 173.7



(C), 153.5 (C=O), 150.7 (HC=C), 146.5 (C), 142.8 (Ar-C), 142.0 (Ar-C), 140.5 (Ar-C), 133.7 (Ar-C), 129.1 (Ar-CH), 128.9 (Ar-CH), 125.4 (Ar-CH), 125.3 (Ar-CH), 122.6 (HC=C), 122.1 (Ar-CH), 122.0 (Ar-CH), 119.6 (Ar-CH), 111.6 (Ar-CH), 110.8 (Ar-CH), 105.0 (HC=C), 104.7 (HC=C), 75.4 (CH), 63.5 (CH<sub>2</sub>), 56.5 (CH), 56.5 (CH), 56.2 (CH), 55.8 (CH), 50.8 (CH<sub>2</sub>), 49.0 (C), 48.8 (C), 44.4 (CH<sub>2</sub>), 42.7 (C), 42.0 (CH), 40.5 (CH), 40.2 (CH<sub>2</sub>), 35.8 (CH), 35.6 (CH), 35.3 (CH), 35.1 (CH<sub>2</sub>), 34.6 (C), 32.8 (NCH<sub>3</sub>), 32.6 (CH<sub>2</sub>), 31.9 (CH<sub>2</sub>), 29.7 (CH<sub>2</sub>), 29.4 (CH<sub>2</sub>), 28.3 (CH<sub>2</sub>), 28.2 (CH<sub>3</sub>), 27.2 (CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 27.0 (CH<sub>2</sub>), 26.4 (CH<sub>2</sub>), 24.2 (CH<sub>2</sub>), 23.4 (CH<sub>3</sub>), 23.1 (CH<sub>2</sub>), 20.9 (CH<sub>2</sub>), 18.7 (CH<sub>3</sub>), 18.7 (CH<sub>3</sub>), 12.1 (CH<sub>3</sub>). HRMS (TOF-ESI +ve) (*m/z*): [M-I]<sup>+</sup> calcd for C<sub>85</sub>H<sub>120</sub>N<sub>7</sub>O<sub>5</sub>, 1318.9351; found 1318.9374.

#### 5.18.6 Synthesis of bischolesterol Cy5 alcohol

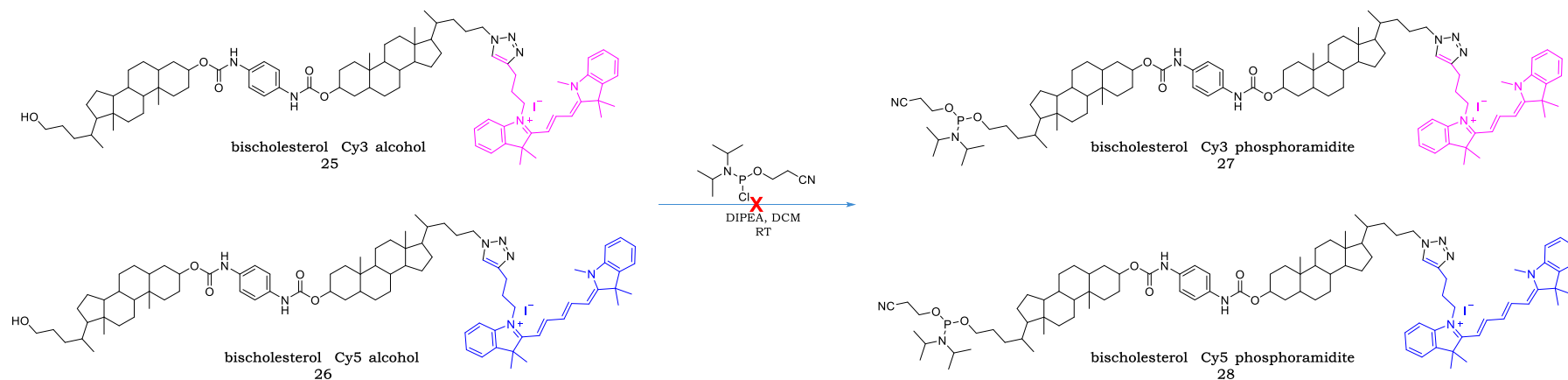


26

Bischolesterol Cy5 alcohol **26** was synthesised in accordance to the general procedure, using bischolesterol monoazide/alcohol **14** (223mg, 0.25mmol), Cy5 monoalkyne **22** (138mg, 0.25mmol), DIPEA (40μl, 0.25mmol), and copper iodide (230mg, 1.23mmol) in dry THF (20ml). The crude was purified by flash chromatography with an eluent system of 95:5 DCM:MeOH to give the final product as a blue solid (71.5mg, 22%).

$^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  8.06-7.97 (m, 2H), 7.76 (s, 1H), 7.38-7.30 (m, 8H), 7.23-7.18 (m, 3H), 7.09 (d,  $J$  = 7.9Hz, 1H), 6.95-6.83 (m, 3H), 6.55 (d,  $J$  = 13.4Hz, 1H), 6.26 (d,  $J$  = 13.4Hz, 1H), 4.68 (s, 2H), 4.37-4.28 (m, 2H), 4.23 (s, 2H), 3.62 (s, 5H), 3.03 (t,  $J$  = 6.5Hz, 2H), 2.28-2.25 (m, 2H), 1.95-0.91 (m, 89H), 0.64 (s, 3H), 0.61 (s, 3H).  $^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ )  $\delta$  173.3 (C), 172.7 (C), 153.8 (HC=C), 153.5 (C=O), 153.0 (HC=C), 146.1 (C=C), 142.8 (Ar-C), 142.0 (Ar-C), 141.2 (Ar-C), 140.8 (Ar-C), 133.7 (Ar-C), 128.8 (Ar-CH), 128.6 (Ar-CH), 126.6 (HC=C), 125.3 (Ar-CH), 124.9 (Ar-CH), 122.2 (HC=C/Ar-CH), 119.5 (Ar-CH), 111.2 (Ar-CH), 110.3 (Ar-CH), 104.6 (HC=C), 103.6 (HC=C), 75.3 (CH), 63.5 ( $\text{CH}_2$ ), 56.5 (CH), 56.4 (CH), 56.2 (CH), 55.8 (CH), 50.8 ( $\text{CH}_2$ ), 49.5 (C), 49.1 (C), 44.1 ( $\text{CH}_2$ ), 42.7 (C), 42.0 (CH), 40.5 (CH), 40.2 ( $\text{CH}_2$ ), 35.8 (CH), 35.6 (CH), 35.3 (CH), 35.1 ( $\text{CH}_2$ ), 34.6 (C), 32.6 ( $\text{CH}_2$ ), 31.9 ( $\text{NCH}_3/\text{CH}_2$ ), 29.7 ( $\text{CH}_2$ ), 29.4 ( $\text{CH}_2$ ), 28.3 ( $\text{CH}_2$ ), 28.1 ( $\text{CH}_3$ ), 28.1 ( $\text{CH}_3$ ), 27.1 ( $\text{CH}_2$ ), 27.0 ( $\text{CH}_2$ ), 26.8 ( $\text{CH}_2$ ), 26.4 ( $\text{CH}_2$ ), 24.2 ( $\text{CH}_2$ ), 23.4 ( $\text{CH}_3$ ), 22.8 ( $\text{CH}_2$ ), 20.9 ( $\text{CH}_2$ ), 18.7 ( $\text{CH}_3$ ), 18.6 ( $\text{CH}_3$ ), 12.1 ( $\text{CH}_3$ ). HRMS (TOF-ESI +ve) ( $m/z$ ):  $[\text{M-I}]^+$  calcd for  $\text{C}_{87}\text{H}_{122}\text{N}_7\text{O}_5$ , 1344.9507; found 1344.9508.

## 5.19 Synthesis of bischolesterol dye phosphoramidites (failed)



**Scheme 5.8** Failed synthesis of bischolesterol dye phosphoramidites.

The assigned bischolesterol dye alcohol (20-130mg) was azeotroped with dry DCM (3 x 5ml) and then redissolved in dry DCM (5ml) for reaction. The solution was degassed for five minutes before and during the addition of DIPEA (2.5 eq.) and 2-cyanoethyldiisopropylchlorophosphoramidite (1.2 eq.), the latter of which was added dropwise to the reaction. The reaction was left stirring under argon until completion was shown *via* TLC, which was typically 2 hours. Once complete, the solution was diluted with degassed DCM (10ml) and washed with degassed sat. sodium bicarbonate solution (20ml). The organics were isolated, dried over sodium sulphate, and solvent removed to leave either a pink or blue solid. Analysis of all crudes by  $^{31}\text{P}$  NMR showed signals at ~14ppm, highlighting the presence of only oxidised phosphorus and not of the desired phosphoramidite product.

For repeated phosphitylation reactions, bischolesterol dye alcohols **25** and **26** were treated with EDTA prior to reaction. The compounds were dissolved in DCM (50ml) and washed with an aqueous EDTA solution (0.25M, 150ml). The organics were isolated, dried over magnesium sulphate, and the solvent removed to regenerate the bischolesterol dye alcohol for phosphitylation.

## 5.20 Synthesis and purification of transmembrane molecules

### 5.20.1 Strategy 1- coupling alkyne modified DNA to bischolesterol azide derivatives with copper(I) iodotriethylphosphite

*Adapted from Wilks<sup>12</sup>*

DMF (100ml) was degassed for at least 30 minutes before use. 1mM stock solutions of bischolesterol dye azides **23** & **24**, bischolesterol diazide **10**, and CuI.P(OEt<sub>3</sub>) in degassed DMF (1ml) were prepared. Alkyne modified DNA (200μM) in Milli-Q water (1ml) stocks were also prepared.

For 1ml conjugation reactions in 2ml vials was added:

- 750μl degassed DMF
- 100μl of the assigned azide stock
- 100μl of CuI.P(OEt<sub>3</sub>) stock
- 50μl of DNA alkyne stock

Reactions contained 10μM DNA alkyne, 100μM azide, and 100μM CuI.P(OEt<sub>3</sub>). Argon was bubbled into the solution throughout additions and the vial was immediately sealed after the final addition. Vials were then left shaking overnight. Solvent was removed on a speed vac and the residue redissolved in Milli-Q water (1ml). Solutions were then desalted with a NAP-10 column and concentrated in preparation for HPLC purification.

### 5.20.2 Strategy 2- coupling DBCO modified DNA to bischolesterol dye azides

**Excess bischolesterol dye azide reactions:** 100μl solutions were prepared in 2ml vials by adding 50μl DBCO modified DNA stock in Milli-Q water (150μM) to 50μl bischolesterol dye stock in DMF (1mM). The vials were then sealed and left

standing overnight in either a fridge, on a lab bench at room temperature, or in a heating block set at 50°C. Solvent was removed on a speed vac and the residue redissolved in Milli-Q water (1ml). Solutions were then desalted with a NAP-10 column and concentrated in preparation for HPLC purification.

**Excess DBCO modified DNA reactions:** 200µl solutions were prepared in 2ml vials. The vials were then sealed and left shaking at room temperature overnight.

**2x excess DNA recipe:** 100µl DBCO modified DNA stock (120µM), 99µl DMF, and 1µl bischolesterol dye stock (1mM)

**10x excess DNA recipe:** 100µl DBCO modified DNA stock (120µM), 95µl DMF, and 5µl bischolesterol dye stock (1mM)

### 5.20.3 Purification of transmembrane molecules

The following HPLC methods were used to successfully separate transmembrane molecules from unreacted DNA in strategy 1 and strategy 2 reactions. However, they were unable to separate the transmembrane molecules from unreacted bischolesterol dye azide, and significant product carryover was observed with various C18 columns. Samples were monitored at UV absorbances of 260nm for DNA, 550nm for Cy3 containing samples, and 645nm for Cy5 containing samples. 10-100µl of sample was injected.

**Table 5.5** Solvent gradients used in the purification of transmembrane molecules on the Kinetex C18 column.

Time/mins	HPLC water/%	methanol/%
0	100	0
30	0	100
40	0	100
40.1	100	0

45	100	0
----	-----	---

**Table 5.6** Solvent gradients used in the purification of transmembrane molecules on the F5 column.

Time/mins	HPLC water/%	methanol/%
0	100	0
10	100	0
30	0	100
40	0	100
40.1	100	0
45	100	0

The table below describes the HPLC purification of the transmembrane molecules using Thermo Fisher's Oligo RP column. This column was the only one found to separate all starting materials from the transmembrane molecule product without causing significant carryover of product.

**Table 5.7** Solvent gradients used in the purification of transmembrane molecules on the Thermo Oligo RP column.

Time/mins	0.1M TEAA in HPLC water/%	acetonitrile/%
0	95	5
10	80	20
15	80	20
20	0	100
25	0	100
25.1	95	5
30	95	5

## 5.21 List of references

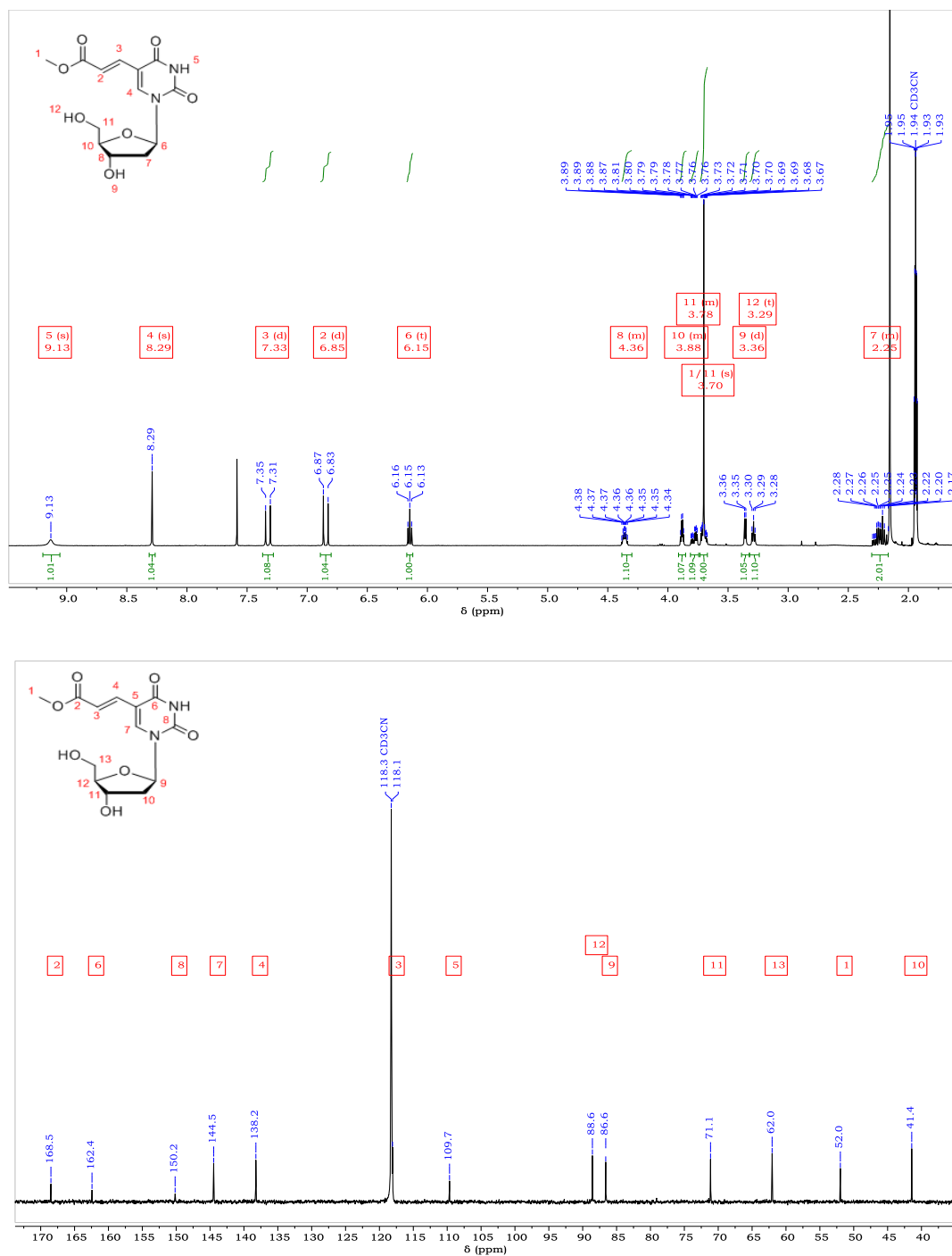
1. T. Ami and K. Fujimoto, *ChemBioChem*, 2008, **9**, 2071-2074.
2. G. Deglane, F. Morvan, F. Debart and J.-J. Vasseur, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 951-954.
3. L. Ma, M. Melegari, M. Colombini and J. T. Davis, *J. Am. Chem. Soc.*, 2008, **130**, 2938-2939.
4. Y. Kobuke and T. Nagatani, *J. Org. Chem.*, 2001, **66**, 5094-5101.
5. C. Goto, M. Yamamura, A. Satake and Y. Kobuke, *J. Am. Chem. Soc.*, 2001, **123**, 12152-12159.
6. R. Sharma, F. Majer, V. K. Peta, J. Wang, R. Keaveney, D. Kelleher, A. Long and J. F. Gilmer, *Biorg. Med. Chem.*, 2010, **18**, 6886-6895.
7. G. M. Sheldrick, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 2008, **64**, 112-122.
8. G. M. Sheldrick, *Acta Crystallogr., Sect. C: Struct. Chem.*, 2015, **71**, 3-8.
9. O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339-341.
10. M. Gerowska, L. Hall, J. Richardson, M. Shelbourne and T. Brown, *Tetrahedron*, 2012, **68**, 857-864.
11. E. A. Owens, N. Bruschi, J. G. Tawney and M. Henary, *Dyes Pigm.*, 2015, **113**, 27-37.
12. T. R. Wilks, J. Bath, J. W. de Vries, J. E. Raymond, A. Herrmann, A. J. Turberfield and R. K. O'Reilly, *ACS Nano*, 2013, **7**, 8561-8572.



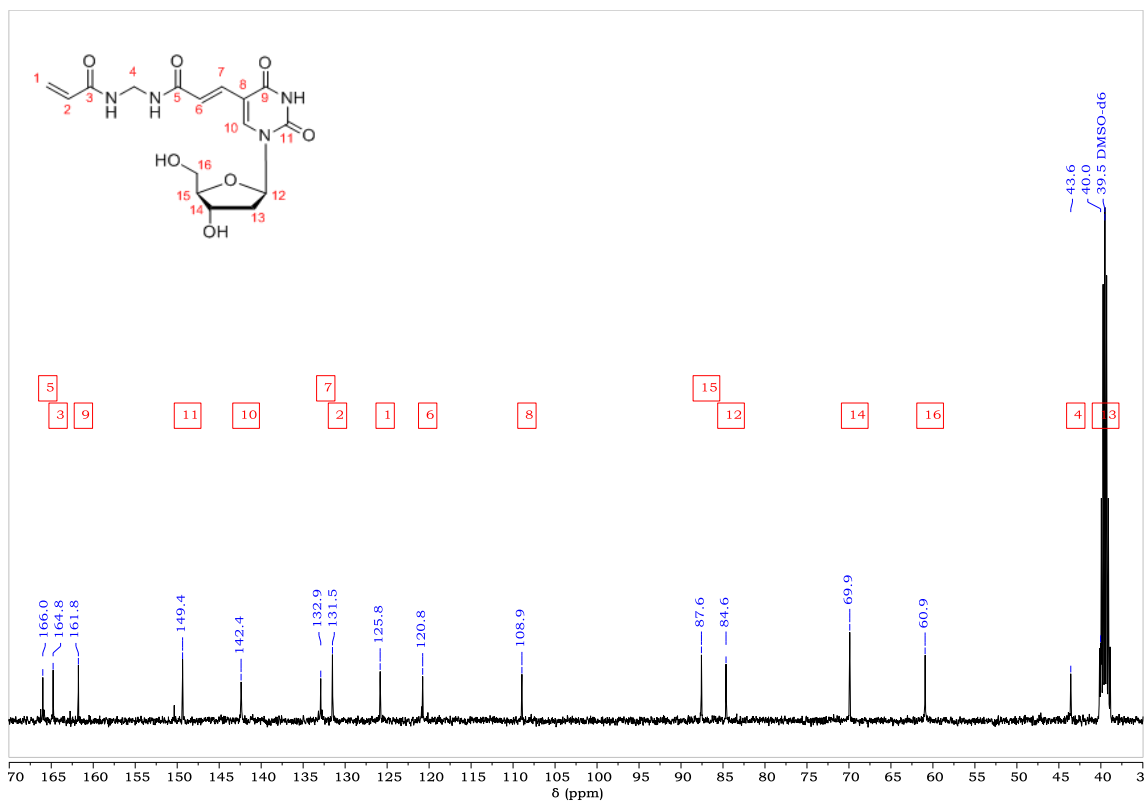
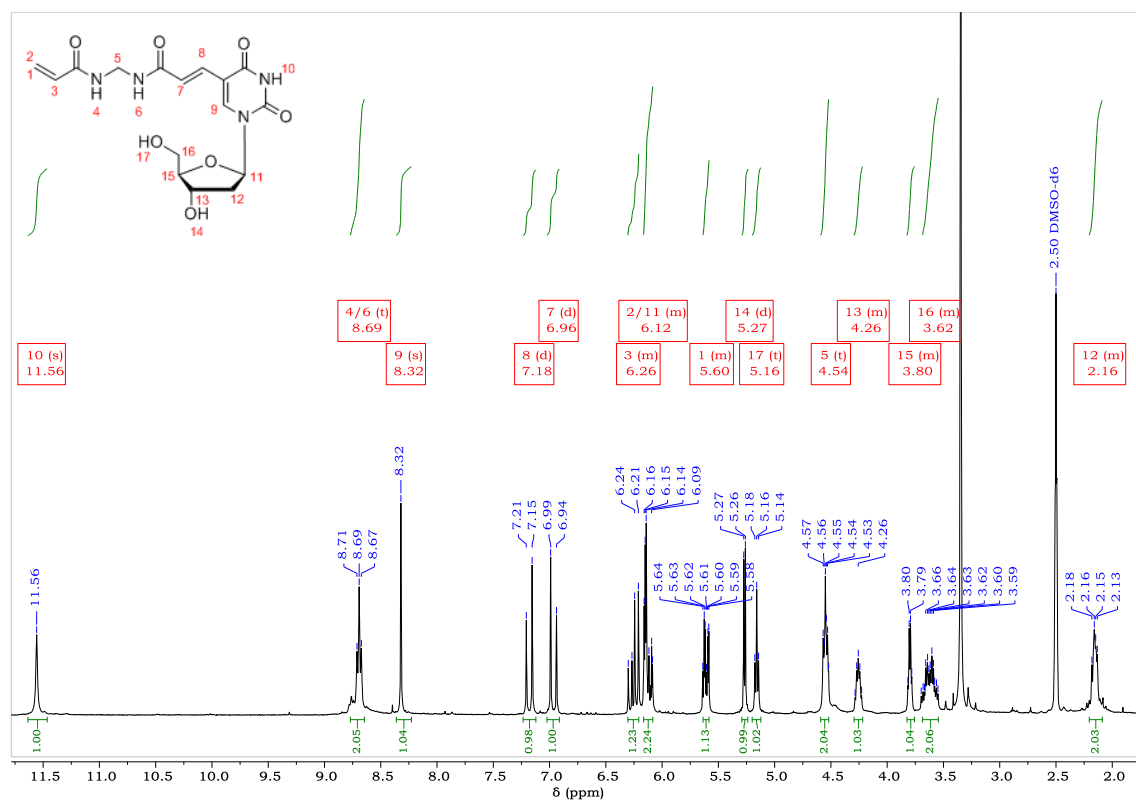
# Chapter 6 Appendix

## 6.1 NMR data for thymidine acrylate, thymidine acrylamide, and thymidine acrylamide derivatives

### 6.1.1 Compound 1- thymidine acrylate $^1\text{H}$ and $^{13}\text{C}$ NMR

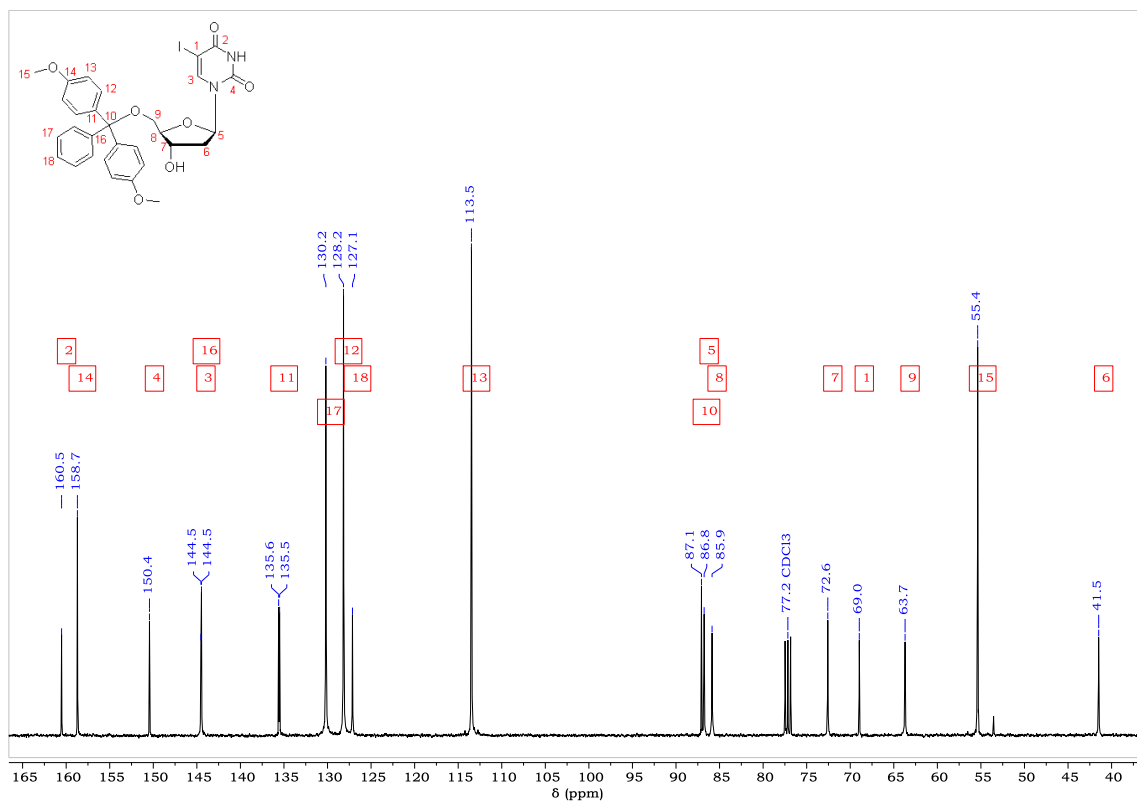
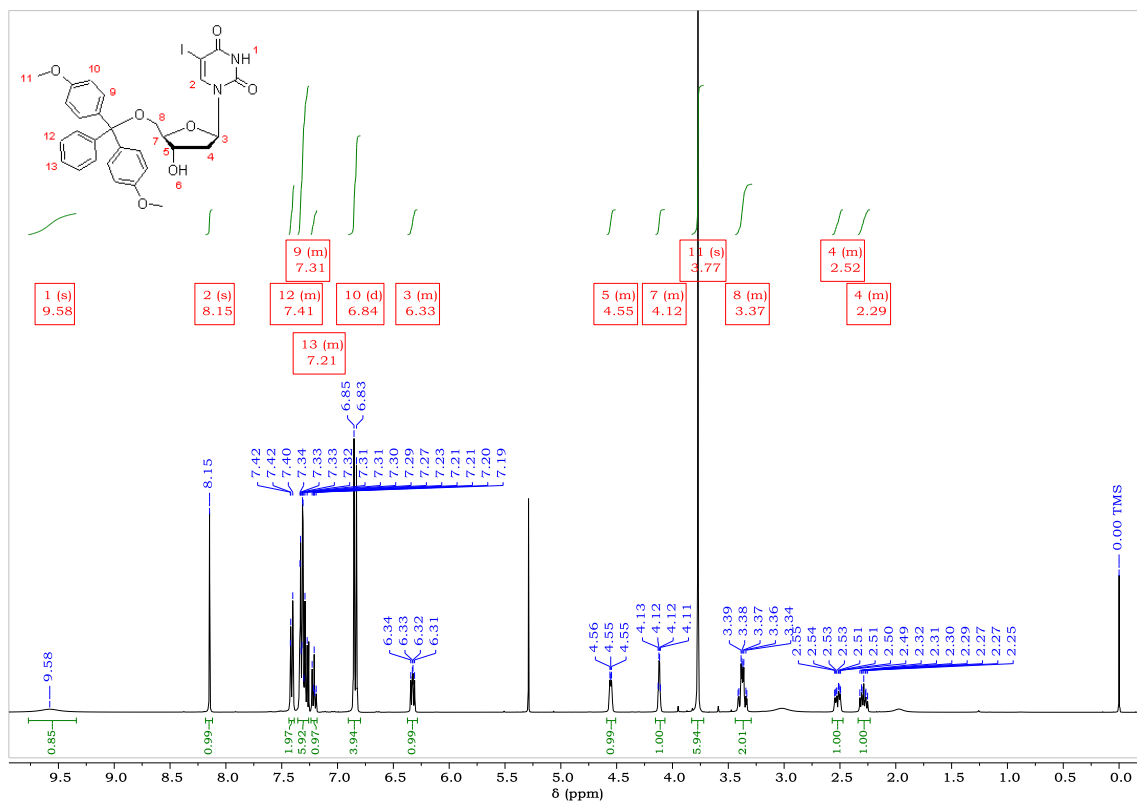


## 6.1.2 Compound **2**- thymidine acrylamide <sup>1</sup>H and <sup>13</sup>C NMR

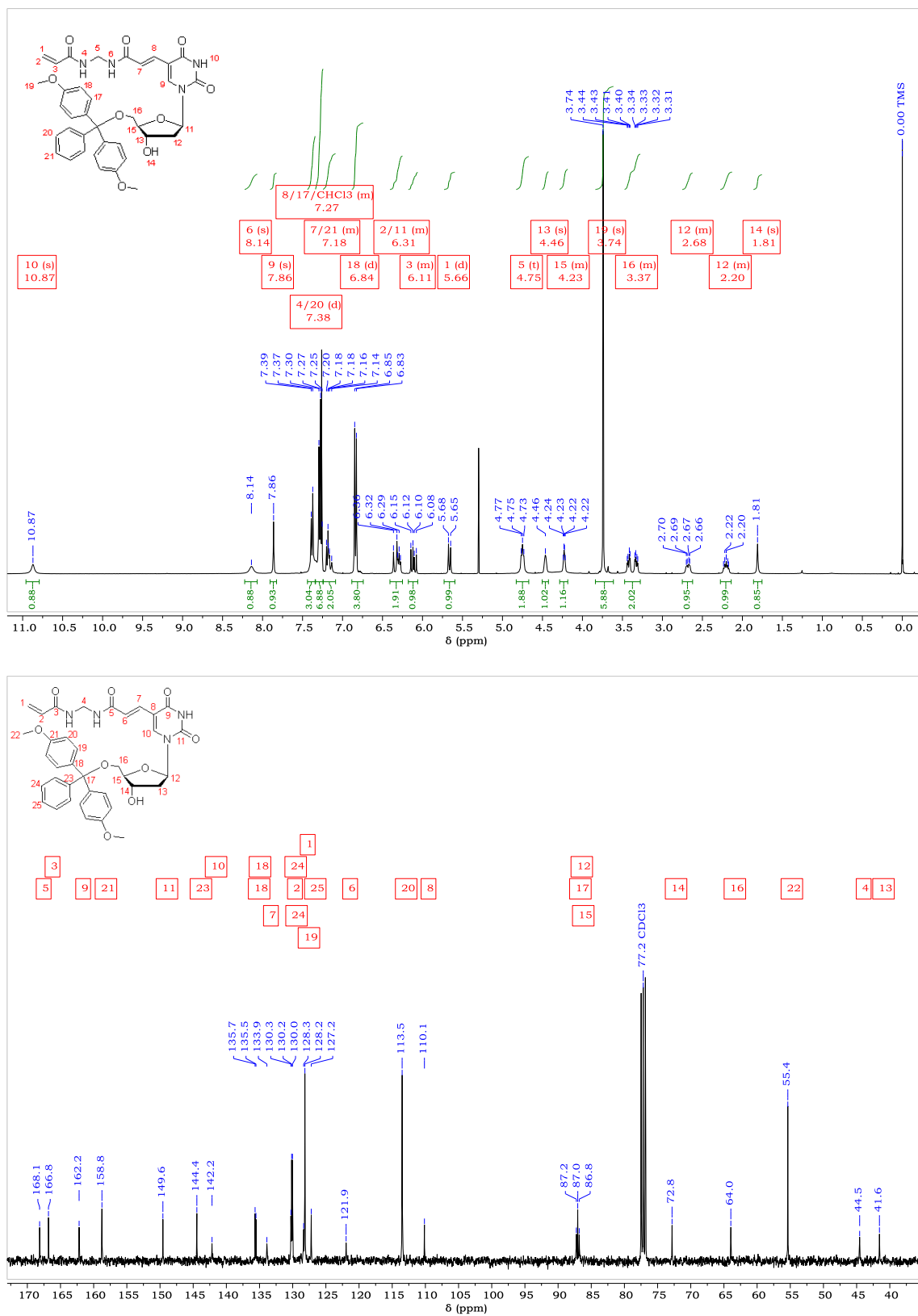


### 6.1.3 Compound **3**- 5'-O-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine

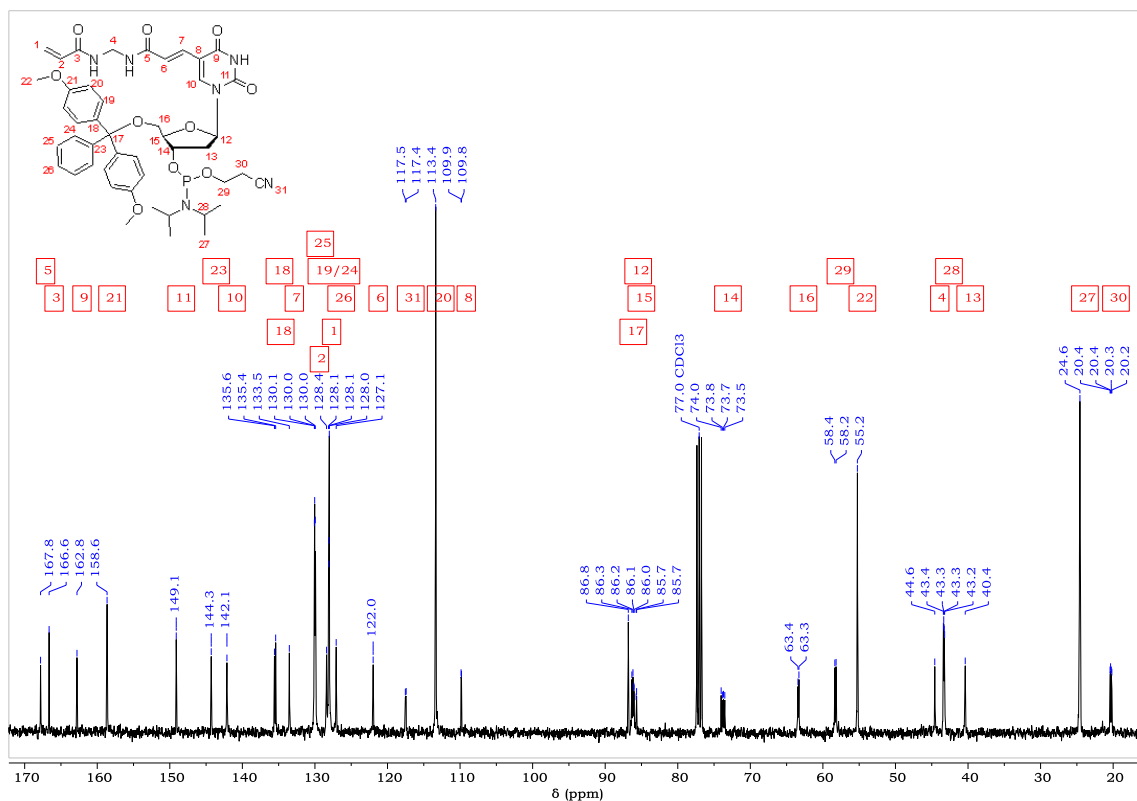
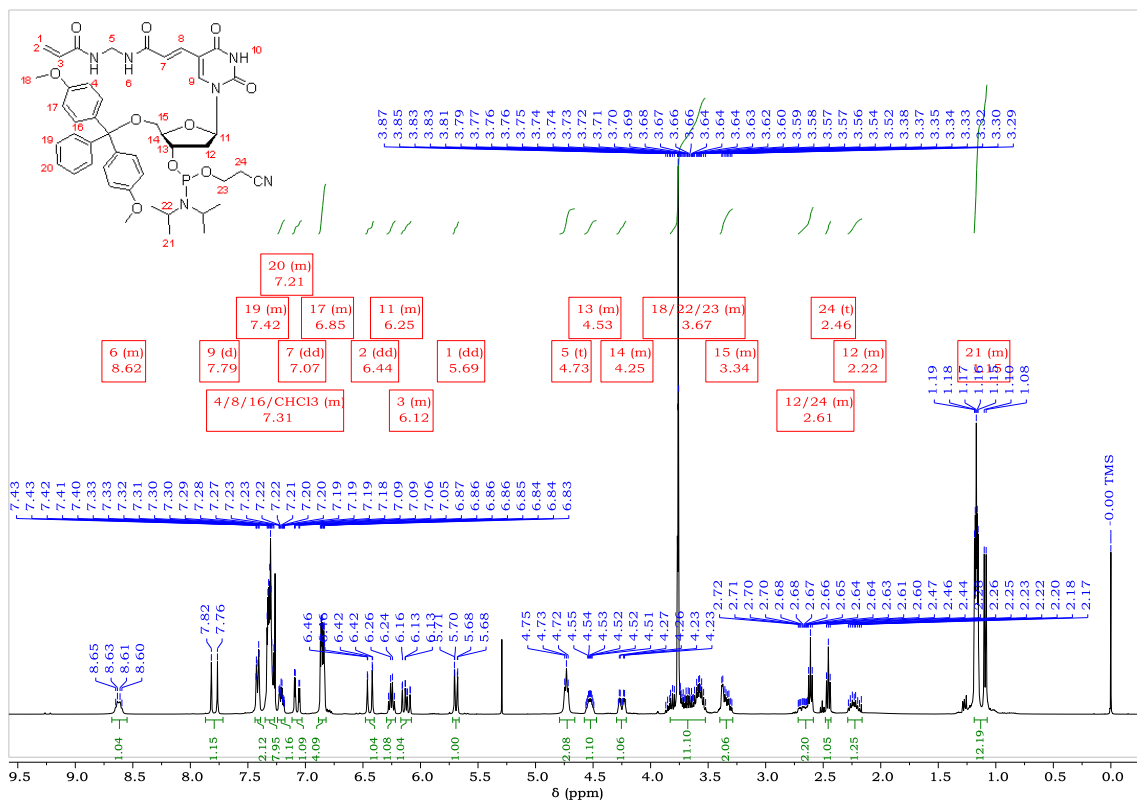
#### <sup>1</sup>H and <sup>13</sup>C NMR

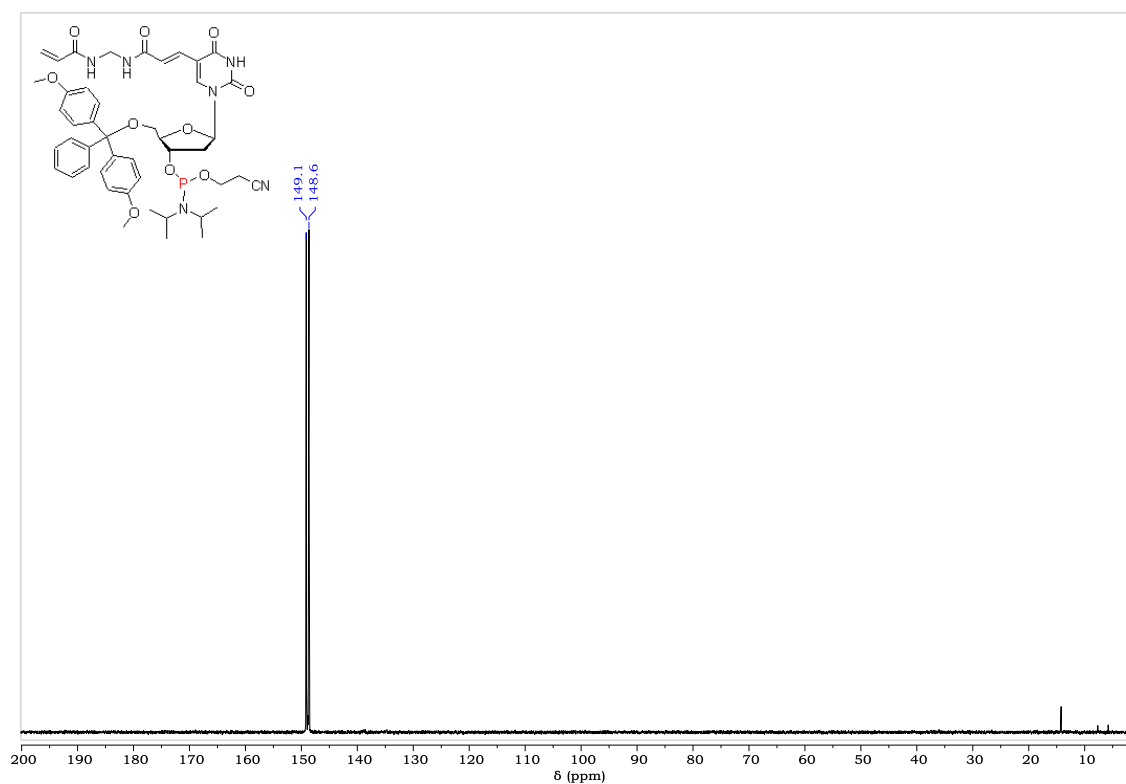


6.1.4 Compound **4**- 5'-O-(4,4'-dimethoxytrityl)-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine <sup>1</sup>H and <sup>13</sup>C NMR



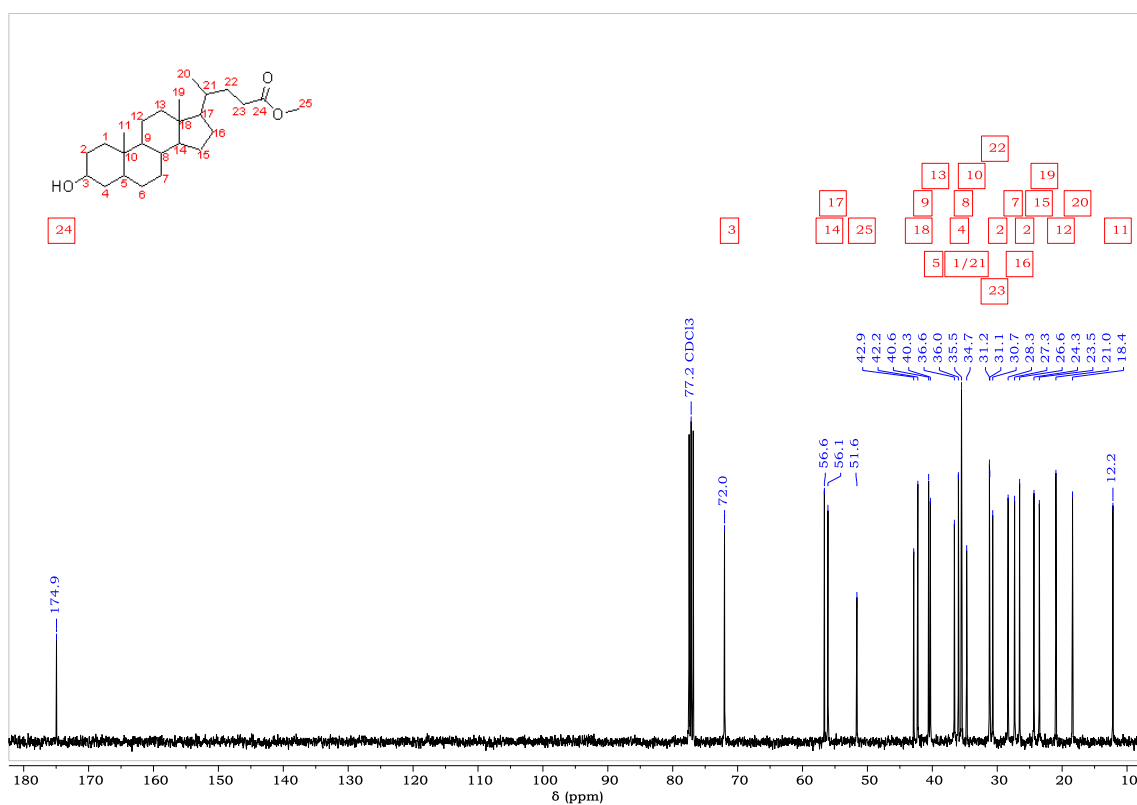
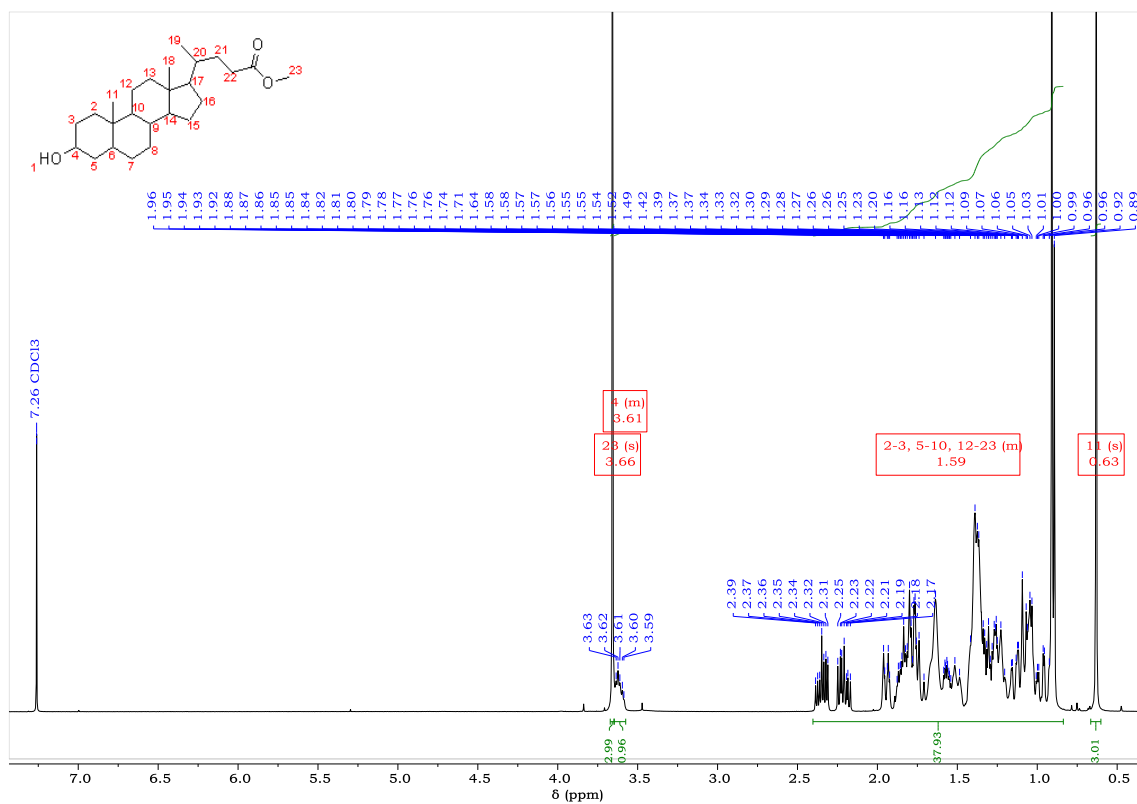
6.1.5 Compound **5**- 5'-O-(4,4'-dimethoxytrityl)-3'-O-[2-cyano ethoxy-(N,N'-diisopropylamino)-phosphino]-5-(N,N'-methylenebisacrylamido)-2'-deoxyuridine <sup>1</sup>H, <sup>13</sup>C, and <sup>31</sup>P NMR



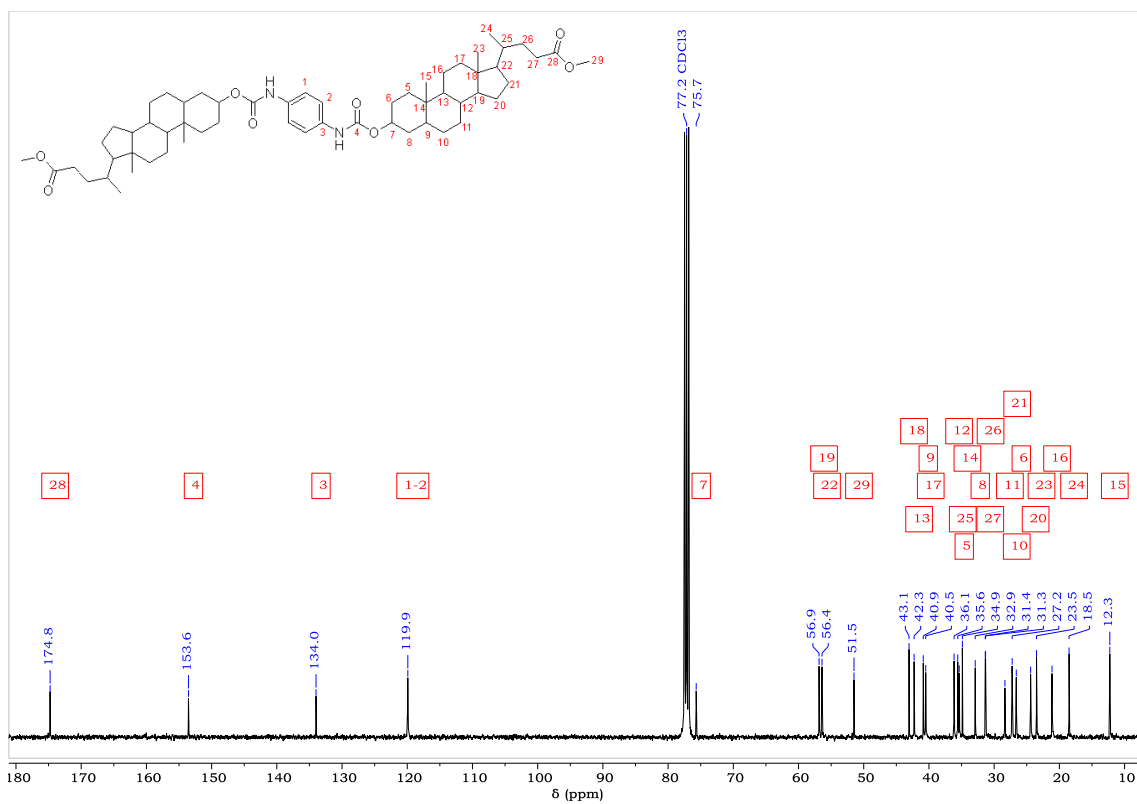
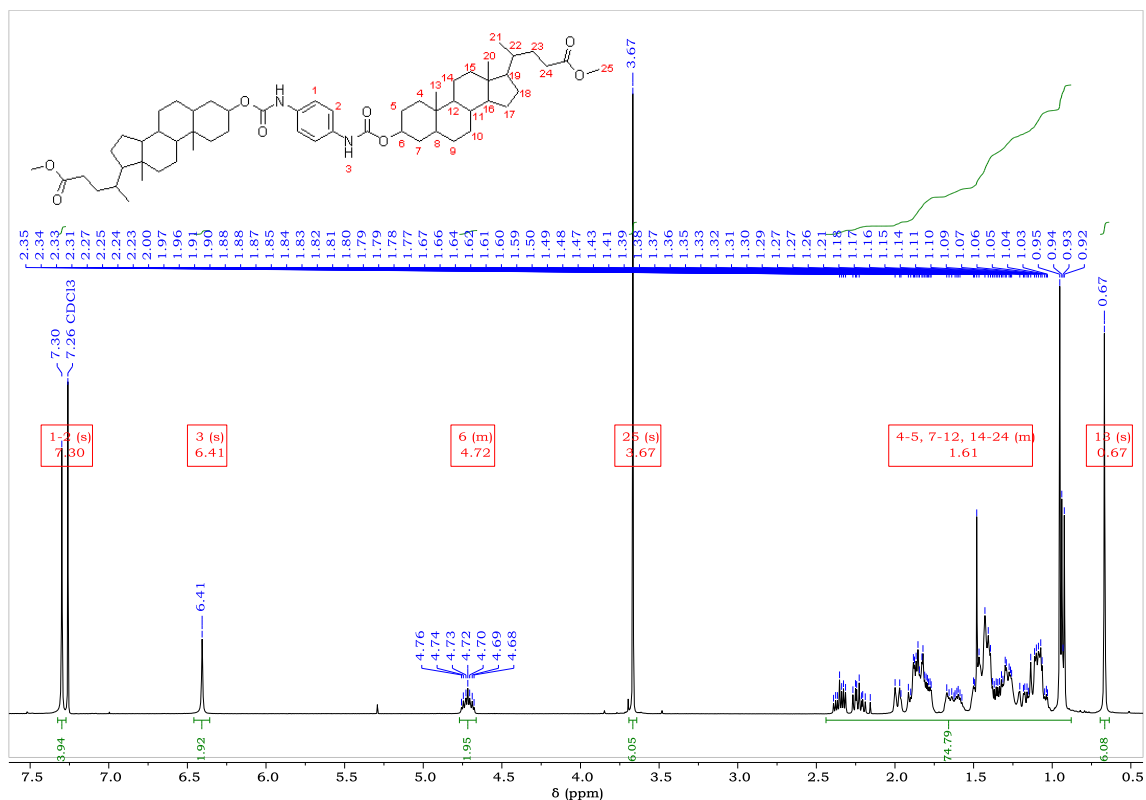


## 6.2 NMR data for bischolesterol derivatives

### 6.2.1 Compound **6**- lithocholic acid methyl ester <sup>1</sup>H and <sup>13</sup>C NMR

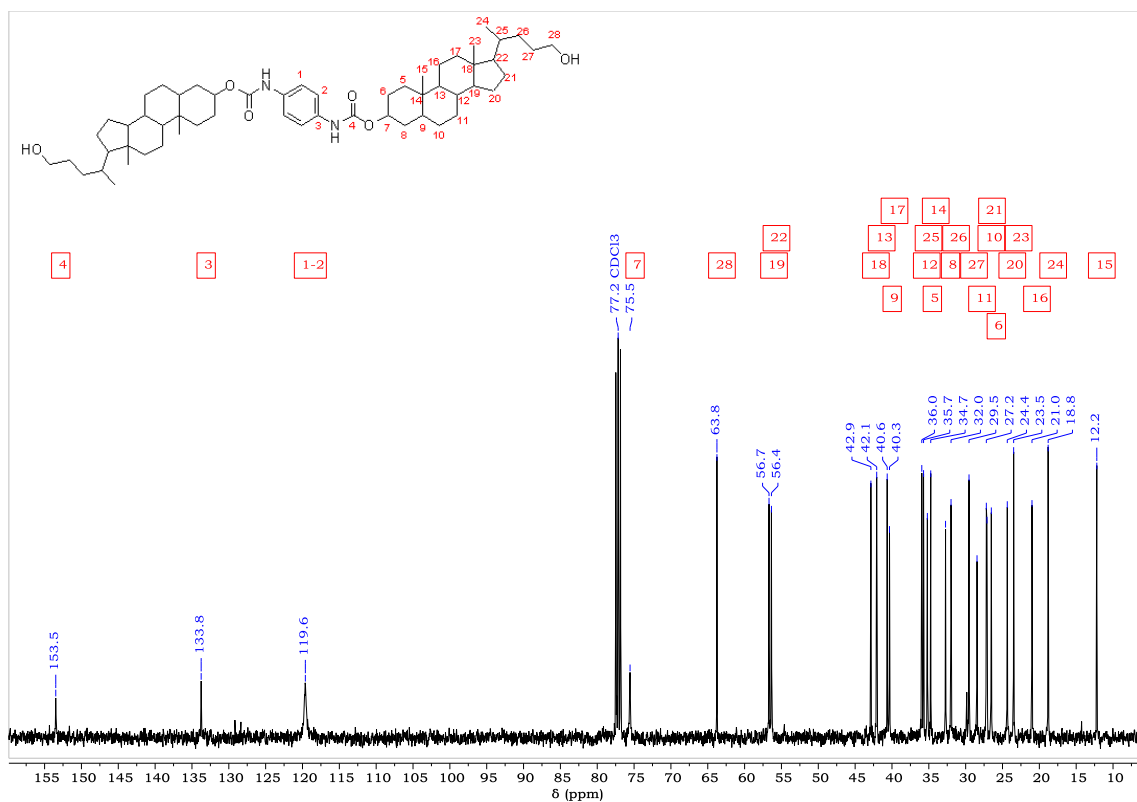
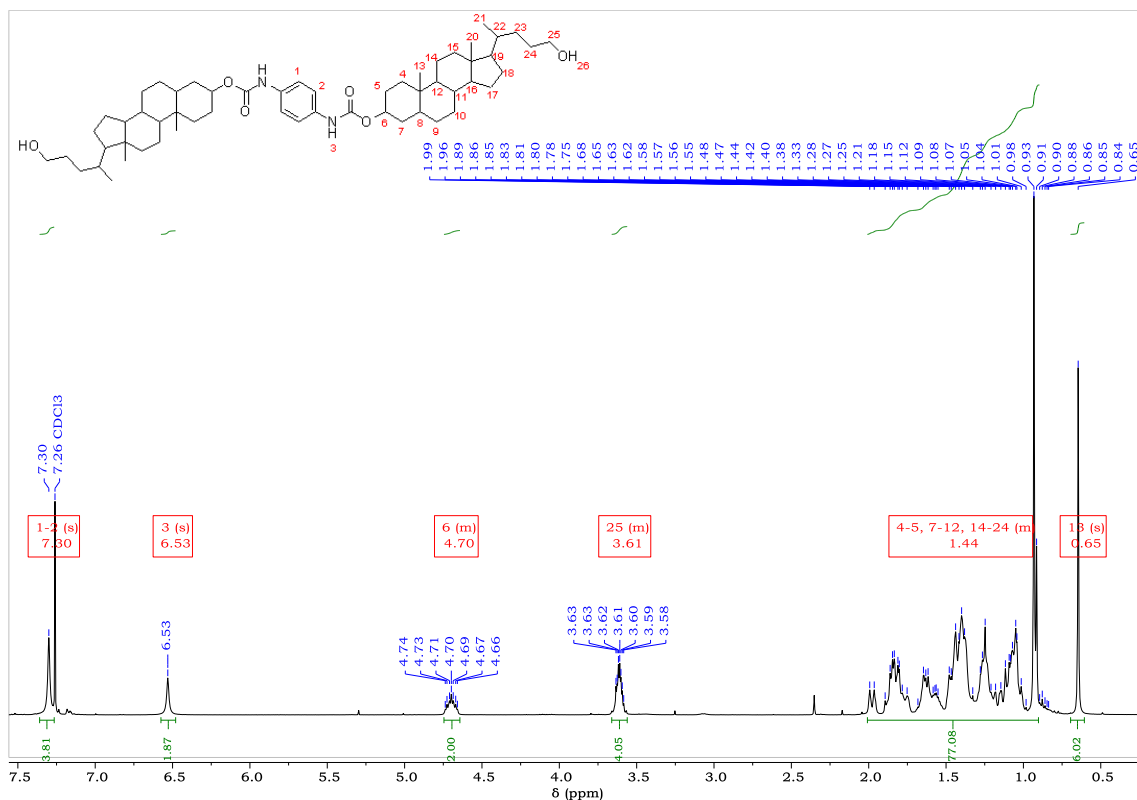


## 6.2.2 Compound 7- bischolesterol diester $^1\text{H}$ and $^{13}\text{C}$ NMR

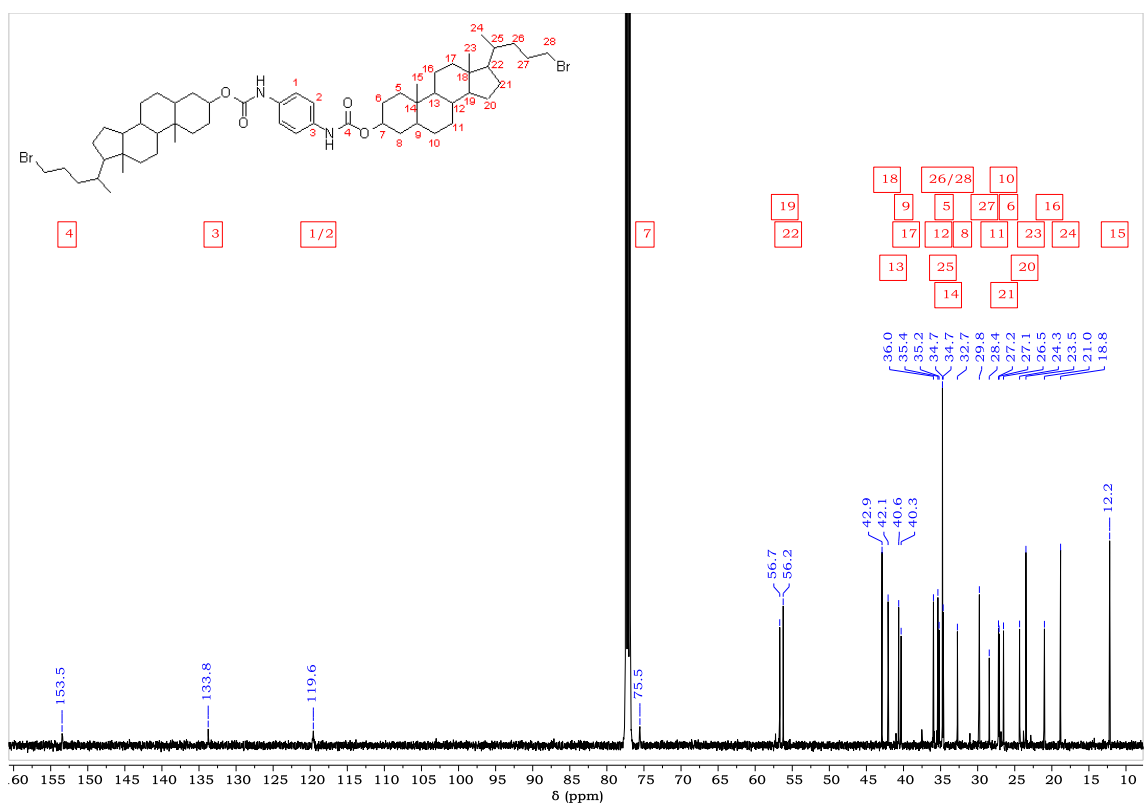
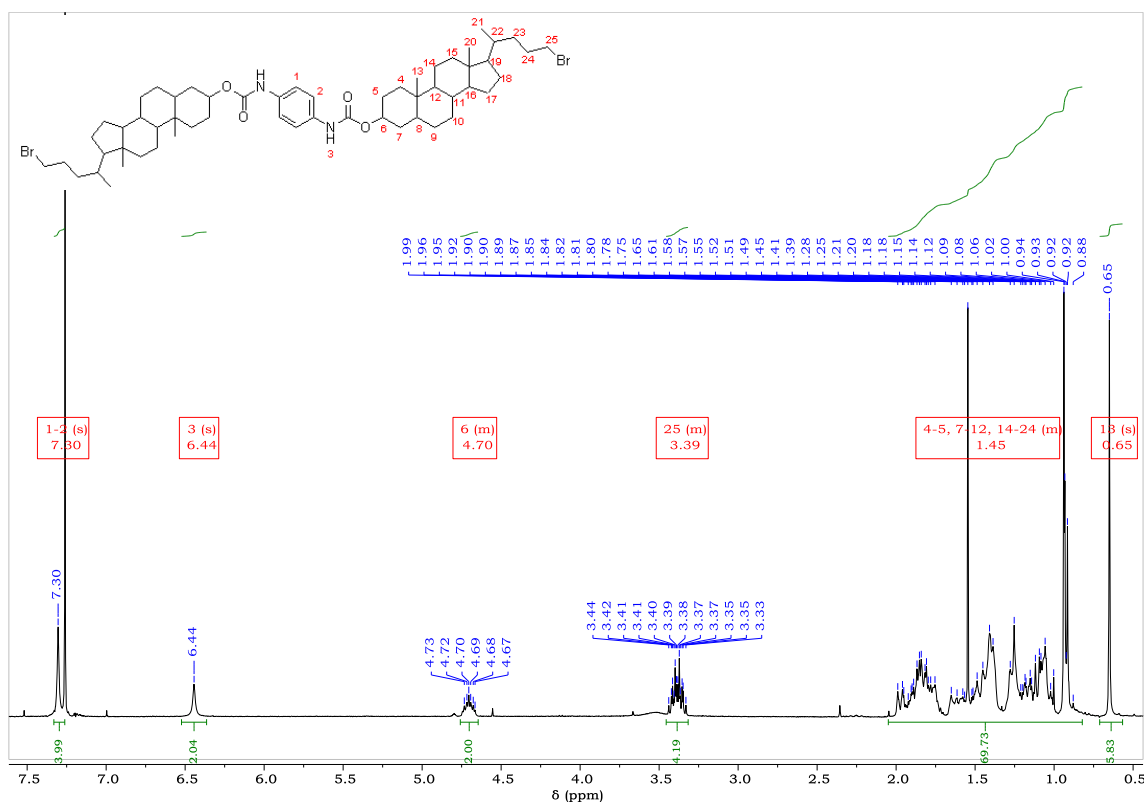




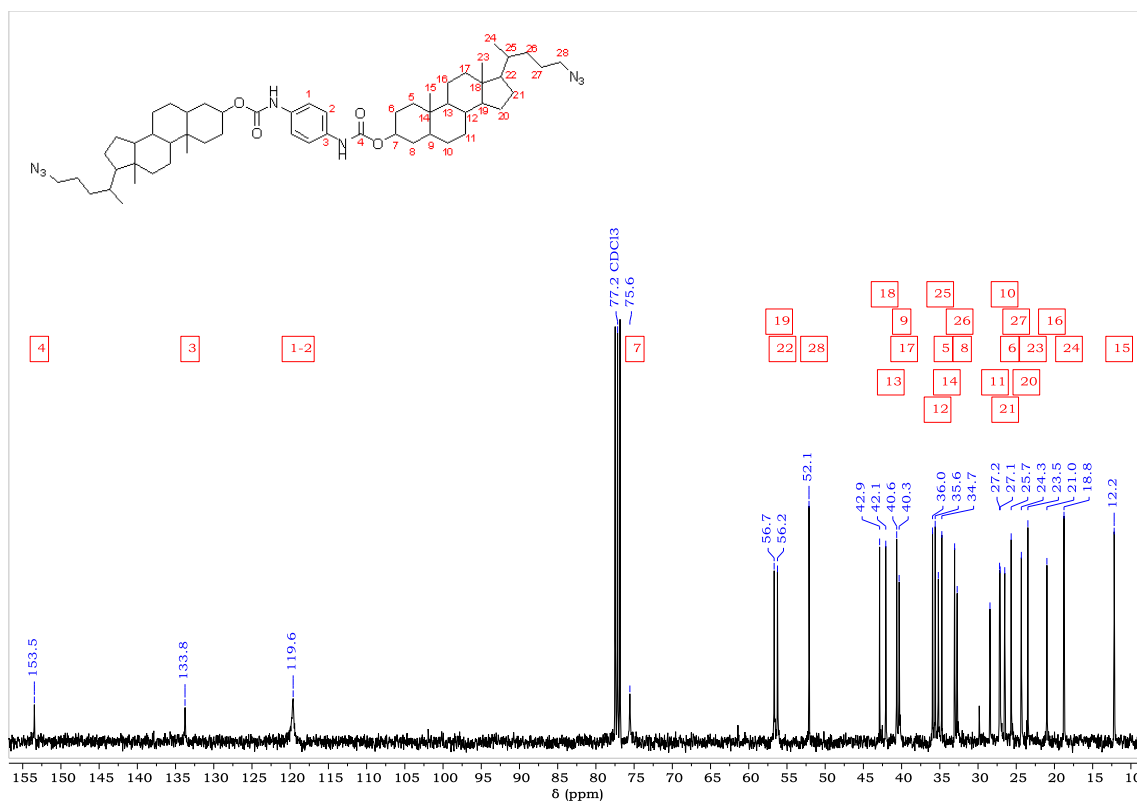
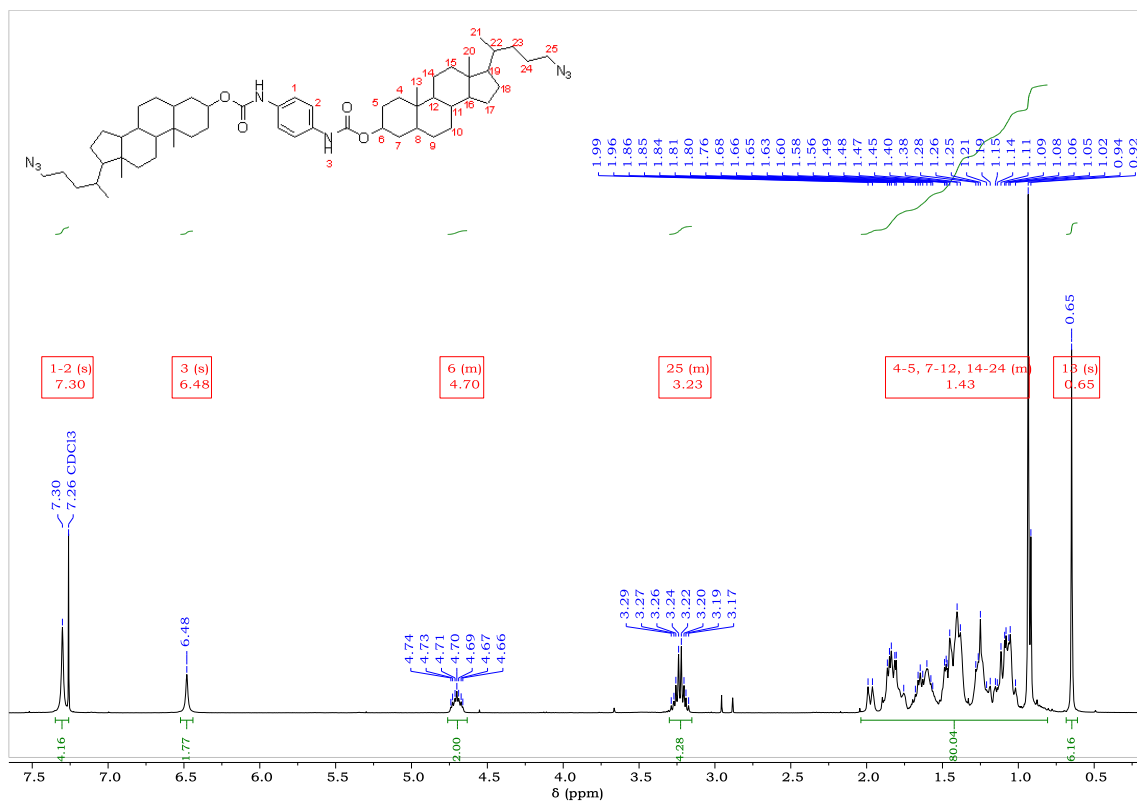
### 6.2.3 Compound 8- bischolesterol dialcohol $^1\text{H}$ and $^{13}\text{C}$ NMR



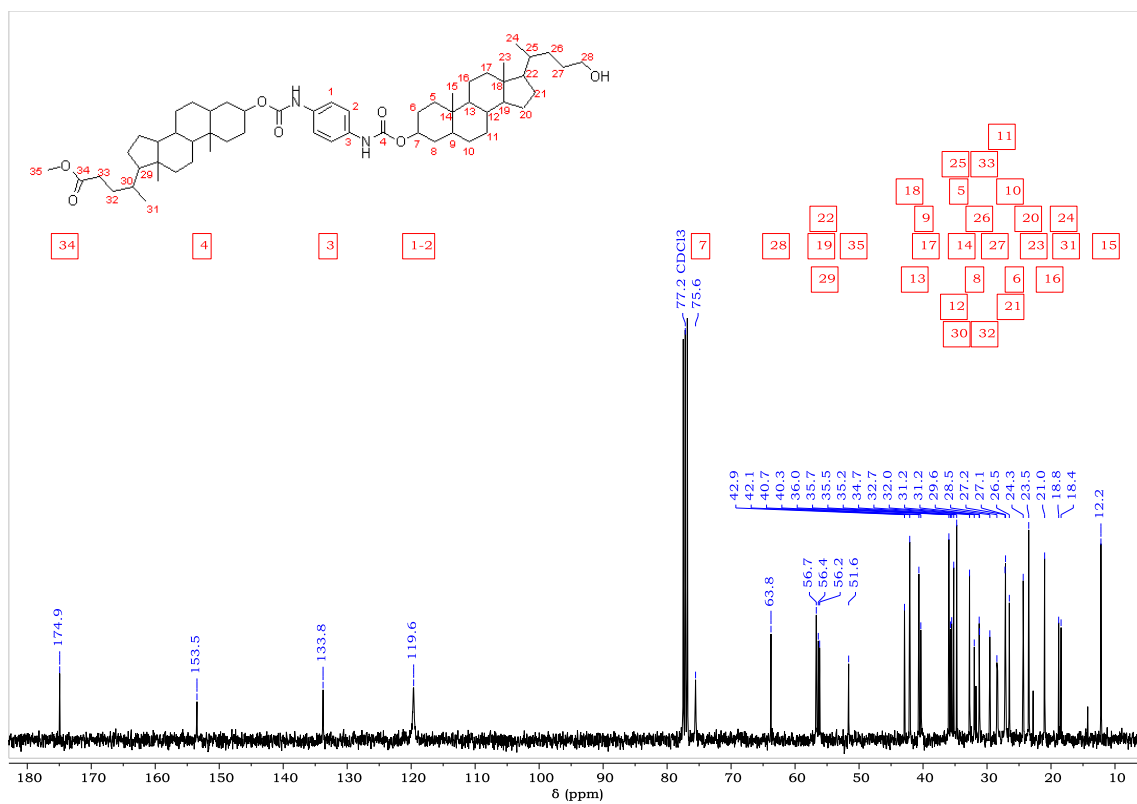
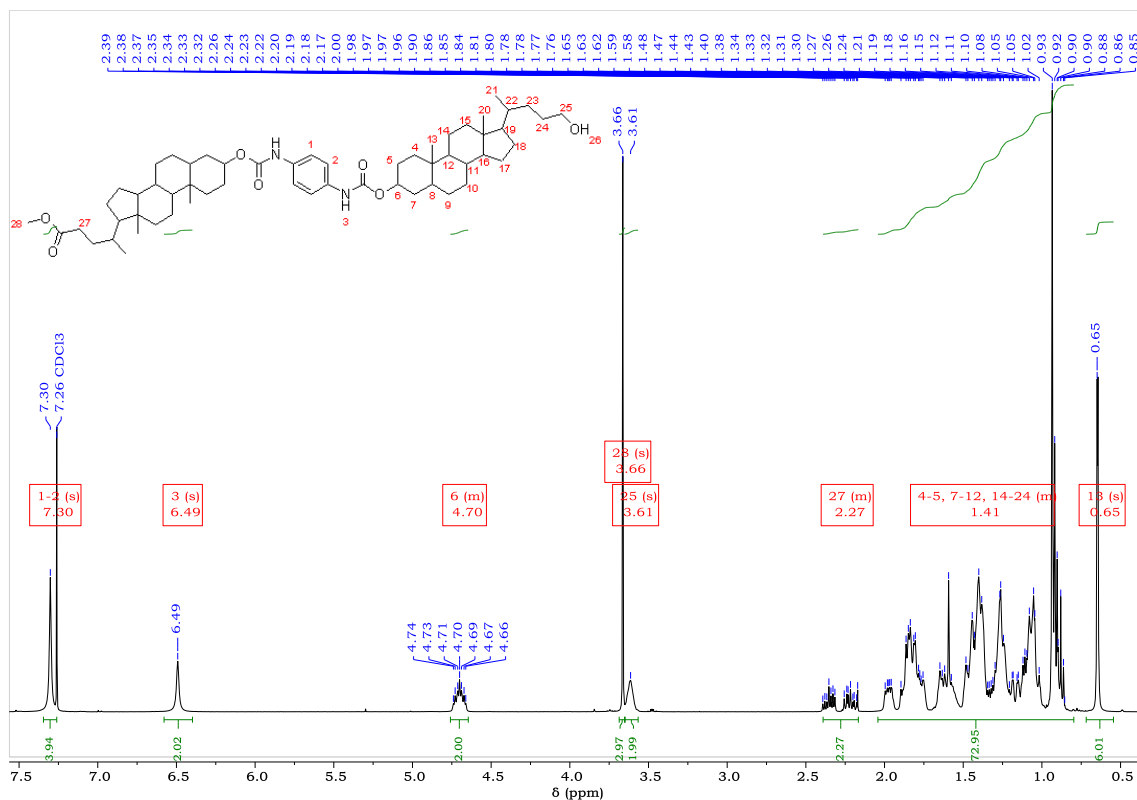
## 6.2.4 Compound 9- bischolesterol dibromide $^1\text{H}$ and $^{13}\text{C}$ NMR



## 6.2.5 Compound **10**- bischolesterol diazide <sup>1</sup>H and <sup>13</sup>C NMR

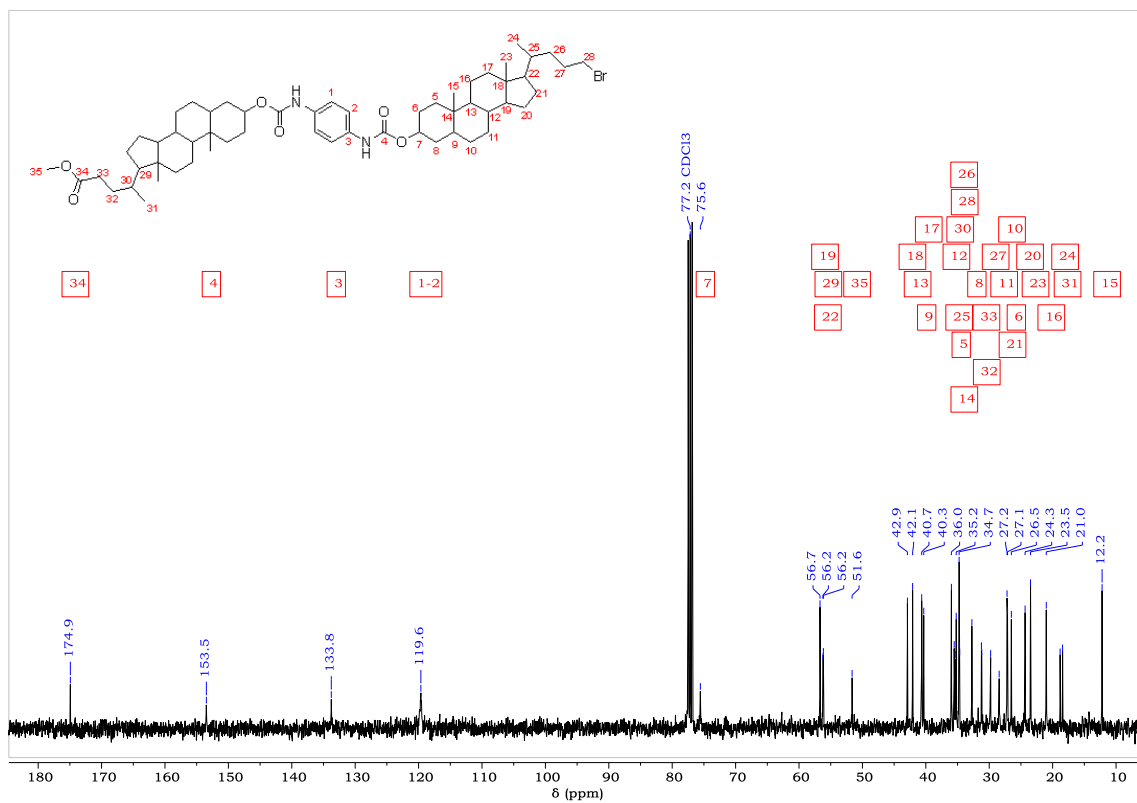
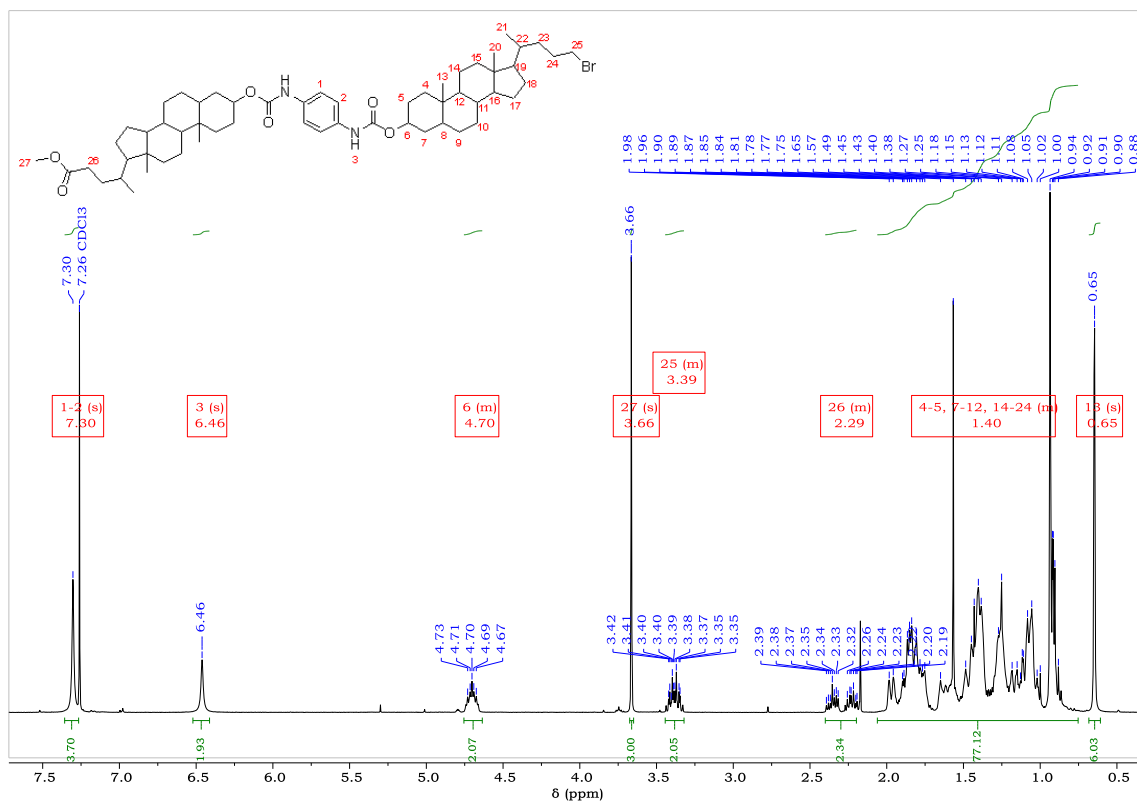


## 6.2.6 Compound **11**- bischolesterol monoester/alcohol $^1\text{H}$ and $^{13}\text{C}$ NMR

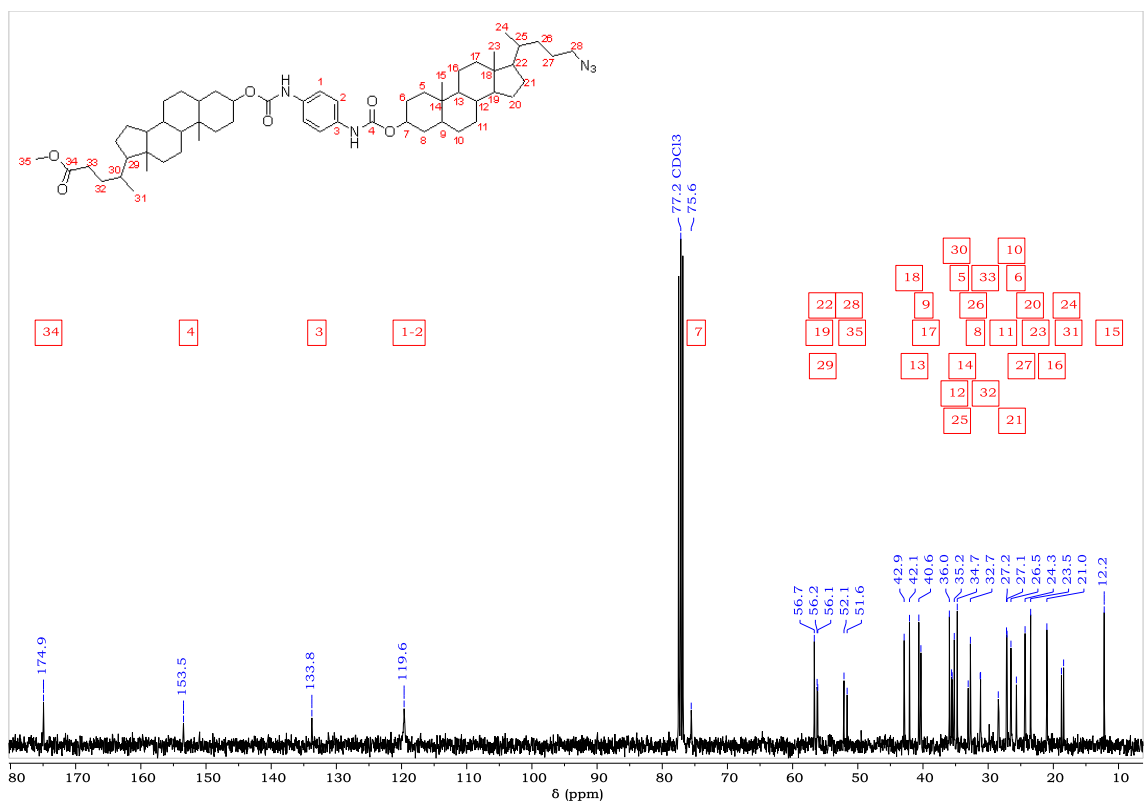
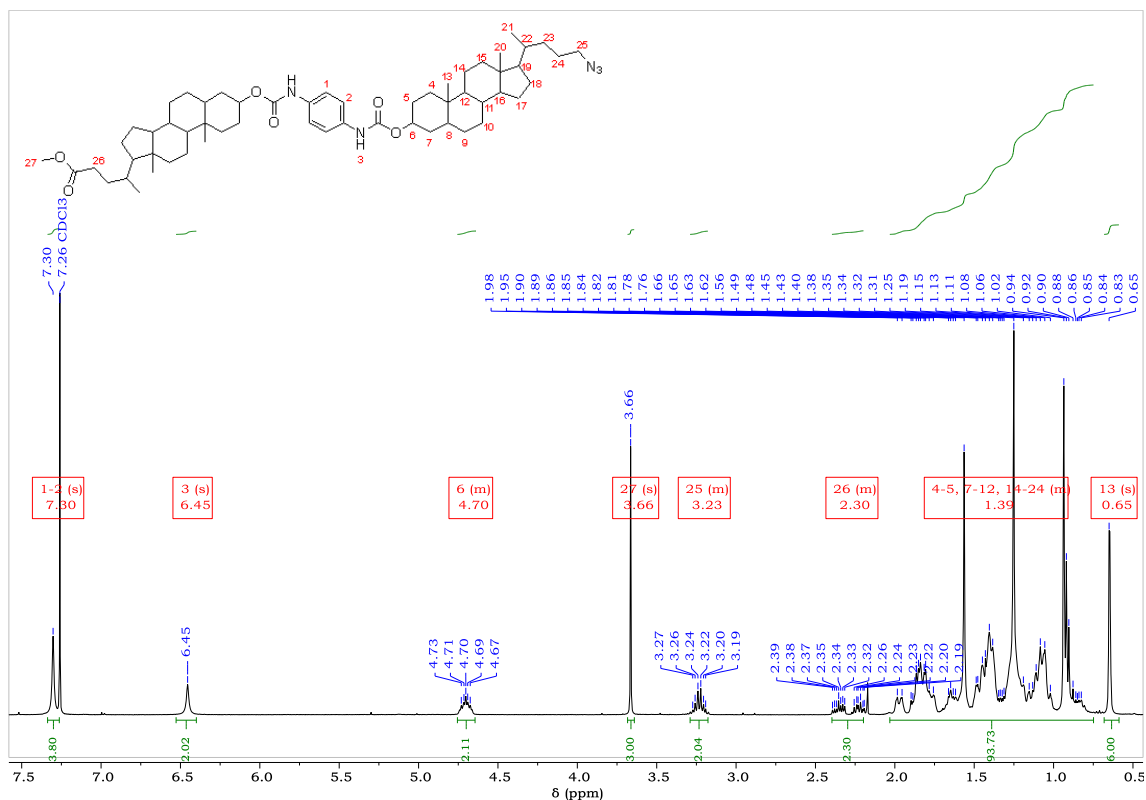


## 6.2.7 Compound **12**- bischolesterol monoester/bromide $^1\text{H}$ and $^{13}\text{C}$

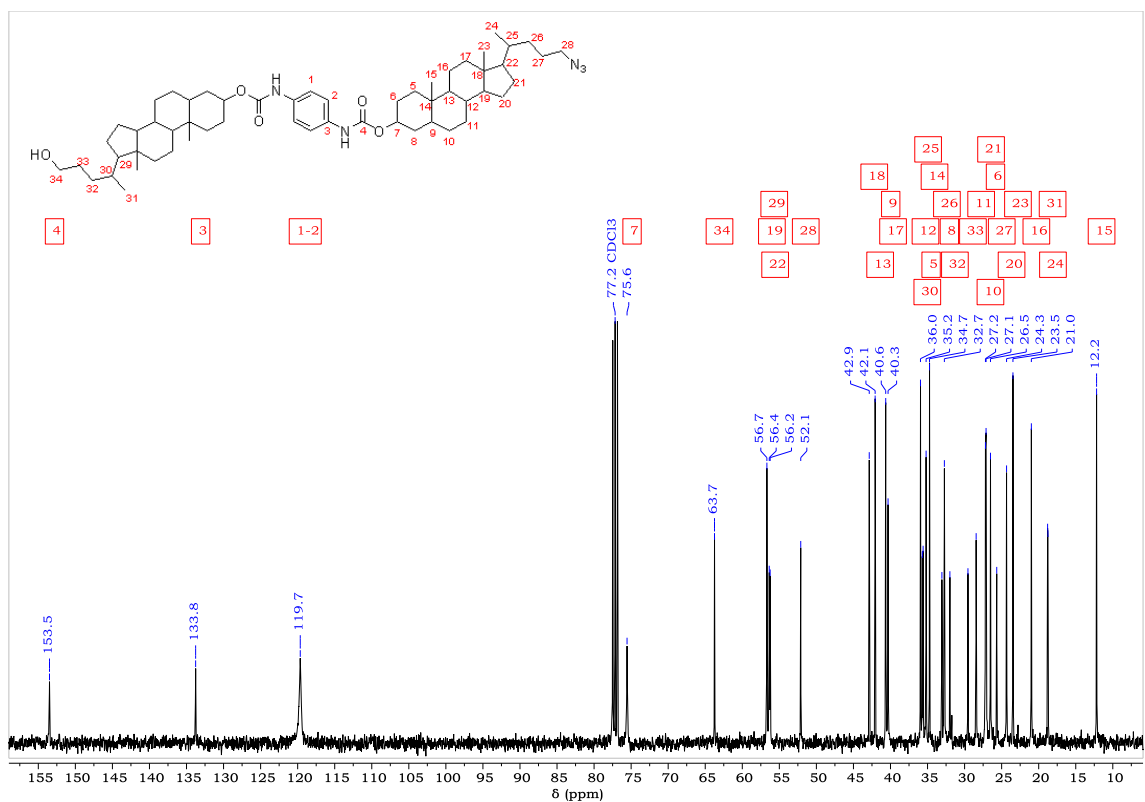
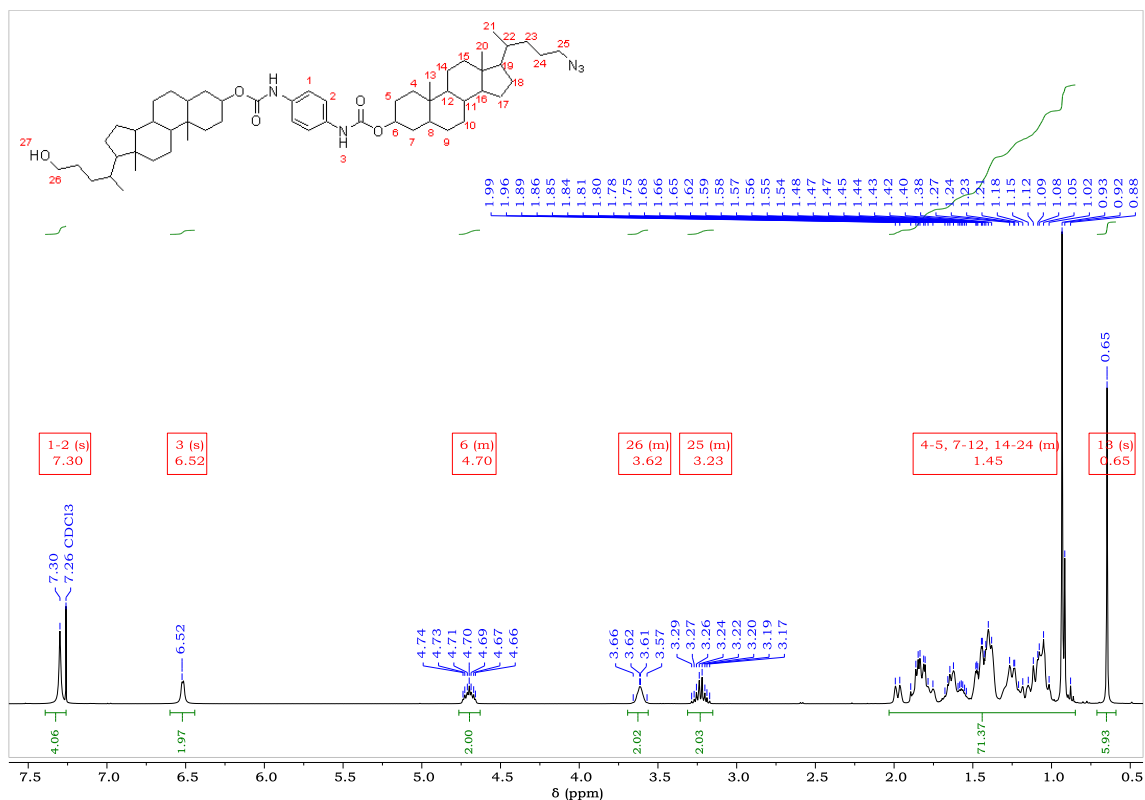
### NMR



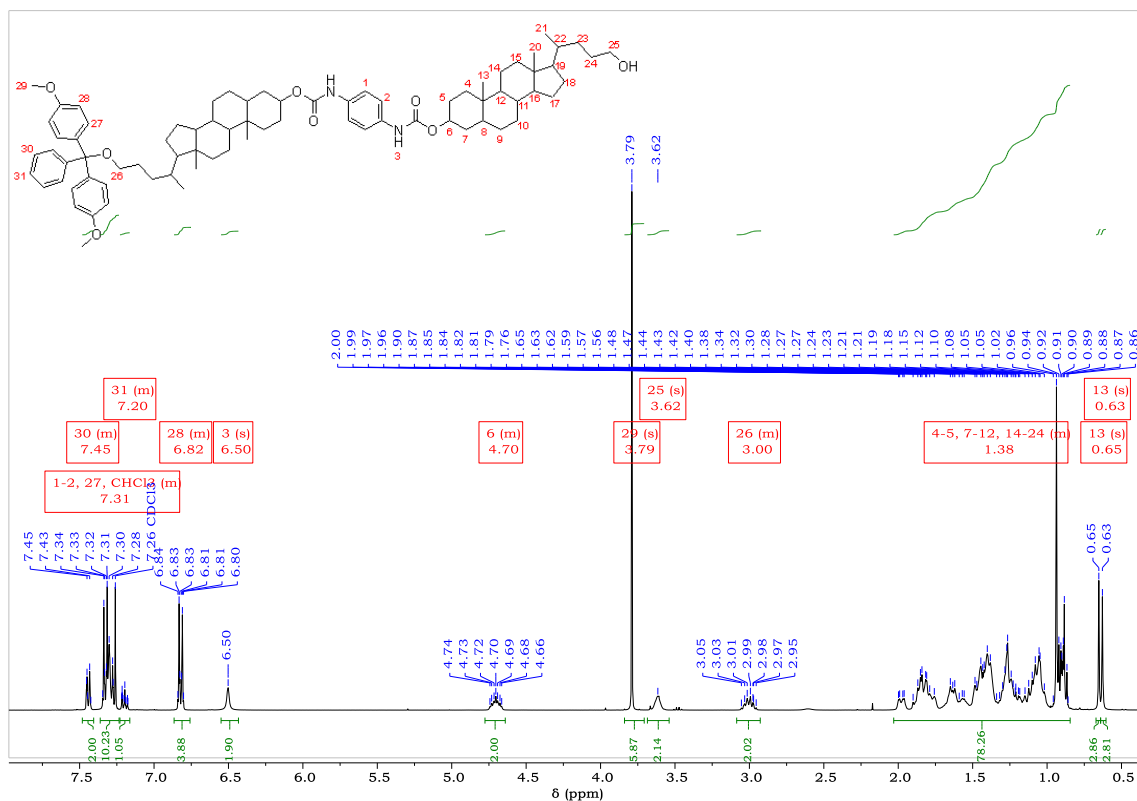
## 6.2.8 Compound **13**- bischolesterol monoester/azide $^1\text{H}$ and $^{13}\text{C}$ NMR



## 6.2.9 Compound **14**- bischolesterol monoazide/alcohol $^1\text{H}$ and $^{13}\text{C}$ NMR

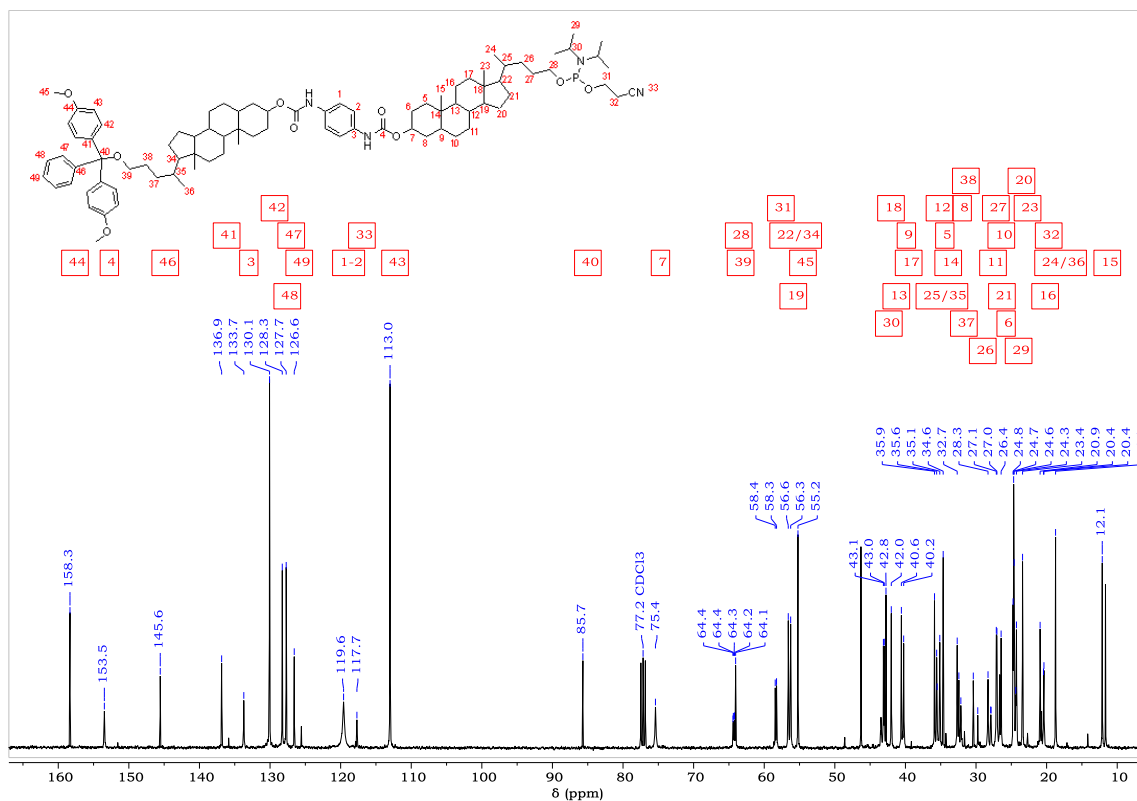


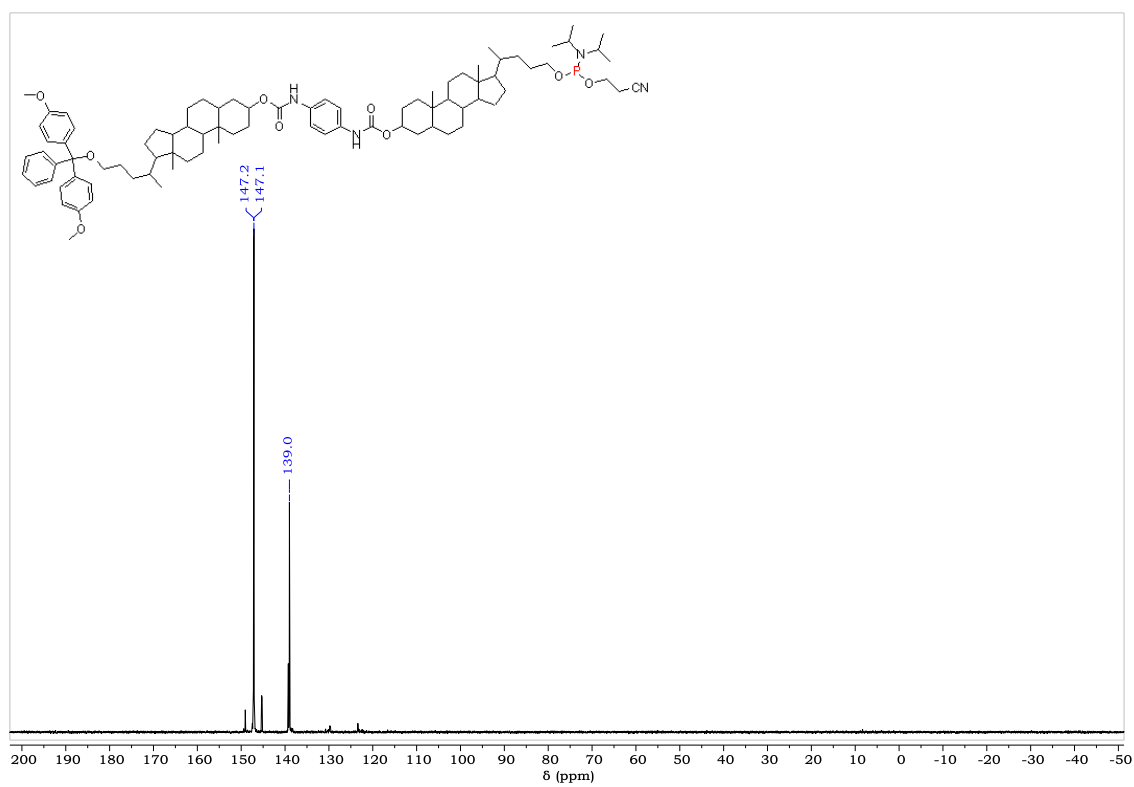
## 6.2.10 Compound **15**- mono DMT protected bischolesterol alcohol $^1\text{H}$ and $^{13}\text{C}$ NMR





## $^1\text{H}$ , $^{13}\text{C}$ , and $^{31}\text{P}$ NMR

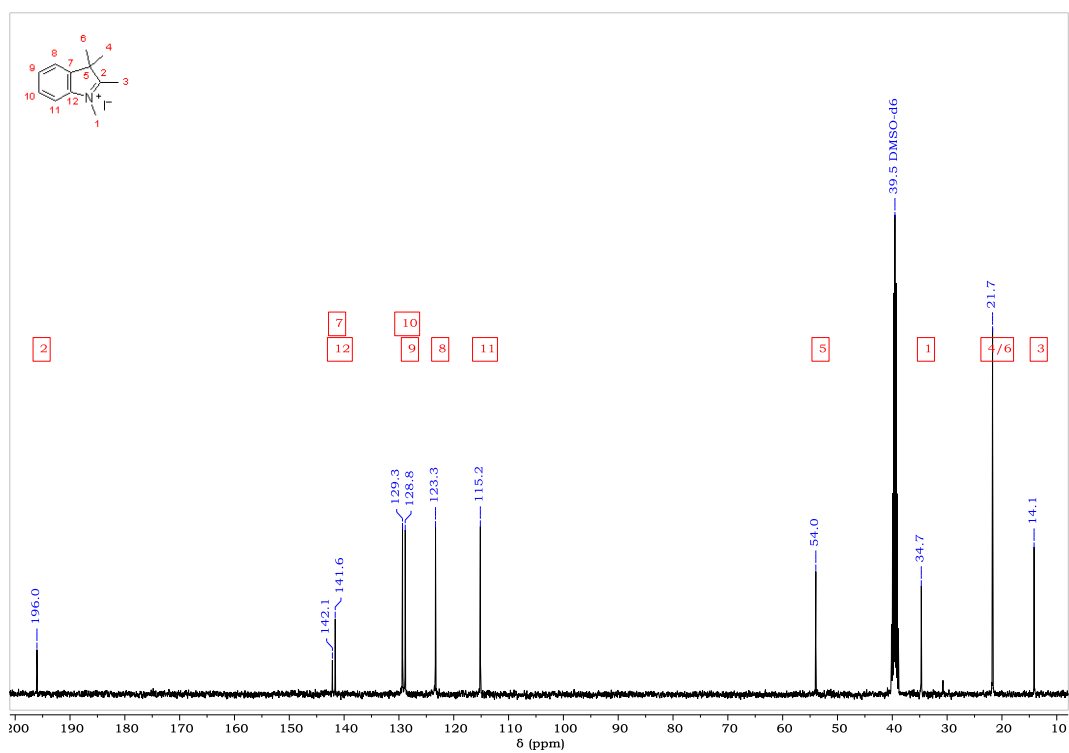
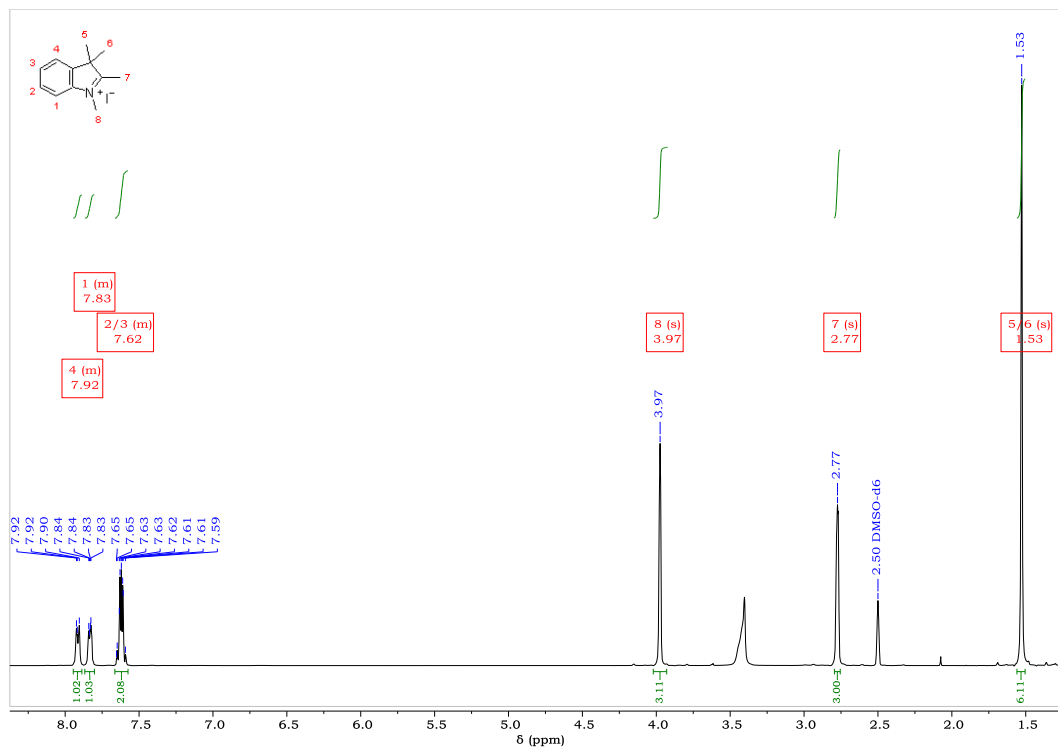




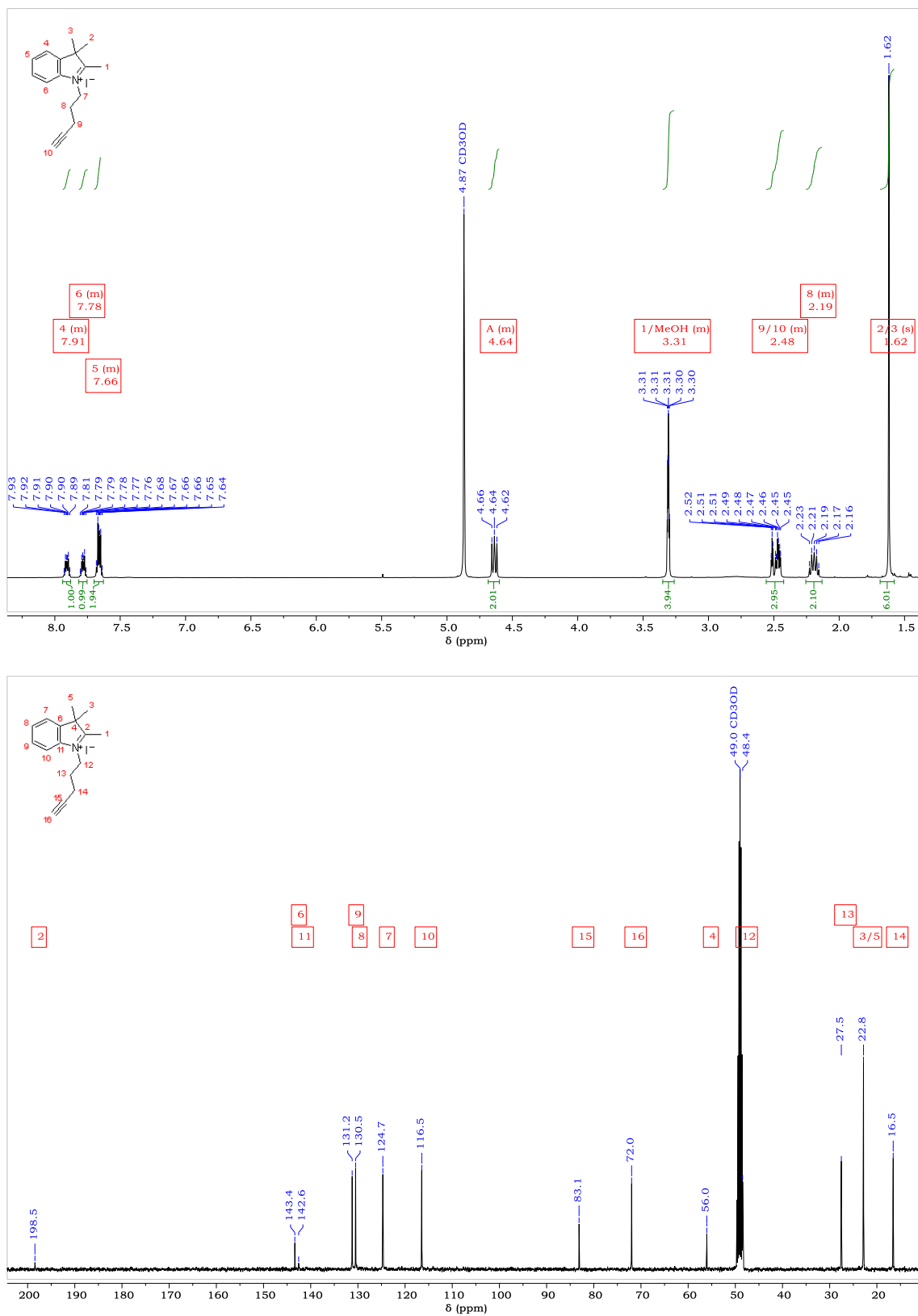
## 6.3 NMR data for cyanine dye derivatives

### 6.3.1 Compound **17**- 1,2,3,3-tetramethyl-3H-indolinium iodide <sup>1</sup>H and

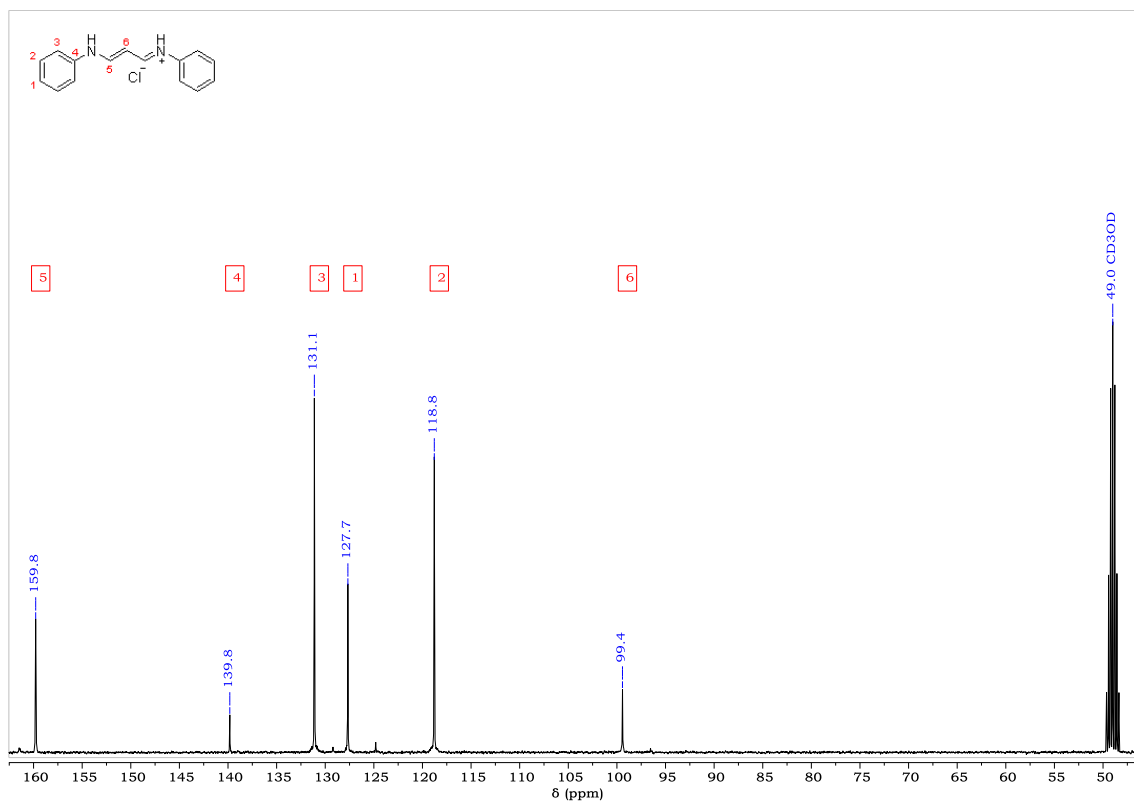
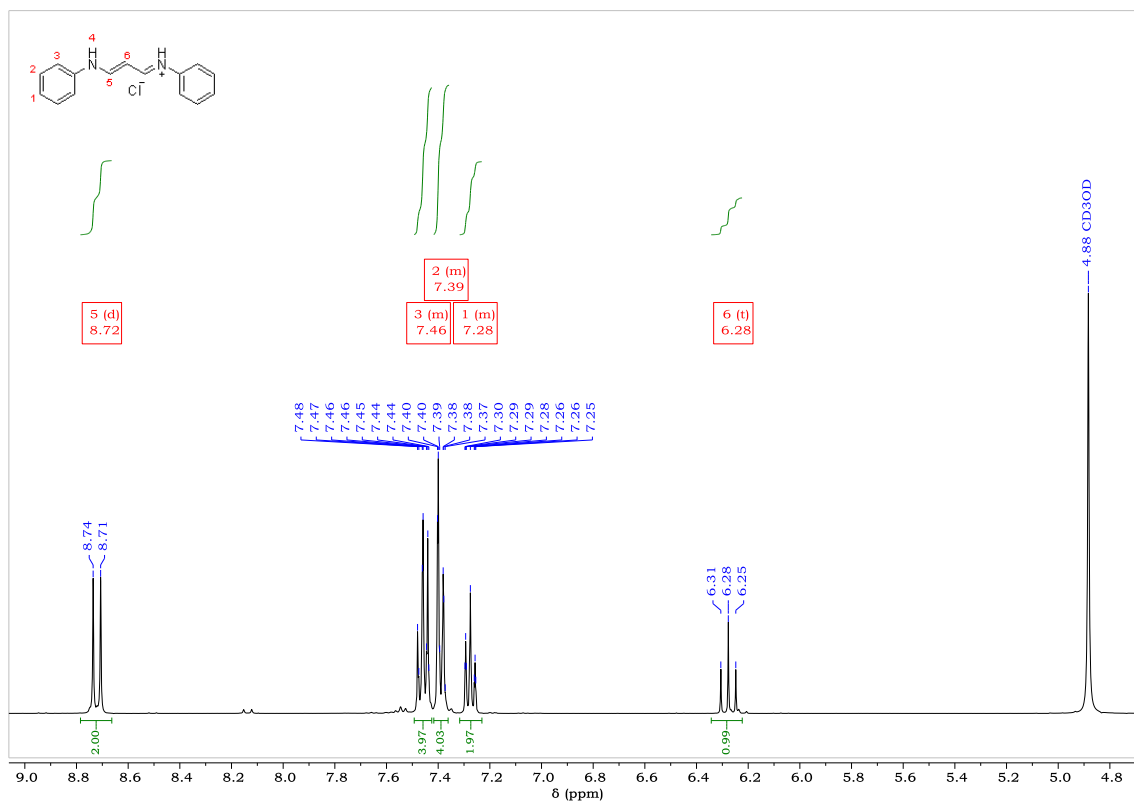
#### <sup>13</sup>C NMR



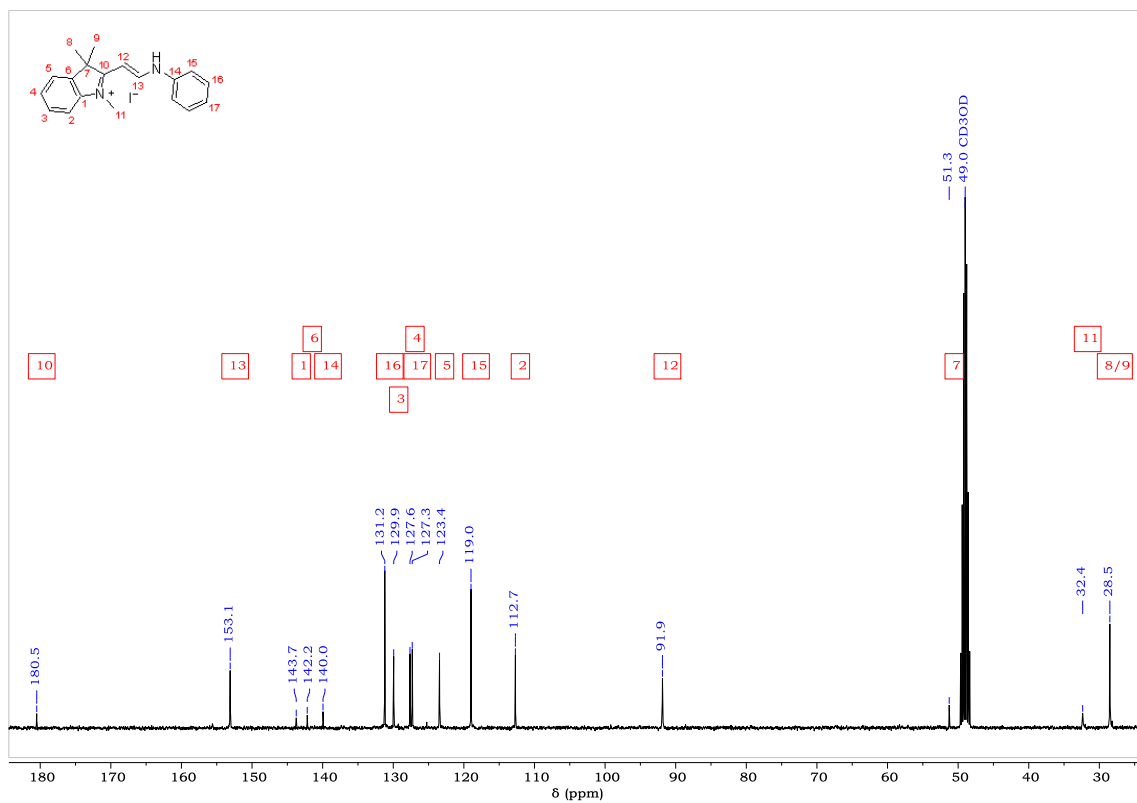
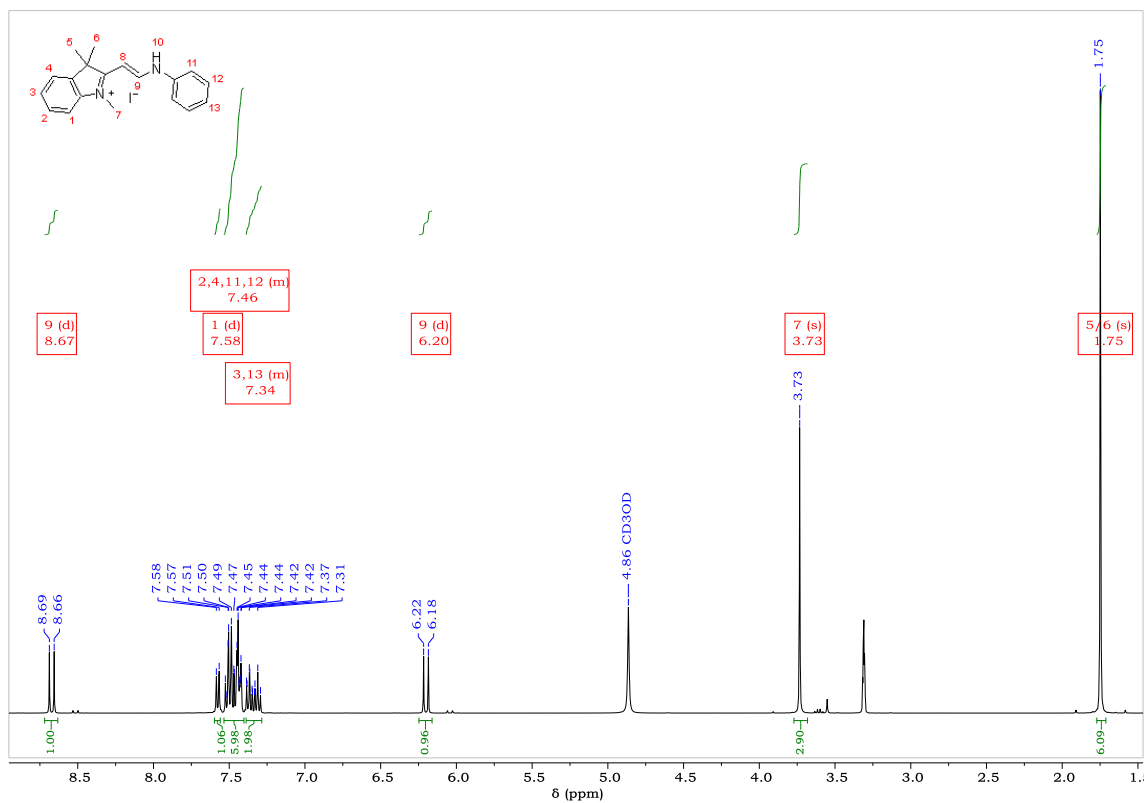
### 6.3.2 Compound **18**- 1-(4-pentynyl)-2,3,3-trimethyl-3H-indolinium iodide $^1\text{H}$ and $^{13}\text{C}$ NMR



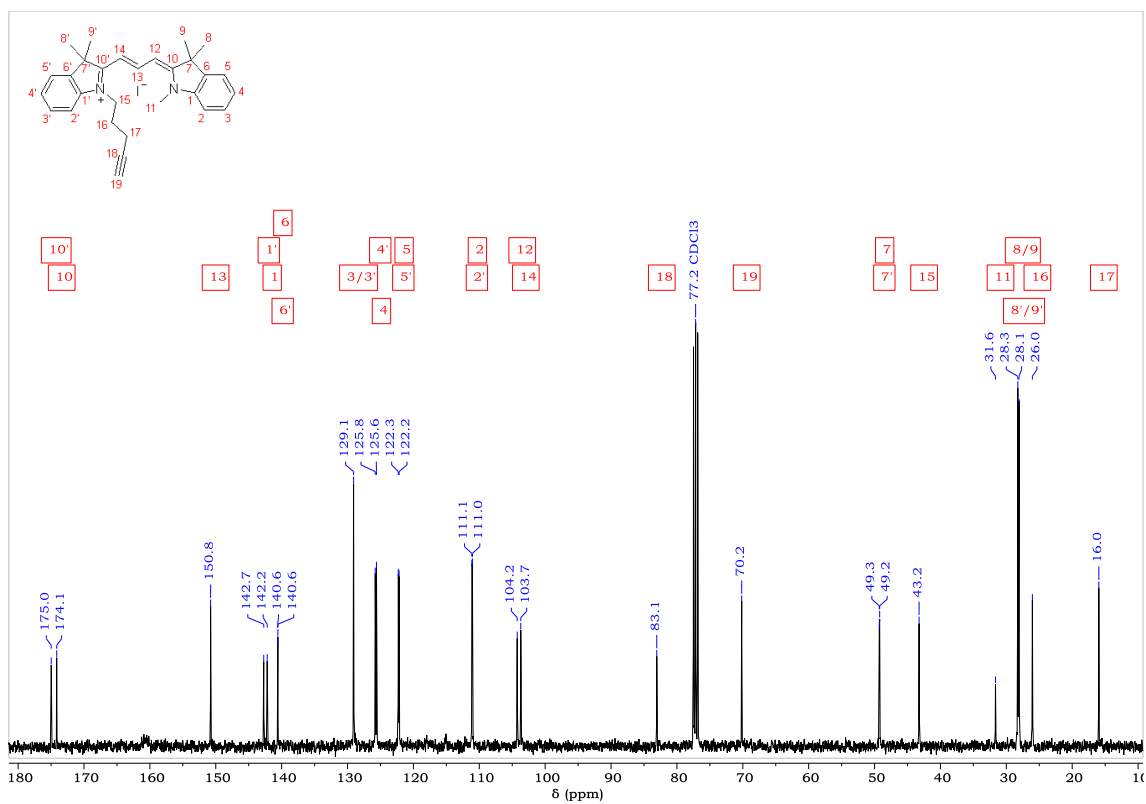
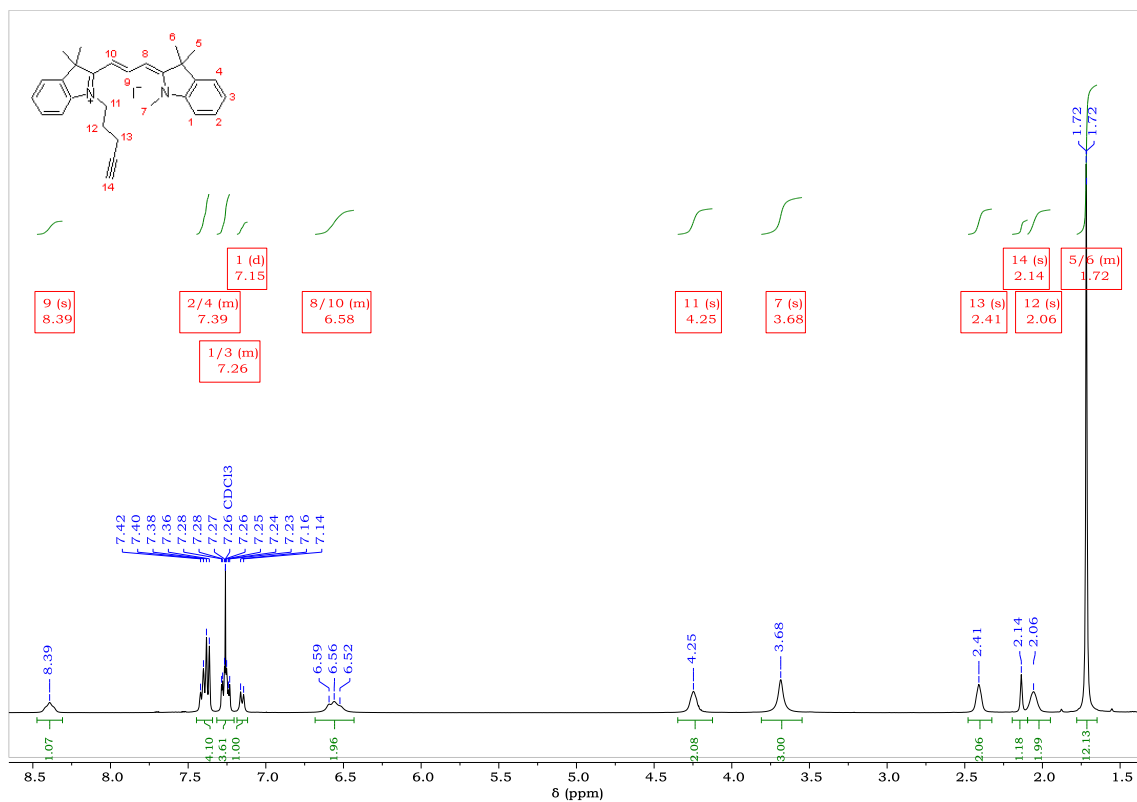
### 6.3.3 Compound **19**- malondialdehyde bis(phenylimine) monohydrochloride $^1\text{H}$ and $^{13}\text{C}$ NMR



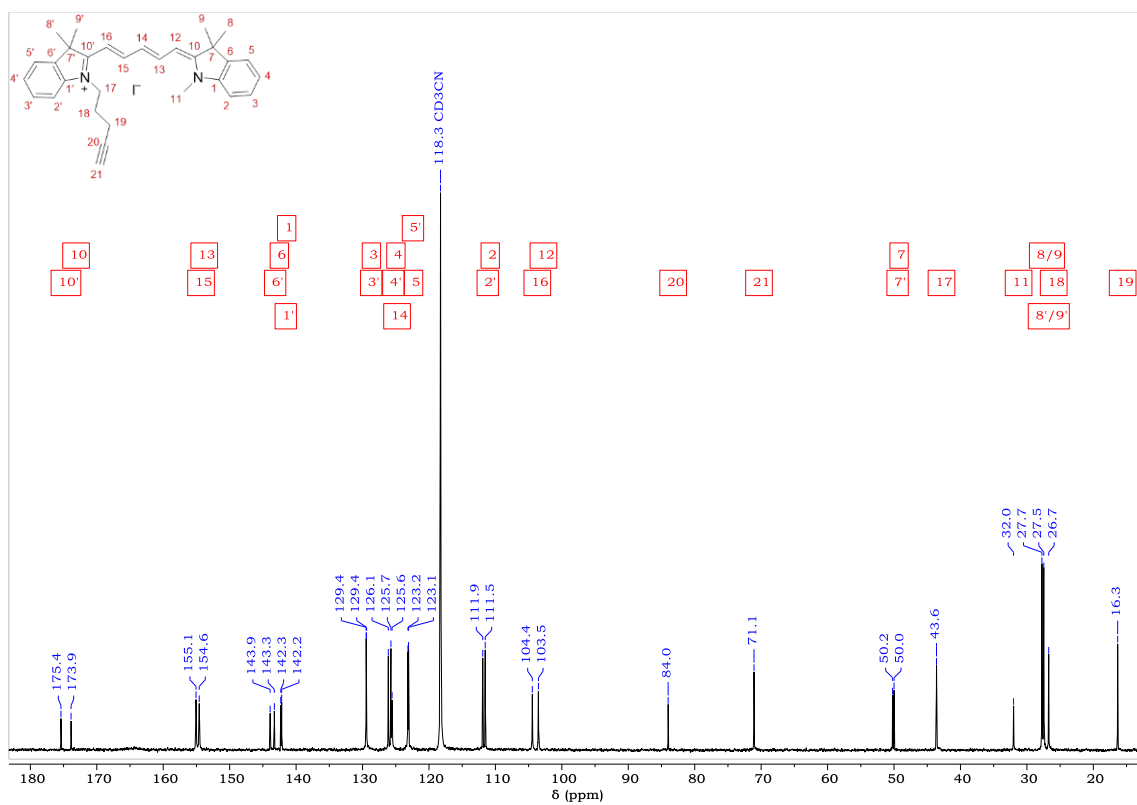
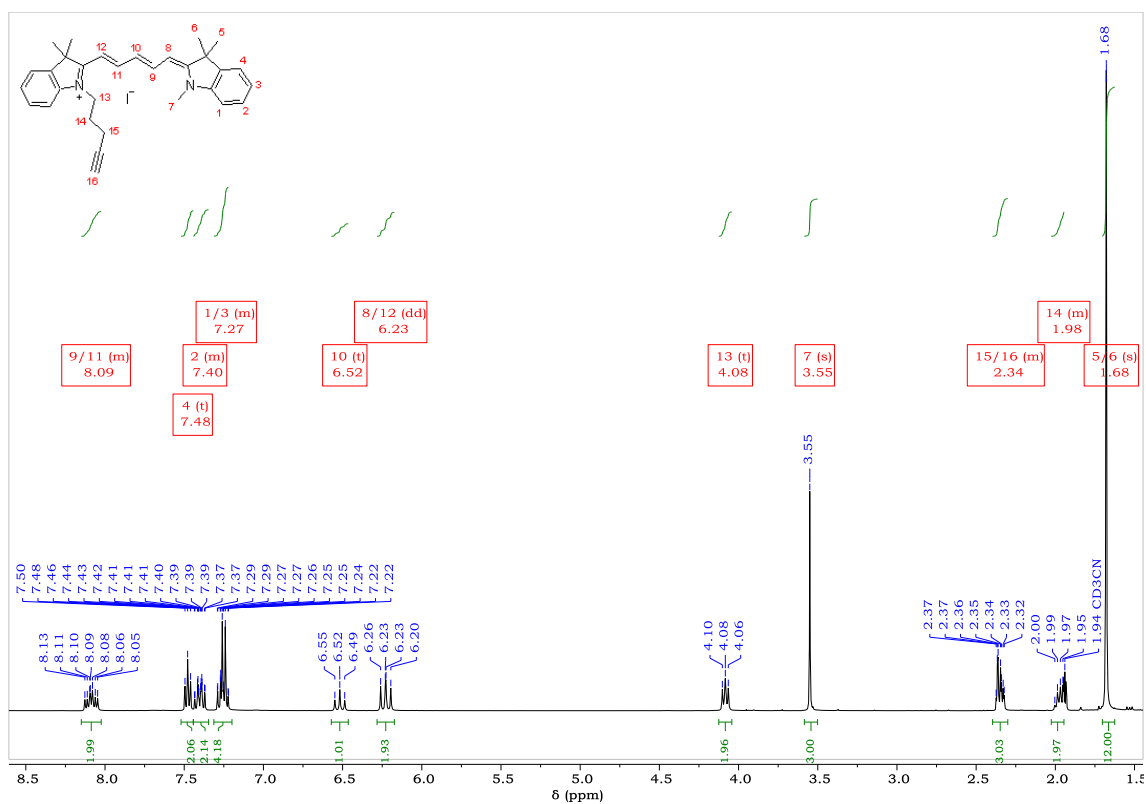
### 6.3.4 Compound **20**- 1,3,3-trimethyl-2-(2-(phenylamino)vinyl)-3H-indolium iodide <sup>1</sup>H and <sup>13</sup>C NMR



### 6.3.5 Compound **21**- Cy3 monoalkyne <sup>1</sup>H and <sup>13</sup>C NMR



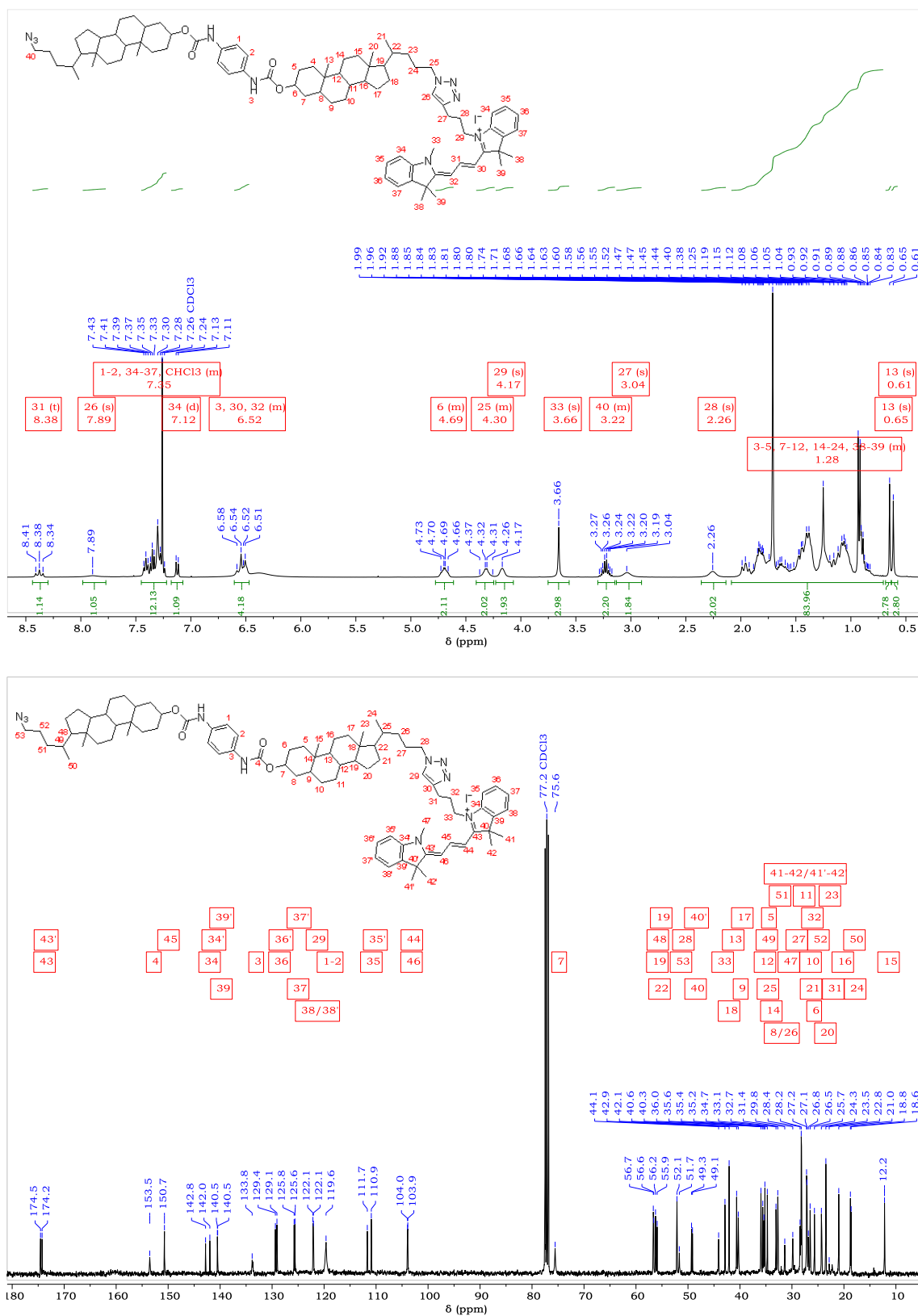
### 6.3.6 Compound **22**- Cy5 monoalkyne <sup>1</sup>H and <sup>13</sup>C NMR



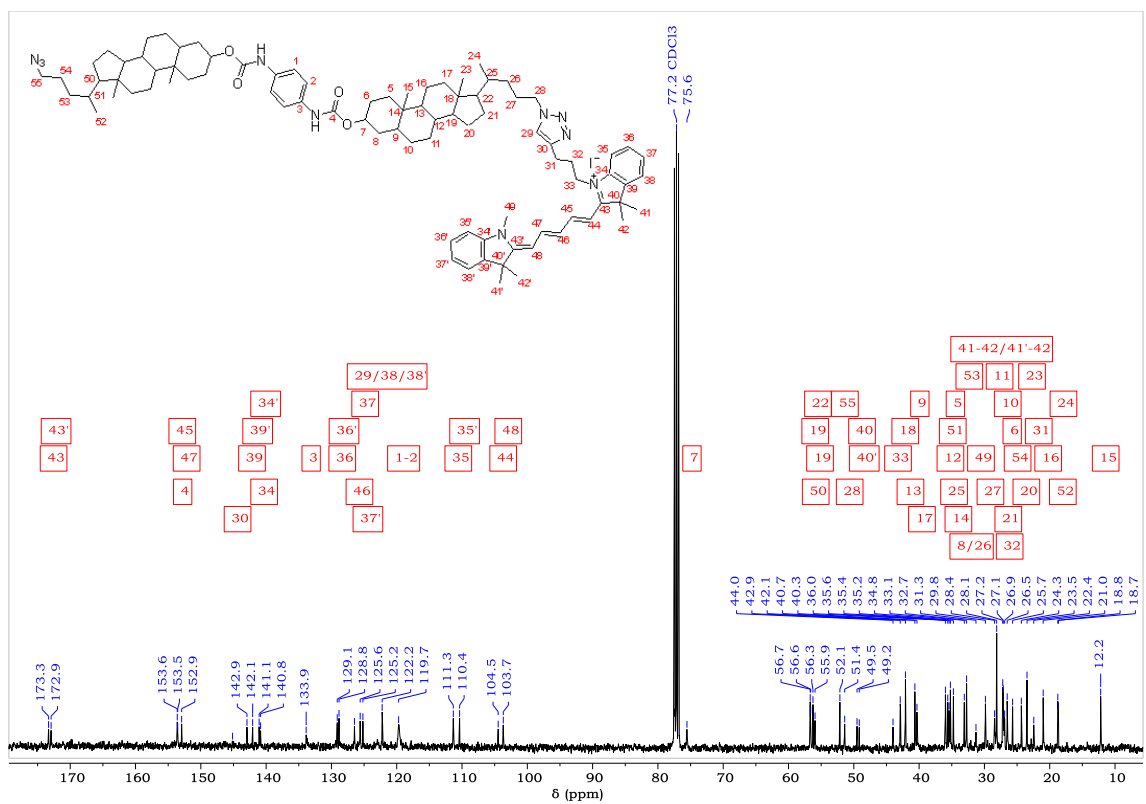
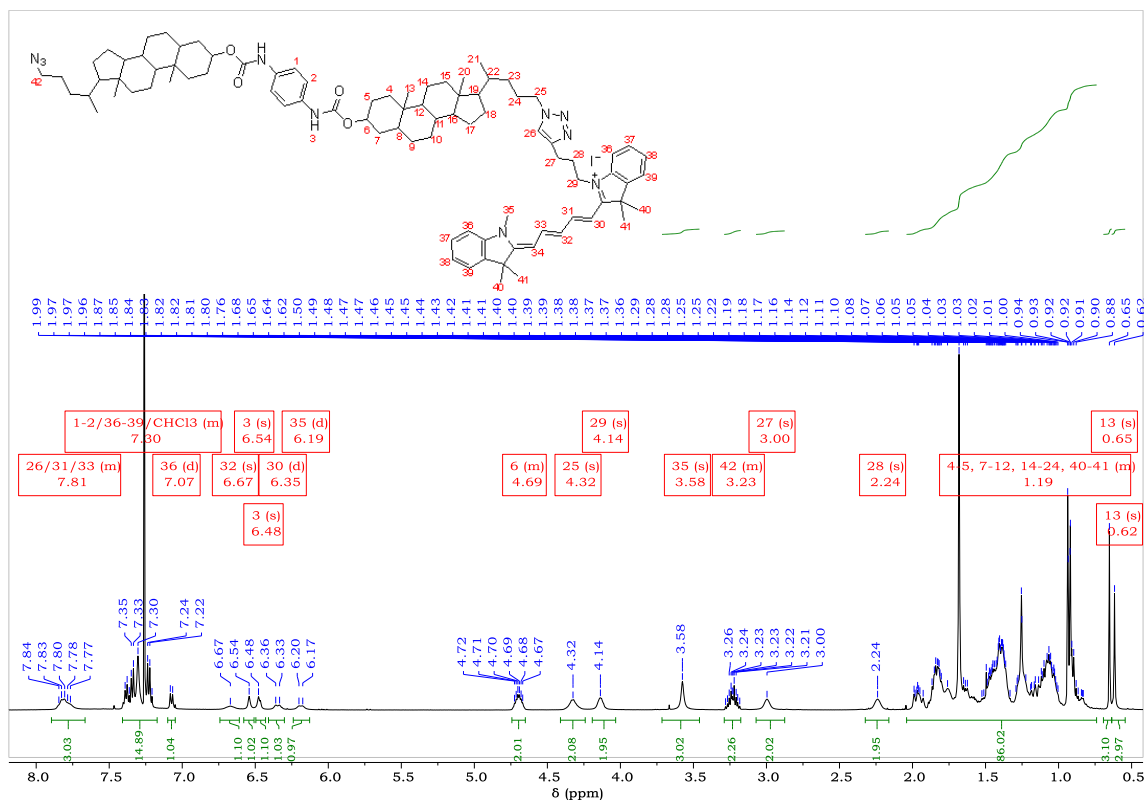


## 6.4 NMR data for bischolesterol-dye conjugates

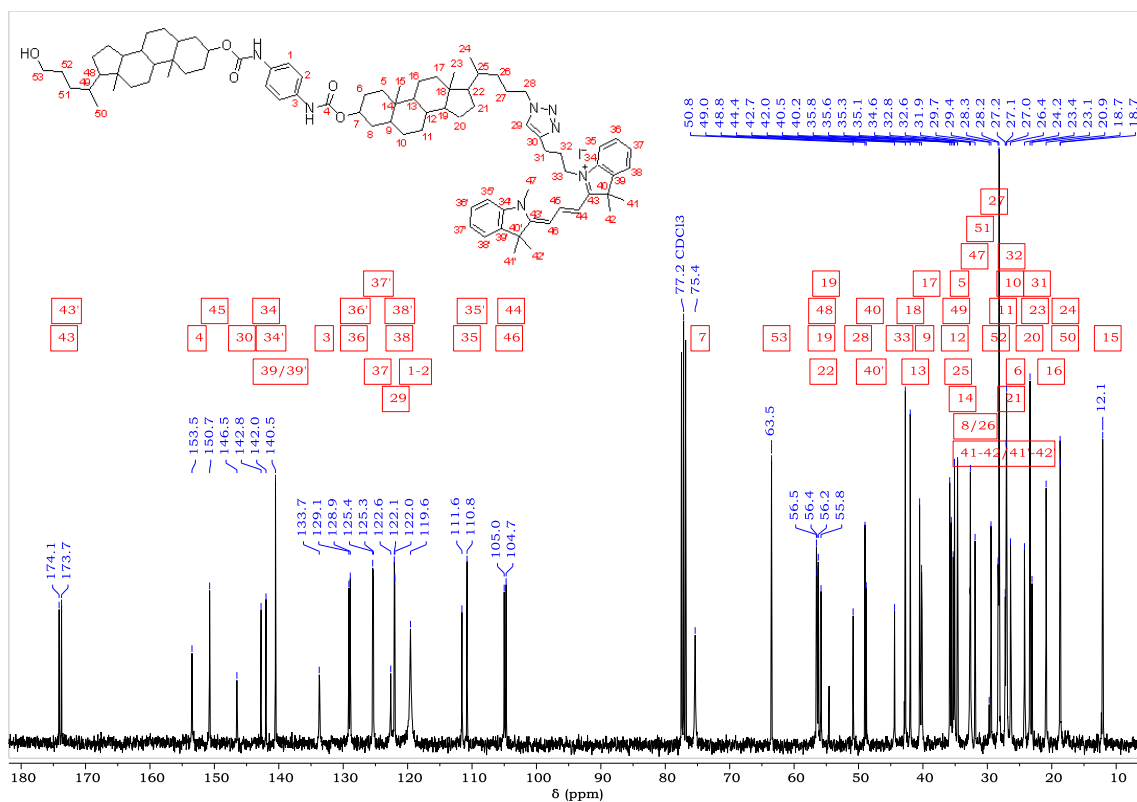
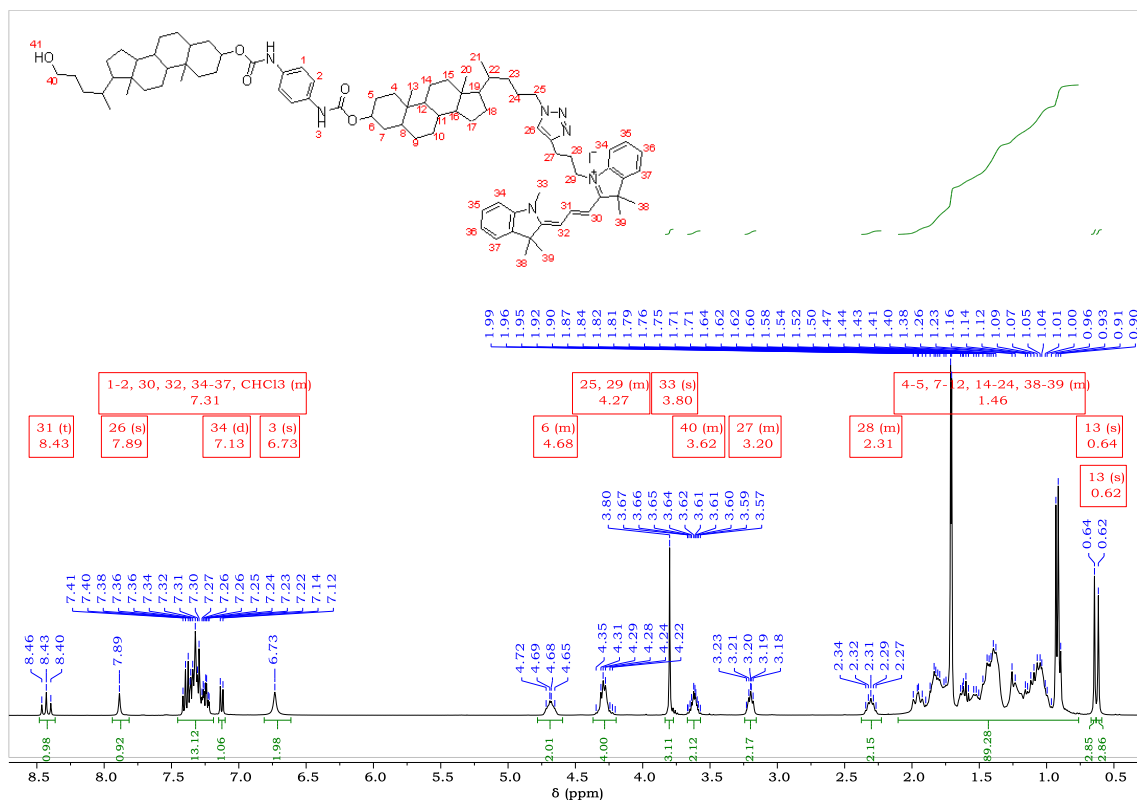
### 6.4.1 Compound **23**- bischolesterol Cy3 azide $^1\text{H}$ and $^{13}\text{C}$ NMR



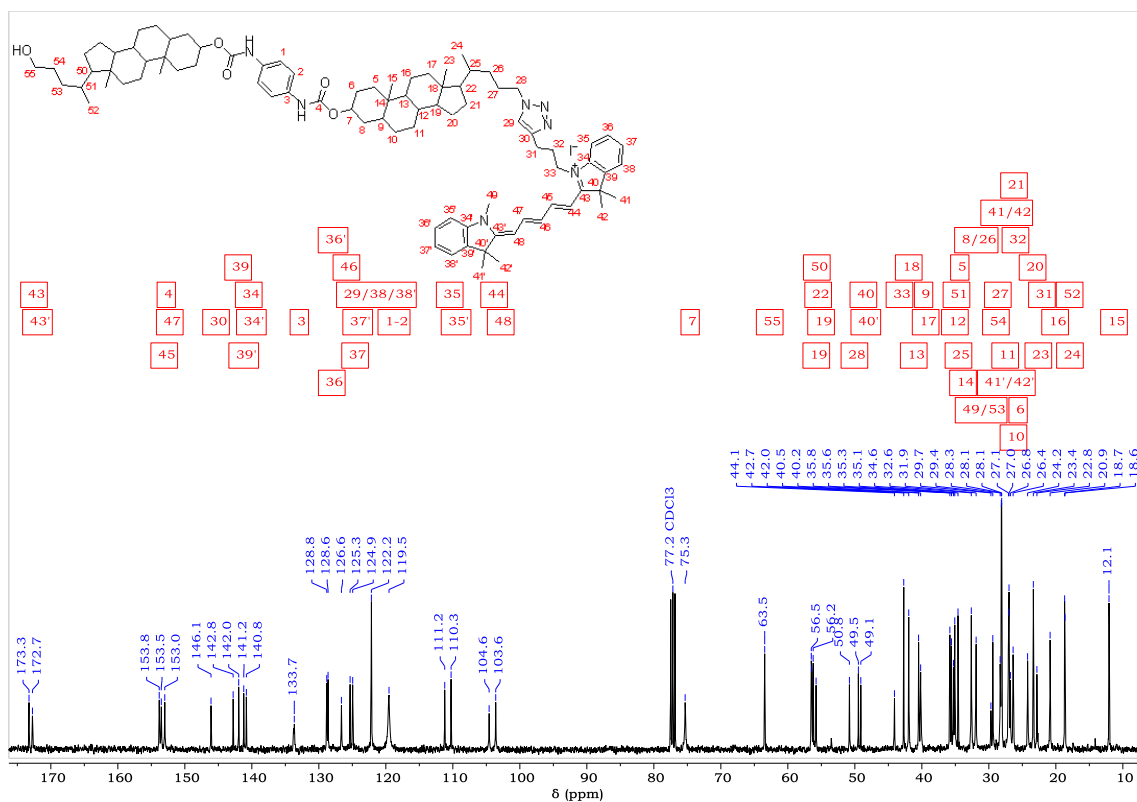
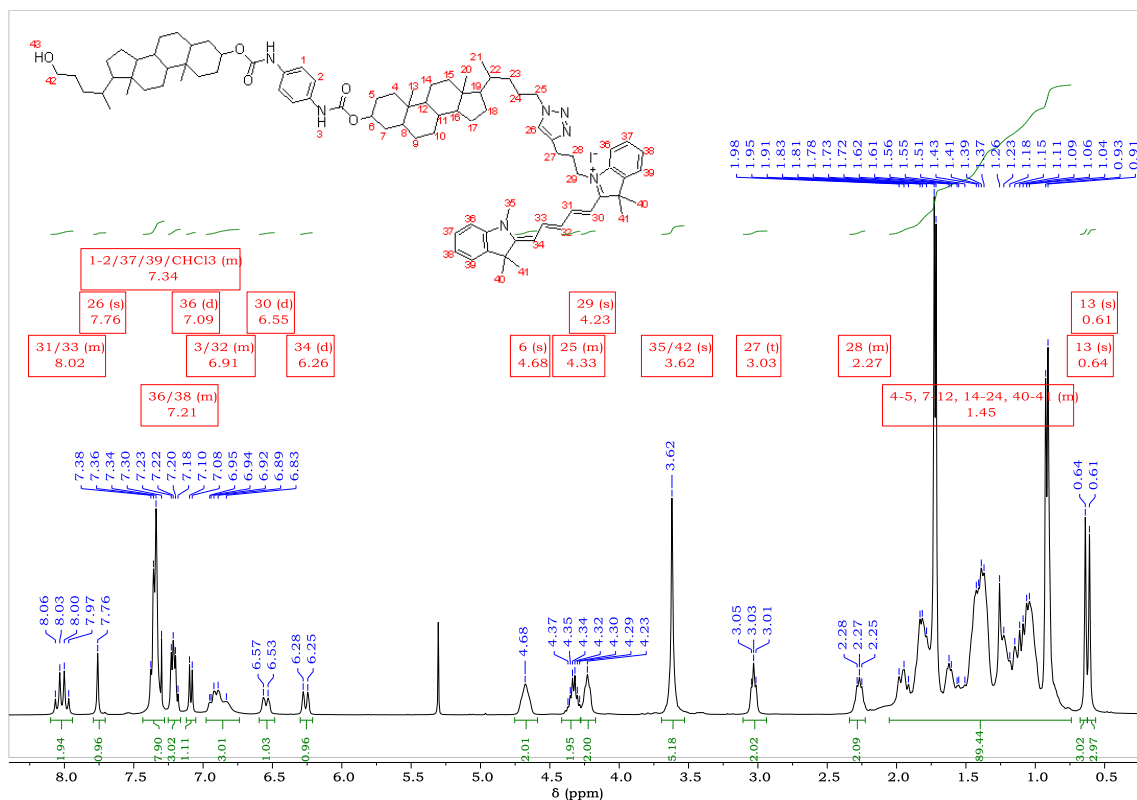
## 6.4.2 Compound **24**- bischolesterol Cy5 azide <sup>1</sup>H and <sup>13</sup>C NMR



### 6.4.3 Compound **25**- bischolesterol Cy3 alcohol $^1\text{H}$ and $^{13}\text{C}$ NMR

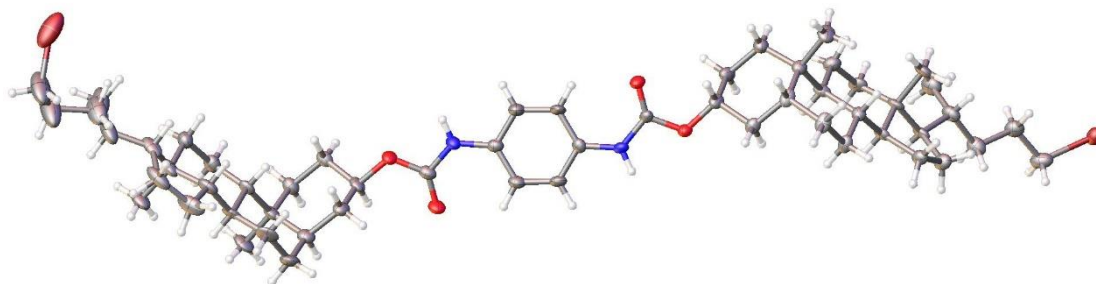


#### 6.4.4 Compound **26**- bischolesterol Cy5 alcohol <sup>1</sup>H and <sup>13</sup>C NMR



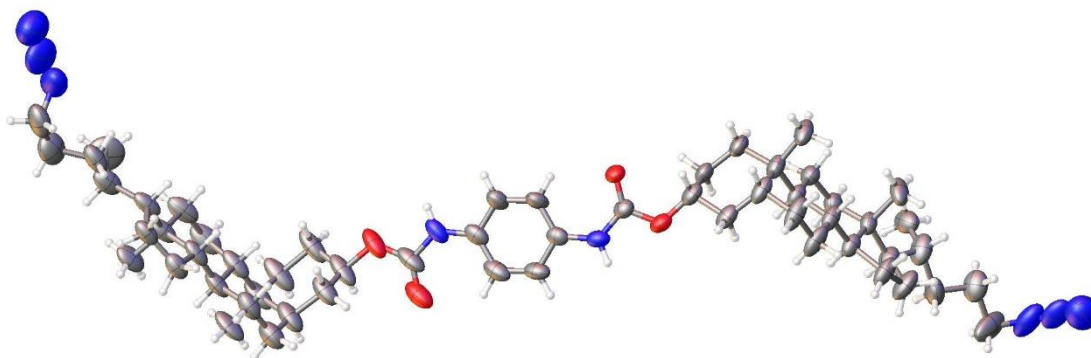
## 6.5 Crystal structure images of bischolesterol dibromide and bischolesterol diazide

### 6.5.1 Compound **9**- bischolesterol dibromide crystal structure image



Crystal structure of molecule 1 of compound **9** with ellipsoids drawn at the 50% probability level. The structure contains three crystallographically-independent molecules and six molecules of DMF of which only molecule 1 is shown.

### 6.5.2 Compound **10**- bischolesterol diazide crystal structure image



Crystal structure of compound **10** with ellipsoids drawn at the 50% probability level. The structure contains two molecules of DMF which have been omitted for clarity.

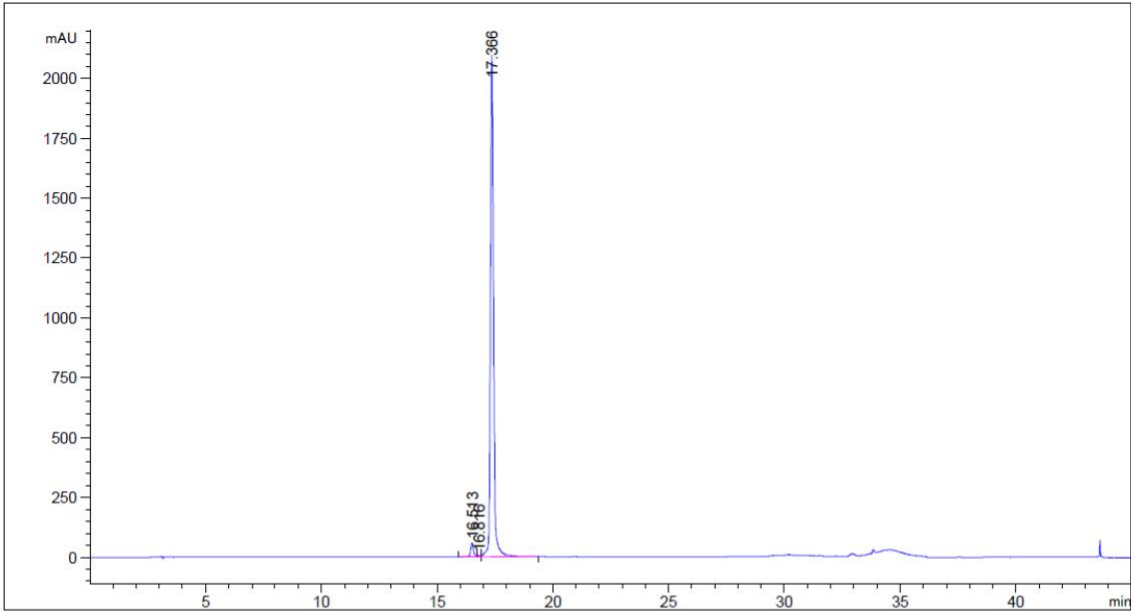
## 6.6 Mass spectrometry characterisation of oligonucleotides

Oligo name	Oligo sequence (5'→3')	Predicted mass ( <i>m/z</i> )	Observed mass ( <i>m/z</i> )
<b>Acryl-T16</b>	T*GG TTG GTG TGG TTG G	5168	5168
<b>Acryl-T7/T9</b>	GGT TGG T*GT* GGT TGG	5002	5002
<b>Acryl T12/T13</b>	GGT* T*GG TGT GGT TGG	5002	5002
<b>Acryl-T3/T4</b>	GGT TGG TGT GGT* T*GG	5002	5001
<b>Acryl-T3/T12</b>	GGT T*GG TGT GGT T*GG	5002	5002
<b>Acryl-T4/T13</b>	GGT* TGG TGT GGT* TGG	5002	5002
<b>Acryl-T7/T9/T16</b>	T*GG TTG GT*G T*GG TTG G	5442	5481
<b>TBA15</b>	GGT TGG TGT GGT TGG	4726	4726
<b>TBA15 complementary</b>	CCA ACC ACA CCA ACC	4420	4419
<b>TBA29</b>	AGT CCG TGG TAG GGC AGG TTG GGG TGA CT	9086	9086
<b>TBA15-A</b>	GGT TGG TG	2497	2496
<b>TBA15-B</b>	TGG TTG G	2168	2167
<b>TBA15-Aex</b>	AAA AAA GGT TGG TG	4376	4375
<b>TBA15-Bex</b>	TGG TTG GTT TTT T	3993	3991
<b>TBA15-alkyne</b>	alkyne-GGT TGG TGT GGT TGG	4886	4886
<b>TBA15-DBCO</b>	DBCO-GGT TGG TGT GGT TGG	5296	5296
<b>TBA29-alkyne</b>	alkyne-AGT CCG TGG TAG GGC AGG TTG GGG TGA CT	9246	9246
<b>TBA29-DBCO</b>	DBCO-AGT CCG TGG TAG GGC AGG TTG GGG TGA CT	9656	9655
<b>nina-alkyne</b>	alkyne-TGG ACT CTC TCA ATG	4703	4703
<b>nina-DBCO</b>	DBCO-TGG ACT CTC TCA ATG	5113	5113

T\* - acrylamide modified T

# 6.7 Analytical HPLC chromatograms of oligonucleotides

## 6.7.1 Acryl-T16



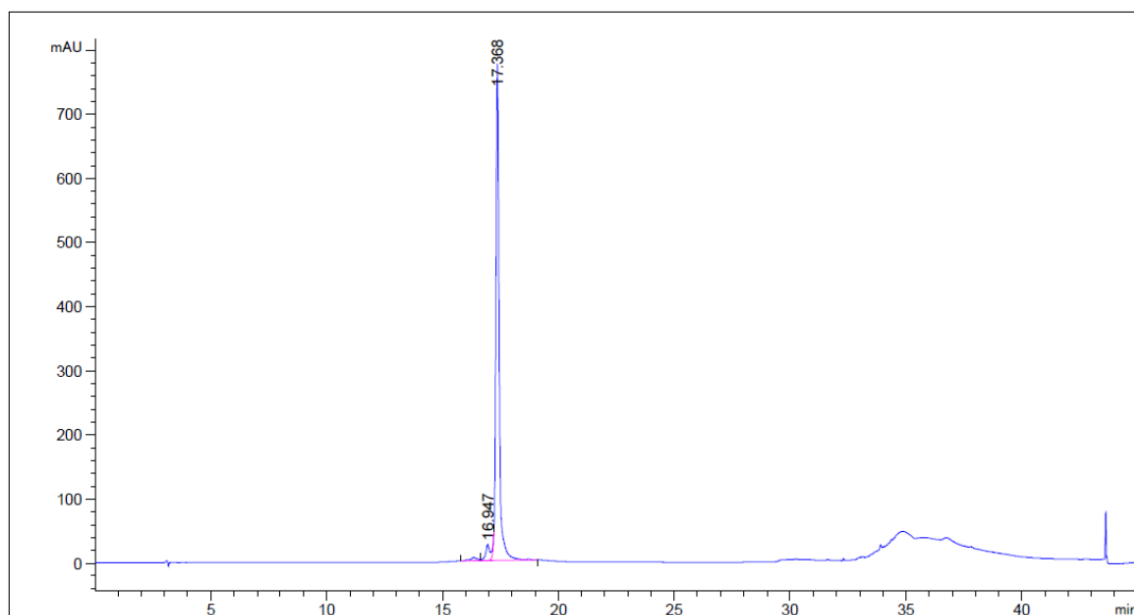
=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	16.513	BV R	0.1435	550.02069	56.54726	2.7464
2	16.816	VV E	0.1036	12.60930	1.75927	0.0630
3	17.366	VV R	0.1404	1.94646e4	2095.81494	97.1907

## 6.7.2 Acryl-T7/T9



### Area Percent Report

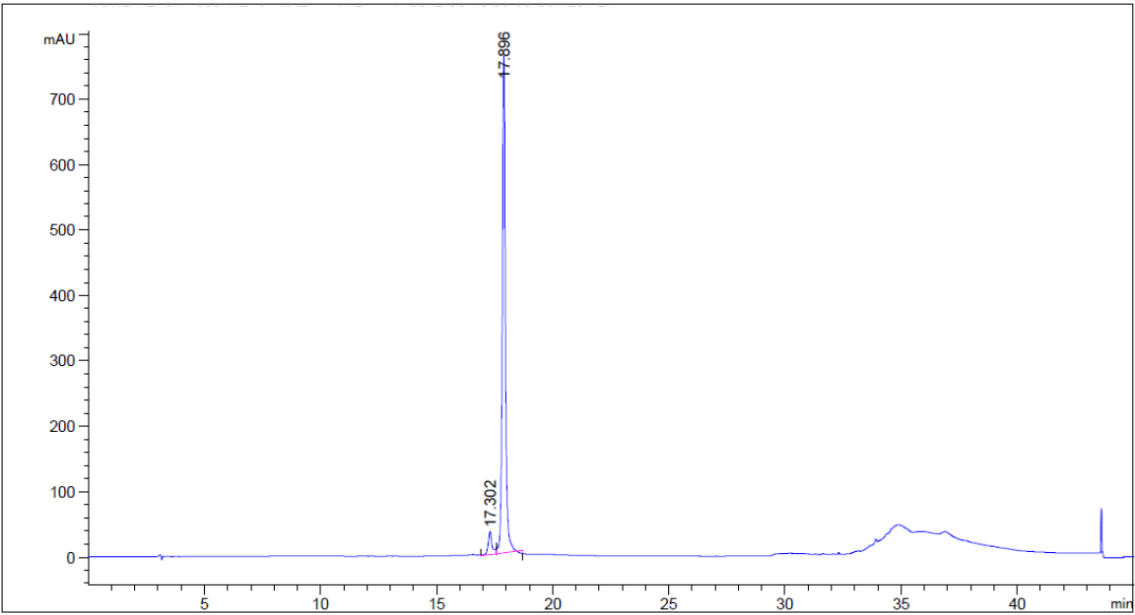
Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	16.947	VV E	0.2009	366.27063	25.60049	4.4879
2	17.368	VV R	0.1455	7795.06494	775.17017	95.5121



6.7.3 Acryl-T12/T13



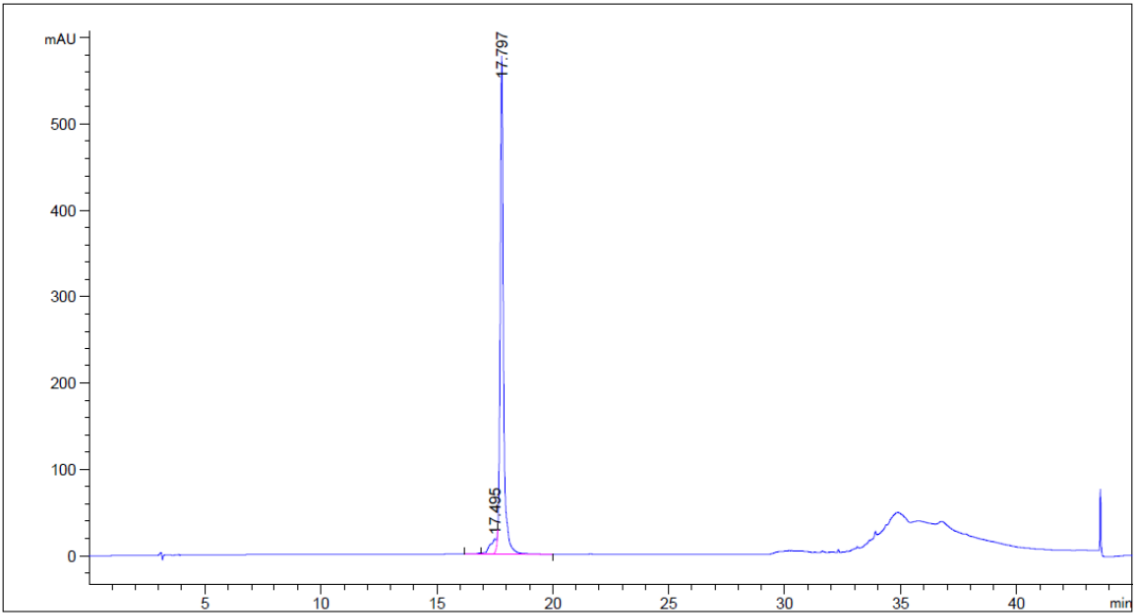
=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	17.302	MF	0.2019	429.82382	35.48186	5.5963
2	17.896	FM	0.1590	7250.71191	760.21503	94.4037

6.7.4 Acryl-T3/T4



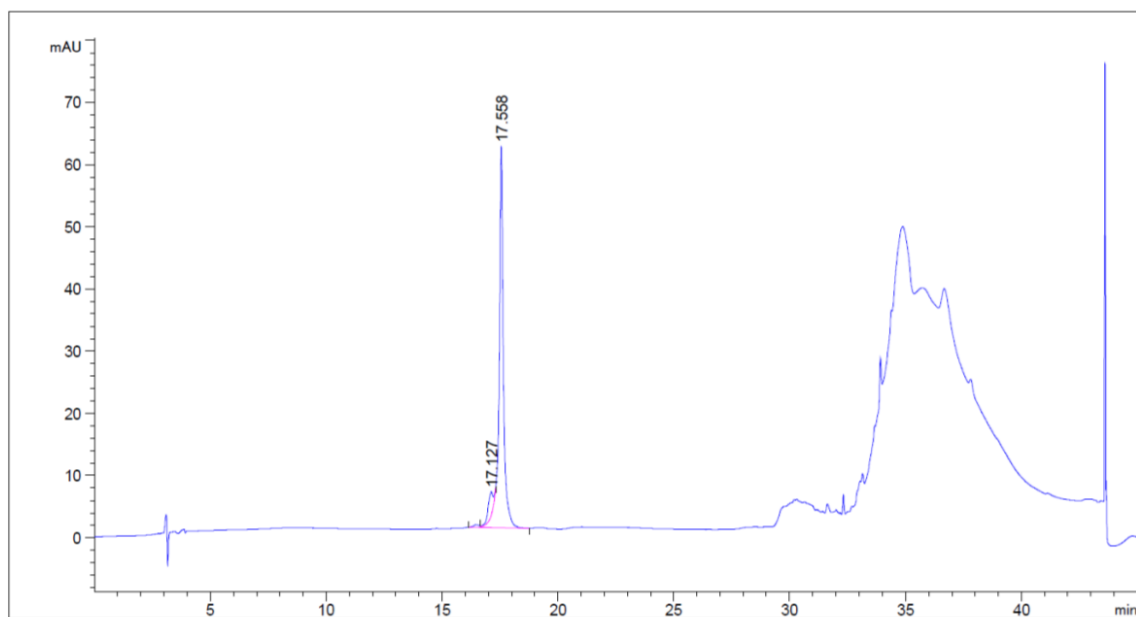
=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	17.495	VV E	0.2556	317.93710	16.15706	5.0455
2	17.797	VB R	0.1550	5983.52393	576.79681	94.9545

### 6.7.5 Acryl-T3/T12



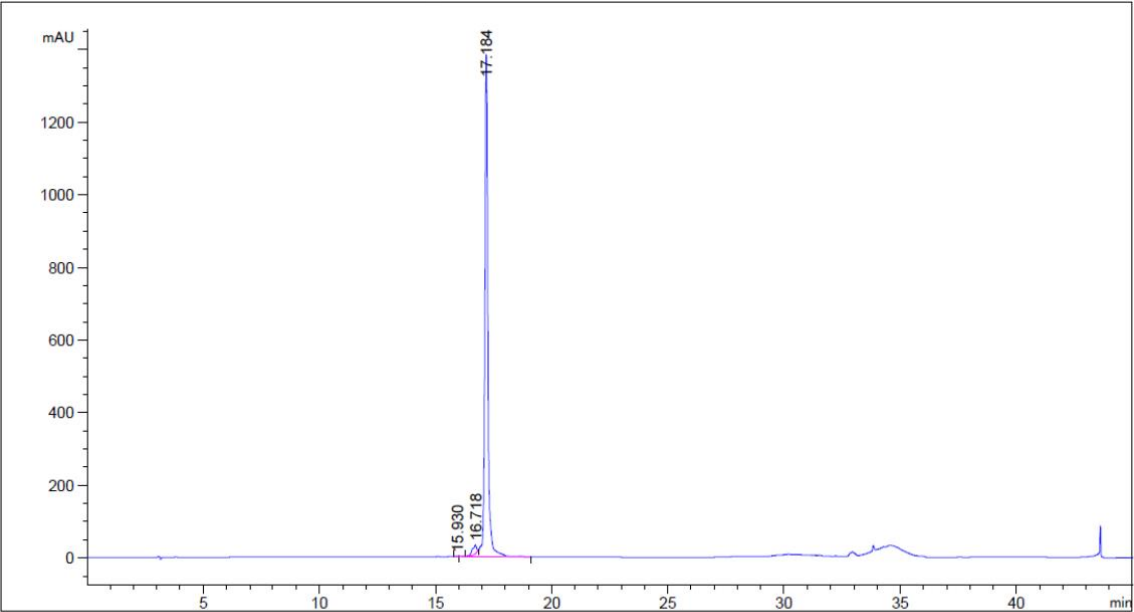
#### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	17.127	VV E	0.2371	57.08802	3.76125	6.4385
2	17.558	VB R	0.1918	829.58411	61.33638	93.5615

6.7.6 Acryl-T4/T13



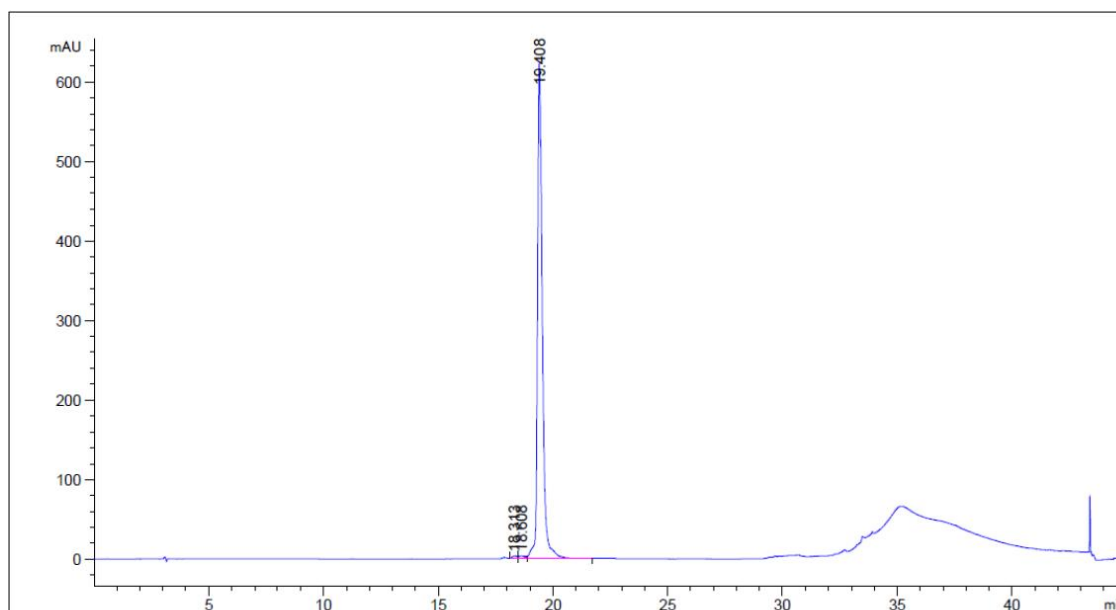
=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	15.930	BV E	0.1135	15.03663	2.00598	0.1122
2	16.718	VV E	0.1875	322.22189	23.27284	2.4034
3	17.184	VV R	0.1401	1.30694e4	1384.45923	97.4844

### 6.7.7 Acryl-T7/T9/T16



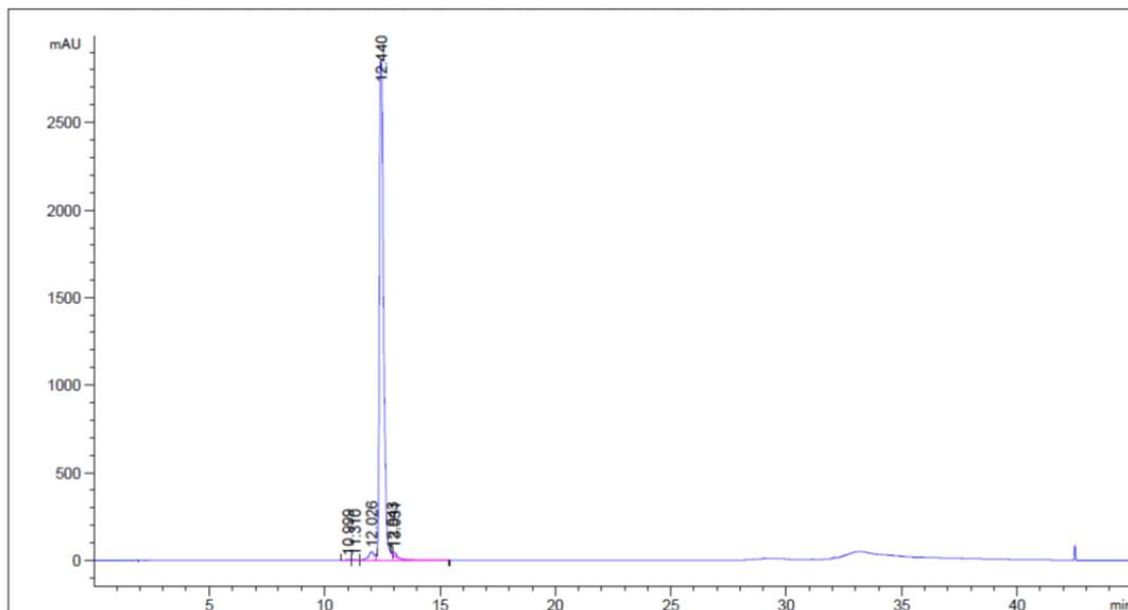
#### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	18.313	BV E	0.2067	42.35224	3.11375	0.4540
2	18.608	VV E	0.2198	53.16344	3.30259	0.5699
3	19.408	VB R	0.2290	9232.51855	622.73004	98.9760

## 6.7.8 TBA15



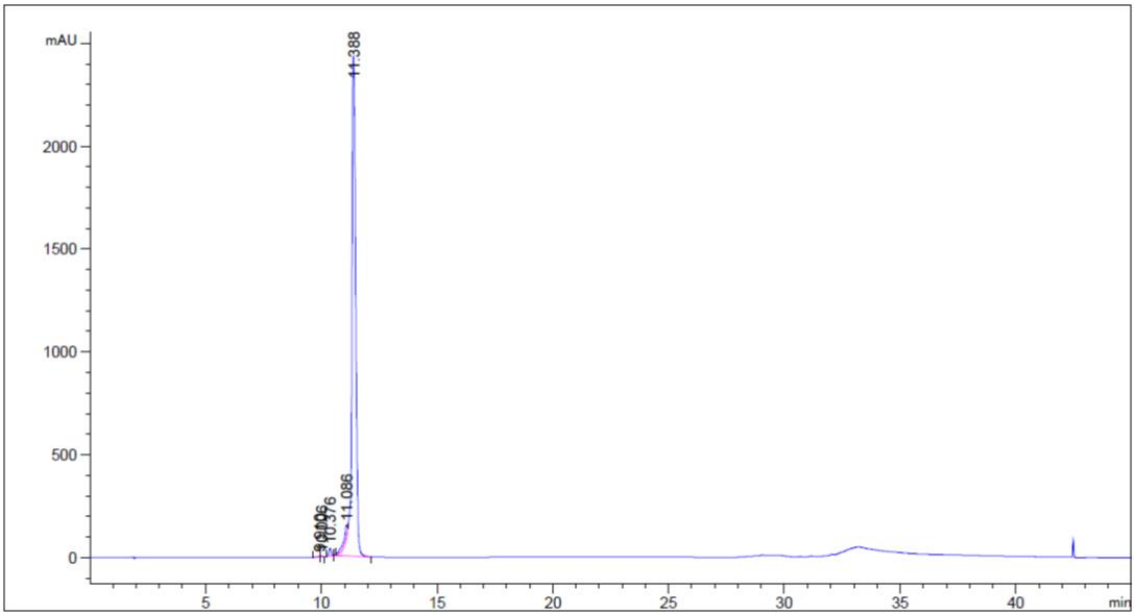
### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	10.999	BV	0.1414	39.91587	4.17932	0.1045
2	11.310	VB	0.1507	35.89089	3.58824	0.0940
3	12.026	BV E	0.2617	871.17682	47.55590	2.2805
4	12.440	VV R	0.1992	3.69481e4	2850.73071	96.7195
5	12.943	VV E	0.0721	51.30489	9.95322	0.1343
6	13.031	VB E	0.1729	254.88455	20.52154	0.6672

6.7.9 TBA15 complementary



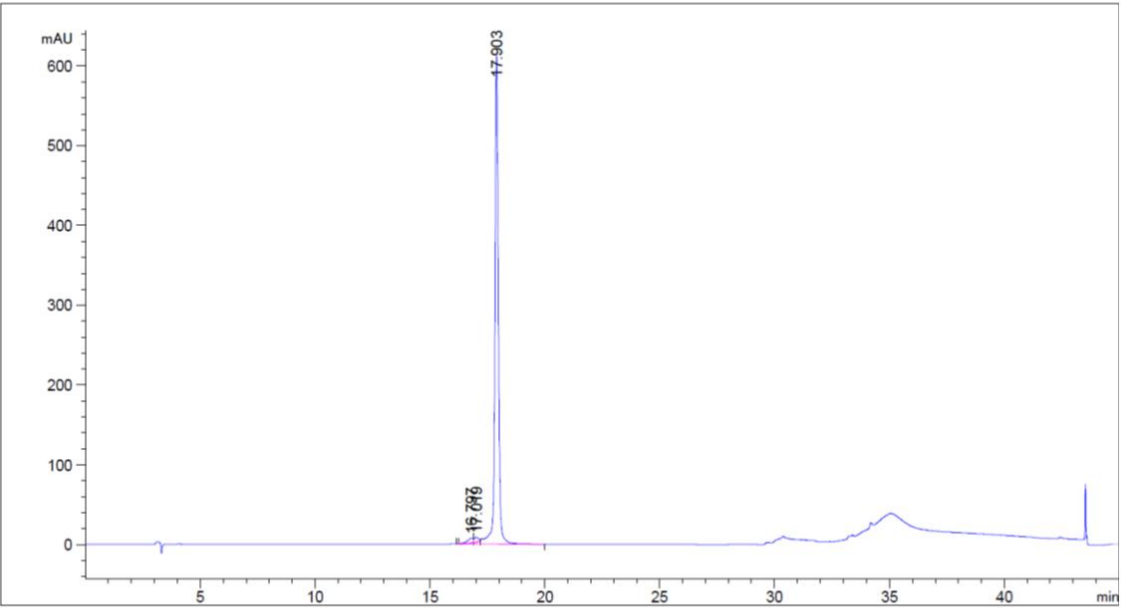
=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	9.910	BV	0.1096	25.88356	3.29750	0.0866
2	10.006	VB	0.1043	29.47282	4.28074	0.0986
3	10.376	BB	0.1068	255.39516	37.79216	0.8544
4	11.086	BV E	0.1833	792.41296	58.02243	2.6509
5	11.388	VB R	0.1802	2.87891e4	2429.93213	96.3095

6.7.10 TBA29



=====  
Area Percent Report  
=====

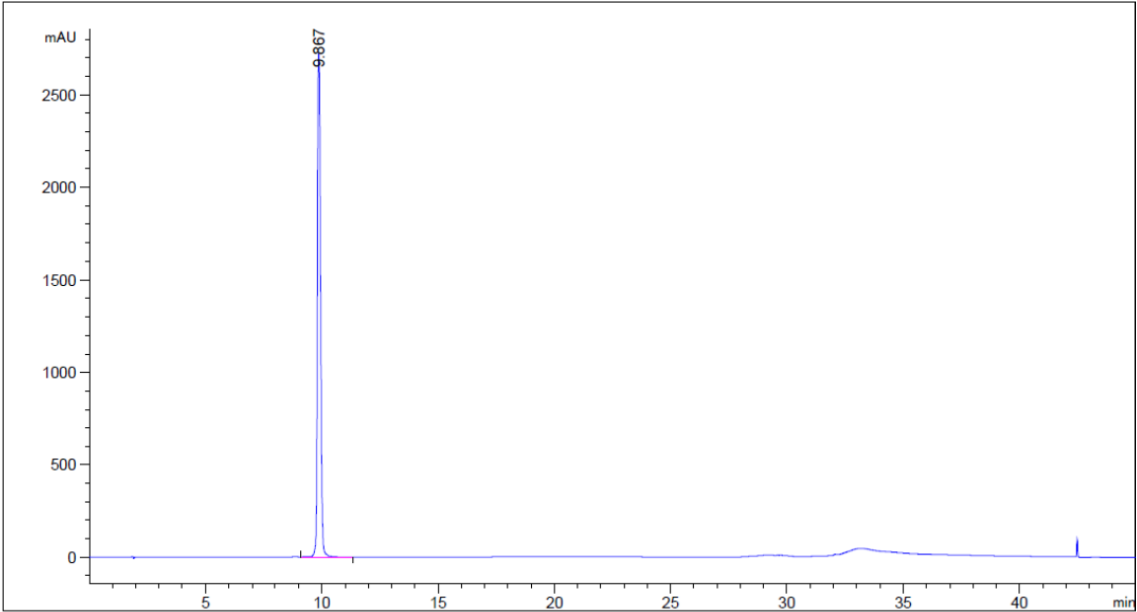
Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	16.797	BV E	0.2350	106.78830	6.32494	1.6035
2	17.019	VV E	0.1830	80.17532	6.36255	1.2039
3	17.903	VB R	0.1572	6472.79150	613.02246	97.1926



6.7.11 TBA15-A



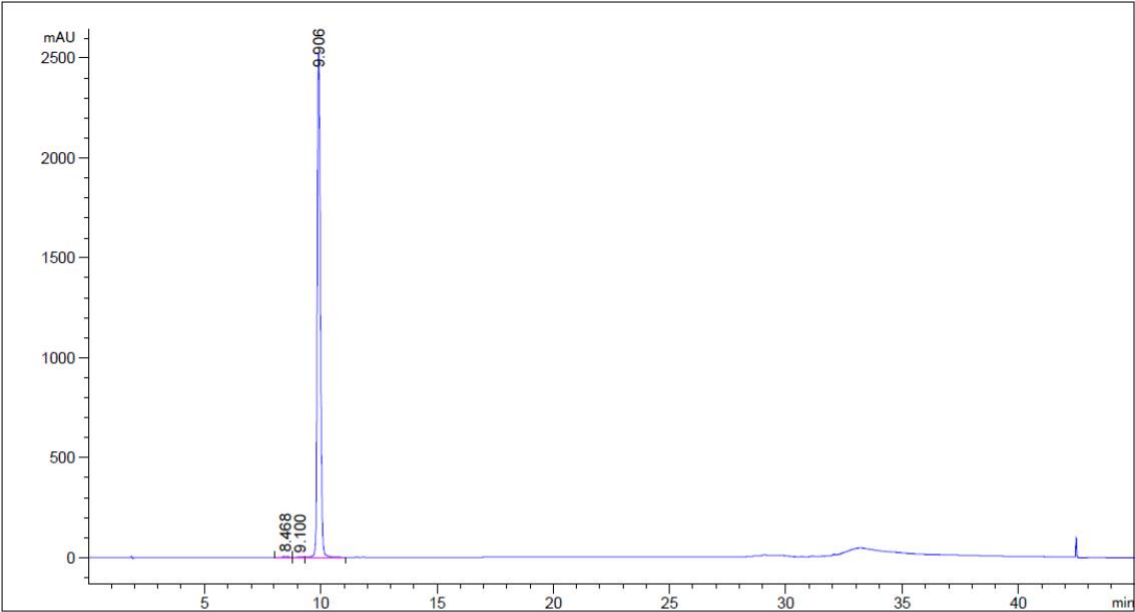
=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	9.867	VB R	0.1493	2.55043e4	2722.28784	100.0000

6.7.12 TBA15-B



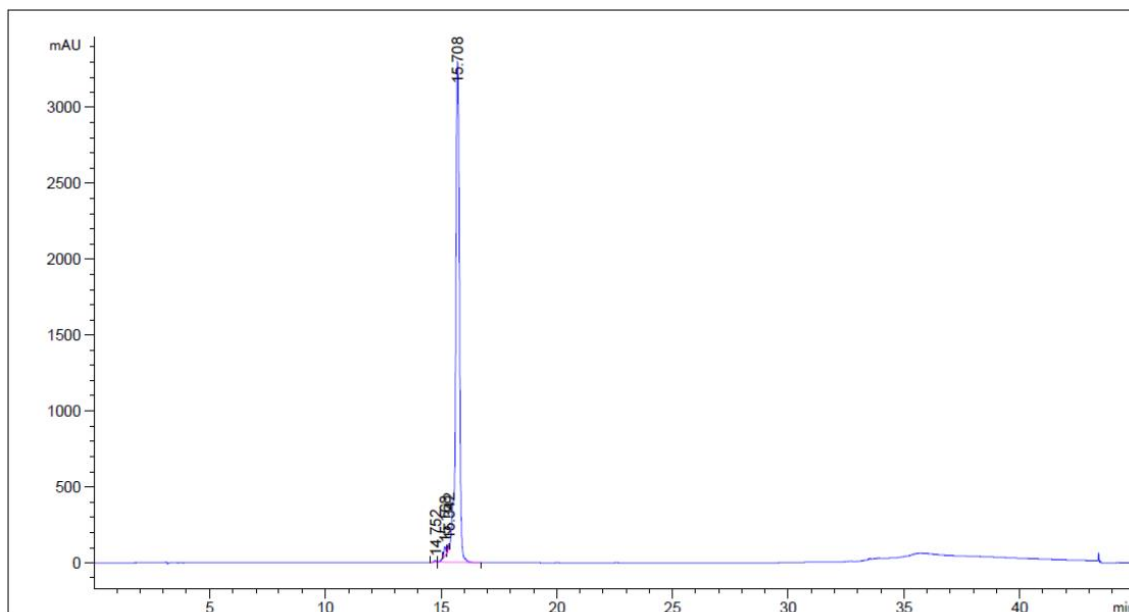
=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	8.468	BV	0.1727	62.47841	5.03726	0.2648
2	9.100	VV E	0.1566	18.69323	1.80823	0.0792
3	9.906	VB R	0.1467	2.35102e4	2522.53345	99.6559

### 6.7.13 TBA15-Aex



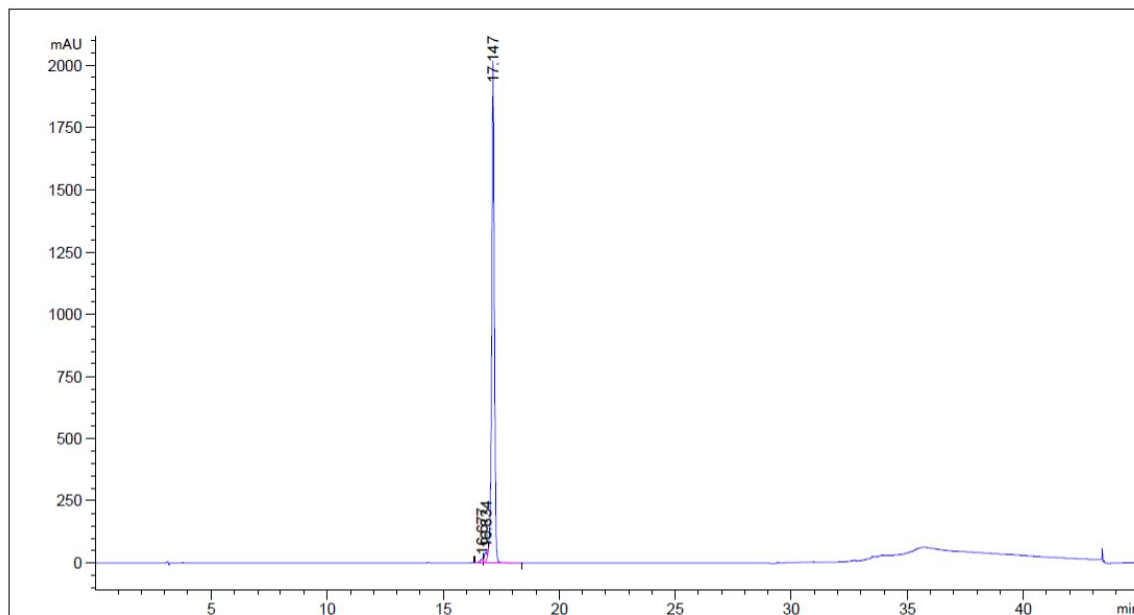
#### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	14.752	BB	0.1084	79.87394	11.59521	0.1994
2	15.168	BV E	0.0970	355.59683	58.46492	0.8877
3	15.342	VV E	0.1043	151.33327	21.43991	0.3778
4	15.708	VB R	0.1796	3.94717e4	3297.88550	98.5351

### 6.7.14 TBA15-Bex



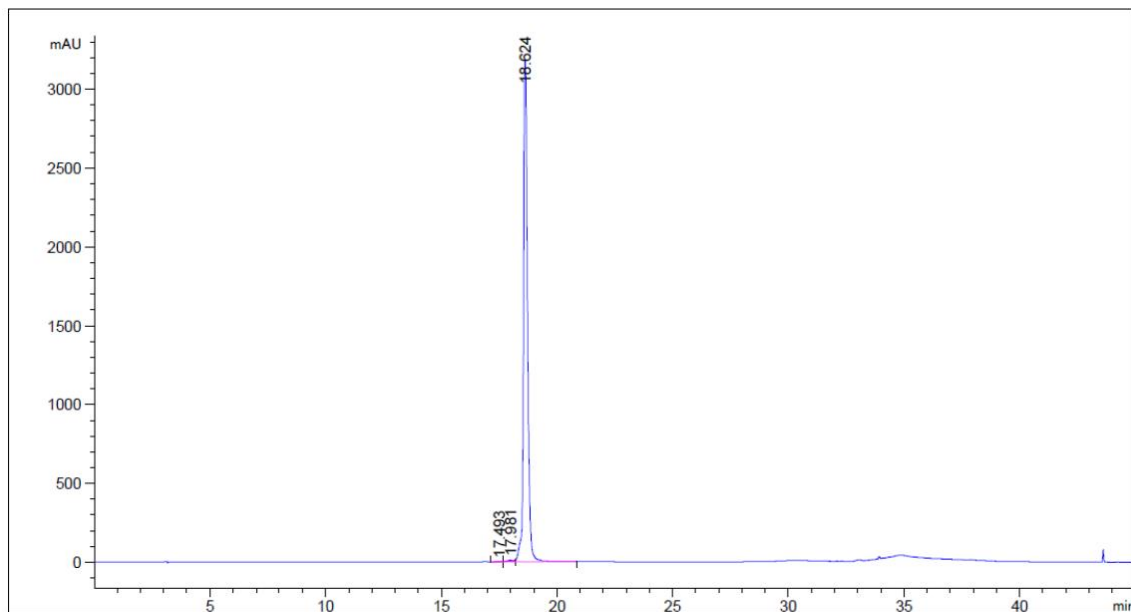
#### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	16.677	BV E	0.1118	115.46568	15.34862	0.6793
2	16.834	VV E	0.1203	299.58453	36.29262	1.7625
3	17.147	VB R	0.1239	1.65831e4	2017.75635	97.5583

### 6.7.15 TBA15-alkyne



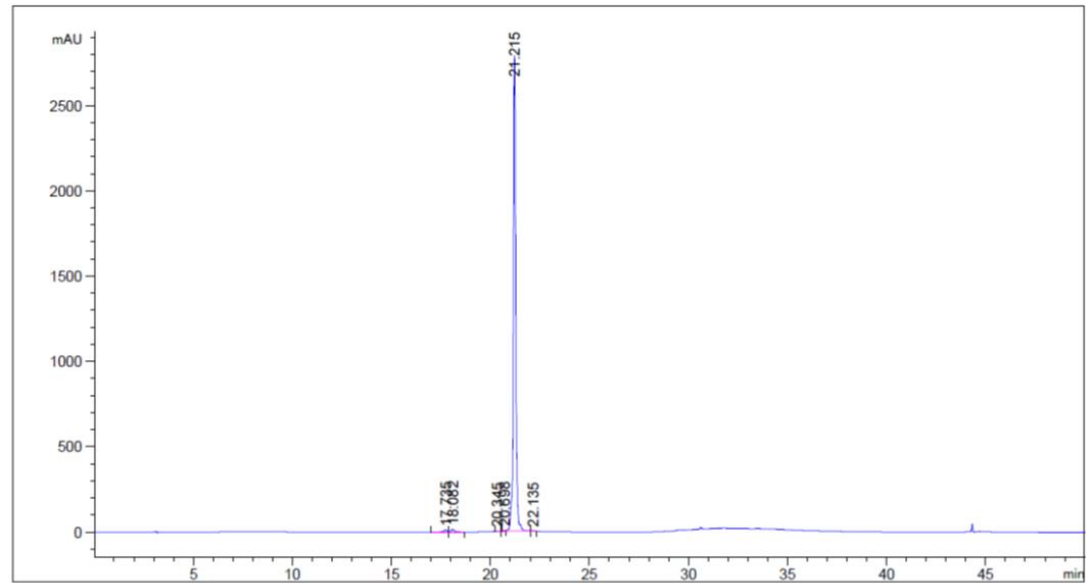
#### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	17.493	BV E	0.1769	24.07919	1.99437	0.0627
2	17.981	VV E	0.2494	172.03165	9.14680	0.4477
3	18.624	VB R	0.1823	3.82291e4	3179.43628	99.4896

6.7.16 TBA15-DBCO



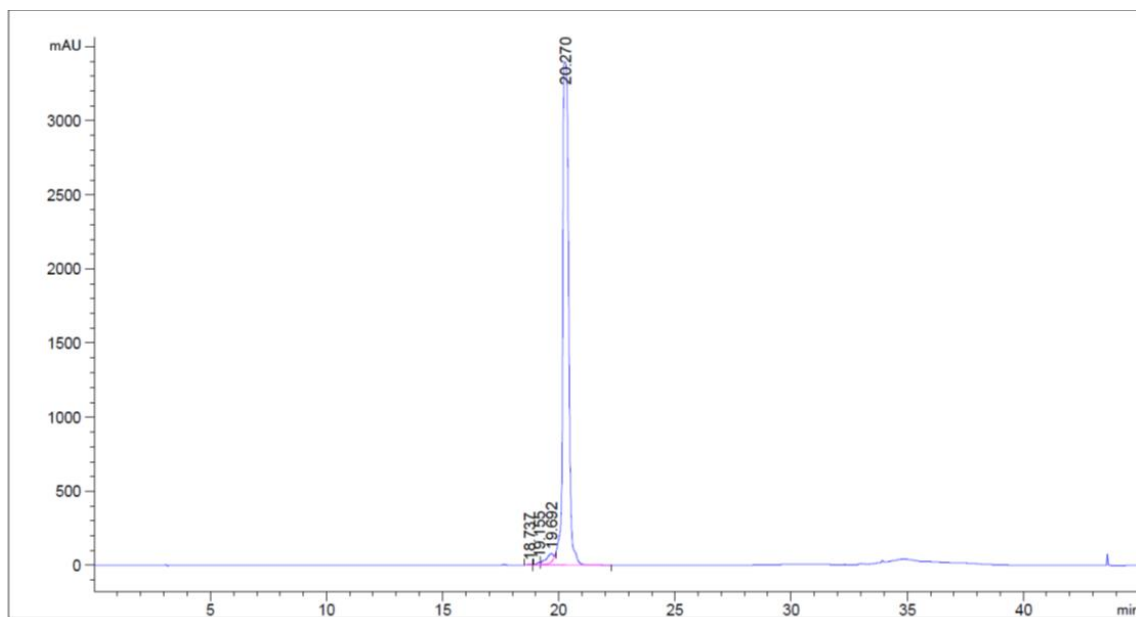
=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 B, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	17.735	BV	0.1683	124.64729	11.16322	0.5119
2	18.082	VB	0.1685	163.39066	14.61020	0.6710
3	20.345	VB	0.1277	19.08033	2.23201	0.0784
4	20.698	BV E	0.1158	22.34124	3.03616	0.0918
5	21.215	VV R	0.1305	2.39994e4	2786.91528	98.5651
6	22.135	BB	0.1295	19.92305	2.43408	0.0818

### 6.7.17 TBA29-alkyne



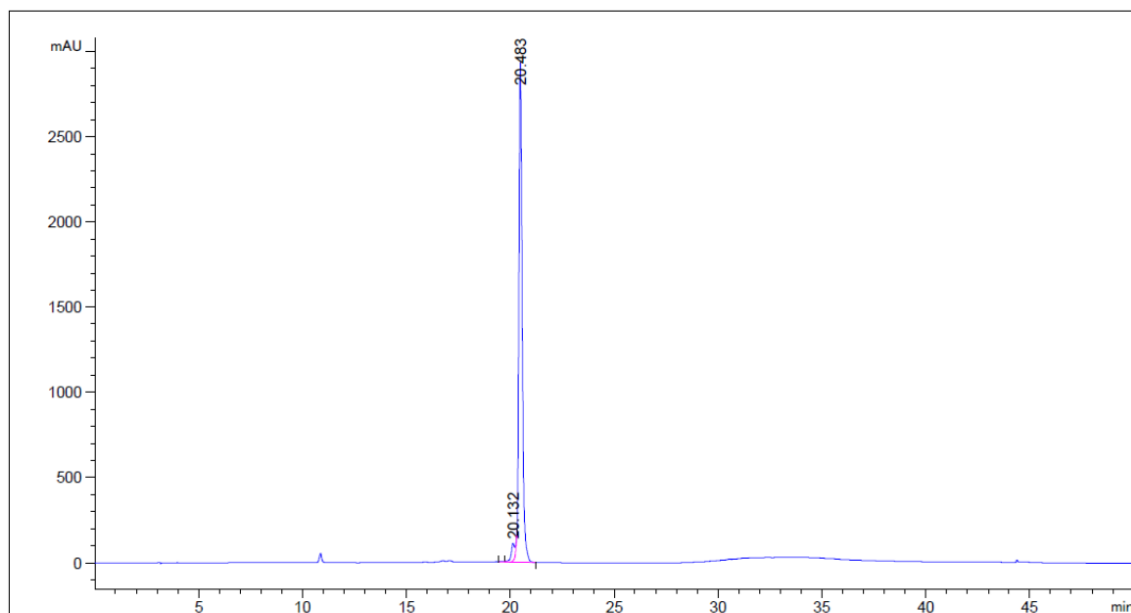
#### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	18.737	BB	0.1362	36.72422	4.19364	0.0585
2	19.155	BV E	0.1362	135.33910	15.16028	0.2154
3	19.692	VV E	0.3022	1302.25952	55.81950	2.0730
4	20.270	VB R	0.2170	6.13451e4	3392.05493	97.6531

## 6.7.18 TBA29-DBCO



### Area Percent Report

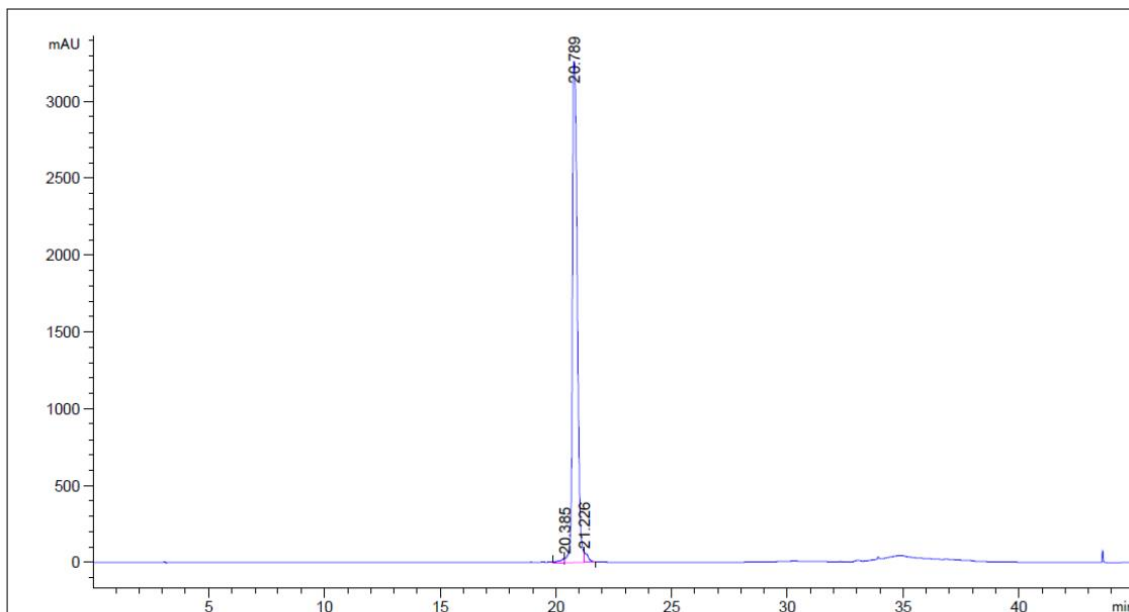
Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 B, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	20.132	VV E	0.1781	1268.68152	102.73675	3.6698
2	20.483	VB R	0.1725	3.33024e4	2928.60840	96.3302



## 6.7.19 nina-alkyne



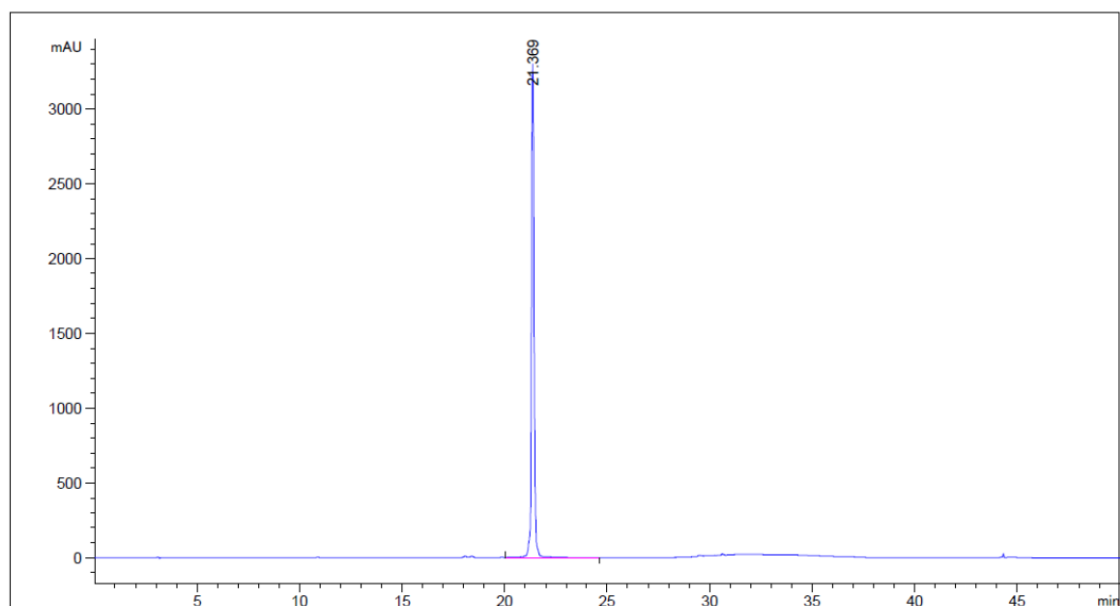
### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 F, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	20.385	MF	0.2555	426.00830	27.78962	0.8600
2	20.789	FM	0.2468	4.83689e4	3266.11621	97.6385
3	21.226	FM	0.1982	743.82953	62.53381	1.5015

## 6.7.20 nina-DBCO



### Area Percent Report

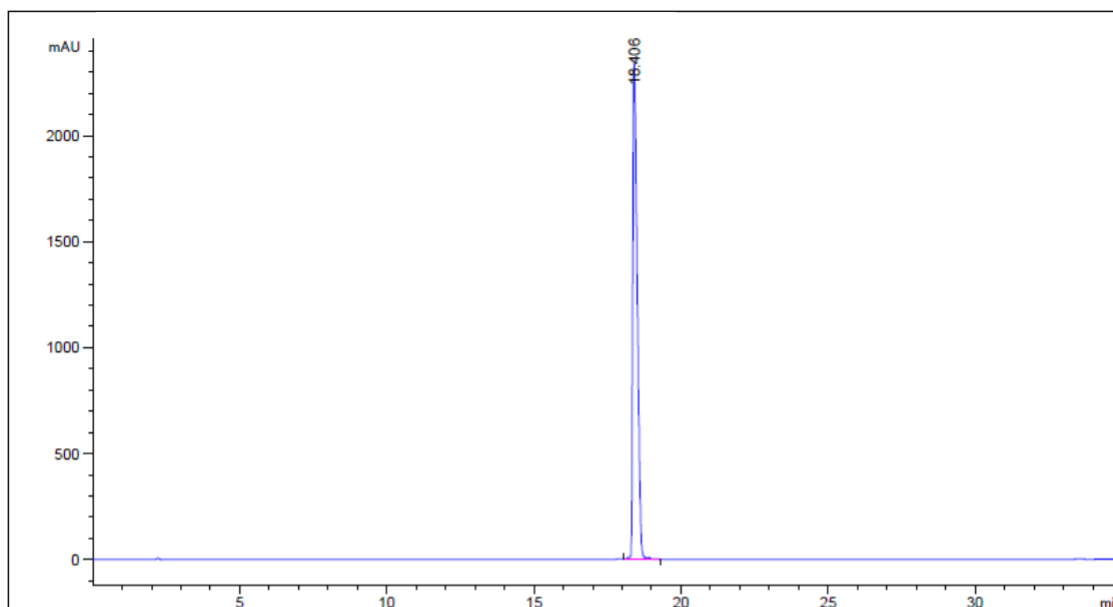
Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Do not use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 B, Sig=260,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	21.369	VV R	0.1396	2.99033e4	3304.60083	100.0000

## 6.8 Analytical HPLC chromatograms of bischolesterol dye azides

### 6.8.1 Compound **23**- bischolesterol Cy3 azide



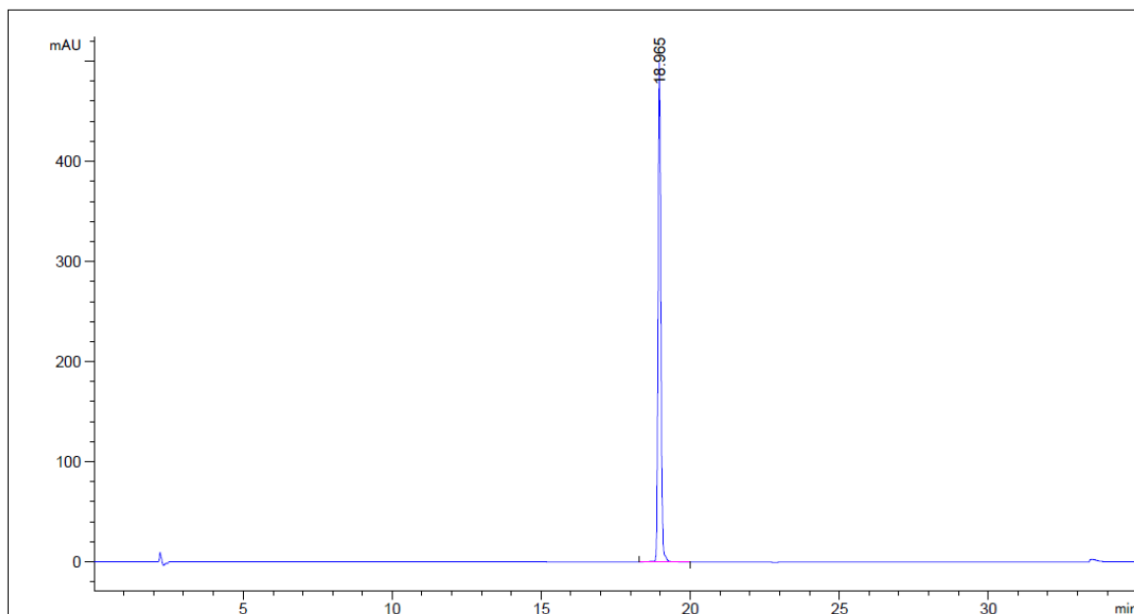
#### Area Percent Report

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 C, Sig=550,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	18.406	VV R	0.1642	2.46958e4	2344.33301	100.0000

### 6.8.2 Compound **24**- bischolesterol Cy5 azide



=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000  
Use Multiplier & Dilution Factor with ISTDs

Signal 1: MWD1 D, Sig=645,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	18.965	VB R	0.1013	3317.41479	500.49576	100.0000

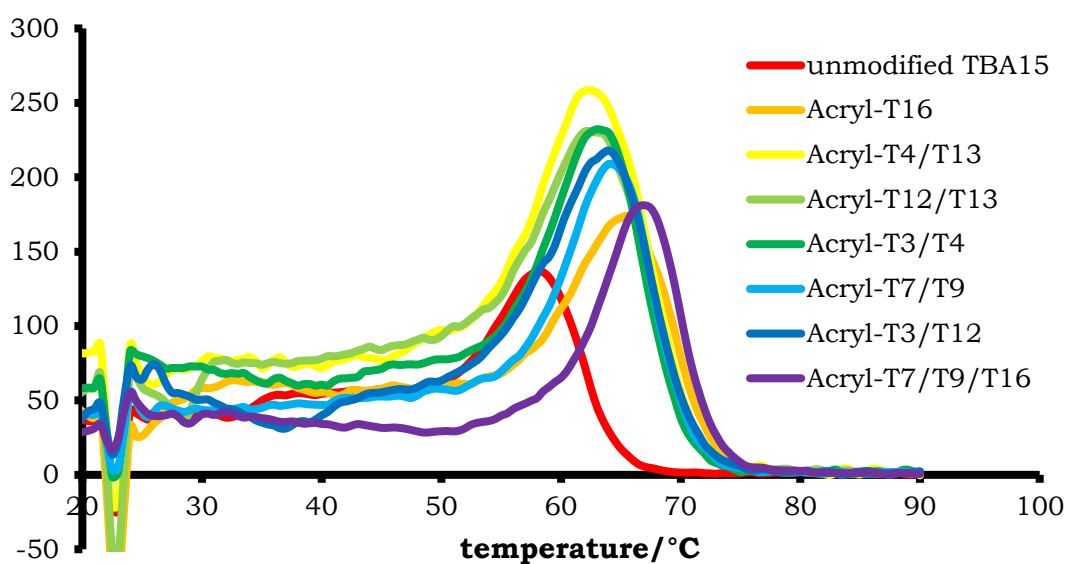
## 6.9 Mass spectrometry characterisation of transmembrane molecules

Synthetic strategy	Name	Predicted mass* ( <i>m/z</i> )	Observed mass* ( <i>m/z</i> )
Strategy one	TBA15-bischolesterol-Cy3	6230	6230
	TBA15-bischolesterol-Cy5	6257	6255
	TBA29-bischolesterol-Cy3	10589	10588
	nina-bischolesterol-Cy3	6047	6047
	nina-bischolesterol-Cy5	6074	6072
Strategy two	TBA15-bischolesterol-Cy3	6640	6640
	TBA15-bischolesterol-Cy5	6667	6666
	TBA29-bischolesterol-Cy5	11026	11025
	nina-bischolesterol-Cy3	6457	6457
	nina-bischolesterol-Cy5	6484	6484
Strategy four	TBA15 bischolesterol Cy5	6204	6206
	nina-bischolesterol-Cy3	5996	5995
	nina-bischolesterol-Cy5	6022	6022
DNA- bischolesterol <i>via</i> triazole	TBA15-bischolesterol	5821	5821
	nina-bischolesterol	5638	5638
DNA- bischolesterol <i>via</i> phosphate	nina-bischolesterol middle	5186	5185
	nina-bischolesterol end	5490	5490

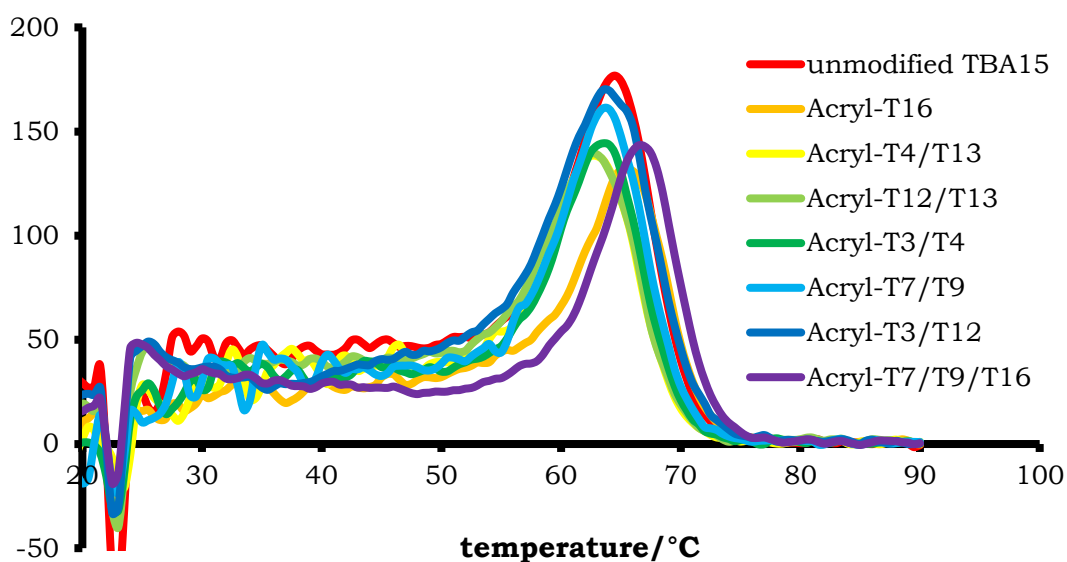
\*masses without dye negative counterion predicted and found.

## 6.10 $T_m$ data for thermal melting experiments of acrylamide modified thrombin binding aptamer strands with complementary sequence

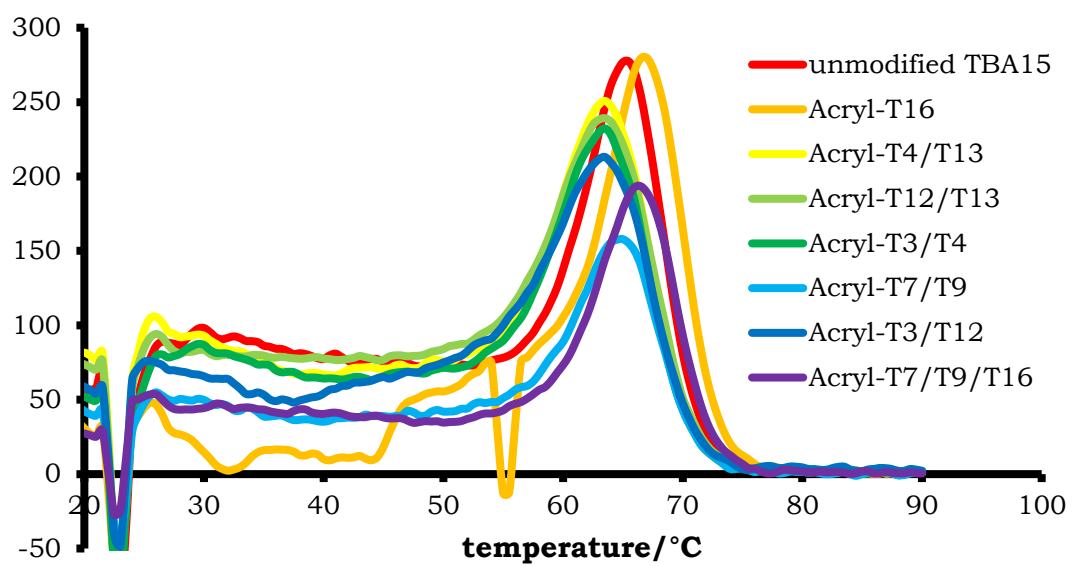
### 6.10.1 First thermal melting experiment



### 6.10.2 Second thermal melting experiment

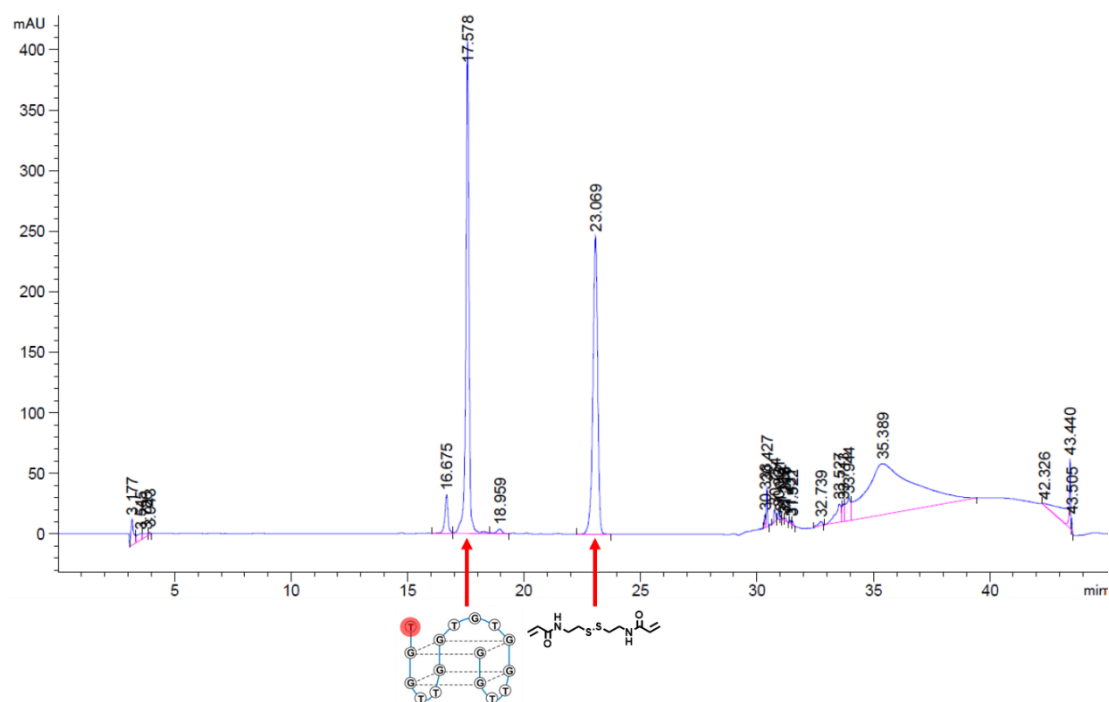


### 6.10.3 Third thermal melting experiment

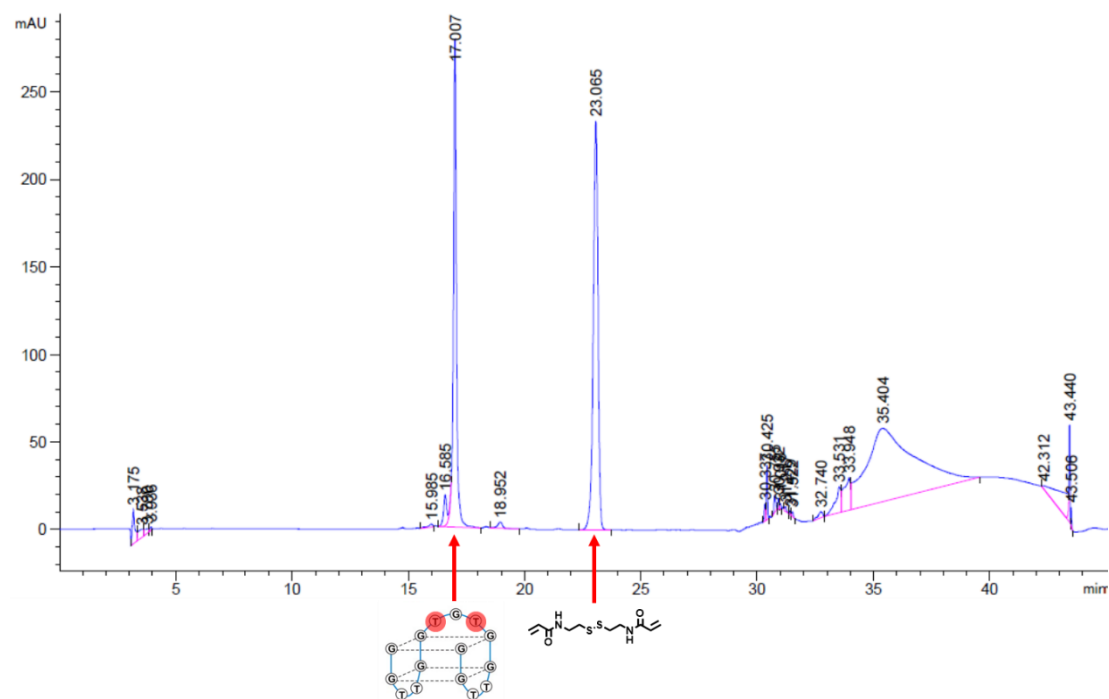


## 6.11 Data for control HPLC experiments with acrylamide DNA and BAC with/without TEMED and APS

### 6.11.1 BAC and Acryl-T16

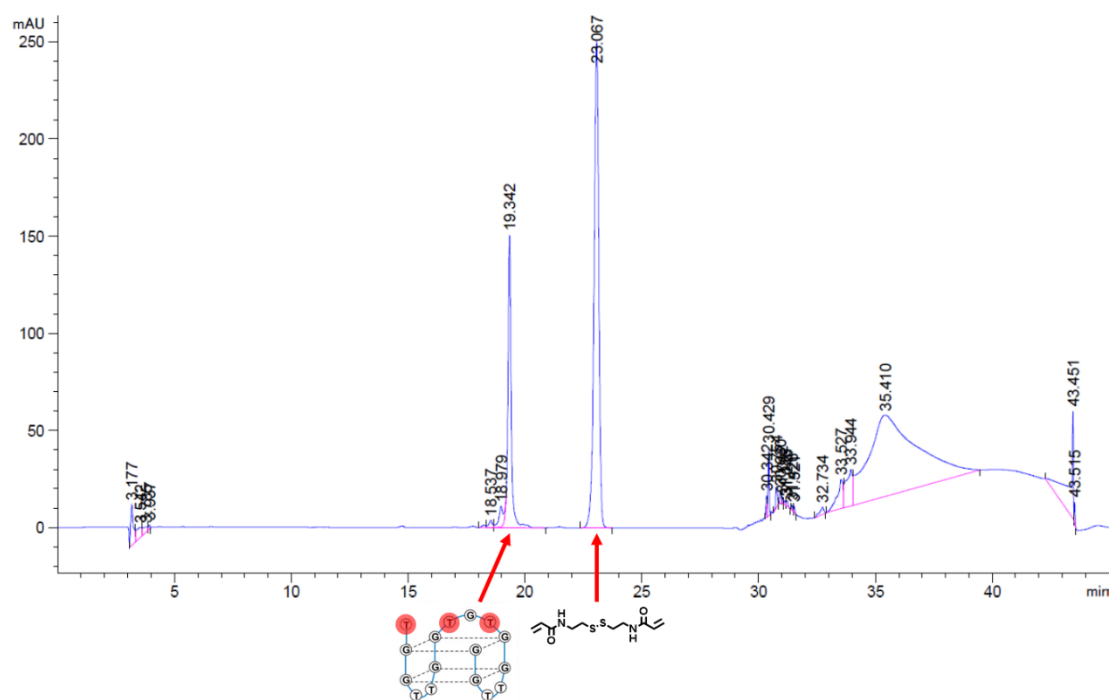


### 6.11.2 BAC and Acryl- T7/T9

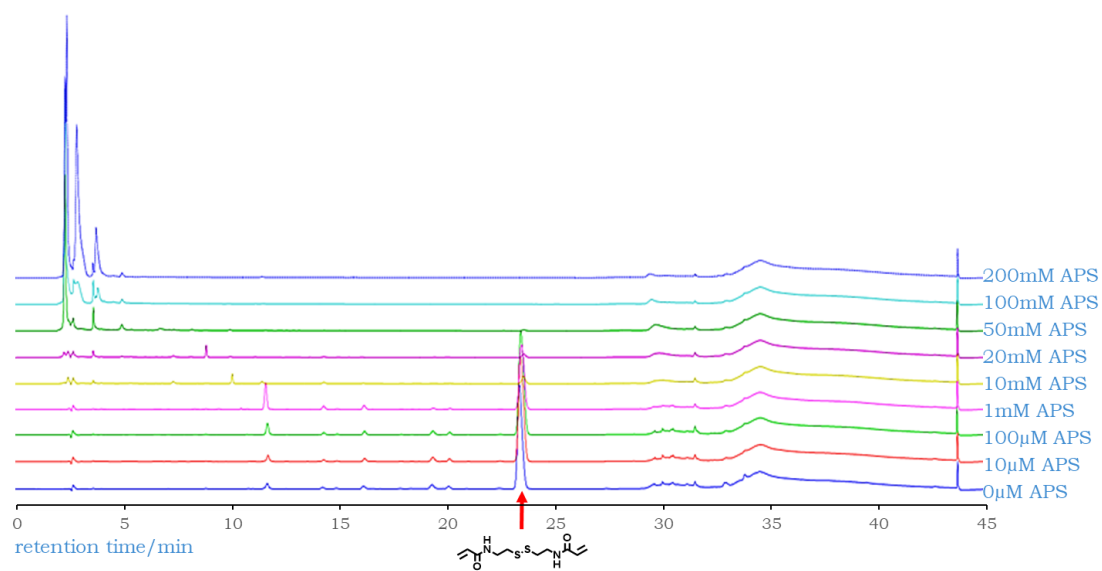




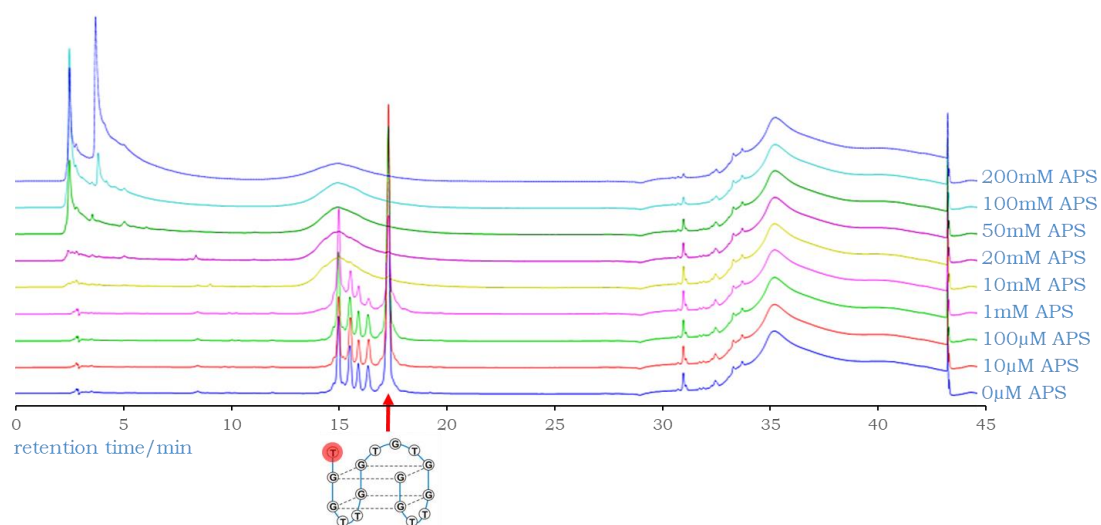
### 6.11.3 BAC and Acryl-T7/T9/T16



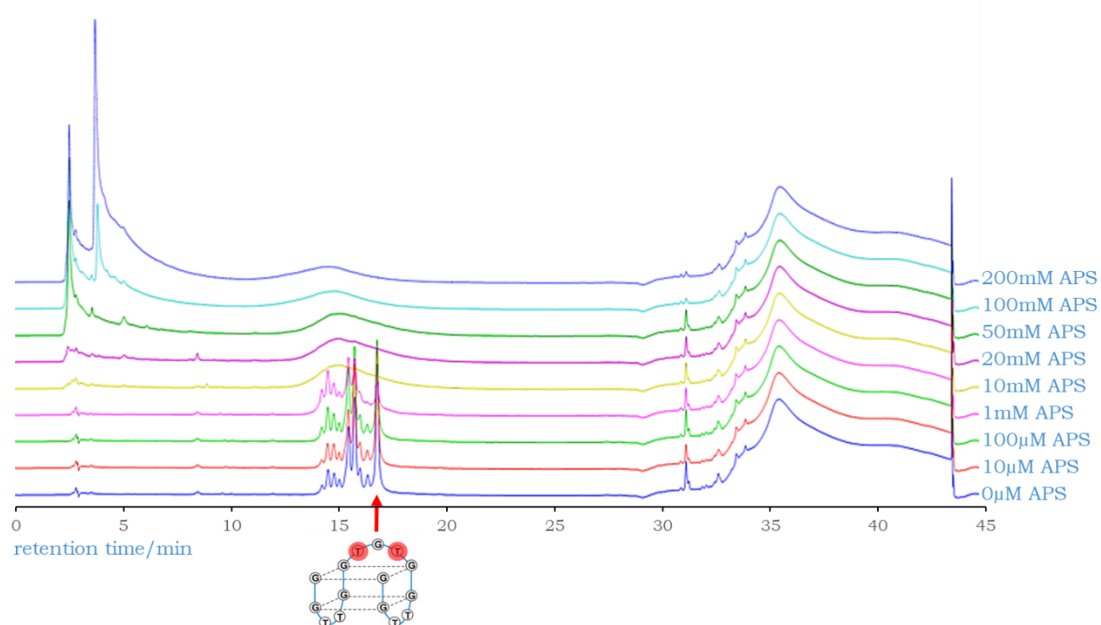
### 6.11.4 BAC and TEMED with increasing concentrations of APS



#### 6.11.5 Acryl-T16 and TEMED with increasing concentrations of APS



#### 6.11.6 Acryl- T7/T9 and TEMED with increasing concentrations of APS



### 6.11.7 Acryl-T7/T9/T16 and TEMED with increasing concentrations of APS

