

**PNA-protein conjugates for nano-scale modeling of protein
aggregates**

Zahra Gholami

School of Science and Technology

Nottingham Trent University

Director of Study:

Dr. Quentin S. Hanley

Thesis submitted for the Degree of Doctor of Philosophy (PhD)

November 2013

This research is part of “Nanoactuatuion of plasma membrane protein clustres” project

Acknowledgment

This research would not have been possible without the support of many people. First and foremost, I would like to thank Dr. Quentin Hanley for giving me the opportunity to work in his group as well as his invaluable support, guidance and supervision. I would also like to appreciate Dr. Elizabetta Verderio Edwards and Dr. Christopher Garner as my second and third supervisors.

I express my gratitude to members of NanoActuate consortium: Prof. dr. ir Luc Brunsveld, the advisor of this research, and his laboratory staff especially Dr. Dung T. Dang and Dr. Hoang D. Nguyen at Eindhoven University of Technology for their kind help and discussions during my stay in Netherlands and for plasmid PHT581 and help with protein expression. I also thank Prof. dr. Vinod Subramaniam (University of Twente) as well as Dr. Peter Verveer and Jenny Ibach at Max Planck Institute of Molecular Physiology of Dortmund for accepting me in their group and for *in vivo* anisotropy experiments. My deepest appreciation is due to my dearest friend, Zahra Zolmajd Haghighi for her kindness and great support at work and in life in Nottingham.

I acknowledge funding from NanoSci-E+ as a part of the NanoActuate consortium. In the UK, NanoActuate was administered by EPSRC as EP/H00694X/1.

I also greatly thank the other members of our group at NTU, Dr. Homanaz Ghafari and Dr. Kathryn Murray for their kind help and useful discussions.

Last but not the least; I would like to deeply thank my loving family for all their kind support, patience and encouragement.

Abstract

Programmable assembly of proteins on molecular frameworks requires the development of facile and orthogonal chemical approaches and molecular scaffolds. In this research, the unique characteristics of PNA were applied to create controllable protein assemblies directed by precise PNA-DNA hybridization. The signatures of assembly were studied via FRET, providing a powerful tool which should be effective in live system imaging. Two model systems were developed in this study. In the first model system, site-selective conjugation of monomeric teal fluorescent protein (mTFP) to PNA was achieved by covalent linkage of mTFP to PNA via expressed protein ligation. The mTFP-PNA conjugates were efficiently aligned on a DNA beacon, to create a hetero-FRET system. The FRET indicated by decrease of fluorescence intensity and lifetime of the donor and an increase of donor anisotropy. The assembly of similar multiple mTFP-PNA constructs on DNA scaffolds provided dimeric and oligomeric forms which were studied by SEC-HPLC and SDS-PAGE. A decrease of anisotropy was exhibited due to homo-FRET following induced formation of dimers and oligomers. In the second model system, fluorescent SNAP-PNA conjugates were controllably assembled on DNA frameworks forming dimers and oligomers. The site-selective covalent linkage of peptide nucleic acid (PNA) to SNAP protein was achieved by modification of PNA with O⁶-benzyl guanine (BG) which is a specific substrate for SNAP. The modified BG-PNA has been labeled with Atto dyes and thereafter, chemo-selectively conjugated to SNAP protein. Efficient assembly into dimeric and oligomeric forms were observed using SEC-HPLC and SDS-PAGE. DNA directed assembly of homo- and hetero-dimers of SNAP-PNA constructs induced homo- and hetero-FRET, respectively. Longer DNA scaffolds controllably aligned similar fluorescent SNAP-PNA constructs into higher oligomers exhibiting homo-FRET. These systems could readily extend from homodimers and oligomers to binary, ternary, and higher oligomer systems containing any number of different dyes or fluorescent proteins in precisely engineered arrangements.

Publication

1. Z. Gholami, L. Brunsveld and Q. Hanley, PNA-Induced assembly of fluorescent proteins using DNA as a framework, *Bioconjugate Chem*, Vol. 24, 2013, 1378-1386.
2. Z. Gholami, Q. Hanley, Fluorescence anisotropy in a protein: DNA system undergoing inducible assembly, *Biophysical Society's 57th Annual meeting*, Philadelphia, USA, 2013.
3. Z. Gholami, Q. Hanley and L. Brunsveld, Induced assembly of fluorescent proteins-PNA conjugate using DNA as a framework, *Meeting of the British Biophysical Society*, Durham, UK, 2012.

Contents

Acknowledgment	i
Abstract.....	ii
Publication	iii
Contents.....	iv
List of Figures	vii
List of Schemes.....	ix
List of Tables	x
List of Abbreviations	xi
CHAPTER 1	1
Assembly of PNA- protein conjugate: Application to study controlled assembly of proteins	1
1.1 Peptide nucleic acid (PNA)	2
1.1.1 PNA conjugation.....	3
1.2 Native chemical ligation.....	5
1.3 Expressed protein ligation	7
1.4 The importance of protein clustering	10
1.5 Clustering based on PNA.....	11
1.5 Protein tags	13
1.6 Aims and outline of thesis.....	16
1.7 References	18
CHAPTER 2	23
Conjugation of fluorescent protein to PNA	23
2.1 Introduction	24
2.2 Results and Discussion	25
2.2.1 Expression and purification of fluorescent proteins with a thioester group at the C-terminus using EPL method.....	25
2.2.2 PNA-mTFP ligation by NCL method.....	27
2.2.3 Kinetic study of mTFP-PNA NCL	28

2.4 Conclusion and outlook	38
2.5 Experimental part	38
2.5.1 Protein expression and purification.....	39
2.5.2 PNA- Protein ligation.....	41
2.6 References	43
CHAPTER 3	46
Induced assembly of fluorescent protein-PNA conjugates using DNA as a framework	46
3.1 Introduction	47
3.1.1 DNA beacon	49
3.1.2 Florescence resonance energy transfer (FRET).....	50
3.1. 3 Fluorescence anisotropy	51
3.1.4 Fluorescence lifetime.....	53
3.2 Results and Discussion	54
3.2.1 Size exclusion HPLC analysis of mTFP-PNA assembly on DNA scaffolds.....	54
3.2.2 Hetero-FRET occurrence after hybridization of PNA-mTFP to DNA beacon.....	59
3.2.3 Homo-FRET system demonstrating dimer and oligomer assembly.....	63
3.3 Conclusion and outlook	66
3.4 Experimental part	66
3.4.1 Hybridization and assembly of mTFP-PNA with DNA	66
3.4.2 Photophysical measurements.....	67
3.5 References	69
CHAPTER 4	72
Coupling of Purified SNAP protein with fluorescent PNA	72
4.1 Introduction	73
4.2 Results and Discussion	75
4.2.1 Labeling of PNA to BG substrates	76
4.2.2 Labeling of BG -PNA conjugates with fluorescent dyes	81
4.2.3 Coupling of BG-PNA-fluorophore to SNAP protein.....	86

4.3 Conclusion and outlook	91
4.4 Experimental Part	92
4.4.1 Conjugation of BG to PNA.....	92
4.4.2 Conjugation of BG-PNA to fluorophore	92
4.4.3 Coupling of purified SNAP protein with BG-PNA-Atto 488 and 532 constructs	93
4.4.4 Dialysis purification of SNAP-PNA-Atto constructs.....	93
4.4.5 SDS-PAGE electrophoresis of SNAP-PNA-Atto 488 and 532 conjugates.....	93
4.4.6 MALDI-TOF mass spectrometry analysis.....	94
4.4.7 RP-HPLC of BG-PNA and BG-PNA-Atto 488 and 532 conjugates	94
4.4.8 SEC-HPLC of SNAP-PNA-Atto 488 and SNAP-PNA-Atto 532 conjugates	94
4.5 References	94
CHAPTER 5	96
Hetero- and Homo-FRET systems created by assembly of fluorescent SNAP-PNA units in dimer and higher oligomer forms.....	96
5.1 Introduction	97
5.2 Results and Discussion	98
5.2.1 Assembly of fluorescent SNAP-PNA in dimer and higher oligomer forms	98
5.2.2 Hetero-FRET system.....	104
5.2.3 Homo-FRET system	108
5.3 Conclusion and outlook.....	112
5.4 Experimental part	113
5.4.1 Hybridization and assembly of fluorescent SNAP-PNA with DNA	113
5.4.2 SDS-PAGE electrophoresis analysis.....	113
5.4.3 SEC-HPLC analysis	113
5.4.4 Photophysical measurements.....	114
5.5 References	114
SUMMARY	116
6.1 References	119

List of Figures

Figure 2.1: SDS-PAGE result of purified FP-thioesters.....	26
Figure 2.2: Mass spectrometry results of mTFP ligation to PNA.....	28
Figure 2.3 Fitting of the final mTFP-PNA yield in a pseudo first order growth model.....	32
Figure 2.4: MTFP-PNA ligation yield vs. catalyst: MPAA by MALDI-TOF mass spectrometry.....	35
Figure 2.5: MTFP-PNA Ligation yield vs. catalyst: MPAA by UV.Vis spectrophotometry.....	36
Figure 3.1: SDS-PAGE results of mTFP-PNA dimer.....	54
Figure 3.2: SDS-PAGE results of mTFP-PNA tetramer.....	55
Figure 3.3: SEC-HPLC calibration curve.....	56
Figure 3.4: SEC-HPLC results of mTFP-PNA dimer assembly on DNA1.....	57
Figure 3.5: SEC-HPLC results of mTFP-PNA tetramer assembly on DNA2.....	58
Figure 3.6: SEC analysis of titrating DNA2 with mTFP-PNA.....	59
Figure 3.7: Spectral overlap between mTFP and 6-FAM.....	60
Figure 3.8: Energy transfer efficiency of mTFP-PNA: DNA beacon hetero-FRET system.....	61
Figure 3.9: Polar co-ordinate presentation of mTFP-PNA: Beacon: PNA2 complexes.....	62
Figure 3.10: Anisotropy changes of the DNA1-mTFP-PNA homo-FRET system.....	64
Figure 3.11: Anisotropy changes of the DNA2-mTFP-PNA homo-FRET system.....	65
Figure 4.1a: Mass spectrometry data of BG-PNA1.....	78
Figure 4.1b: Mass spectrometry data of BG-PNA2.....	78
Figure 4.1c: Mass spectrometry data of BG-PNA3.....	79
Figure 4.2: RP-HPLC of BG-PNA1 and BG-PNA3.....	80
Figure 4.3: Spectra and chemical structure of Atto dyes and Alexaflour488.....	81
Figure 4.4a: Mass spectrometry of BG-PNA1-Atto488.....	84
Figure 4.4b: Mass spectrometry of BG-PNA3-Atto532.....	84
Figure 4.5: RP-HPLC of BG-PNA3-Atto532 at 532 nm (top).BG-PNA1-Atto488 at 501 nm (bottom).....	85
Figure 4.6: The normalized emission of free Atto488 and SNAP-PNA-Atto 488.....	86
Figure 4.7 a: Mass spectrometry of SNAP-PNA1-Atto488.....	87
Figure 4.7b: Mass spectrometry of SNAP-PNA3-Atto532.....	88
Figure 4.8: SDS-PAGE results SNAP-PNA-Atto conjugates.....	88
Figure 4.9: SEC-HPLC and SDS-PAGE results of SNAP-PNA-Atto532 and SNAP-PNA-Atto488.....	89
Figure 4.10: Effect of different ratio of substrate and incubation time at 37 and 25 °C on SNAP-PNA-Atto488 conjugation.....	90
Figure 5.1: SDS-PAGE results SNAP-PNA-Atto dyes dimers and oligomers.....	100
Figure 5.2: SEC-HPLC and SDS-PAGE results of dimer assembly of SNAP-PNA1-Atto488 on DNA1.....	102
Figure 5.3: SEC-HPLC and SDS-PAGE results of dimer assembly of SNAP-PNA1-Atto488 and SNAP-PNA3-Atto532 on DNA2.....	102
Figure 5.4: SEC-HPLC and SDS-PAGE results of terimer assembly of SNAP-PNA1-Atto488 on dna3.....	103
Figure 5.5: SEC-HPLC and SDS-PAGE results of terimer assembly of SNAP-PNA1-Atto488 on dna4.....	103
Figure 5.6: SEC analysis of titrating DNA4 with SNAP-PNA-Atto488.....	104

Figure 5.7: Spectral overlap between Atto488 emission and Atto532 excitation as a FRET pair.....	105
Figure 5.8: Decreasing the emission intensity of donor (SNAP-PNA1-Atto 488) after addition of the same ratio of acceptor (SNAP-PNA3-Atto532) in hetero-FRET system.....	106
Figure 5.9: Replicates observed for sensitized emission intensity	107
Figure 5.10: Anisotropy changes of the SNAP-PNA-Atto488:DNA1 homo FRET system.....	109
Figure 5.11: Anisotropy changes of the SNAP-PNA-Atto488:DNA3 homo FRET system.....	110
Figure 5.12: Anisotropy changes of the SNAP-PNA-Atto488:DNA4 homo FRET system.....	111

List of Schemes

Scheme 1.1: Structure of PNA compared to DNA.....	2
Scheme 1.2: Thioether bond formation in a Michael type reaction.....	4
Scheme 1.3: Conjugation of PNA-peptide by oxime formation.....	4
Scheme 1.4: Native chemical ligation.....	6
Scheme 1.5: Graphical representation of palsmid for mTFP-Intein-CBD construct	8
Scheme 1.6: Expressed protein ligation.....	9
Scheme 1.7: Traditional and current views of GPCR signaling	11
Scheme 1.8: α -Her2 PNA dimer and tetramer.....	13
Scheme 1.9: Schematic structure of GFP and its chromophore formation	14
Scheme 1.10: Schematic structure of SNAP-tag bound to its substrate.....	15
Scheme 2.1: Schematic depiction of the expressed recombinant FP constructs	26
Scheme 2.2: mTFP-PNA ligation via native chemical ligation (NCL).	29
Scheme 2.3: Reaction of thiol and thioester.	34
Scheme 2.4: C-terminal thioester introduced at the end of the protein.....	39
Scheme 3.1: DNA beacon structure in closed and open forms.	48
Scheme 3.2: Assembly of mTFP-PNA in tetramer form using DNA as a framework	49
Scheme 3.3: Graphical depiction of homo-FRET principle.....	53
Scheme 4.1: Coupling of BG-PNA-Atto dyes to the SNAP protein.....	75
Scheme 4.2: Conjugation of the cysteine residue of PNA to the BG-maleimide.....	76
Scheme 4.3: Conjugation of the lysine residue of PNA to the BG-NHS ester.	77
Scheme 4.4: Conjugation of the lysine residue of BG-PNA1 to the Atto488-NHS ester.....	82
Scheme 4.5: Conjugation of the aspartic acid residue of BG-PNA2 to Alexa fluor488 hydrazide	83
Scheme 5.1: Assembly of SNAP-PNA1-Atto constructs into homo-dimer and hetero-dimer	98
Scheme 5.2: Assembly of SNAP-PNA1-Atto constructs in higher oligomer forms	99
Scheme 5.3: Randomly assembled species at 1:1 stoichiometry of DNA2:monomer.....	108
Scheme 5.4: Different Inter-fluorophore distances between two monomer on DNA3	110

List of Tables

Table 2.1: Kinetic behaviour of PNA-mTFP NCL system	31
Table 2.2 : Errors estimated for the fitted parameters of mTFP-PNA pseudo first order model	33
Table 5.1: Oligonucleotide sequences were used to make assembled SNAP dimer and oligomers	99

List of Abbreviations

6-FAM	6-Carboxyfluorescein
AGT	O ₆ -alkylguanine-DNA alkyltransferase
BG	Benzylguanine
CBD	Chitin binding domain
Cys	Cysteine
DTE	Dithioerythritol
EDC	Carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EPL	Expressed protein ligation
FLIM	Fluorescence lifetime imaging microscopy
Fmoc	9-fluorenylmethyloxycarbonyl
FP	Fluorescent protein
FRET	Fluorescence resonance energy transfer
GFP	Green fluorescent protein
Gly	Glycine
GPCR	G-protein coupled receptor
GPI	Glycosylphosphatidylinositol
His	Histidine
HPLC	High-performance liquid chromatography
LB	Lysogeny broth
Lys	Lysine
MALDI-TOF	Matrix-assisted laser desorption/ionization-time of flight mass spectrometry
mCFP	Monomeric cyan fluorescent protein
mCherry	Monomeric cherry fluorescent protein
MESNA	2-Mercaptoethansulfonic acid
mGFP	Monomeric green fluorescent protein
mKate	Monomeric Kate fluorescent protein
mKO2	Monomeric Kusabira-Orange 2 fluorescent protein
MPAA	4-Mercaptophenylacetic acid
MS	Mass spectrometry
mTFP	Monomeric teal fluorescent protein
mYFP	Monomeric yellow fluorescent protein
NCL	Native chemical ligation
NHS	N-hydroxysuccinimidyl
Ni-NTA	Nickelnitrilotriacetate
PNA	Peptide nucleic acid
RP-HPLC	Reversed-phase HPLC
SDS	Sodium dodecyl sulphate
SEC-HPLC	Size exclusion HPLC
Ser	Serine
tBoc	Tert-butyloxycarbonyl
TC	Tetracysteine
TCEP	Tris (2-carboxyethyl) phosphine
TFA	Trifluoroacetic acid

TIRF	Total internal reflection fluorescence
TMD	Trans membrane domain
TR-FRET	Time-resolved fluorescence energy transfer
Tyr	Tyrosine

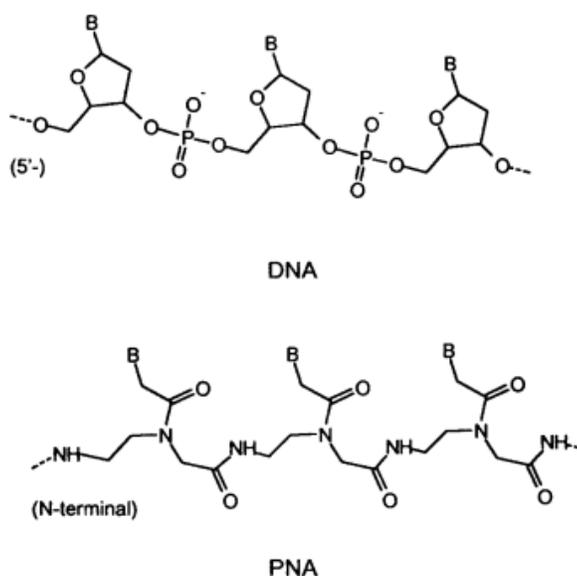
Chapter 1

Assembly of PNA-protein conjugate: Application to study controlled assembly of proteins

The nano scale supramolecular assemblies of proteins, such as those occurring in cell membranes during the first steps in signal transduction, play crucial roles in cell functions. The complex dynamic ordering of cell membrane proteins, such as clustering, typically changes after external stimulation. However, the mechanisms of these processes and their consequences on regulating cell growth, differentiation, shape changes and cell death are still not completely understood. Developing a protein model system to control over aggregation of proteins allows studying of protein clustering and its effect on downstream signaling. In this project, model systems based on assemblies of two popular protein tags, fluorescent proteins and SNAP, were created which can be easily applied to study other proteins by fusing them to these tags using molecular biology techniques. The assembly is based on the specific recognition of PNA for complementary DNA or PNA scaffolds. Peptide nucleic acids (PNA) are functional mimics of DNA with a pseudo-peptide backbone. Compared to DNA, the better biological robustness and unique chemistry of PNA makes it a useful component to controllably assemble biomolecules. The use of fluorescence resonance energy transfer (FRET) together with model systems undergoing PNA induced assembly can provide a new perspective to observe protein clustering.

1.1 Peptide nucleic acid (PNA)

PNA is a DNA mimic in which the regular phosphodiester backbone has been replaced by a pseudopeptide skeleton. The backbone consists of repeating N-(2-aminoethyl) glycine units with purine and pyrimidine nucleobases attached to the aminoethyl glycine nitrogens via a methylene-carbonyl linkage (Scheme 1.1).¹⁻⁵ PNA binds to the complementary PNA, DNA and RNA sequences through Watson-Crick base pairs and preferentially in the antiparallel manner (N-terminus of PNA facing the 3' end of a complementary oligonucleotide) with high affinity and sequence specificity. PNA-DNA and PNA-RNA hybrids are more stable than the equivalent oligonucleotide complexes because of the absence of electrostatic interference between their neutral chain and the polyanionic oligonucleotide.⁶⁻⁸ Due to resistance to nuclease and protease digestion and high thermal stability, PNA has been proposed as an antigen or antisense agent in molecular biology and in gene diagnosis and therapy.^{3,9}

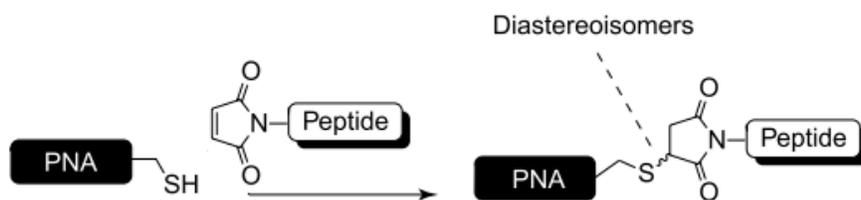


Scheme 1.1: Chemical structure of PNA in comparison to DNA. The backbone consists of repeating N-(2-aminoethyl) glycine units with purine and pyrimidine nucleobases attached to the aminoethyl glycine nitrogens via a methylene-carbonyl linkage⁹.

1.1.1 PNA conjugation

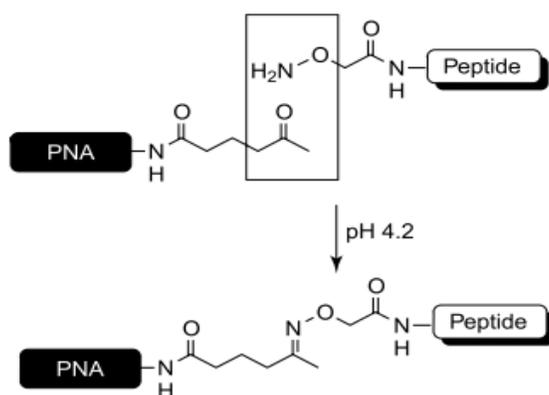
Due to the peptide structure of the PNA backbone, PNA conjugation can be performed using peptide bioconjugation techniques. Several kinds of peptide-peptide conjugation chemistries are available. For example, conjugation is possible through carboxyl, amine and thiol groups. Thiol becomes a powerful nucleophile at physiological pH in aqueous solutions which provides the most specificity and is conveniently introduced in peptides or PNA via a cysteine.^{2,3}

PNA-peptide conjugates can be produced through two strategies: on-line continuous solid phase peptide synthesis ((9-fluorenylmethyloxycarbonyl (Fmoc) or tert-butyloxycarbonyl (tBoc)) or conjugation of fragments through regular peptide-peptide conjugation methods.⁹ PNA-peptide ligation is done under conditions compatible with the amino acid side chains and PNA nucleobases using a variety of methods to link the PNA and peptide moieties.² In one of these, peptide and PNA fragments incorporating a cysteine residue at either the N- or C-terminus are ligated via a disulfide bond.^{4, 10} The direct oxidation of free thiol groups causes formation of desired products as well as symmetric dimers. Using an excess of one of the fragments can improve the ligation yield through increasing conversion selectivity.² However, it is worth noting that an excess of one part can result in the formation of more homo-dimers as well which may interfere during purification steps. A more profitable method to make disulfide bonds utilizes the nucleophilic substitution of a free thiol group of one fragment and an activated thiol group of the other part.¹¹ To prevent the unwanted dimerization of fragments possessing free thiol groups, the coupling reaction should be done in an oxygen free environment.² The nucleophilic substitution of thiol groups can also result in a more stable thioether linkage through reacting a free thiol group of either PNA or peptide part in a Michael type reaction with the double bond of a maleimide functionality of the other part⁹ (Scheme 1.2). For example, Nielsen *et al.* conjugated maleimide-containing PNAs with N-terminal cysteine oligopeptides to produce a variety of bacterial antisense probes.¹²



Scheme 1.2: Thioether bond formation through reacting a free thiol group of either PNA or peptide part in a Michael type reaction with the double bond of a maleimide functionality of the other part ².

Unlike the poor control over the modification of the side chains of cysteine or lysine in classical peptide bioconjugation methods (for example carbodiimide or maleimide crosslinking strategies) which may cause loss of the biological activity of peptide or PNA, novel methods provide controlled site-specific modification and minimize the chance of inactivation. A more chemoselective alternative strategy to produce PNA-peptide conjugates is via oxime formation.¹³ In this method, an amino-oxy group of a peptide was reacted with a ketone functionalized PNA (Scheme 1.3). The condensation was done in an aqueous buffer at pH 4.2 with high efficiency.⁹



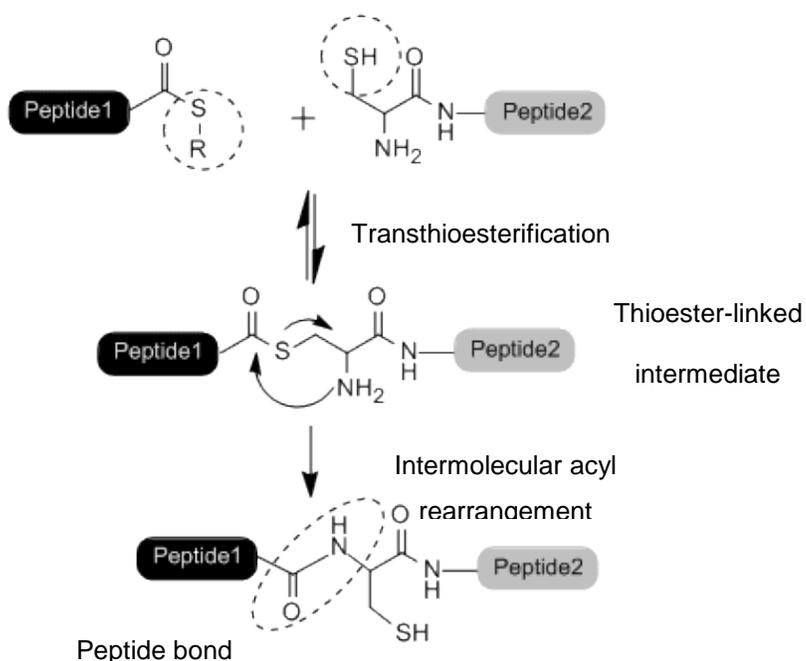
Scheme 1.3: Conjugation of PNA-peptide by oxime formation through reaction of an amino-oxy group of a peptide with a ketone functionalized PNA.²

Assembly of a PNA-peptide conjugate via Kent's native chemical ligation (NCL) strategy¹⁴ is an alternative method which was reported for conjugation of a nuclear localization signal (NLS) peptide-thioester with an N-Cys (tBoc)-PNA fragment in a denaturing (guanidine hydrochloride) aqueous buffer at pH 7-8 with thiophenol and tris (2-carboxyethyl) phosphine (TCEP) as a catalyst and reducing agent, respectively.^{4,9}

1.2 Native chemical ligation

NCL is used extensively for chemoselective ligation of unprotected peptides.¹⁵ In conventional ligation methods, enthalpic activation by coupling reagents needs to use a protection scheme for other competing functional groups. These protections often cause poor solubility of protected fragments and low ligation efficiency. High and specific reactivity between thiol and thioester groups which is often not seen in amino acids and the stability of the native peptide bond are some advantages of NCL. The NCL reaction can be carried out in mild conditions (buffered aqueous solution and neutral pH) ideal for protein chemistry since those conditions have minimal effect on the native structure and function of proteins.¹⁶ Strongly basic conditions decrease the stability of thioester groups and have an impact on some residues such as Lys making them susceptible to react with thioesters.¹⁷ On the other hand, the reactivity of the thiol group and the N-terminal amine is reduced in acidic conditions.¹⁶

NCL is based on the transthioesterification reaction of peptide thioesters with thiol groups of peptides with N-terminal cysteine. The reaction proceeds through a series of reversible thiol-thioester exchanges, initially with an exogenous alkyl or aryl thiol component and then with the thiol groups of cysteine residues. Transthioesterification with the side chain thiol of N-terminal cysteine produces a thioester-linked intermediate¹⁸ which spontaneously undergoes a rapid intramolecular S-N transfer through a favourable intramolecular nucleophilic attack by the α amine group of cysteine. This forms a native peptide bond between the two peptide fragments (Scheme 1.4).



Scheme 1.4: Native chemical ligation consisting of two steps of transthioesterification between a thioester group and a SH group followed by intermolecular acyl rearrangement to form a native peptide bond.

All 20 amino acid thioesters undergo the NCL reaction, but their side chains have a perceptible impact on the reaction rate. Glycine thioesters react quickly while beta-branched side chains and proline react slowly. More reactive thiols assist sterically hindered thioester ligation.¹⁹⁻²¹

Another important factor is the nature of the thioester which can significantly affect the NCL reaction. Aryl thioesters more effectively increase the rate of the thiol-thioester exchange step than alkyl thioesters.

Peptide thioesters are usually synthesized as alkyl derivatives (using 2-mercaptoethansulfonic acid (MESNA))²² and then converted into aryl thioesters by addition of an excess of an aryl thiol such as the odourless and water soluble 4-mercaptophenylacetic acid (MPAA). NCL usually proceeds at low peptide concentration without side reactions and in high yield. NCL reactions can be done in the presence of denaturing agents or detergents such as SDS (sodium dodecyl sulphate) especially for insoluble peptides under standard conditions.²³

It has been shown that high concentrations of MPAA (up to 400 mM) increased the rate of thiol-thioester exchange and do not interfere with the transthioesterification or subsequent rearrangement of the peptide-peptide ligation.¹⁸

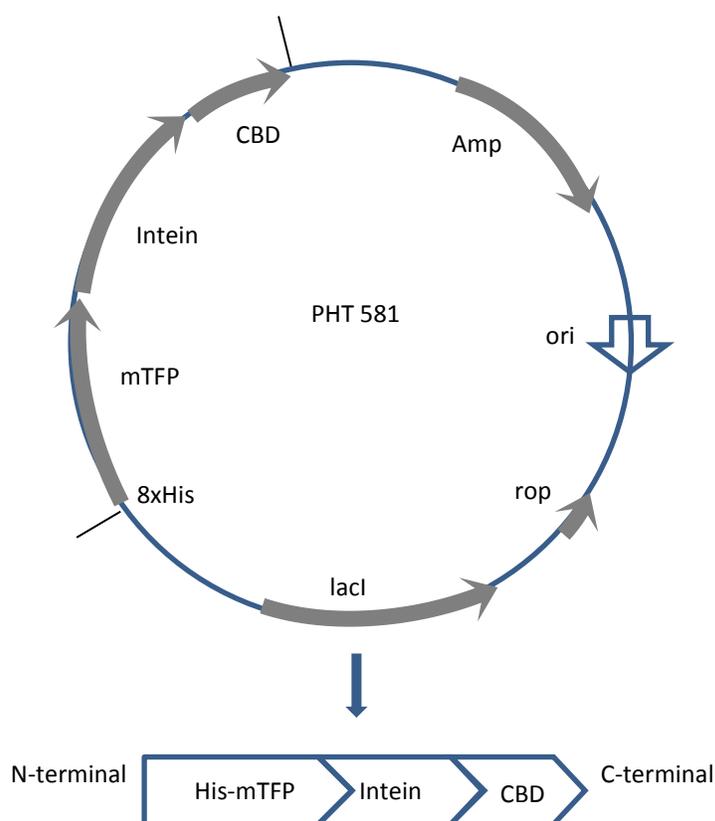
1.3 Expressed protein ligation

NCL has become a popular and successful method for peptide conjugation due to the accessibility and high stability of the starting materials. Cysteine residues and thioester groups can be introduced into peptides through standard solid phase peptide synthesis and the cysteine at the ligation site does not need to be further processed.^{16, 24}

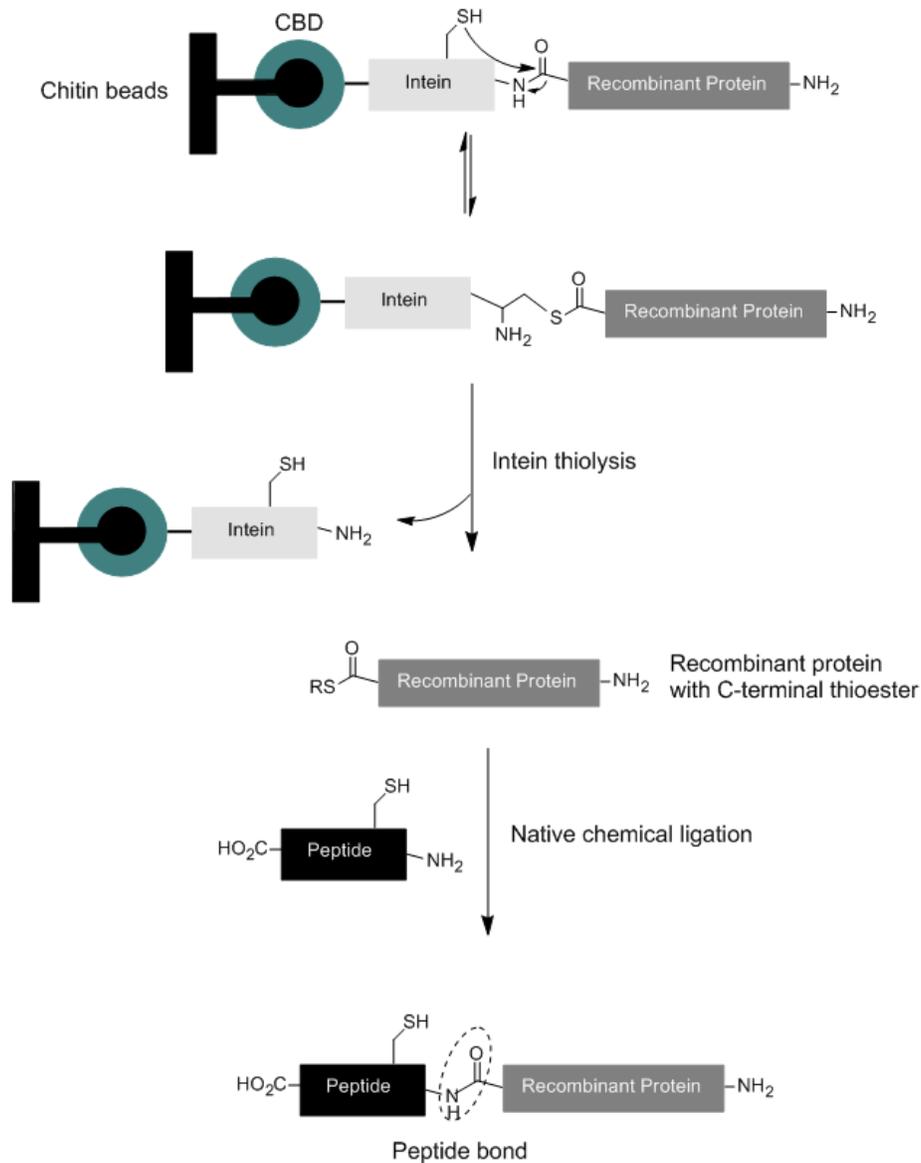
Using NCL for polypeptide ligation and semi-synthesis of proteins is a major advance in protein chemistry and exhibits a powerful unity of biology and chemistry.²⁵ Expressed protein ligation (EPL) and less frequently, intein-mediated protein ligation are semi-synthetic versions of NCL. Since it was reported in 1998, NCL has achieved wide use in protein structure and function analysis, especially in protein modification and labelling.^{26, 27} In this method, a recombinant protein with thioester or cysteine at the C- or N-terminus can be prepared biosynthetically via thiolysis of a corresponding protein-intein fusion (Scheme 1.5).²⁸ Inteins are protein splicing elements that can be introduced using molecular biology methods. They are able to excise themselves from precursor proteins with accompanying fusion of the neighboring regions of the protein. Inteins approximate the self-splicing behavior of RNA introns and can play a powerful role in protein manipulation.²⁹

The modified intein can be cleared by treatment with thiols, generating a thiol or thioester group at the end of the protein. The next steps toward conjugation of this modified protein to a peptide fragment are similar to those of the NCL method.²⁷ EPL has been used previously for the ligation and synthesis of a variety of peptide-protein^{16, 30-32} and PNA- protein conjugates.³³

Purification of a protein fragment possessing a thioester-linked moiety consists of thiol dependent cleavage of an engineered protein containing a modified intein.^{29, 34-36} More than 100 inteins have been reported in unicellular eukaryotic organisms, eubacteria and archeobacteria.^{25, 37} Inteins exhibit some conserved portions in their structure. For example a cysteine or serine residue is very common at the N- terminus of inteins and is responsible for an acyl shift at that splice junction. The nucleophilic sulfohyderyl or hydroxyl group of the cysteine or serine side chain attacks the linkage at the N terminal splice junction leading to a branched intermediate.³⁷ The scissile peptide bond of the intermediate is replaced by a thioester linkage through induced cleavage by a nucleophilic substitution with a thiol reagent such as MESNA which is normally used in standard EPL reaction. Nucleophilic attack on the thioester linkage creates the base point for the production of a C-terminal thioester moiety on a protein of interest (Scheme 1.6).^{22, 34, 38}



Scheme 1.5: Graphical representation of plasmid PHT 581 for expression of mTFP-Intein-CBD constructs.

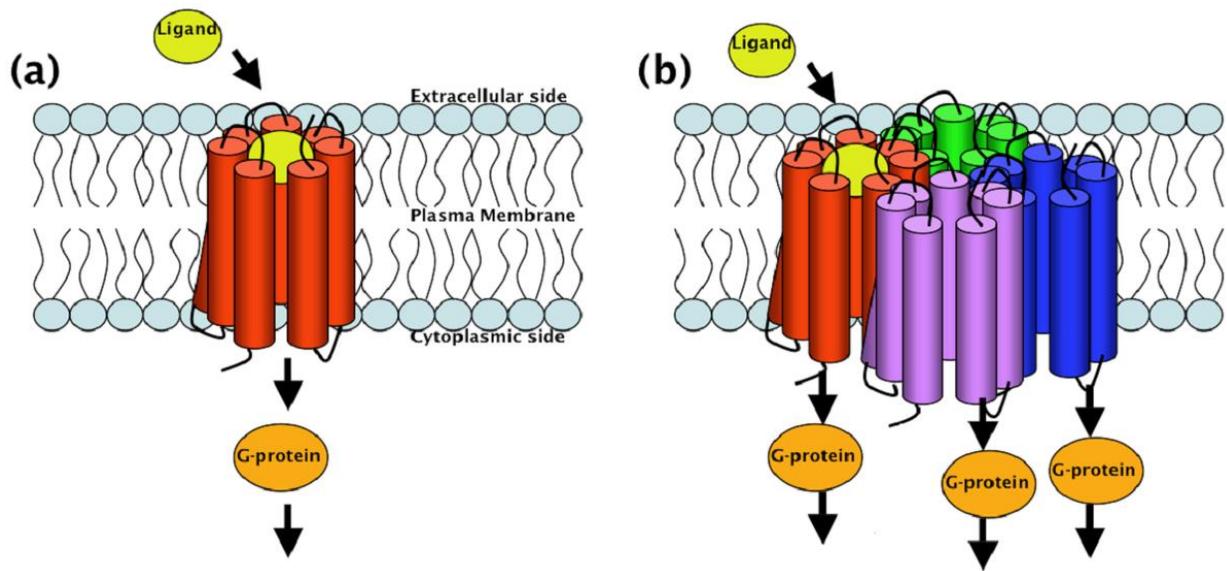


Scheme 1.6: Schematic depiction of expressed protein ligation. Nucleophilic substitution with a thiol reagent such as MESNA induces N-terminal cleavage of intein which generates a thioester group at the C-terminus of target protein. The next steps are similar to NCL in which ligation occurs via chemoselective reaction and a subsequent S-N acyl rearrangement to form a peptide bond between reacting species.

1.4 The importance of protein clustering

Protein oligomerization plays a prominent role for regulating the function of many proteins including enzymes, receptors, ion channels, and transcription factors. However, the unwanted self-assembly of proteins into toxic misfolded aggregates interrupts the formation of their regular functional states and leads to some serious diseases such as Alzheimer's, Huntingdon's, and Parkinson's diseases. Therefore, understanding, predicting and engineering of protein oligomers attracts considerable interest.³⁹⁻⁴³

In signal transduction, transmission of the signal is conducted via protein interactions in the signal transduction chain. Activation of a signaling chain is usually thought to be mediated through dimerization or oligomerization of the same proteins.⁴⁴ Cell membrane proteins which regulate the transfer of small molecules into and out of the cell, are estimated to compose 20-30% of all proteins in the sequenced genome.⁴⁵ The tendency of many of these proteins, such as the G-protein coupled receptor (GPCR) family and the epidermal growth factor receptor (EGFR), to assemble in dimer or higher order complexes has been widely reported. For example, it was traditionally assumed that the monomeric form of GPCR participated in ligand binding and signal transduction. A single ligand induced conformational changes in receptor which led to the activation of receptor followed by the subsequent activation of G protein or effector. This assumption has been challenged recently by the discovery of homo- and/or hetero oligomers of GPCRs suggesting that ligand(s) binding to a single or more receptors might activate neighboring receptors in the oligomeric forms which may significantly affect signaling process (Scheme 1.7).^{46, 47}



Scheme 1.7: Traditional and current views of GPCR signaling. a) In traditional view, it was assumed that one ligand activated one monomer receptor which participated in signal transduction. b) In current view, GPCRs might get together in homomeric and/or heteromeric oligomer forms in which ligand(s) binding may activate neighbouring receptors.⁴⁷

Nanoclusters in the cell membrane are assembled from lipid-lipid, lipid-protein and protein-protein interaction.⁴⁴ Generation of these self-assembled structures is extremely important for regulating signal transduction through the cell membrane and maintaining homeostasis. These features indicate the importance of controlling assembly of these proteins to reduce the risk of converting cells to a cancerous state and many aspects of this conversion remain unclear.^{44, 48-64}

Designing a model system with control over the process of forming nanoclusters could help answer some of the critical questions about these organized biological phenomena.

1.5 Clustering based on PNA

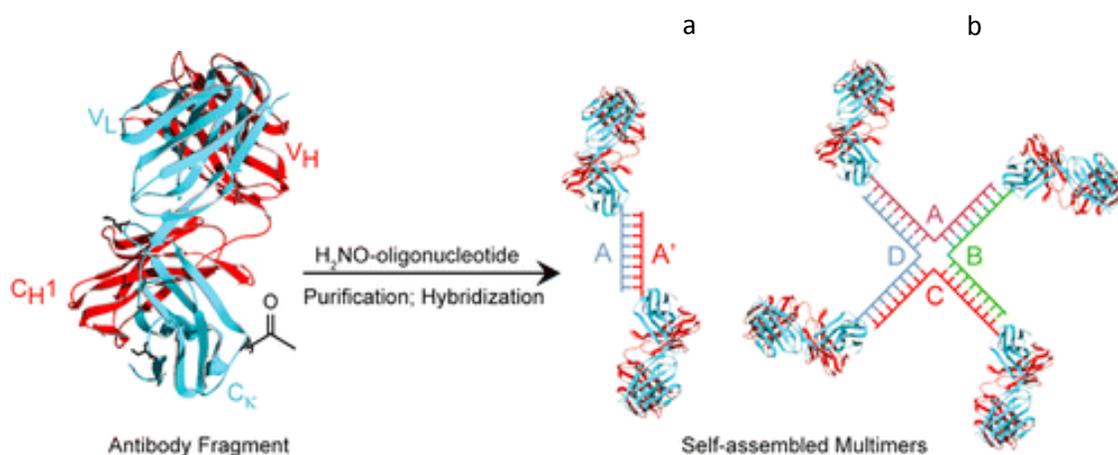
Studying inducible aggregation such as that occurring in cell membrane proteins requires a model system that can be controllably assembled. Biological macromolecules such as DNA and proteins are currently being studied as building blocks of self-assembled nanoarchitectures due

to their size and unique recognition capabilities.⁶⁵ Furthermore self-assembled DNA-protein conjugate systems could be used as a superb controllable template for spatially arraying other molecules with increased relative accuracy and programmability.^{1, 65-67} The programmable hybridization of nucleic acid provides a framework to design nanoscopic assemblies.^{6, 68-76}

PNA provides superior control properties over the dynamics of assembled system including better stability than DNA duplexes, even for short sequences, and higher mismatch sensitivity. Moreover, its neutral net charge allows for tuning the structural and electrostatic characteristics through using other amino acids instead of glycine. Modification with other amino acids is easily accessible through synthesis or conjugation methods and should allow good control over the dynamics of assembly.^{6, 7} The unique properties of PNA gave good results for programmable assembly of nanoparticles^{1, 77} which potentially can be extended to clustering of other molecules such as proteins.

PNA tagged encoding technology has been used to assemble libraries of small molecules,^{78, 79} carbohydrate,⁸⁰ peptide,⁸¹⁻⁸³ and protein fragments^{33, 84} into organized microarrays through hybridization to DNA. Due to the compatibility with standard peptide chemistry, PNA is the only oligonucleotide tag which can be co-synthesized with small molecules by solid phase synthesis. It allows PNA-encoded libraries sensitized by split- and -mix method to be decoded in one step.^{85, 86}

Recently, the self-assembly of PNA- α HER2 antibody Fab fragment conjugates into homodimer, heterodimer, and higher order multimers of defined composition, valency and controlled geometry has been reported. The tetrameric assembly showed superb activity in comparison with parent monoclonal antibody. Site specific modification of antibody using genetically encoded unnatural amino acids allowed precise control of PNA-antibody conjugation (Scheme 1.8).^{70, 84, 87}



Scheme 1.8: α Her2 PNA dimer (a) and PNA tetramer (b). In dimer, the complementary oligonucleotides A and A' are coupled to Fab fragments and allow formation of dimer. In tetramer, four complementary oligonucleotides A, B, C, D are coupled to Fab fragments leading to the formation of a cruciform⁸⁴.

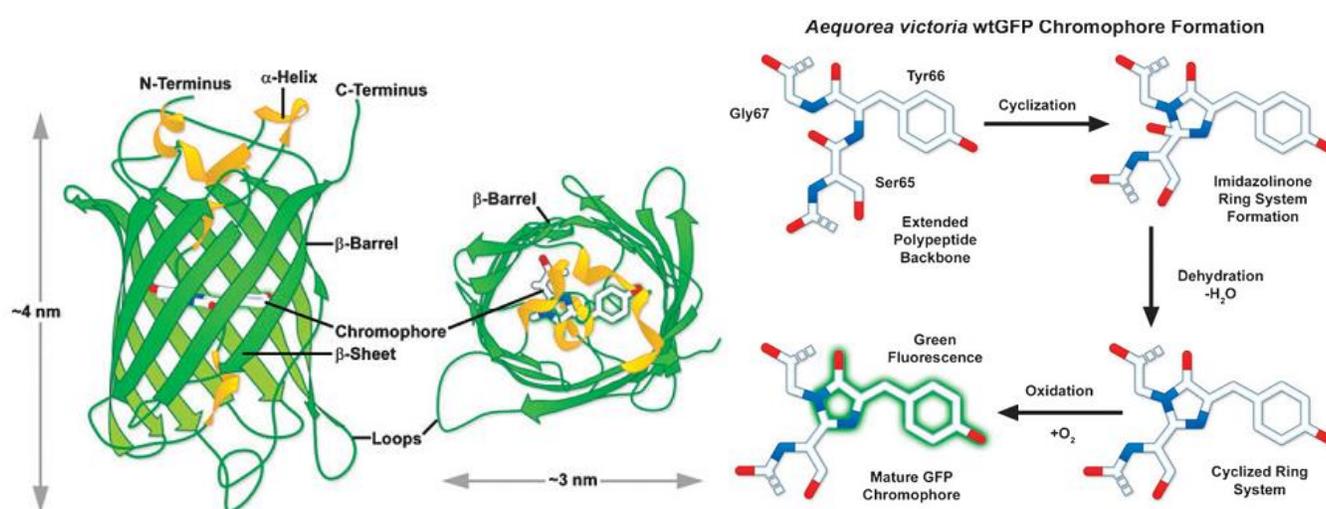
The oligomeric self-assembly of PNA- tagged carbohydrates with controlled topology has also been reported.^{80, 88} Scheibe and Seitz used the hybridization of a PNA-sugar conjugate with complementary DNA as a powerful tool to create well defined spatial arrangement of carbohydrates which can be applied for precise spatial screening of carbohydrate-lectin interactions.^{8, 89-91}

1.5 Protein tags

The site-specific labeling of protein with synthetic molecules such as PNA can provide an intriguing and versatile tool to study the function and structure of proteins and their behavior in clustering forms. The non-invasive imaging of the dynamics of proteins in living systems can be obtained by fusion of proteins of interest with protein or peptide tags as a means for subsequent attachment of a fluorophore or other biophysical probes.^{92, 93}

The introduction of fluorescent proteins in 1994 as selective, genetic tags revolutionized live cell imaging.⁹⁴ Green fluorescent protein (GFP) and its variants showed great fluorescent properties and compatibility with living systems. GFP is a stable, 27kDa protein with distinctive

spectral characteristics. GFP is a β -barrel composed of eleven antiparallel β -strands forming a compact cylinder. There is an α -helix inside the barrel which contains a chromophore in the middle. The cylinder has a diameter and length of about 30 Å and 40 Å, respectively.⁹⁵⁻⁹⁷ The chromophore forms by rearrangement and oxidation of three amino acid residues of Ser, Tyr and Gly after spontaneous folding of the protein (Scheme 1.9).⁹⁴ Variants of GFP have been used frequently as FRET pairs. For example, they can be easily fused to proteins of interest through cloning methods. Due to the flexibility of the N- and C- termini on the surface of the β -barrel, the structure is maintained in fused proteins.⁹⁸⁻¹⁰¹



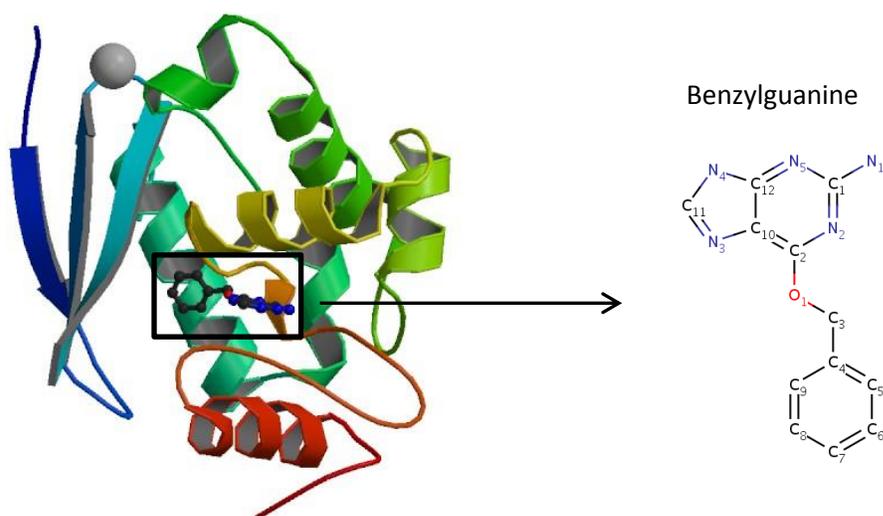
Scheme 1.9: Schematic structure and dimension of *A. Victoria* GFP (left) and formation of its chromophore after rearrangement and oxidation of three amino acids: Ser, Tyr, and Gly after spontaneous folding of protein (right).¹⁰²

Currently, the development of a wide variety of improved monomeric fluorescent proteins with a broad range of spectral properties has provided a potent tool for multiparameter imaging of cellular structure and processes.^{95-98, 101, 103-117}

Self-labeling protein tags introduced another alternative modification of proteins. The advantage of these tags is their high specificity and selectivity.⁹³ Furthermore, a wide range of

colours with unique photophysical properties can be obtained by subsequent modification of self-labeling tags with organic fluorophores which make them a great method to localize and study the fusion protein behaviors in living systems.¹¹⁸⁻¹²⁰

The SNAP tag is a 19-20 kDa self-labeling tag which was developed by mutation of the DNA repair protein O₆-alkylguanine-DNA alkyltransferase (AGT). The labeling of a SNAP-tag is based on the specific reaction of benzylguanine (BG) derivatives with a reactive cysteine residue of AGT leading to an irreversible covalent linkage (Scheme 1.10).¹¹⁹⁻¹²⁷ The BG derivatives can be applied in different conjugation procedures allowing specific labeling to a wide variety of synthetic probes including PNA. It should be noticed that the nature of the ligand attached to BG does not have an impact on the rate of the SNAP-tag reaction with a BG derivative.^{120, 128-130}



Scheme 1.10: Schematic structure of SNAP-tag bound to its substrate benzylguanine. The labeling of a SNAP-tag is based on the specific reaction of benzylguanine (BG) derivatives with a reactive cysteine residue of AGT leading to an irreversible covalent linkage (182 amino acid, (PDB entry 3KZZ: DOI:10.2210/pdb3kzy/pdb).

SNAP-tag labeling has been used for many *in vitro* and *in vivo* experiments *such as* localization and trafficking of fusion proteins in cell membranes, labeling of antibody fragments, designing

fluorescent biosensors, controlling yeast transcription and visualization of metabolite signaling.¹³¹⁻¹³⁸

Both FP and SNAP tags can be easily fused to the proteins of interest to study their localization, clustering and dynamics in living systems. The combination of the unique recognition characteristics of PNA with the wide application of these tags can provide versatile, well defined units to create programmed self-assembled protein model systems. The induced assembly by DNA templates can be studied through fluorescent techniques and especially using fluorescence resonance energy transfer (FRET) methods.

Protein clustering is normally studied via co-immunoprecipitation or chemical cross-linking which are inclined to artifacts since the conditions of experiments may induce protein clustering.¹³⁹ Lately, microscopy methods based on FRET have been developed to investigate clustering of proteins.^{57, 140}

The nonradiative nature of FRET makes it a powerful tool to follow protein dynamics and protein-protein interaction in physiological conditions. Moreover, it can be applied to study cell membrane protein aggregation *in vivo*. The development of analytical methods and instrumentation is on-going particularly in the case of homo-FRET. The extent of FRET (and its reversibility) can be assessed as a function of time, thus revealing the dynamics of donor and acceptor proximity.^{114, 139, 141, 142}

1.6 Aims and outline of thesis

PNA with its unique properties is well suited to use as a recognition tag for programmable assembly of proteins. Conjugation of PNA to protein tags (FP and SNAP tags) provides an opportunity to combine the high recognition specificity and stability of PNA with the versatile applications of protein tags to create protein self-assembly model systems. This model system allows alignment of PNA-protein tag conjugates in a controllable manner which can be used to

mimic other protein self-assembled systems in biological processes and study their behaviors. The compatibility and nonradiative characteristics of FRET in living systems makes it a powerful tool to understand the signatures of protein assembly via the designed PNA-protein model system.

The aim of this thesis is to develop such model systems based on DNA directed controllable self-assembly of PNA-protein conjugates to study protein clustering behavior. Fluorescent protein and SNAP tags are used to represent programmable self-assembly of proteins. Assembly directed by PNA-PNA hybridization was originally proposed but discarded for this part of the project mostly for cost reason. However, they may still be needed in case of using model systems in membrane.

In **Chapter 2**, expressed protein ligation as a semisynthetic version of native chemical ligation is used to conjugate PNA to a fluorescent protein tag. The recombinant fluorescent protein with a thioester group at the C-terminus is expressed in *E. coli* and subsequently purified as a precursor form to conjugate an N-terminal thiol group of a PNA. The final FP-PNA conjugate provides a building block to create a model system of protein assembly.

In **Chapter 3**, DNA scaffolds direct programmable assembly of FP-PNA conjugates. One DNA template is a fluorescent beacon with 6-FAM and Dabcyl at its ends. This beacon directs assembly of the FP-PNA conjugate to create an assembled hetero-FRET system. Using two other DNA scaffolds allows assembly of multiple FP-PNA constructs in dimer and tetramer forms. Assembly of multiple similar FP-PNAs in a row induced homo-FRET. The visualization of self-assembled FRET systems can be obtained through fluorescence techniques such as intensity, frequency domain lifetime and anisotropy measurements. This model system provides control over self-assembly of FP-PNA through precise recognition of PNA on the DNA framework. The well-defined characteristics of the model system can be extended to any other proteins through fusion to FPs.

Chapter 4 reports the use of the self-labeling SNAP-tag to create PNA-protein conjugate units for developing self-assembly. The model system based on a SNAP-PNA conjugate can provide another permissive approach to study behaviors of other proteins of interest which can be fused to a SNAP tag. This chapter describes creating PNA-SNAP tag conjugates which require modification of PNA with a specific substrate for SNAP (O⁶-benzyl guanine (BG)). The subsequent conjugation of BG-PNA with fluorophores allows the study of the model system with fluorescent techniques.

The controllable self-assembly of SNAP-PNA on DNA scaffolds is discussed in **Chapter 5**. Three DNA scaffolds are used as frameworks to create dimer and oligomer forms. Using different fluorophores creates a hetero dimer that exhibits hetero-FRET system. Additional assembled models feature homo-dimer and homo-oligomer forms by assembly of similar fluorescent SNAP-PNAs in a row which can be studied by fluorescent techniques.

1.7 References

1. G. Prencipe, S. Maiorana, P. Verderio, M. Colombo, P. Fermo, E. Caneva, D. Prospero and E. Licandro, *Chemical communications (Cambridge, England)*, 2009, **40**, 6017-6019.
2. M. C. de Koning, G. A. van der Marel and M. Overhand, *Current opinion in chemical biology*, 2003, **7**, 734-740.
3. M. C. de Koning, D. V. Filippov, N. Meeuwenoord, M. Overhand, G. A. van der Marel and J. H. van Boom, *Tetrahedron letters*, 2002, **43**, 8173-8176.
4. M. C. de Koning, D. V. Filippov, N. Meeuwenoord, M. Overhand, d. M. van and J. H. van Boom, *Synlett*, 2001, **10**, 1516,1518.
5. V. Menchise, G. De Simone, T. Tedeschi, R. Corradini, S. Sforza, R. Marchelli, D. Capasso, M. Saviano and C. Pedone, *Proceedings of the National Academy of Sciences of the United States of America*, 2003, **100**, 12021-12026.
6. Z. L. Pianowski and N. Winssinger, *Chemical Society Reviews*, 2008, **37**, 1330-1336.
7. E. Rozners, *Journal of Nucleic Acids*, 2012, **2012**, 8.
8. N. Spinelli, E. Defrancchi and F. Morvan, *Chemical Society Reviews*, 2013, **42**, 4557-4573.
9. P. E. Nielsen, *Peptide Nucleic Acids: Methods and Protocols (Methods in Molecular Biology)*, Humana Press Inc., USA, 2002.
10. J. J. Turner, D. Williams, D. Owen and M. J. Gait, in *Current Protocols in Nucleic Acid Chemistry*, John Wiley & Sons, Inc., Editon edn., 2001.
11. P. J. Roth, C. Boyer, A. B. Lowe and T. P. Davis, *Macromolecular Rapid Communications*, 2011, **32**, 1123-1143.
12. L. Good, S. K. Awasthi, R. Dryselius, O. Larsson and P. E. Nielsen, *Nature biotechnology*, 2001, **19**, 360-364.
13. P. Neuner, P. Gallo, L. Orsatti, L. Fontana and P. Monaci, *Bioconjugate chemistry*, 2003, **14**, 276-281.
14. P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776-779.
15. J. P. Tam, Q. Yu and Z. Miao, *Biopolymers*, 1999, **51**, 311-332.
16. C. P. Hackenberger and D. Schwarzer, *Angewandte Chemie (International ed.in English)*, 2008, **47**, 10030-10074.
17. J. Dheur, N. Ollivier and O. Melnyk, *Organic letters*, 2011, **13**, 1560-1563.

18. E. C. Johnson and S. B. Kent, *Journal of the American Chemical Society*, 2006, **128**, 6640-6646.
19. C. Wang, Q.-X. Guo and Y. Fu, *Chemistry -An Asian Journal*, 2011, **6**, 1241-1251.
20. H. Rohde and O. Seitz, *Biopolymers*, 2010, **94**, 551-559.
21. S. B. Pollock and S. B. Kent, *Chemical communications (Cambridge, England)*, 2011, **47**, 2342-2344.
22. M. M. Bastings, I. van Baal, E. W. Meijer and M. Merkx, *BMC Biotechnology*, 2008, **8**, 76.
23. Z. Tan, S. Shang and S. J. Danishefsky, *Angewandte Chemie International Edition*, 2010, **49**, 9500-9503.
24. L. De Rosa, A. Russomanno, A. Romanelli and L. Andrea, *Molecules*, 2013, **18**, 440-465.
25. T. W. Muir, D. Sondhi and P. A. Cole, *Proceedings of the National Academy of Sciences of the United States of America*, 1998, **95**, 6705-6710.
26. B. L. Nilsson, M. B. Soellner and R. T. Raines, *Annual Review of Biophysics and Biomolecular Structure*, 2005, **34**, 91-118.
27. R. R. Flavell and T. W. Muir, *Accounts of Chemical Research*, 2009, **42**, 107-116.
28. R. M. Hofmann and T. W. Muir, *Current Opinion in Biotechnology*, 2002, **13**, 297-303.
29. T. C. Evans, Jr., J. Benner and M. Q. Xu, *The Journal of biological chemistry*, 1999, **274**, 18359-18363.
30. R. Mhida, A. Vallin, N. Ollivier, A. Blanpain, G. Shi, R. Christiano, L. Johannes and O. Melnyk, *Bioconjugate Chemistry*, 2010, **21**, 219-228.
31. X. Duburcq, C. Olivier, F. Malingue, R. Desmet, A. Bouzidi, F. Zhou, C. Auriault, H. Gras-Masse and O. Melnyk, *Bioconjugate Chemistry*, 2004, **15**, 307-316.
32. M. A. Gauthier and H. A. Klok, *Chemical Communications* 2008, **23**, 2591-2611.
33. M. Lovrinovic, R. Seidel, R. Wacker, H. Schroeder, O. Seitz, M. Engelhard, R. S. Goody and C. M. Niemeyer, *Chemical communications (Cambridge, England)*, 2003, **7**, 822-823.
34. M. G. Paulick, M. B. Forstner, J. T. Groves and C. R. Bertozzi, *Proceedings of the National Academy of Sciences of the United States of America*, 2007, **104**, 20332-20337.
35. M. Q. Xu and T. C. Evans, Jr., *Methods (San Diego, Calif.)*, 2001, **24**, 257-277.
36. J. Kalia and R. T. Raines, *ChemBiochem*, 2006, **7**, 1375-1383.
37. R. David, M. P. Richter and A. G. Beck-Sickinger, *European Journal of Biochemistry / FEBS*, 2004, **271**, 663-677.
38. I. van Baal, H. Malda, S. A. Synowsky, J. L. van Dongen, T. M. Hackeng, M. Merkx and E. W. Meijer, *Angewandte Chemie (International ed.in English)*, 2005, **44**, 5052-5057.
39. J. E. Dayhoff, B. A. Shoemaker, S. H. Bryant and A. R. Panchenko, *Journal of Molecular Biology*, 2010, **395**, 860-870.
40. T. Perica, C. Chothia and S. A. Teichmann, *Proceedings of the National Academy of Sciences*, 2012, **109**, 8127-8132.
41. M. A. Ross Ca Fau - Poirier and M. A. Poirier, *Nature Medicine*, 2004, **10**, S10-S17.
42. F. Bemporad and F. Chiti, *Chemistry & Biology*, 2012, **19**, 315-327.
43. M. H. Ali and B. Imperiali, *Bioorganic & Medicinal Chemistry*, 2005, **13**, 5013-5020.
44. A. S. Harding and J. F. Hancock, *Trends in Cell Biology*, 2008, **18**, 364-371.
45. A. Rath and C. M. Deber, *FEBS Letters*, 2007, **581**, 1335-1341.
46. S. R. George, B. F. O'Dowd and S. P. Lee, *Nat Rev Drug Discov*, 2002, **1**, 808-820.
47. L. Skrabanek, M. Murcia, M. Bouvier, L. Devi, S. R. George, M. J. Lohse, G. Milligan, R. Neubig, K. Palczewski, M. Parmentier, J. P. Pin, G. Vriend, J. A. Javitch, F. Campagne and M. Filizola, *BMC Bioinformatics*, 2007, **8**, 177.
48. L. Albizu, M. Cottet, M. Kralikova, S. Stoev, R. Seyer, I. Brabet, T. Roux, H. Bazin, E. Bourrier, L. Lamarque, C. Breton, M.-L. Rives, A. Newman, J. Javitch, E. Trinquet, M. Manning, J.-P. Pin, B. Mouillac and T. Durroux, *Nature Chemical Biology*, 2010, **6**, 587-594.
49. F. Ciruela, J. P. Vilardaga and V. Fernandez-Duenas, *Trends in Biotechnology*, 2010, **28**, 407-415.
50. O. Cottet M Fau - Faklaris, D. Faklaris O Fau - Maurel, P. Maurel D Fau - Scholler, E. Scholler P Fau - Doumazane, E. Doumazane E Fau - Trinquet, J.-P. Trinquet E Fau - Pin, T. Pin Jp Fau - Durroux and T. Durroux, *Frontiers in Endocrinology*, 2012, **3**, 1664-2392
51. E. Doumazane, P. Scholler, J. M. Zwier, E. Trinquet, P. Rondard and J.-P. Pin, *The FASEB Journal*, 2011, **25**, 66-77.
52. S. Ganguly, A. H. A. Clayton and A. Chattopadhyay, *Biophysical Journal*, 2011, **100**, 361-368.
53. R. M. Hanyaloglu Ac Fau - Seeber, T. A. Seeber Rm Fau - Kohout, R. J. Kohout Ta Fau - Lefkowitz, K. A. Lefkowitz Rj Fau - Eidne and K. A. Eidne, *J Biol Chem*, 2002, **277**, 50422-50430.

54. G. Milligan, *Molecular Pharmacology*, 2013, **84**, 158-169.
55. L. Pin Jp Fau - Comps-Agrar, D. Comps-Agrar L Fau - Maurel, C. Maurel D Fau - Monnier, M. L. Monnier C Fau - Rives, E. Rives MI Fau - Trinquet, J. Trinquet E Fau - Kniazeff, P. Kniazeff J Fau - Rondard, L. Rondard P Fau - Prezeau and L. Prezeau, *The Journal of Physiology*, 2009, **587**, 5337-5344.
56. M.-L. Prezeau L Fau - Rives, L. Rives MI Fau - Comps-Agrar, D. Comps-Agrar L Fau - Maurel, J. Maurel D Fau - Kniazeff, J.-P. Kniazeff J Fau - Pin and J. P. Pin, *Current Opinion in Pharmacology*, 2010, **10**, 6-13
57. A. H. Clayton, F. Walker, S. G. Orchard, C. Henderson, D. Fuchs, J. Rothacker, E. C. Nice and A. W. Burgess, *The Journal of Biological Chemistry*, 2005, **280**, 30392-30399.
58. T. W. Gadella, Jr. and T. M. Jovin, *The Journal of Cell Biology*, 1995, **129**, 1543-1558.
59. J. Ichinose, M. Murata, T. Yanagida and Y. Sako, *Biochemical and Biophysical Research Communications*, 2004, **324**, 1143-1149.
60. M. Martin-Fernandez, D. T. Clarke, M. J. Tobin, S. V. Jones and G. R. Jones, *Biophysical Journal*, 2002, **82**, 2415-2427.
61. G. Brockhoff, P. Heiss, J. Schlegel, F. Hofstaedter and R. Knuechel, *Cytometry*, 2001, **44**, 338-348.
62. P. Monaghan, M. G. Ormerod and M. J. O'Hare, *International journal of cancer. Journal international du cancer*, 1990, **46**, 935-943.
63. N. Normanno, A. De Luca, C. Bianco, L. Strizzi, M. Mancino, M. R. Maiello, A. Carotenuto, G. De Feo, F. Caponigro and D. S. Salomon, *Gene*, 2006, **366**, 2-16.
64. Y. Qin, E. M. E. Verdegaal, M. Siderius, J. P. Bebelman, M. J. Smit, R. Leurs, R. Willemze, C. P. Tensen and S. Osanto, *Pigment Cell & Melanoma Research*, 2010, **24**, 207-218.
65. C. M. Niemeyer, *Angewandte Chemie (International ed. in English)*, 2010, **49**, 1200-1216.
66. C. M. Niemeyer, *Biochemical Society*, 2004, **32**, 51-53.
67. H. Lu, O. Schops, U. Woggon and C. M. Niemeyer, *Journal of the American Chemical Society*, 2008, **130**, 4815-4827.
68. P. W. K. Rothmund, *Nature*, 2006, **440**, 297-302.
69. M. Röthlingshöfer, K. Gorska and N. Winssinger, *Journal of the American Chemical Society*, 2011, **133**, 18110-18113.
70. K. K. Sadhu, M. Röthlingshöfer and N. Winssinger, *Israel Journal of Chemistry*, 2013, **53**, 75-86.
71. N. C. Seeman, *Nature*, 2003, **421**, 427-431.
72. U. Feldkamp and C. M. Niemeyer, *Angewandte Chemie International Edition*, 2006, **45**, 1856-1876.
73. Z. Dadon, M. Samiappan, E. Y. Safranchik and G. Ashkenasy, *Chemistry – A European Journal*, 2010, **16**, 12096-12099.
74. P. K. Lo, K. L. Mettera and H. F. Sleiman, *Current Opinion in Chemical Biology*, 2010, **14**, 597-607.
75. C. M. Niemeyer, *Angewandte Chemie International Edition*, 2001, **40**, 4128-4158.
76. A. Roloff and O. Seitz, *Chemical Science*, 2013, **4**, 432-436.
77. R. Chakrabarti and A. M. Klibanov, *Journal of the American Chemical Society*, 2003, **125**, 12531-12540.
78. J. P. Dagher, M. Ciobanu, S. Alvarez, S. Barluenga and N. Winssinger, *Chemical Science*, 2011, **2**, 625-632.
79. J. L. Harris and N. Winssinger, *Chemistry – A European Journal*, 2005, **11**, 6792-6801.
80. K.-T. Huang, K. Gorska, S. Alvarez, S. Barluenga and N. Winssinger, *ChemBioChem*, 2011, **12**, 56-60.
81. J. J. Diaz-Mochon, L. Bialy and M. Bradley, *Chemical Communications*, 2006, **0**, 3984-3986.
82. N. Svensen, Juan J. Díaz-Mochón and M. Bradley, *Chemistry & Biology*, 2011, **18**, 1284-1289.
83. D. Pouchain, J. J. Díaz-Mochón, L. Bialy and M. Bradley, *ACS Chemical Biology*, 2007, **2**, 810-818.
84. S. A. Kazane, J. Y. Axup, C. H. Kim, M. Ciobanu, E. D. Wold, S. Barluenga, B. A. Hutchins, P. G. Schultz, N. Winssinger and V. V. Smider, *Journal of the American Chemical Society*, 2013, **135**, 340-346.
85. H. D. Urbina, F. Debaene, B. Jost, C. Bole-Feysot, D. E. Mason, P. Kuzmic, J. L. Harris and N. Winssinger, *ChemBioChem*, 2006, **7**, 1790-1797.
86. N. Winssinger, *Artificial DNA: PNA & XNA*, 2012, **3**, 105-108.
87. B. M. Hutchins, S. A. Kazane, K. Stafflin, J. S. Forsyth, B. Felding-Habermann, P. G. Schultz and V. V. Smider, *Journal of Molecular Biology*, 2011, **406**, 595-603.
88. K. Gorska, K.-T. Huang, O. Chaloin and N. Winssinger, *Angewandte Chemie International Edition*, 2009, **48**, 7695-7700.

89. C. Scheibe, A. Bujotzek, J. Dervede, M. Weber and O. Seitz, *Chemical Science*, 2011, **2**, 770-775.
90. C. Scheibe and O. Seitz, in *Pure & Applied Chemistry*, International Union of Pure & Applied Chemistry (IUPAC), Editon edn., 2012, vol. 84, pp. 77-85.
91. C. Scheibe, S. Wedepohl, S. B. Riese, J. Dervede and O. Seitz, *ChemBioChem*, 2013, **14**, 236-250.
92. H. M. O'Hare, K. Johnsson and A. Gautier, *Current Opinion in Structural Biology*, 2007, **17**, 488-494.
93. M. J. Hinner and K. Johnsson, *Chemical Biotechnology & Pharmaceutical Biotechnology*, 2010, **21**, 766-776.
94. C. Jing and V. W. Cornish, *Accounts of Chemical Research*, 2011, **44**, 784-792.
95. J. K. Sanders and S. E. Jackson, *Chemical Society Reviews*, 2009, **38**, 2821-2822.
96. M. Chalfie, *Photochemistry and Photobiology*, 1995, **62**, 651-656.
97. M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher, *Science (New York, N.Y.)*, 1994, **263**, 802-805.
98. R. A. Cinelli, A. Ferrari, V. Pellegrini, M. Tyagi, M. Giacca and F. Beltram, *Photochemistry and Photobiology*, 2000, **71**, 771-776.
99. G. H. Patterson, D. W. Piston and B. G. Barisas, *Analytical Biochemistry*, 2000, **284**, 438-440.
100. V. Subramaniam, Q. S. Hanley, A. H. Clayton and T. M. Jovin, *Methods in Enzymology*, 2003, **360**, 178-201.
101. N. Hayes, E. Howard-Cofield and W. Gullick, *Cancer Letters*, 2004, **206**, 129-135.
102. R. N. Day and M. W. Davidson, *Chemical Society Reviews*, 2009, **38**, 2887-2921.
103. T. Kobayashi, N. Morone, T. Kashiyama, H. Oyamada, N. Kurebayashi and T. Murayama, *PLoS ONE*, 2008, **3**, e3822.
104. M. Zaccolo and T. Pozzan, *IUBMB life*, 2000, **49**, 375-379.
105. R. Wombacher and V. W. Cornish, *Journal of Biophotonics*, 2011, **4**, 391-402.
106. J. Lippincott-Schwartz and G. H. Patterson, *Science*, 2003, **300**, 87-91.
107. D. M. Chudakov, M. V. Matz, S. Lukyanov and K. A. Lukyanov, *Physiological Reviews*, 2010, **90**, 1103-1163.
108. R. Miyake, T. Uchimura, X. Li and T. Imasaka, *Chemical and Pharmaceutical Bulletin*, 2013, **61**, 82-84.
109. C. Stadler, E. Rexhepaj, V. R. Singan, R. F. Murphy, R. Pepperkok, M. Uhlen, J. C. Simpson and E. Lundberg, *Nature Methods*, 2013, **10**, 315-323.
110. H. Abe, Y. Yanagawa, K. Kanbara, K. Maemura, H. Hayasaki, H. Azuma, K. Obata, Y. Katsuoka, M. Yabumoto and M. Watanabe, *Journal of Andrology*, 2005, **26**, 568-577.
111. I. Gautier, M. Tramier, C. Durieux, J. Coppey, R. B. Pansu, J. C. Nicolas, K. Kemnitz and M. Coppey-Moisan, *Biophysical Journal*, 2001, **80**, 3000-3008.
112. C. S. Lisenbee, S. K. Karnik and R. N. Trelease, *Traffic*, 2003, **4**, 491-501.
113. T. Nagai, K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba and A. Miyawaki, *Nature Biotechnology*, 2002, **20**, 87-90.
114. B. A. Pollok and R. Heim, *Trends in Cell Biology*, 1999, **9**, 57-60.
115. R. Y. Tsien, *Annual Review of Biochemistry*, 1998, **67**, 509-544.
116. B. Zhang, *Biochemical and Biophysical Research Communications*, 2004, **323**, 674-678.
117. G. Crivat and J. W. Taraska, *Trends in Biotechnology*, 2012, **30**, 8-16.
118. A. Juillerat, C. Heinis, I. Sielaff, J. Barnikow, H. Jaccard, B. Kunz, A. Terskikh and K. Johnsson, *ChemBioChem*, 2005, **6**, 1263-1269.
119. N. Johnsson and K. Johnsson, *ACS Chemical Biology*, 2007, **2**, 31-38.
120. K. Johnsson, *Nature Chemical Biology*, 2009, **5**, 63-65.
121. A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel and K. Johnsson, *Nat Biotech*, 2003, **21**, 86-89.
122. A. Keppler, H. Pick, C. Arrivoli, H. Vogel and K. Johnsson, *Proceedings of the National Academy of Sciences of the United States of America*, 2004, **101**, 9955-9959.
123. C. Keppler A Fau - Arrivoli, L. Arrivoli C Fau - Sironi, J. Sironi L Fau - Ellenberg and J. Ellenberg, *Biotechniques*, 2006, **41**, 167-170, 172, 174-165.
124. M. Keppler A Fau - Kindermann, S. Kindermann M Fau - Gendreizig, H. Gendreizig S Fau - Pick, H. Pick H Fau - Vogel, K. Vogel H Fau - Johnsson and K. Johnsson, *Methods*, 2004, **32**, 437-444.
125. M. Z. Lin and L. Wang, *Physiology*, 2008, **23**, 131-141.
126. D. Jung, K. Min, J. Jung, W. Jang and Y. Kwon, *Molecular BioSystems*, 2013, **9**, 862-872.
127. T. Gronemeyer, C. Chidley, A. Juillerat, C. Heinis and K. Johnsson, *Protein Engineering Design and Selection*, 2006, **19**, 309-316.
128. T. Komatsu, K. Johnsson, H. Okuno, H. Bito, T. Inoue, T. Nagano and Y. Urano, *Journal of the American Chemical Society*, 2011, **133**, 6745-6751.

- 129.D. Maurel, S. Banala, T. Laroche and K. Johnsson, *ACS Chemical Biology*, 2010, **5**, 507-516.
- 130.C. Campos, M. Kamiya, S. Banala, K. Johnsson and M. González-Gaitán, *Developmental Dynamics*, 2011, **240**, 820-827.
- 131.M. A. Brun, K.-T. Tan, E. Nakata, M. J. Hinner and K. Johnsson, *Journal of the American Chemical Society*, 2009, **131**, 5873-5884.
- 132.C. Donovan and M. Bramkamp, *Microbiology*, 2009, **155**, 1786-1799.
- 133.A. F. Hussain, F. Kampmeier, V. von Felbert, H. F. Merk, M. K. Tur and S. Barth, *Bioconjugate Chemistry*, 2011, **22**, 2487-2495.
- 134.F. Kampmeier, M. Ribbert, T. Nachreiner, S. Dembski, F. Beaufils, A. Brecht and S. Barth, *Bioconjugate Chemistry*, 2009, **20**, 1010-1015.
- 135.M. Bannwarth, I. R. Correia, M. Sztretye, S. Pouvreau, C. Fellay, A. Aebischer, L. Royer, E. Riños and K. Johnsson, *ACS Chemical Biology*, 2009, **4**, 179-190.
- 136.M. Kamiya and K. Johnsson, *Analytical Chemistry*, 2010, **82**, 6472-6479.
- 137.H. Haruki, M. R. Gonzalez and K. Johnsson, *PLoS ONE*, 2012, **7**, e37598.
- 138.J. P. Pin, L. Comps-Agrar, D. Maurel, C. Monnier, M. L. Rives, E. Trinquet, J. Kniazeff, P. Rondard and L. Prézeau, *The Journal of Physiology*, 2009, **587**, 5337-5344.
- 139.A. N. Bader, E. G. Hofman, J. Voortman, P. M. en Henegouwen and H. C. Gerritsen, *Biophysical Journal*, 2009, **97**, 2613-2622.
- 140.J. Szollosi, P. Nagy, Z. Sebestyen, S. Damjanovitcha, J. W. Park and L. Matyus, *Journal of Biotechnology*, 2002, **82**, 251-266.
- 141.E. A. Jares-Erijman and T. M. Jovin, *Analytical Techniques / Mechanisms*, 2006, **10**, 409-416.
- 142.Y. Chen, J. P. Mauldin, R. N. Day and A. Periasamy, *Journal of Microscopy*, 2007, **228**, 139-152.

Chapter 2

Conjugation of fluorescent protein to PNA

The specific binding characteristics of oligonucleotides make them an attractive platform to direct programmable assembly of biomolecules. PNA, as a functional analogue of DNA, provides unique chemistry and robustness to conveniently conjugate to proteins and induce their assembly through precise recognition of complementary DNA scaffolds. The pseudo-peptide backbone of PNA facilitates attachment of PNA to proteins through protein conjugation methods. In this chapter, a site-specific conjugation of PNA to monomeric teal fluorescent protein (mTFP) was achieved using EPL and NCL, producing a native peptide bond at the ligation site. Recombinant mTFP was expressed in *E. coli* and purified thereafter. The purity and integrity of the expressed protein was confirmed by mass spectrometry, SDS-PAGE analysis and UV-Vis spectrophotometry. The conjugation was assessed by mass spectrometry and spectrophotometry, showing almost complete conversion of mTFP to the ligated form. The resulting conjugates can be used as units to create a model system to assemble fluorescent proteins in a controllable manner (Chapter 3).

2.1 Introduction

Alignment of biomolecules in a precisely programmed manner is of a great interest in the study of their function and interaction. One of the attractive molecular building blocks to assemble controllable architectures is nucleic acid. The programmable hybridization of nucleic acid provides a framework to design nanoscopic assemblies.¹⁻¹⁰ PNA, as a DNA mimic, has a unique chemistry which makes it a better building block than DNA for programming self-assembly. The main advantage of PNA is that it achieves more specific and stable duplexes with shorter sequences during the hybridization than DNA while providing less sensitivity to the ionic strength of the solution.^{11, 12}

During the past two decades, the development of fluorescent proteins as genetically encoded markers which can be fused to virtually any protein has revolutionized the investigation and manipulation of proteins involved in complex biochemical processes in living systems. Today, a vast number of improved monomeric FPs are available with a broad range of spectral properties and these provide a valuable tool for multiparameter imaging of cellular structure and processes.¹³⁻³²

Efficient approaches for labelling, assembling and tracking proteins are needed to understand their function and interaction in living systems.³³ Numerous methods have been developed for protein conjugation. Novel methods of bioconjugation which are based on bioorthogonal chemoselective approaches are more efficient in terms of the site-specificity of the conjugation reaction while minimizing perturbations to the structure and the function of the target biomolecules.³⁴⁻³⁸ One of the most successful semisynthetic protein ligation techniques is expressed protein ligation (EPL) in which a recombinant protein containing a C-terminal α -thioester group is ligated to the thiol group of a cysteine residue at the N-terminus of another protein via native chemical ligation (NCL) generating a native peptide bond at the ligation site. EPL has been successfully used on a wide variety of proteins for conjugation of proteins to peptides and to an oligonucleotide.^{36, 39-43}

In this chapter, EPL enables ligation of a recombinant FP possessing a thioester group at the C-terminus with the N-terminal cysteine residue of a PNA, producing a native peptide bond at the ligation site. Creating FP-PNA conjugates via EPL provides an ideal system to study programmable protein assembly. The unique recognition characteristics of PNA provide precise control over the inducible DNA-directed assembly of FP-PNA conjugates. Also, it allows the study of protein assembly by observing the photophysical properties of assembled FPs.

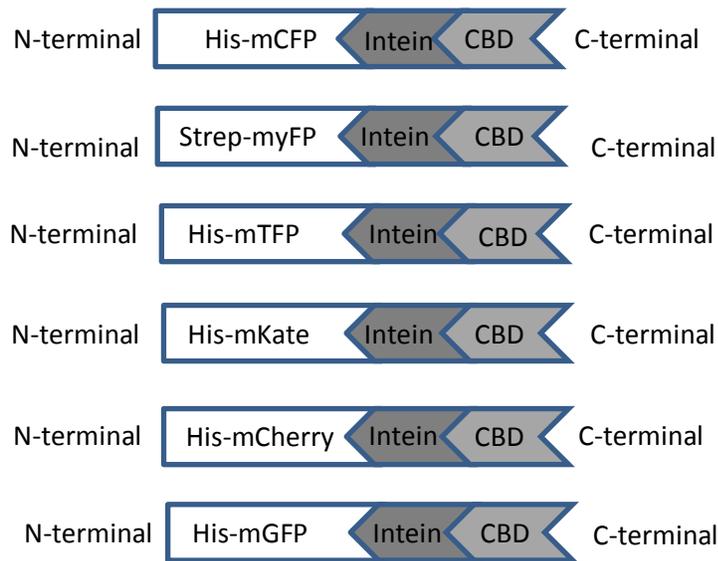
2.2 Results and Discussion

2.2.1 Expression and purification of fluorescent proteins with a thioester group at the C-terminus using EPL method

Recombinant plasmids of mCFP, mYFP, mKate, mTFP, mGFP and mCherry were expressed through EPL in *E. coli*. The expressed constructs included a chitin binding domain (CBD) used for protein purification through CBD affinity chromatography. The FPs were expressed in *E. coli* with additional histidine or streptavidin tag residues at the N-terminus which can be used for extra purification.

Incubation of the FP-intein-CBD construct bound to a chitin column with MESNA resulted in cleavage of thioester-FP from the rest of the construct (Scheme 2.1).

The proteins were isolated in different final concentrations (mYFP: 15.2 mg/ml (530 μ M), mTFP: 8 mg/ml (287 μ M), mCFP: 22.8 mg/ml (798 μ M), and mKate: 11.2 mg/ml (395 μ M)). Expression was poor for mGFP and mCherry. SDS-PAGE and mass spectrometry showed single purified proteins for mYFP and mCFP. However, impurities were observed with mTFP and mKate. These two proteins were further purified via a Ni-NTA His-tag chromatography. SDS-PAGE electrophoresis indicated good purity of the final products (Fig 2.1).



Scheme 2.1: Schematic depiction of expressed recombinant FP constructs. There is a polyhistidine (8XHis) at the N-terminal of all constructs except for mYFP which instead possesses an N-terminal-strep-tag (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys).

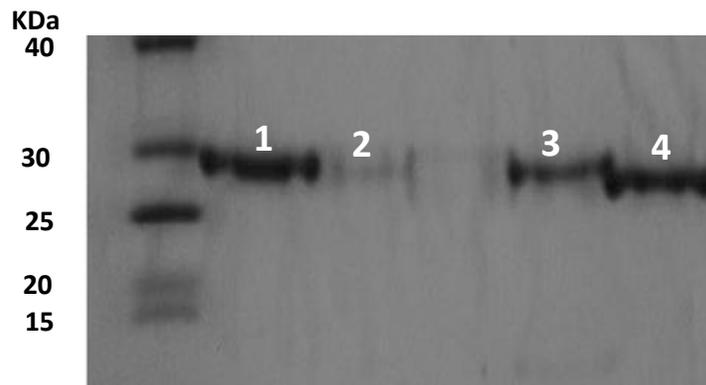


Figure2.1: SDS-PAGE electrophoresis result of purified mCFP-thioester (1), mYFP-thioester (2), mKate-thioester (3) and mTFP-thioester (4) after staining with Coomassie Blue. In all cases only single bands were observed indicating high purity.

Among those expressed FPs, the newly developed monomeric teal fluorescent protein (mTFP) was chosen for the rest of the study due to its greater brightness and photostability compared to CFP, YFP and mKate. It has also been reported as a good donor in different FRET pairs.^{27, 44-49}

Mass Spectrometry showed the integrity of the purified mTFP-thioester (measured m/z 27975 \pm 4 Da, calculated m/z 27990 Da) and SDS-PAGE (28 kDa) indicated the purity of the resulting mTFP-thioester.

2.2.2 PNA-mTFP ligation by NCL method

Purified mTFP (100 μ M) with a thioester group at the C-terminus was conjugated with excess cysteine-PNA (Cys-ACGTAC) (400 μ M) through NCL in the presence and absence of 50 mM MPAA as a catalyst at pH 7. Solutions were incubated for 20-485 min and overnight (\sim 18hrs) at room temperature. Mass spectrometry results showed that the ligation of PNA to mTFP occurred only in the presence of MPAA as catalyst. The ligation did not proceed using MESNA as a catalyst or with DTE as a reducing agent even after overnight incubation.

Three peaks appeared in the mass spectrum (Fig 2.2). The peaks were related to hydrolysed mTFP at m/z 27833 \pm 7 Da (calculated m/z 27844 Da), the mTFP-PNA product at m/z 29560 \pm 6 Da (calculated m/z 29555 Da), and partial double PNA attachment to mTFP at m/z 31293 \pm 7 Da (calculated m/z 31286 Da). The trace of second PNA attachment was efficiently removed by incubation with the reducing agent TCEP followed by centrifugation assisted dialysis. Since native mTFP does not feature a cysteine and the ligation solution was not oxygen free, the appearance of the second peak might be attributed to disulphide bond formation between the second PNA and the newly introduced SH group at the ligation site. HPLC analysis of the final product confirmed purity and integrity of the mTFP-PNA following with treatment with TCEP with a yield near 100% (Chapter 3).

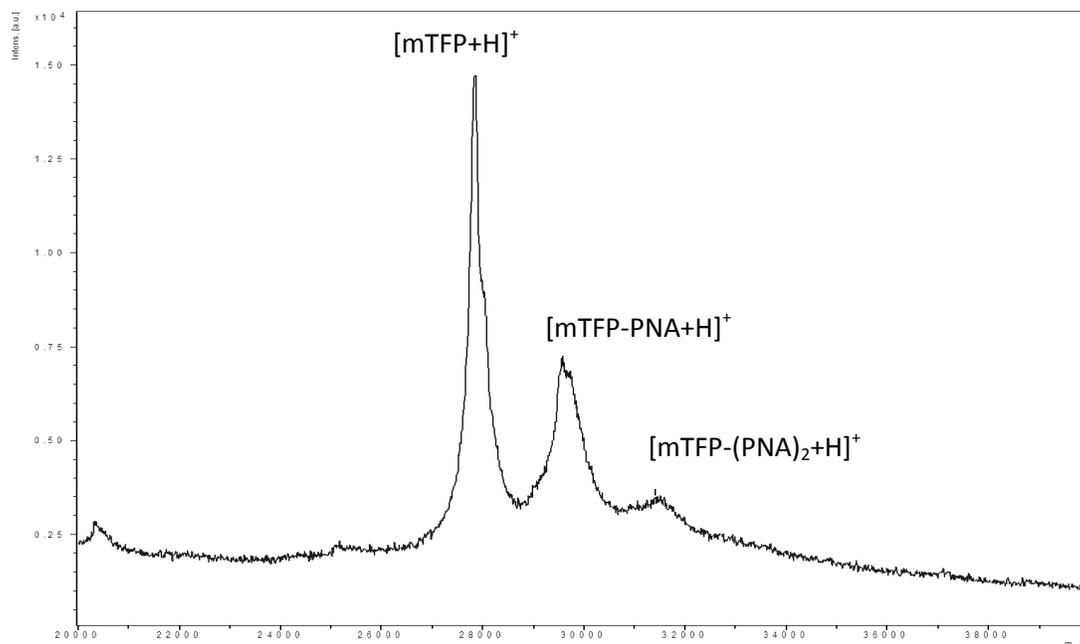
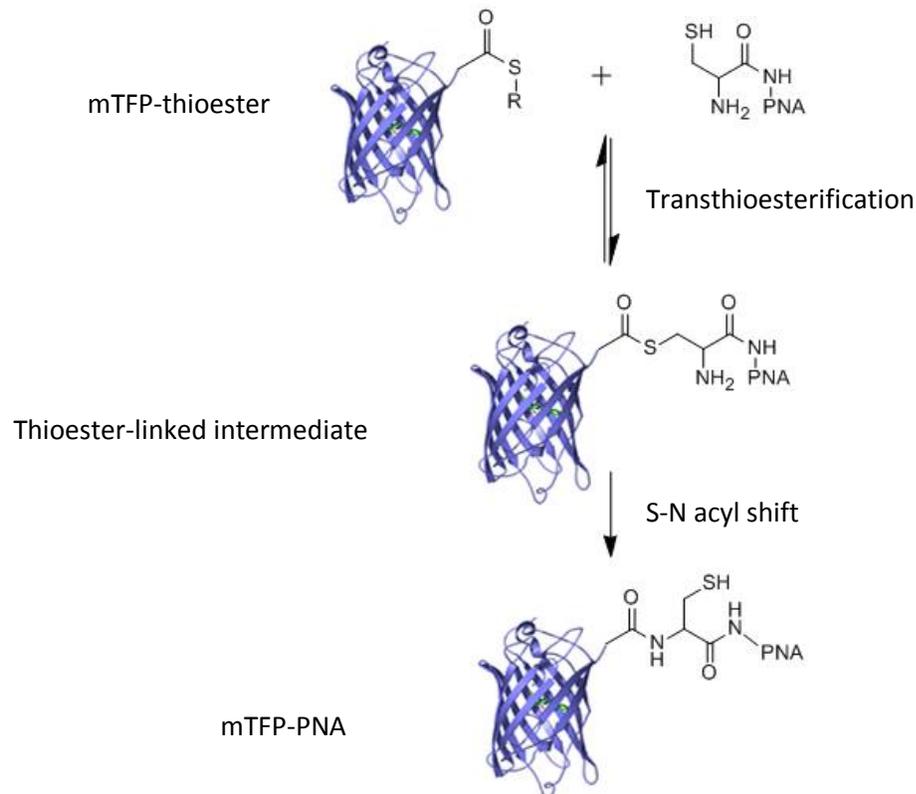


Figure 2.2: Mass spectrometry results of mTFP ligation to PNA (1734 Da) in the presence of MPAA (50 mM) and TCEP (70 mM) at pH 7 after 4 hrs incubation. Assignments: mTFP: 27833 Da and mTFP-PNA: 29560 Da. The third peak at 31293 Da is consistent with attachment of a second PNA to mTFP which was efficiently removed by incubation with the reducing agent TCEP followed by centrifugation assisted dialysis.

The mTFP-PNA ligation condition was chosen after optimising the ligation yield in terms of the effect of catalyst concentration (5-300mM), ligation time (20-485 min and overnight) and the ratio of PNA to mTFP (400-600 and 700 μ M). Based on the data, an optimised 1:4 ratio of mTFP to PNA in the presence of 50 mM MPAA and 70 mM TCEP at pH 7 and 60 min ligation time was used in all ligation experiments. Excess PNA and MPAA and TCEP were removed using centrifugation.

2.2.3 Kinetic study of mTFP-PNA NCL

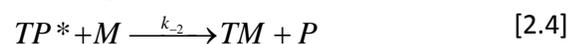
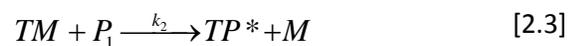
The kinetic behaviour of mTFP-PNA ligation consists of three key steps involving i) reversible thiol-thioester exchange, ii) reversible transthoesterification reaction and iii) fast and irreversible intramolecular acyl-transfer rearrangement⁵⁰⁻⁵² (Scheme 2.2).



Scheme 2.2: mTFP-PNA ligation via native chemical ligation (NCL). NCL is based on the transthoesterification reaction of mTFP- thioesters with thiol groups of PNA N-terminal cysteine. The transthoesterification reaction is followed by a fast and irreversible intramolecular acyl-transfer rearrangement forming a native peptide bond between mTFP and PNA.

Native chemical ligation is also thought to be a concerted anionic S_N2 substitution reaction. This reaction is second order overall, and first order on each of the reactants.⁵³

In theory, the following steps represent a complex set of reactions which, depending on the conditions, can be reduced to pseudo first order behaviour (Table 2.1).



In the above steps of mTFP-PNA ligation, P is PNA, T is mTFP, M is MPAA, TM is the mTFP-MPAA, TP* is the mTFP-PNA intermediate, and TP is mTFP-PNA with the native peptide bond. The eqn 2.1 and 2.2 are the reversible thiol-thioester exchange of mTFP-thioester with MPAA. The reversible transtioesterification with PNA is shown in eqn 2.3 and 2.4. The last eqn (2.5) is related to irreversible intramolecular acyl-transfer rearrangement.

Considering a variety of treatments,⁵³⁻⁵⁵ both pre-equilibrium and steady state approximation result in a pseudo first order kinetics for PNA (P) and mTFP (T) consumption and yield of ligated mTFP (TP) in our experiment conditions which is in accord with the observed data (Fig 2.3). The reaction summarized in Table 2.1.

Table 2.1 Kinetic behaviour of PNA-mTFP NCL system.

Assumptions ⁵³⁻⁵⁵	Saturated concentration	Proposed NCL kinetic reaction
Assumption 1: $k_1 > k_2 + k_{-1}$ Rate limiting step: Transthioesterification step Kinetic approximation: 1. Pre-equilibrium approximation for thiol-thioester exchange reaction (k_1 and k_{-1}) 2. Steady state approximation for intermediate TP* ($k_3 + k_{-2} > k_2$)	Catalyst MPAA (M): $[M] \gg [T]$ so $[M] \approx [M_0]$ (initial Concentration of M) PNA (P): $[P] > [T]$ so $[P] \approx [P_0]$ (initial Concentration of P)	$\frac{dTP}{dt} = \frac{k_1 k_2 k_3 [P_0] [M_0]}{k_{-1} (k_{-2} [M_0] + k_3)} [T]$ $\frac{k_1 k_2 k_3 [P_0] [M_0]}{k_{-1} (k_{-2} [M_0] + k_3)} = Z$ $\frac{dTP}{dt} = Z [T]$ Pseudo first order reaction $\Rightarrow [TP] = T_0 (1 - e^{-Zt})$ Thus: i. If $k_2 [M_0] < k_3$ (before consuming all of [T] as a reactant) $\Rightarrow Z = \frac{k_1 k_2 [P_0] [M_0]}{k_{-1}}$ ii. If $k_2 [M_0] > k_3$ (after consuming all of [T] as a reactant) $\Rightarrow Z = \frac{k_1 k_2 k_3 [P_0]}{k_{-1} k_{-2}}$
Assumption 2: $k_1 < k_2 + k_{-1}$ Rate limiting step: Thiol-thioester exchange step Kinetic approximation: 1. Steady state approximation for intermediate TM 2. Steady state approximation for intermediate TP* ($k_3 + k_{-2} > k_2$)	Catalyst MPAA (M): $[M] \gg [T]$ so $[M] \approx [M_0]$ (initial Concentration of M) PNA (P): $[P] > [T]$ so $[P] \approx [P_0]$ (initial concentration of P)	$\frac{dTP}{dt} = \frac{k_1 k_2 k_3 [P_0] [M_0]}{(k_{-1} + k_2 [P_0]) (k_{-2} [M_0] + k_3)} [T]$ If $k_1 \ll k_2 [P_0]$: $\frac{k_1 k_3 [M_0]}{(k_{-2} [M_0] + k_3)} = K$ $\frac{dTP}{dt} = K [T]$ Pseudo first order reaction $\Rightarrow [TP] = T_0 (1 - e^{-Kt})$ Thus: i. If $k_2 [M_0] < k_3$ (before consuming all of [T] as a reactant) $\Rightarrow K = k_1 [M_0]$ ii. If $k_2 [M_0] > k_3$ (after consuming all of [T] as a reactant) $\Rightarrow K = \frac{k_1 k_3}{k_{-1}}$

The kinetic study of mTFP-PNA ligation was performed using 50 mM MPAA , 70 mM TCEP at pH 7 with 20-485 min incubation time. The ligation yield was estimated through spectrophotometry at 260 nm as an indicator of the ligated PNA concentration and it was compared with the calculated absorbance of 100% ligation PNA to mTFP (1:1 ratio) to obtain the final percent yield. The results showed that the yield increased with time to a maximal value of ~90%.

The plot of yield as a function of time (Fig 2.3) showed a good fit to a first order rate law for mTFP-PNA ligation. Thus, in good agreement with theory, it can be concluded that the behaviour of native chemical ligation of mTFP-PNA is consistent with a pseudo-first order reaction.

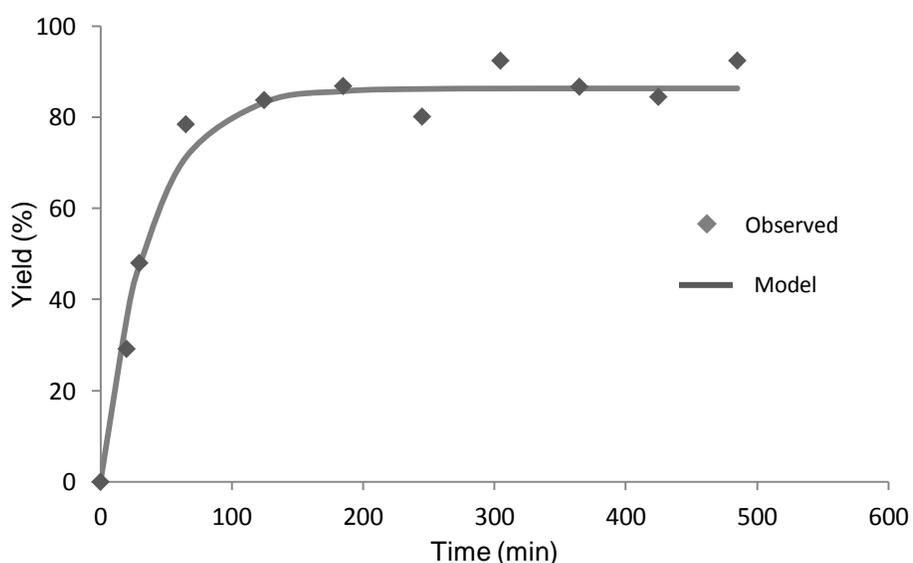


Figure 2.3 fitting of final mTFP-PNA yield in a pseudo first order growth model.

To test whether the addition of growth and decay terms in a kinetic model will explain sufficient variance to justify their inclusion; the F-test was used:⁵⁶

$$F = \frac{\frac{SS_1 - SS_2}{k - g}}{\frac{SS_2}{n - k}} \quad [2.6]$$

where ss_1 and ss_2 are the sum of squared residuals for the reduced (fewer parameters) and complete (more parameters) models. k is the number of parameters of complete model, g is the number of parameters of reduced model, and n is the number of data points.

The sum of squared residuals ss_1 and ss_2 are calculated based on following equation:

$$ss_i = \sum_{i=1}^n (y_{i_{observed}} - y_{i_{calculated}})^2 \quad [2.7]$$

In which $y_{observed}$, $y_{calculated}$ are the observed and calculated values for PNA according to a particular model and n is the number of data points.

The null hypothesis (H_0) is that the additional parameters do not explain sufficient variance to justify their inclusion. If the F-statistics is larger than the critical F value with $k-g$ and $n-k$ degree of freedoms, H_0 is rejected and it can be concluded that the addition of more parameters is justified.

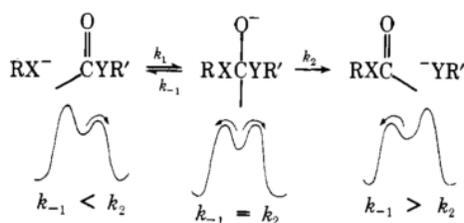
The F-test showed that the data significantly fit in a pseudo first order growth model and there is no evidence of any competing reaction. The errors of fitted parameters were estimated using macro SolvStat.xls⁵⁷ and are shown in Table 2.2 for mTFP-PNA production. The estimated pseudo-first order rate constant would be $0.026 \pm 0.003 \text{ min}^{-1}$.

Table 2.2 : Errors estimated for the fitted parameters of a pseudo first order growth model mTFP-PNA production

Parameters	Minimized value	R^2 (coefficient of determination)	SE(y) (the standard error of y(t))
k_1	$0.026 \pm 0.003 \text{ (min}^{-1}\text{)}$	0.97	4.88
A_1	$86 \pm 2 \text{ (}\mu\text{M)}$		

The rate of NCL depends on a number of variables including pH, the identity of C-terminal amino acid in the thioester segment and the nature of the thiol leaving group.^{51, 58}

Hupe and Jencks reported that the reaction of thiol RS^- with thioester is a completely symmetrical reaction. Therefore, as a tetrahedral intermediate is produced it will convert into products and reactants equally. When RS^- is more basic, it is a stronger nucleophile and worse leaving group than the thiol part of the thioester. In this condition, the intermediate will convert to products preferentially and the attack of RS^- will be rate limiting. When RS^- is less basic and therefore a better leaving group the intermediate will convert to reactants preferentially. Thus, intermediate production will be at equilibrium and expulsion of the leaving group will be the rate limiting step (scheme 2.3).⁵⁹



Scheme 2.3: Reaction of thiol and thioester⁵⁹.

The above explanation in conjunction with the absence of an observable thioester-linked intermediate ligation product in other NCL reports suggested that the rate limiting-step in NCL is the transthioesterification with the thiol group of N-terminal cysteine residue.⁵¹ In contrast, Johnson and co-workers demonstrated that the addition of aryl thiols such as MPAA as a catalyst (instead of alkyl thiols) can increase the rate of NCL and the rate limiting step of Cys-NCL catalysed by low concentration of added aryl thiols is the thiol-thioester exchange at the C-terminus of the thioester moiety with catalyst. They showed that aryl thiols increase the NCL rate since they are good thioester leaving groups and the transthioesterification of an aryl thioester is more rapid than the reverse thiol-thioester reaction which is followed by a fast and irreversible intramolecular acyl-transfer rearrangement generating a peptide bond at the ligation site.⁵⁰⁻⁵²

In our experiment, the initial MALDI-TOF mass spectrometry data of mTFP-PNA ligation samples (containing 5-300 mM MPAA with 240 min ligation time at the similar conditions mentioned above) exhibited an increase of ligation yield up to using 50 mM MPAA followed by a decrease of yield (Fig 2.4).

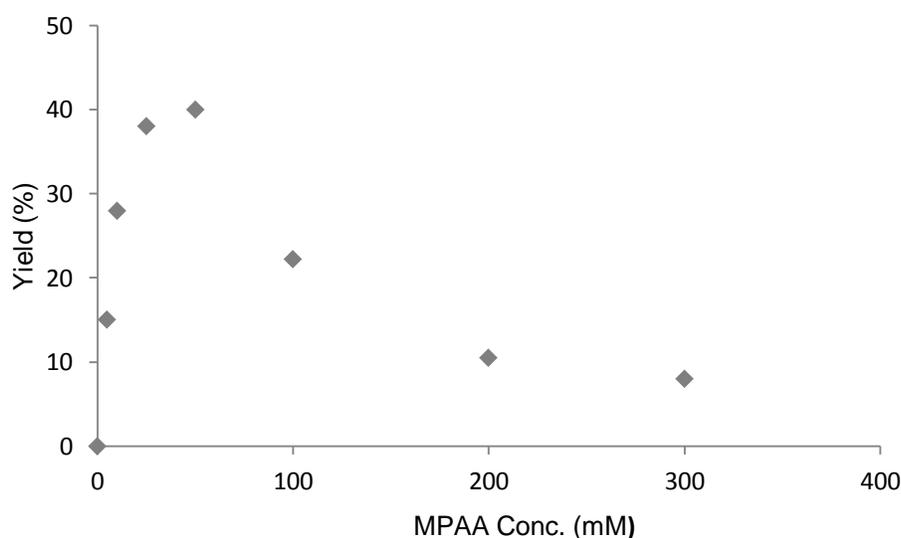


Figure 2.4: The graph of the ligation yield vs. different concentration of MPAA (5-300mM) after 240 min ligation time obtained by MALDI-TOF mass spectrometry.

It should be noticed that MALDI-TOF is a powerful method to recognize and investigate ligation product but there are issues about the lack of peak reproducibility and fragmentation⁶⁰⁻⁶². Although this kind of data is not normally used for quantitative applications like this, the decreasing yield when the concentration exceeds 50 mM in a roughly $1/[MPAA]$ function (based on fitting) in conjunction with the above discussion on NCL rate limiting step may imply that the rate limiting step of mTFP-PNA ligation is transthioesterification of MPAA-thioester with the thiol group of N-terminal cys-PNA. This may be due to MPAA molecules competing with PNA in transthioesterification and also by catalyst molecules preventing the bulky PNA molecules from reaching the mTFP. Competition between MPAA and PNA could lead to the decrease of yield For fixed time.

The decreasing trend of MS data was not observed when the effect of catalyst concentration on mTFP–PNA ligation was investigated by varying concentration of MPAA (5-200mM) at fixed pH (7) and incubation time (60 min) using UV-visible spectrophotometry.

The UV-visible spectrophotometry results were obtained by using the estimated absorbance of the mTFP-PNA at 260 nm as an indicator of PNA concentration and it was compared with the estimated absorbance of 1:1 ratio of PNA conjugation to mTFP (100%) to calculate the percentage of final ligation yield. The absorbance of the same concentration of pure mTFP was subtracted from the total absorbance. The data showed more than 90% yield. The yield increased monotonically with added catalyst up to 100 mM (Fig 2.5).

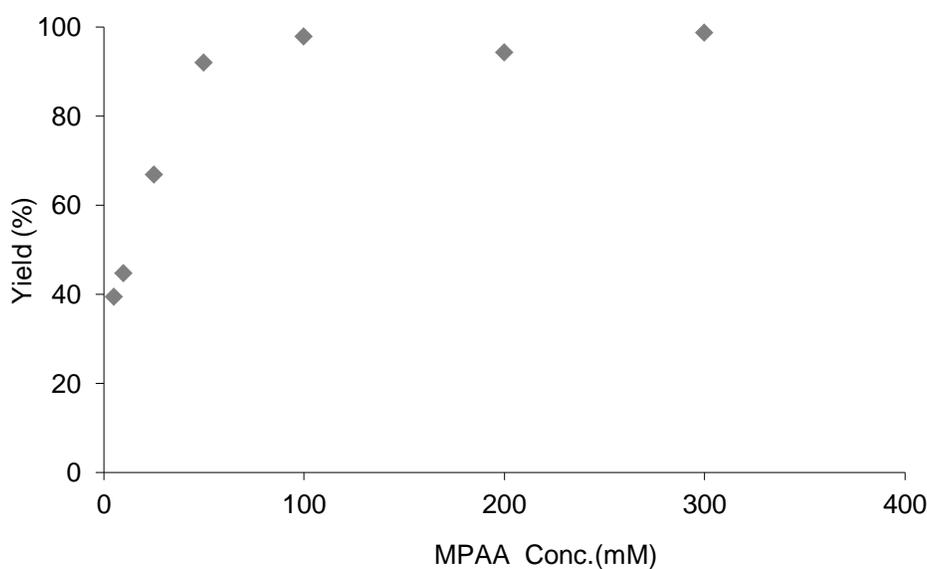


Figure 2.5: Ligation yield (μM) vs. catalyst: MPAA (5-300mM) with a 60 min reaction time at room temperature measured by UV.Vis spectrophotometry.

The difference of results between mass spectrometry and spectrophotometry is believed to be due to the effect of pH on NCL rate since the pH of two sets of experiments was slightly different. The MS experiments were done early in the project with less pH control. The MPAA stock was not prepared fresh for the MS experiment and the pH paper that was used to adjust

pH of MPAA and TCEP to 7 was less accurate (pH paper 1-14) compared to the spectrophotometry experiment. MPAA is soluble in water in pH 7 so the pH of the stock solution should be adjusted to pH 7 by adding NaOH. Using more accurate pH paper 6.4- 8 for that solution showed higher pH (maximum 0.6 differences).

As mentioned above, the rate of NCL is affected by the pH of the ligation solution since the rate of thiol-thioester exchange and hydrolysis for thioester is pH-dependent.^{51, 63} For the MS experiment, using higher concentration of MPAA could increase the pH of the ligation solution more than similar solutions used in UV-Vis experiment. For example, the observed pH of the ligation solution was 7.6 using 300 mM of not fresh MPAA. Therefore, the concentration of OH⁻ could be 4 times more in that ligation solution. More OH⁻ could lead to increase of base hydrolysis of thioester group of the protein via the reaction below:



Consequently, the decrease of thioester reactant leads to lower yield like what was observed in mass results using 100-300 mM MPAA. Moreover, the increase of pH for example from 7 to 7.6 can result in increase of the MPAA anionic form by a factor of 4 based on the Hendersson-Hasselbach equation (pK_a MPAA= 6.6):

$$pH = pK_a + \log \frac{[S^-]}{[SH]}$$

The active form of MPAA in NCL reaction is the anionic form and its concentration can increase in higher pH.⁵² Therefore, the competition of MPAA with PNA could be even more in this situation.

Therefore, even slight difference in pH adjustment may decrease the amount of active thioester and increasing the competition of MPAA and PNA, leading to the reduction of mTFP-PNA yield in a fixed time (Ms data, Fig 2.4). A full exploration of the effects of pH was beyond the scope of the current study.

2.4 Conclusion and outlook

In this chapter, expressed protein ligation as a semisynthetic version of native chemical ligation provided a site-specific and facile conjugation of fluorescent protein bearing C-terminal thioester to N-terminal thiol group of PNA.

Four different fluorescent proteins: mTFP, mYFP, mKate and mCFP with thioester group at C terminus were successfully expressed and purified using molecular biology techniques. Mass spectrometry and SDS-PAGE analysis confirmed the purity of the final FPs. Among those purified FPs, mTFP was chosen for further study because of its greater brightness and photostability compared to mCFP, mYFP, and mKate and its good characteristic to be a donor in a FRET pair.

The ratios of PNA, MPAA and TCEP were optimized for the ligation and the trace of the second PNA attached to the final mTFP-PNA product was removed using TCEP. The conjugation was assessed by mass spectrometry and spectrophotometry and showed complete conversion of FP to ligated form. The reaction rate (k) of the ligation of 4:1 ratio of PNA:mTFP in the presence of MPAA (50 mM) and TCEP (70 mM) at pH 7 and room temperature was estimated as $0.026 \pm 0.003 \text{ min}^{-1}$.

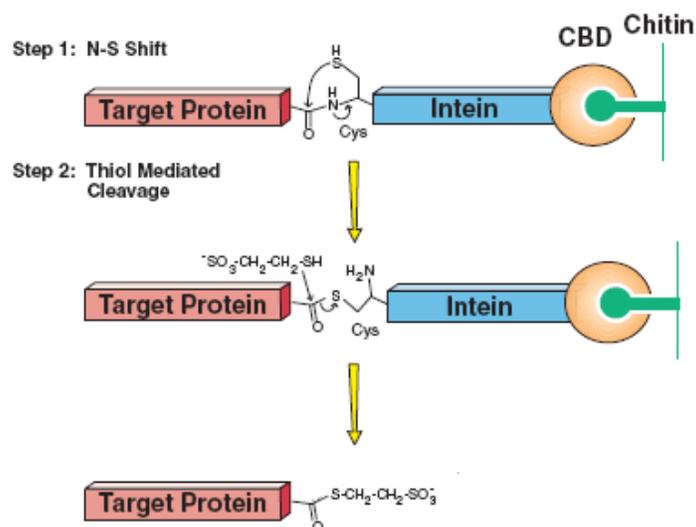
It was shown that EPL as a chemo-selective ligation strategy could enable the production of FP-PNA conjugates for potential use as self-assembly units. These units can be used as models of protein aggregation through specific recognition of PNA for DNA frameworks. Fluorescent proteins can be fused to proteins of interest via cloning methods allowing the progress of assembly to be measured. Together with EPL, FPs and PNA can provide a framework of self-assembling units to study aggregation using model systems.

2.5 Experimental part

All reagents unless specifically noted were purchased from Sigma-Aldrich.

2.5.1 Protein expression and purification

The fluorescent proteins including monomeric teal fluorescent protein (mTFP), monomeric yellow fluorescent protein (mYFP), mCherry, monomeric green fluorescent protein (mGFP), mKate and monomeric cyan fluorescent protein (mCFP) modified plasmids were generated in the group of Prof. Brunsveld⁶⁴ (University of Eindhoven, Netherlands, especial thanks to H. D. Nguyen and D. T. Dang for plasmid pHT581 and help with protein expression) using IMPACT™ system (New England Biolabs (NEB), USA) to express and purify thioesters linked to recombinant proteins via intein-mediated cleavage and an affinity based chitin binding domain purification tag (Scheme 2.4).



Scheme 2.4: C-Terminal thioester introduced at the end of the protein using intein modified IMPACT™ system (NEB).

The modified plasmids were transformed into competent *E. coli* BL21 cells. The recombinant plasmids expressed FP-modified intein-chitin binding domain protein constructs.

For expression experiments, 2 litre LB (lysogeny broth) media containing ampicillin (100 µg/ml) were inoculated with 20 ml of overnight (18 hrs) grown culture for each plasmid. The cultures

were grown at 37 °C at 200 rpm shaking for 2-3 hours. When the OD₆₀₀ reached between 0.5-0.7, the protein expression was induced with 500 µl IPTG (Isopropyl β-D-1-thiogalactopyranoside) (1M) and incubated at 15 °C, 180 rpm overnight. Cells were harvested by centrifugation at 10000 g, 4 °C for 10 min to collect the cell pellets. The pellets were resuspended in 40 ml Bugbuster (Novagen, Germany) reagent and 40µl Benzonase (25 unit/ml) (Novagen, Germany) with several times vortexing. The cell suspension was incubated on a shaking platform for 20 min at room temperature. The insoluble cell debris was removed by centrifugation at 40000g for 40 min at 4°C. The clear supernatant was collected for protein purification with chitin beads (New England Biolab, UK).

Four millilitre of chitin resin (New England Biolabs) was transferred to a 10 ml disposable chromatography column (Biorad) and equilibrated with 10 volumes of washing buffer (Na₂HPO₄ (19.3 mM), NaH₂PO₄ (25 mM), NaCl (500 mM), EDTA (0.5 mM) in distilled deionised water (ddH₂O) , pH 7). Cleavage of the C-terminal intein-tag on the fluorescent proteins was performed through incubation of loaded proteins with 20 ml elution buffer (Na₂HPO₄ (23.2 mM), NaH₂PO₄ (25mM), NaCl (100mM), EDTA (0.5 mM), MESNA (400 mM) in ddH₂O, pH 7.5) overnight in the dark with slow shaking at room temperature. MESNA elution buffer was always prepared fresh immediately before use. Afterwards, the proteins were eluted from the column with another 20 ml of elution buffer and washed with two column volumes of storage buffer (Na₂HPO₄ (19.3 mM), NaH₂PO₄ (25mM), NaCl (50 mM), EDTA 0.1 mM, in ddH₂O pH 7). The collected proteins were concentrated by centrifuging several times at 3700 g for 10-16 min to remove the MESNA and stored at -80°C. The chitin beads were regenerated by washing with 3 column volumes of 0.3 M NaOH solution and soaking with that for 30 min at room temperature. The sodium hydroxide then was removed by washing with 20 column volumes of distilled water and 5 column volumes of 20% ethanol. The regenerated chitin beads then were stored at 4 °C.

The purified protein was checked on a sodium dodecyl sulfate (SDS)-PAGE gel (12% resolving polyacrylamide gel and 5% stacking gel). The resulting mTFP and mKate were further purified

using His tag affinity column chromatography. To prepare the Ni-NTA resin, it was washed with 2-3 column volumes of distilled water. About 20 ml nickel sulphate solution (50 mM) was added to the column followed by washing with distilled water and His-tag buffer (Na_2HPO_4 (19.3 mM), NaH_2PO_4 (25 mM), NaCl (50 mM), EDTA (0.1 mM), imidazole (20 mM) in ddH_2O , pH 7). After adding protein, the column was washed again with 10 ml of His-tag washing buffer and the proteins eluted using 4 ml His-tag elution buffer (Na_2HPO_4 (19.3 mM), NaH_2PO_4 (25 mM), NaCl (50 mM), EDTA (0.1 mM), imidazole (250 mM) in ddH_2O , pH 7). The His-tag elution and washing buffers with imidazole were made fresh. The final concentration of purified proteins was determined by UV-visible spectrophotometer at 280 nm (Perkin Elmer, Lambda 25).

2.5.1.1 SDS-PAGE electrophoresis

10 μl of each sample containing 10 μM of purified protein was heated at 90°C for 5 min and mixed with 10 μl of SDS sample loading buffer (Biorad). The denatured protein was applied to a precast acrylamide gel (Biorad)(12% resolving polyacrylamide gel and 5% stacking gel) and the electrophoresis was run in premixed Tris/Glycine/SDS Running Buffer (Biorad) for 70 min at 100 V. The gels were stained with Coomassie Brilliant Blue (Sigma) was used.

2.5.2 PNA- Protein ligation

Ligation reactions were run on a 100 μl scale using a 4-fold excess of PNA ((Cys-ACGTAC)(Advanced Peptides, USA). Solutions containing 100 μM mTFP in storage buffer (Na_2HPO_4 (19.3 mM), NaH_2PO_4 (25mM), NaCl (50mM), EDTA (0.1 mM), in ddH_2O , pH 7), 400 μM PNA1 (ACGTAC), 50 mM MPAA as a catalyst, and 70 mM aqueous TCEP (pH 7) as a reducing agent were incubated at room temperature for 60 min. These conditions were chosen after optimising the effects of MPAA (0-300 mM), PNA (400, 600, and 700 μM) and incubation time (20-485 min and overnight) on the ligation reaction while other conditions were maintained constant.

Isolation of the product from the remaining reactants was performed using centrifugal filter tubes (Amicon 10000 MW, Millipore) at 13000 rpm. The retained portion was resuspended in 70 mM TCEP, incubated, and centrifuged several times to remove any second PNA attached via a disulphide bond generated at the ligation site. The mTFP-PNA conjugate was subsequently stored at 4°C (short times) or -20°C (long term storage).

The effect of other thiol catalysts (such as MESNA) and DTE as a reducing agent were investigated using the same conditions and for 8, 24 and 48 hrs reaction time.

The ligation results of each experiment were investigated by MALDI-TOF mass spectrometry and UV-Vis spectrophotometry.

2.5.2.1 MALDI-TOF Mass spectrometry analysis

Ligation samples were desalted using C18 Ziptip (Millipore) and the dried-droplet method with 1 µl of sample with 1µl of sinapinic acid (Bruker) matrix put on the sample plate spot. The mass spectrometry analysis was performed using Bruker Ultraflex III MALDI TOF/TOF system. The Bruker flexAnalysis software was used for analysis of the data.

2.5.2.1.1 C18 ZipTip Protocol

The ZipTip protocol consists of three steps. In the first step, the C18 Ziptip was washed 3 times with 10 µL of acetonitrile followed by 3 washes with 80% acetonitrile/0.1% TFA and 4 times with 0.1% TFA. In the second step, samples were bound to the tip by 20 times cycling (aspirate-dispense-aspirate-dispense) of 2-10 µL of sample followed by 3-4 washes with 10 µL 0.1% TFA to remove salts. In the final step, the sample was eluted from the Ziptip with 1.5-5µL of 80% acetonitrile/0.1% TFA to use for MALDI –TOF analysis (Bruker Ultraflex III MALDI TOF/TOF).

2.5.2.2 UV-Vis Spectrophotometry

Ligation samples incubated for 20-485 minutes were diluted after purification and their absorbance at 260 nm measured to estimate PNA concentration. Pure mTFP or mYFP absorbance at 260 nm was measured and subtracted from the mTFP-PNA and mYFP-PNA absorbance. The absorbance of the supernatant containing the unreacted PNA was also measured. The extinction coefficient of the PNA was estimated based on the published values for the molar extinction coefficients of PNA bases. Following this method the overall extinction coefficient is the sum of the individual absorbances from the bases comprising the oligomer :

$$\epsilon_{260} = \sum_{i=1}^n \epsilon_i$$

Where ϵ_i is the molar extinction coefficient of i th base at 260 nm (ϵ_{260} (A) =13.7, ϵ_{260} (G) = 11.7, ϵ_{260} (C) = 6.6, ϵ_{260} (T) = 8.6 mL/(μ mole x cm), obtained from PANAGENE) and n is the number of PNA bases.

2.6 References

1. P. W. K. Rothmund, *Nature*, 2006, **440**, 297-302.
2. Z. L. Pianowski and N. Winssinger, *Chemical Society Reviews*, 2008, **37**, 1330-1336.
3. M. Rothlingshofer, K. Gorska and N. Winssinger, *Journal of the American Chemical Society*, 2011, **133**, 18110-18113.
4. K. K. Sadhu, M. R othlingsh ofer and N. Winssinger, *Israel Journal of Chemistry*, 2013, **53**, 75-86.
5. N. C. Seeman, *Nature*, 2003, **421**, 427-431.
6. U. Feldkamp and C. M. Niemeyer, *Angewandte Chemie International Edition*, 2006, **45**, 1856-1876.
7. Z. Dadon, M. Samiappan, E. Y. Safranchik and G. Ashkenasy, *Chemistry – A European Journal*, 2010, **16**, 12096-12099.
8. P. K. Lo, K. L. Metera and H. F. Sleiman, *Current Opinion in Chemical Biology*, 2010, **14**, 597-607.
9. C. M. Niemeyer, *Angewandte Chemie International Edition*, 2001, **40**, 4128-4158.
10. A. Roloff and O. Seitz, *Chemical Science*, 2013, **4**, 432-436.
11. K. Gorska, K.-T. Huang, O. Chaloin and N. Winssinger, *Angewandte Chemie International Edition*, 2009, **48**, 7695-7700.
12. K. Gorska, J. Beyrath, S. Fournel, G. Guichard and N. Winssinger, *Chemical Communications*, 2010, **46**, 7742-7744.
13. T. Kobayashi, N. Morone, T. Kashiyaama, H. Oyamada, N. Kurebayashi and T. Murayama, *PLoS ONE*, 2008, **3**, e3822.
14. M. Zaccolo and T. Pozzan, *IUBMB life*, 2000, **49**, 375-379.
15. R. Wombacher and V. W. Cornish, *Journal of Biophotonics*, 2011, **4**, 391-402.
16. M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher, *Science (New York, N.Y.)*, 1994, **263**, 802-805.
17. J. Lippincott-Schwartz and G. H. Patterson, *Science*, 2003, **300**, 87-91.

18. D. M. Chudakov, M. V. Matz, S. Lukyanov and K. A. Lukyanov, *Physiological Reviews*, 2010, **90**, 1103-1163.
19. R. Miyake, T. Uchimura, X. Li and T. Imasaka, *Chemical and Pharmaceutical Bulletin*, 2013, **61**, 82-84.
20. C. Stadler, E. Rexhepaj, V. R. Singan, R. F. Murphy, R. Pepperkok, M. Uhlen, J. C. Simpson and E. Lundberg, *Nature Methods*, 2013, **10**, 315-323.
21. H. Abe, Y. Yanagawa, K. Kanbara, K. Maemura, H. Hayasaki, H. Azuma, K. Obata, Y. Katsuoka, M. Yabumoto and M. Watanabe, *Journal of Andrology*, 2005, **26**, 568-577.
22. M. Chalfie, *Photochemistry and Photobiology*, 1995, **62**, 651-656.
23. R. A. Cinelli, A. Ferrari, V. Pellegrini, M. Tyagi, M. Giacca and F. Beltram, *Photochemistry and Photobiology*, 2000, **71**, 771-776.
24. I. Gautier, M. Tramier, C. Durieux, J. Coppey, R. B. Pansu, J. C. Nicolas, K. Kemnitz and M. Coppey-Moisan, *Biophysical Journal*, 2001, **80**, 3000-3008.
25. N. Hayes, E. Howard-Cofield and W. Gullick, *Cancer Letters*, 2004, **206**, 129-135.
26. C. S. Lisenbee, S. K. Karnik and R. N. TRELEASE, *Traffic*, 2003, **4**, 491-501.
27. T. Nagai, K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba and A. Miyawaki, *Nature Biotechnology*, 2002, **20**, 87-90.
28. B. A. Pollok and R. Heim, *Trends in Cell Biology*, 1999, **9**, 57-60.
29. J. K. Sanders and S. E. Jackson, *Chemical Society Reviews*, 2009, **38**, 2821-2822.
30. R. Y. Tsien, *Annual Review of Biochemistry*, 1998, **67**, 509-544.
31. B. Zhang, *Biochemical and Biophysical Research Communications*, 2004, **323**, 674-678.
32. G. Crivat and J. W. Taraska, *Trends in Biotechnology*, 2012, **30**, 8-16.
33. N. Labòria, R. Wieneke and R. Tampé, *Angewandte Chemie International Edition*, 2013, **52**, 848-853.
34. E. M. Sletten and C. R. Bertozzi, *Angewandte Chemie International Edition*, 2009, **48**, 6974-6998.
35. L. I. Willems, N. Li, B. I. Florea, M. Ruben, G. A. van der Marel and H. S. Overkleeft, *Angewandte Chemie International Edition*, 2012, **51**, 4431-4434.
36. L. De Rosa, A. Russomanno, A. Romanelli and L. Andrea, *Molecules*, 2013, **18**, 440-465.
37. N. Stephanopoulos and M. B. Francis, *Nature Chemical Biology*, 2011, **7**, 876-884.
38. E. Baslé, N. Joubert and M. Pucheault, *Chemistry & Biology*, 2010, **17**, 213-227.
39. M. Vila-Perello, Z. Liu, N. H. Shah, J. A. Willis, J. Idoyaga and T. W. Muir, *Journal of the American Chemical Society*, 2013, **135**, 286-292.
40. C. P. Hackenberger and D. Schwarzer, *Angewandte Chemie (International ed. in English)*, 2008, **47**, 10030-10074.
41. X. Duburcq, C. Olivier, F. Malingue, R. Desmet, A. Bouzidi, F. Zhou, C. Auriault, H. Gras-Masse and O. Melnyk, *Bioconjugate Chemistry*, 2004, **15**, 307-316.
42. C. M. Niemeyer, *Biochemical Society*, 2004, **32**, 51-53.
43. C. M. Niemeyer, *Angewandte Chemie (International ed. in English)*, 2010, **49**, 1200-1216.
44. Y. Sun, C. F. Booker, S. Kumari, R. N. Day, M. Davidson and A. Periasamy, *Journal of Biomedical Optics*, 2009, **14**, 054009.
45. Y. Sun, H. Wallrabe, C. F. Booker, R. N. Day and A. Periasamy, *Biophysical Journal*, 2010, **99**, 1274-1283.
46. Y. Sun, H. Wallrabe, S. A. Seo and A. Periasamy, *Chemphyschem : a European Journal of Chemical Physics and Physical Chemistry*, 2011, **12**, 462-474.
47. D. W. Piston and G. J. Kremers, *Trends in Biochemical Sciences*, 2007, **32**, 407-414.
48. R. N. Day, C. F. Booker and A. Periasamy, *Journal of Biomedical Optics*, 2008, **13**, 031203.
49. M. L. Markwardt, G.-J. Kremers, C. A. Kraft, K. Ray, P. J. C. Cranfill, K. A. Wilson, R. N. Day, R. M. Wachter, M. W. Davidson and M. A. Rizzo, *PLoS ONE*, 2011, **6**, e17896.
50. Z. Tan, S. Shang and S. J. Danishefsky, *Angewandte Chemie International Edition*, 2010, **49**, 9500-9503.
51. E. C. Johnson and S. B. Kent, *Journal of the American Chemical Society*, 2006, **128**, 6640-6646.
52. C. Wang, Q.-X. Guo and Y. Fu, *Chemistry - An Asian Journal*, 2011, **6**, 1241-1251.
53. J. H. Espenson, *Chemical Kinetics and Reaction Mechanisms*, MacGraw-Hill, Inc., USA, 1995.
54. B. G. Cox, *Modern Liquid Phase Kinetics*, Oxford Press, USA, 1994.
55. K. A. Connors, *Chemical Kinetics: The Study of Reaction Rates in Solution*, VCH Publishers, Inc., USA, 1990.
56. T. S. Forde and Q. S. Hanley, *Photochemical & Photobiological Sciences*, 2005, **4**, 609-616.
57. E. Joseph Billo, *Excel for Chemists : A Comprehensive Guide*, JohnWiley&Sons, Inc., USA, 2001.
58. S. B. Pollock and S. B. Kent, *Chemical Communications (Cambridge, England)*, 2011, **47**, 2342-2344.
59. D. J. Hupe and W. P. Jencks, *Journal of the American Chemical Society*, 1977, **99**, 451-464.

60. J. Albrethsen, *Journal of Proteomics*, 2011, **74**, 765-773.
61. J. Albrethsen, *Clinical Chemistry*, 2007, **53**, 852-858.
62. M. Bucknall, K. Y. Fung and M. W. Duncan, *Journal of the American Society for Mass Spectrometry*, 2002, **13**, 1015-1027.
63. P. Bracher, P. Snyder, B. Bohall and G. Whitesides, *Origins of Life and Evolution of Biospheres*, 2011, **41**, 399-412.
64. D. A. Uhlenheuer, D. Wasserberg, H. Nguyen, L. Zhang, C. Blum, V. Subramaniam and L. Brunsveld, *Chemistry -A European Journal*, 2009, **15**, 8779-8790.

Chapter 3

Induced assembly of fluorescent protein-PNA conjugates using DNA as a framework

The prominent role of protein clustering in biological processes such as signal transduction has been widely studied. Some aspects of protein behavior in clusters are still unclear. Further effort in this area is vital for understanding the mechanism of disease. For example, membrane protein clustering behavior is thought to be involved in the conversion of cells to a cancerous state. Developing a model system for studying protein clustering could be useful in elucidating the role of protein aggregation in cells. As described in previous chapters, PNA can provide unique functionality to controllably direct the assembly of proteins. In chapter 2, PNA was conjugated to mTFP to provide a fluorescent monomer which could be inducibly assembled and studied by fluorescent techniques. In this chapter, DNA scaffolds are shown to programmably align FP-PNA conjugates to create hetero-FRET or homo-FRET systems. Directed assembly on a DNA beacon with 6-FAM and Dabcyl at its ends, creates an assembled hetero-FRET system with the mTFP-PNA conjugate. Using fluorescence techniques such as intensity, frequency domain lifetime and anisotropy measurements, the assembled system exhibited decreased donor intensity, changes in frequency domain lifetime, and increased anisotropy as indicators of hetero-FRET. Using a DNA scaffold allowed for the assembly of multiple mTFP-PNA constructs exhibiting homo-FRET. Efficient assembly of protein in dimers and oligomers forms on the DNA-PNA frameworks was confirmed with size exclusion chromatography (SEC) and SDS-PAGE. Assembly of multiple proteins in a row induced homo-FRET among the mTFP-PNAs assembled on the DNA scaffolds. The PNA directed assembly on DNA provides an induced and controllable model system of protein clustering suitable for studying protein behavior.

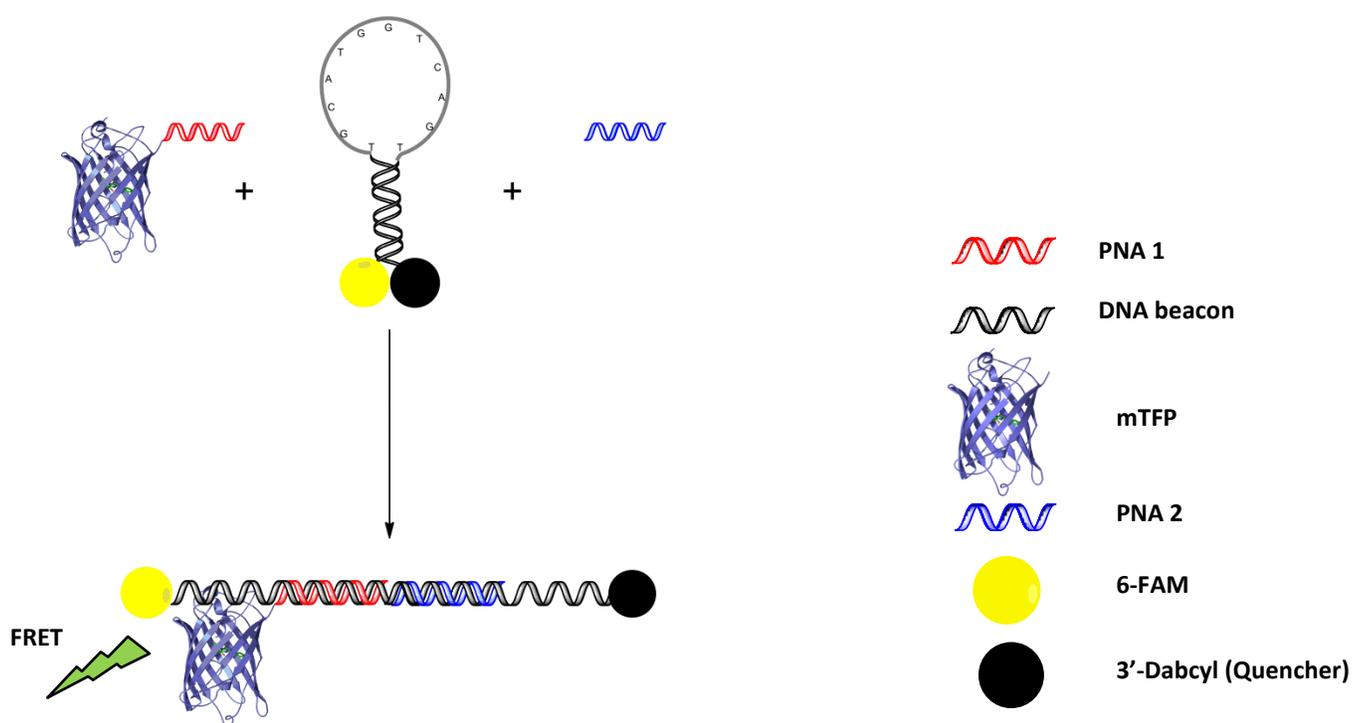
3.1 Introduction

Previous chapters have discussed the need for getting control over protein assembly. Template directed assembly would allow a connection to be made between protein aggregation and the crucial functionality of aggregates in vital biological processes. This could assist in understanding and targeting dysfunctional aggregation associated with a variety of diseases including cancers and neurodegenerative disorders.¹⁻¹⁴

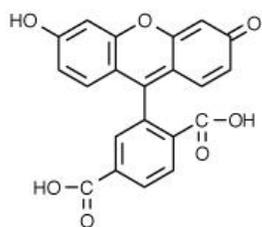
The ability of nucleic acid to controllably self-assemble supramolecular architectures can provide a versatile tool to study protein function and structure.^{15, 16} PNA with its flexible chemistry and biological robustness has been applied as an oligonucleotide tag to be organized into microarrays by self-assembling on a DNA platform.¹⁵⁻²¹ Additionally, it has been reported that PNA could be successfully used to self-assemble carbohydrate ligands, receptor ligands and antibody fragments in programmed dimeric or multimereic complexes.^{16, 22-25}

Clustering of receptors along the cell membrane is a fundamental feature of signal transduction and cellular recognition.^{24, 26, 27} Insight into the behavior of cell-membrane receptors in clusters can be achieved by a study of a model of controllable, well-organized and assembled protein oligomers. In this chapter, the FP-PNA conjugates prepared by NCL (Chapter 2) will be used as building blocks to assemble in dimer and oligomer forms using DNA scaffolds as frameworks. The first model system consisted of a PNA-fluorescent protein (FP) conjugate hybridized to a DNA beacon with a quencher and a fluorophore at its ends, proving the principle of PNA directed inducible assembly (Scheme 3.1). A second model system was engineered consisting of DNA strands containing repeating motifs complementary to the PNA sequence on a fluorescent protein. Such a system will exhibit homo-FRET on assembly and could in theory assemble any number of engineered building blocks (Scheme 3.2).

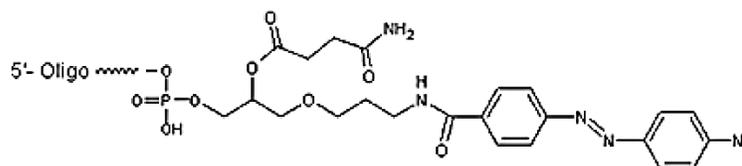
[Pick the date]



6-FAM

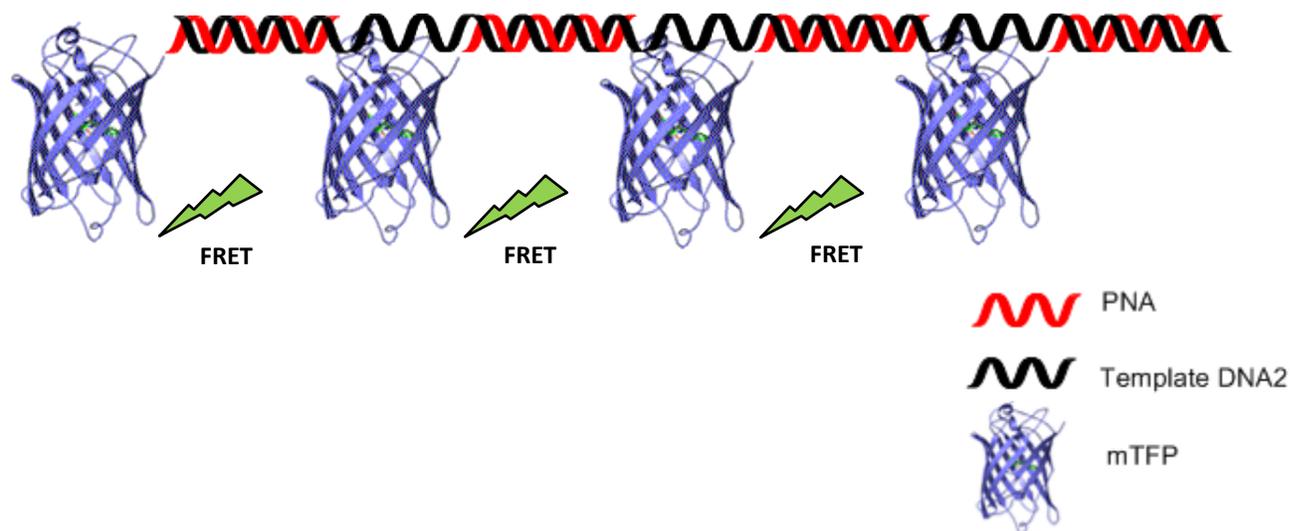


3'-Dabcyll



Scheme 3.1: DNA beacon structure in closed and open forms (top) and chemical structures of 6-FAM (Life Technologies) and 3'-dabcyll (Gene Link™) (bottom). In the assembled system hetero-FRET can occur from mTFP to either the 6-FAM or the Dabcyll group in DNA beacon.

[Pick the date]



Scheme 3.2: Assembly of mTFP-PNA in tetramer form using DNA2 as a framework. Shorter or longer strands may be used to assemble dimer, trimer & higher oligomers.

3.1.1 DNA beacon

NCL and EPL allow specific PNA sequences to be attached to peptides and proteins (Chapter 2). Assembly of the ligated molecules can be induced by combining with complementary PNA or DNA oligonucleotides. One framework for monitoring assembly is with the assistance of a molecular beacon (Scheme3.1).

Molecular beacons have been widely used for biomolecular recognition in biology, chemistry, medical sciences and biotechnology since they were reported in 1996.²⁸ A molecular beacon is a synthetic DNA molecule with a basic stem-loop or hairpin structure. The loop sequence consists of a 15-30 base oligonucleotide complementary to a target sequence while the 5-7 base stem sequence is complementary to itself in the closed state. A fluorophore donor is covalently attached to one end and the other end has an acceptor/quencher. The donor and acceptor are in close proximity when the stem region is in its closed form. For donor-quencher molecular beacons the system is thought to quench fluorophore via static quenching.²⁸ In the presence of target sequences, the longer and more stable target-loop hybrid is formed leading

[Pick the date]

to a conformational change to an open form. Due to a longer distance between fluorophore and quencher, the fluorescence signal of the beacon becomes observable indicating target binding.^{29, 30} Choosing an appropriate fluorophore-quencher pair is an important part in designing a DNA beacon to reduce the background fluorescence of the beacon in the absence of target. Another concern is the selection of the nucleotide base positioned adjacent to the fluorophore due to quenching by the nucleotide. The most strongly quenching base is guanine followed by adenosine, cytosine and thymidine. The electron donating capability of guanine is responsible for its quenching properties, which allows for charge transfer to the fluorophore. Therefore, the sequence of the stem region has to be designed carefully in terms of neighboring nucleotide and stability.²⁸

A variety of different molecular beacons based on chimeric DNA-PNA and purely PNA beacons have been reported to provide stronger binding to templates.^{31, 32} Molecular beacons have the ability to detect targets in living cells. Fluorescence resonance energy transfer (FRET) has been studied in a DNA beacon.³⁰ FRET was demonstrated between fluorophores of two beacons when they hybridized to adjacent regions on a target or between two dyes coupling at the ends of a similar beacon in its closed form.^{28, 29}

3.1.2 Fluorescence resonance energy transfer (FRET)

An assembly induced by hybridization with a beacon or fluorescently labeled protein can be studied using fluorescent methods and FRET. Fluorescence occurs when excited state electrons return to the ground state with emission of light at a longer wavelength.³³ Fluorescence is one of the most sensitive spectroscopic methods and can be used down to the single molecule level. This sensitivity is much lower than NMR, EPR, CD and many other spectroscopic techniques. The fluorescence signal provides a variety of information such as intensity, lifetime, rotational diffusion (polarization and anisotropy) and energy (wavelength). These parameters allow the study of molecular structure, environment and proximity.^{34, 35} Moreover, fluorescence is a nondestructive phenomenon allowing kinetic studies by measuring the signal changes as a function of time.³⁶ Excitation energy is sometimes transferred from one fluorophore to another

[Pick the date]

through resonance energy transfer.^{37, 38} The efficiency of energy transfer depends on the distance between two fluorophores (1-10nm), the relative orientation of the transition dipoles and overlap of their spectra.^{39, 40}

FRET is often observed via reduction of donor emission intensity or lifetime and sensitized acceptor emission. Measuring acceptor intensity increase is more difficult because of the overlap of donor emission spectra. A review of FRET methods with an extensive discussion is available in Jares-Erijman and Jovin.^{35, 41}

FRET is most often implemented in donor-acceptor hetero-systems. When donor and acceptor are the same, FRET still happens and is called homo-FRET. In this case, the best way to assess energy transfer is by reduction of the fluorescence anisotropy of the emitted light.⁴²⁻⁴⁴ Due to the nonradiative nature of FRET, it is a powerful tool to study protein dynamics and protein-protein interaction in physiological conditions. FRET measurements are now being performed *in vivo* using fluorescent proteins fused to other proteins or domains. The development of analytical methods and instrumentation is on-going particularly in the case of homo-FRET. FRET can be applied to measure the aggregation of plasma membrane proteins in living cells. Moreover, the extent of FRET (and its reversibility) can be assessed as a function of time, thus revealing the dynamics of donor and acceptor proximity.^{41, 42, 45, 46}

3.1.3 Fluorescence anisotropy

Aggregation of like molecules, such as might be expected from induced protein assembly, leaves a FRET signature which can be assessed with fluorescent anisotropy measurements. Normally, anisotropy measurements provide information on the shape and size of proteins or the rigidity of the different molecular environments.³³ The basis of anisotropy measurements is photoselective excitation of fluorophores by polarized light. Preferentially, photons whose electric vectors are aligned parallel to the transition moment of the fluorophore are absorbed. In the absence of motion or homo-FRET, when fluorophores are excited with the polarized light they will emit light which is polarized with respect to the polarization of the incoming beam.

[Pick the date]

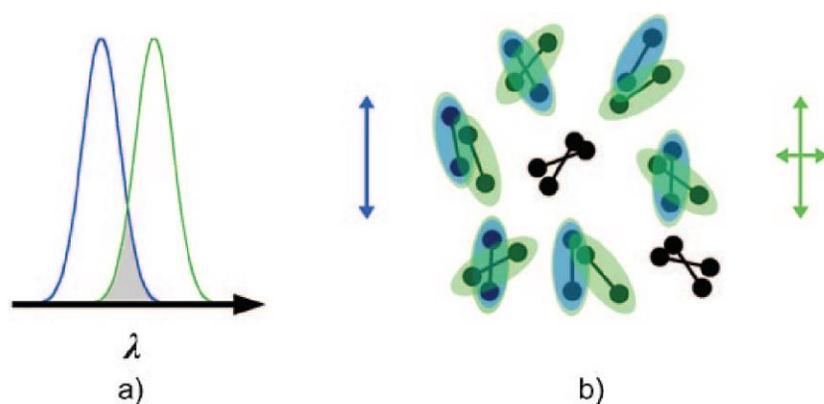
The maximum fluorescence anisotropy (r) is determined by the relative angles between excitation and emission dipole moments of the molecule. Fluorescence anisotropy is measured by assessing the intensity of fluorescence through polarizers oriented parallel ($I_{||}$) and perpendicular (I_{\perp}) to the excitation light.³³

$$r = \frac{I_{||} - I_{\perp}}{I_{||} + 2I_{\perp}} \quad [3.1]$$

Due to the fact that molecules are moving in a solution, the polarized emission light direction is different from the excited light. A large protein molecule moves slowly in the solution and if another molecule binds to that protein a more stable complex is generated which moves even slower and further increases the polarization of the emission light.³³ Another reason to increase anisotropy of a donor is hetero-FRET which happens between two different fluorophores. In this event, the time that the donor spends in the excited state is decreased relative to its rotational motion resulting in an increase of donor anisotropy.⁴⁷

When homo-FRET occurs, the fluorescence emission of randomly oriented dyes becomes depolarized relative to the linearly polarized excitation light. Due to the reversible energy transfer among similar dyes in homo-FRET, the fluorescence emission is not completely depolarized since it contains contributions from both acceptor and donor (Scheme 3.3). The total reduction of anisotropy in homo-FRET depends on the number of participating dyes and their relative orientation. This effect may enable cluster size quantification by measuring the degree of depolarization.^{48, 49}

[Pick the date]



Scheme 3.3: Graphical depiction of homo-FRET principle. a) The overlap between the absorption and emission spectra of a single type of fluorophore results in homo-FRET. b) In randomly oriented dimeric forms, the same fluorophore acts as both donor and acceptor in homo-FRET resulting in the reversible energy transfer which give rise to the donor emission with a polarization parallel to the original excitation light polarization direction. Therefore, when homo-FRET occurs the final emitted fluorescence is not completely depolarized since it consists of both direct donor emission and sensitized acceptor emission. Homo-FRET depolarization increases with the number of participating fluorophores (the left arrow is related to the polarized excitation light direction and the right vertical and horizontal arrows are respectively related to I_{\parallel} and I_{\perp} (eqn 3.1)).⁴⁸

3.1.4 Fluorescence lifetime

Detecting FRET through fluorescence intensity methods leads to uncertain interpretation such as detector noise effects, autofluorescence, photobleaching and spectral bleedthrough. As a result fluorescence lifetime methods are preferred.⁵⁰ Fluorescence lifetime can be assessed by the complementary time domain and frequency domain techniques. In the time domain, a fluorophore is excited with a series of short light pulses and the resulting fluorescence emission is recorded as a function of time.⁵¹ In the frequency domain technique, the fluorophore is excited with a sinusoidally modulated light source and the lifetime measurement relies on assessing the phase shift and modulation depth of the fluorescence emission. The modulation depth and phase shift of the emission light can be extracted from intensity images using Fourier methods or sinusoidal fitting procedures.⁵²⁻⁵⁵

Polar coordinate methods (AB-plots, Phasors, etc.) aid in the analysis and visualization of frequency domain lifetime data in terms of two and three component models.⁵⁶ Via this method, the phase and modulation become linear functions of species population. This has led to the development of global analysis algorithms for frequency domain FLIM (Fluorescence lifetime imaging microscopy) data.^{44, 53, 55, 57-59}

3.2 Results and Discussion

3.2.1 Size exclusion HPLC analysis of mTFP-PNA assembly on DNA scaffolds

To test the capability of model system to assemble mTFP-PNAs, a DNA beacon (Scheme 3.1), DNA1 and DNA 2 (Scheme 3.2) containing two and four complementary parts for mTFP-PNA monomer were chosen as frameworks for directed assembly.

SDS-PAGE analysis of mTFP-PNA assemblies on DNA1 and DNA2 was performed as described in section 3.5. Since the samples were heated before applying the sample to the gel in SDS-PAGE analysis, part of the mTFP-PNA:DNA complex became unhybridized. As a result, three distinct bands at 29, 33 and 62 kDa were observed in SDS-PAGE analysis of 2:1 ratio of mTFP-PNA:DNA1 due to the presence of mTFP-PNA monomer, monomer:DNA1 and dimer:DNA1, respectively (Fig 3.1).

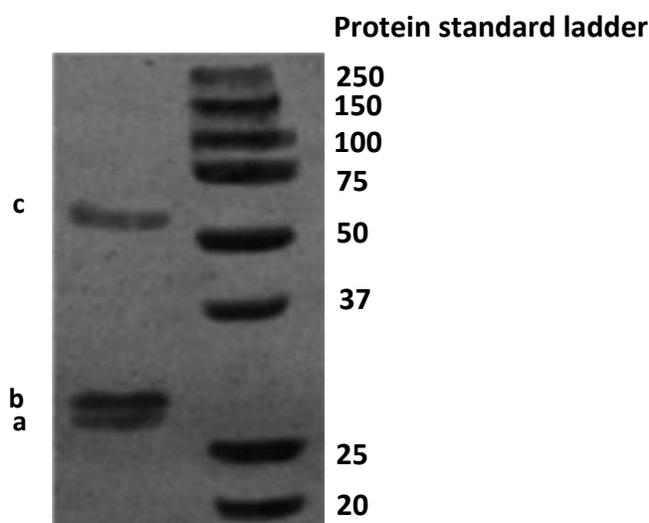


Figure 3.1: SDS-PAGE results for mTFP-PNA dimer. The bands at 29, 33, and 62 kDa, respectively, related to : a) monomer, b) DNA1: monomer, c) DNA1: dimer(left) and standard protein ladder (right).

[Pick the date]

Similarly, five distinct bands at 29, 40, 70, 99 and 129 kD, related to mTFP-PNA monomer, monomer:DNA2, dimer:DNA2, trimer:DNA2 and tetramer:DNA2 were observed in solution featuring 4:1 ratio of mTFP-PNA: DNA2 indicating oligomer assembly (Fig 3.2).

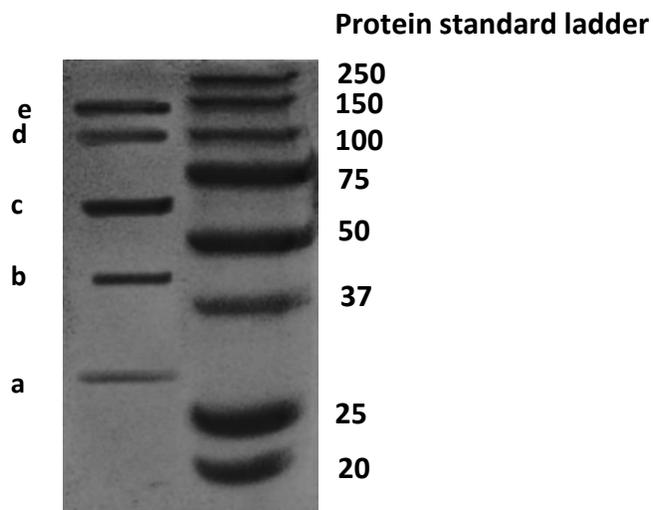


Figure 3.2: SDS-PAGE results for mTFP-PNA tetramer. The bands at 29, 40, 70, 99 and 129 kDa respectively related to: a) monomer, b) monomer:DNA2, c) dimer:DNA2, d) trimer:DNA2 and e) tetramer:DNA2 (left) and standard protein ladder (right).

The SEC-HPLC analysis was calibrated against a molecular weight standard (Fig 3.3). A single peak (8.7min) was observed for mTFP-PNA at 462 nm (maximum absorption of mTFP) and 214 nm (general absorption of all components), in line with the expected molecular mass of a monomeric construct. After adding the DNA beacon to mTFP-PNA solution, a single peak was observed at 492 nm (maximum absorption of 6-FAM) at the same elution time for monomer, showing the hybridization of DNA beacon to mTFP-PNA.

[Pick the date]

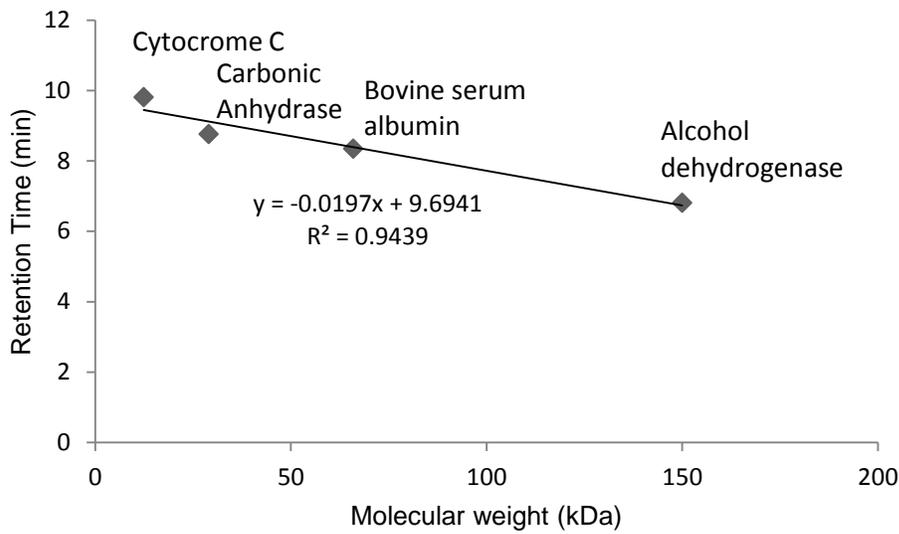


Figure 3.3: SEC-HPLC calibration curve based on the protein standard marker containing Cytochrome c (12.4 kDa), Carbonic Anhydrase (29 kDa), Bovine serum albumin (66 kDa), and Alcohol dehydrogenase (150 kDa).

Efficient formation of assembled mTFP-PNA dimer on DNA1 scaffold was confirmed by the appearance of a new single peak at 8.3 min (predicted to be 8.4 min based on SEC-HPLC calibration curve, Fig 3.3) in the solution of a 2:1 ratio of mTFP-PNA:DNA1 (Fig 3.4).

[Pick the date]

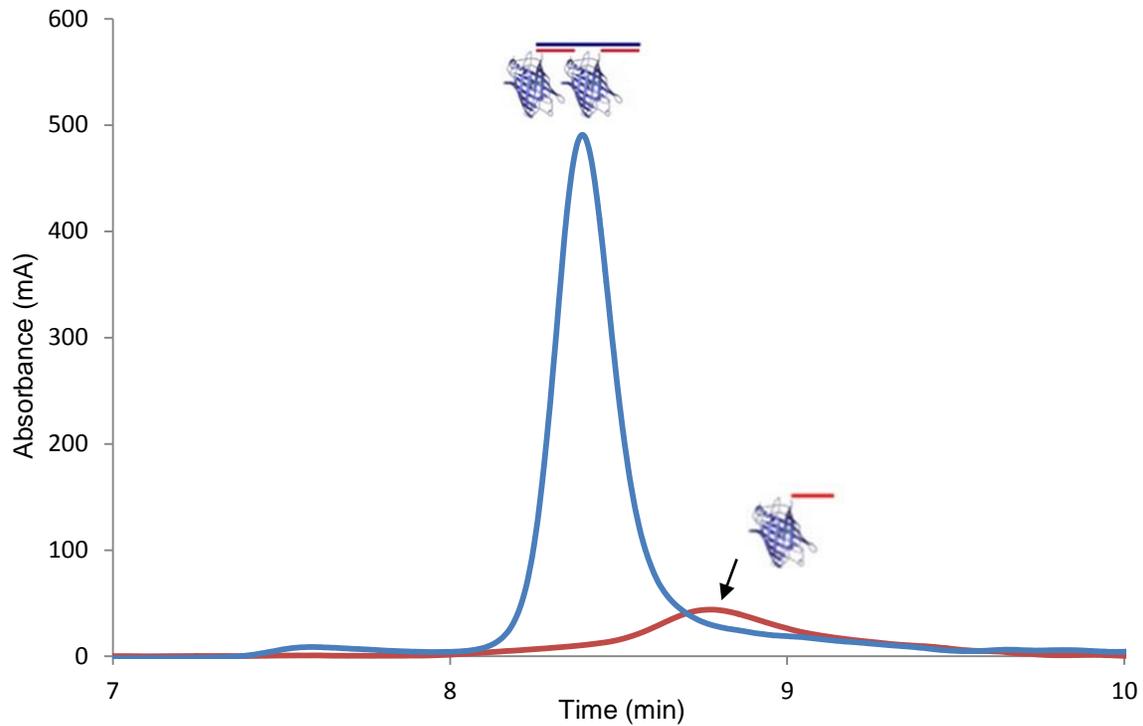


Figure 3.4: SEC-HPLC chromatograms of mTFP-PNA (red line, peak at 8.7 min) and mTFP-PNA-DNA1 (blue line, peak at 8.3 min related to DNA1:dimer).

A chromatogram of the solution featuring a 4:1 ratio of mTFP-PNA:DNA2 showed an additional peak at 7.5 min (predicted to be 7.2 min based on SEC-HPLC calibration curve, Fig 3.3) related to the formation of an assembled mTFP-PNA tetramer (Fig 3.5).

[Pick the date]

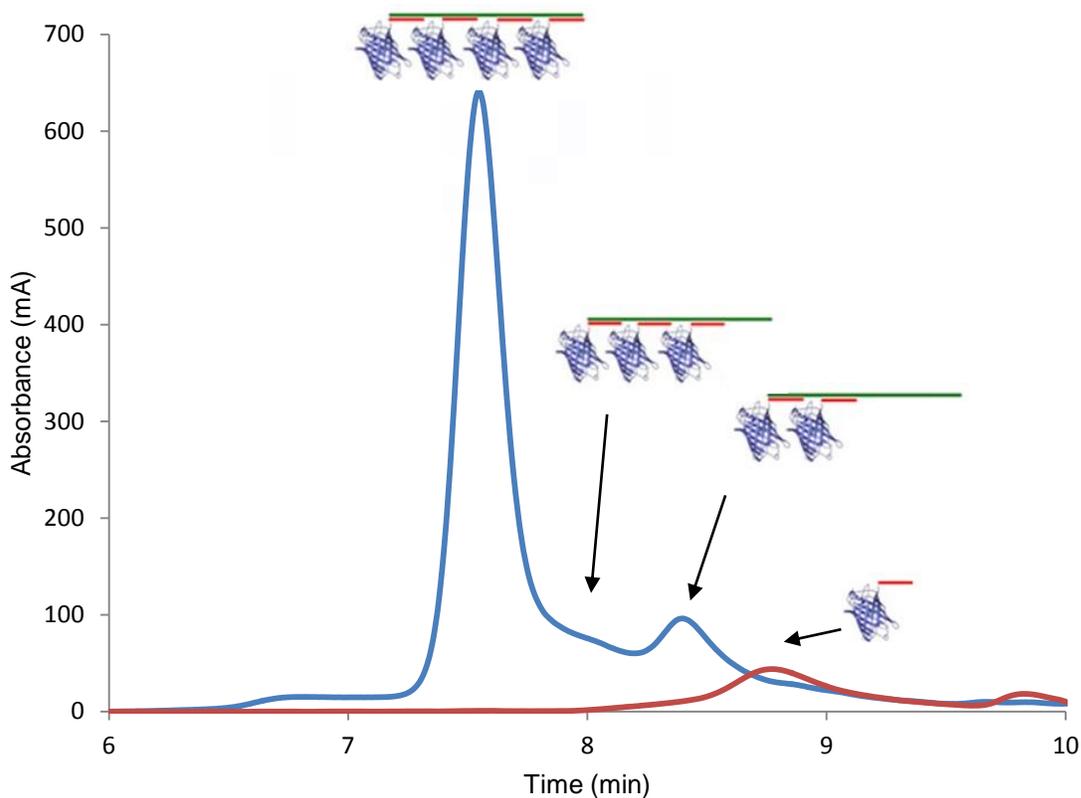


Figure 3.5: SEC-HPLC chromatograms of mTFP-PNA (red line, peak at 8.7 min) and mTFP-PNA:DNA2 (blue line, peak at 7.5 and 8.4 min related to tetramer:DNA2 and dimer:DNA2 , respectively.) .

A gradual decrease of the mTFP-PNA monomer (8.7 min) was observed on titration of DNA2 with increasing amounts of mTFP-PNA monomer (Fig 3.6). At a 4:1 ratio of mTFP-PNA to DNA2, the monomer peak reached zero with a concomitant increase in assembled protein tetramer (7.5 min). The complete disappearance of the monomeric form of mTFP-PNA upon addition of DNA shows that the assembly proceeded very efficiently. Also, the peak related to unbound DNA2 (12 min) could be observed after addition of excess DNA2 to the solution.

[Pick the date]

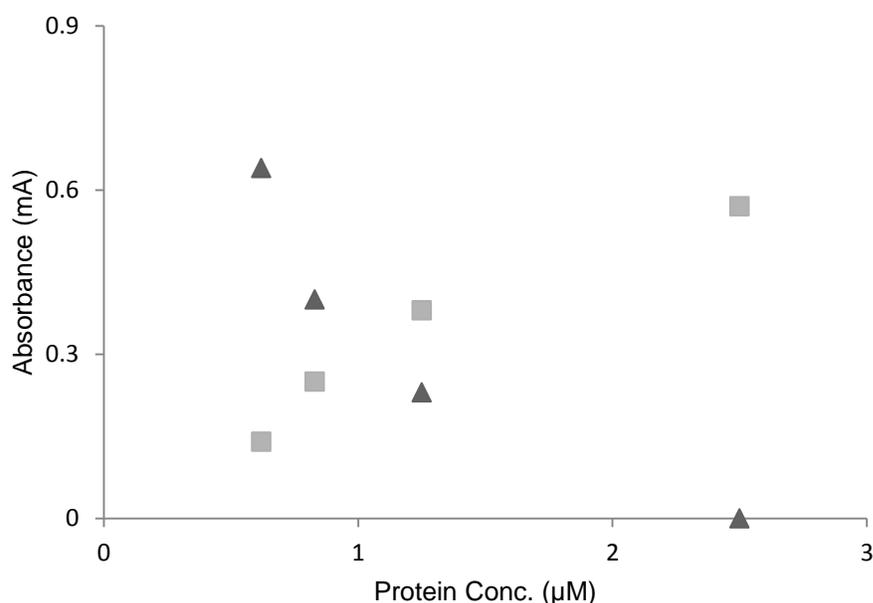


Figure 3.6: SEC analysis of showing the increase of the tetramer peak (7.5 min ■) and decrease of the monomer peak (8.7 min ▲) monitored at 462 nm while titrating DNA2 (0.62 μM) with mTFP-PNA (0.62-2.5 μM).

3.2.2 Hetero-FRET occurrence after hybridization of PNA-mTFP to DNA beacon

A hetero-FRET system was developed based on the hybridization of mTFP-PNA to a DNA beacon (Scheme 3.1). Due to quenching of 6-FAM by Dabcyl group, no fluorescence was observed from the closed form of the beacon. The beacon changed its conformation to the open form and emitted at 521 nm upon addition of mTFP-PNA or PNA2 (CAGTCA), each of which is complementary to the DNA beacon loop sequence (5'-TGATGGTCAGT-3'). Three phenomena indicative of FRET were observed, including: 1) quenched donor emission; 2) frequency domain lifetime in the acceptor region exhibiting τ_{ϕ} greater than τ_m ; and 3) increased donor anisotropy.

3.2.2.1 Intensity measurements

DNA beacon (0-27 μM) was added to 2.5 μM of the mTFP-PNA solution. Due to energy transfer to the acceptors in the beacon (Fig 3.7), donor (mTFP-PNA) emission decreased by addition of DNA beacon.

[Pick the date]

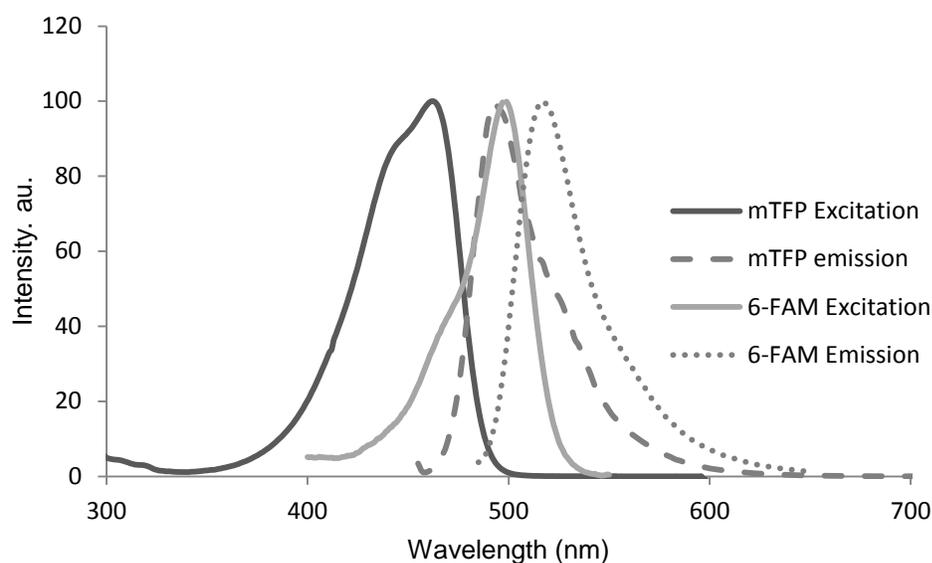


Figure 3.7: Spectral overlap between mTFP emission and 6-FAM excitation as a FRET pair: solid black and grey peaks are related to normalized excitation and emission of mTFP (maximum at 462 and 492 nm, respectively), dashed and dotted peaks is related to normalized excitation and emission of 6-FAM, (maximum at 494 and 521 nm, respectively) (spectra for 6-FAM from Life Technologies).

The mTFP-PNA:DNA beacon assembly showed a 59 ± 4 percent energy transfer over the range of 490-510 nm relative to a control sample of unligated mTFP:DNA beacon in which the estimated distance between 6-FAM and mTFP chromophore was approximately about 4.4 ± 0.1 nm. The control sample showed up to 40% intensity decrease (Fig 3.8). This decrease can be related to the inner filter effect³³ arising from highly colored solution seen with the highest beacon concentrations and quenching by the Dabcyl group on mTFP or FAM (in 7 nm distance) in open form (R_0 : 4.7 nm⁴²).

No inner filter effects were observed with dilute solutions (e.g.: 2.7 μ M DNA beacon). The correction of energy transfer efficiencies (eqn 3.3, Experimental part) has been done by computing relative to the control solutions of mTFP:beacon.

[Pick the date]

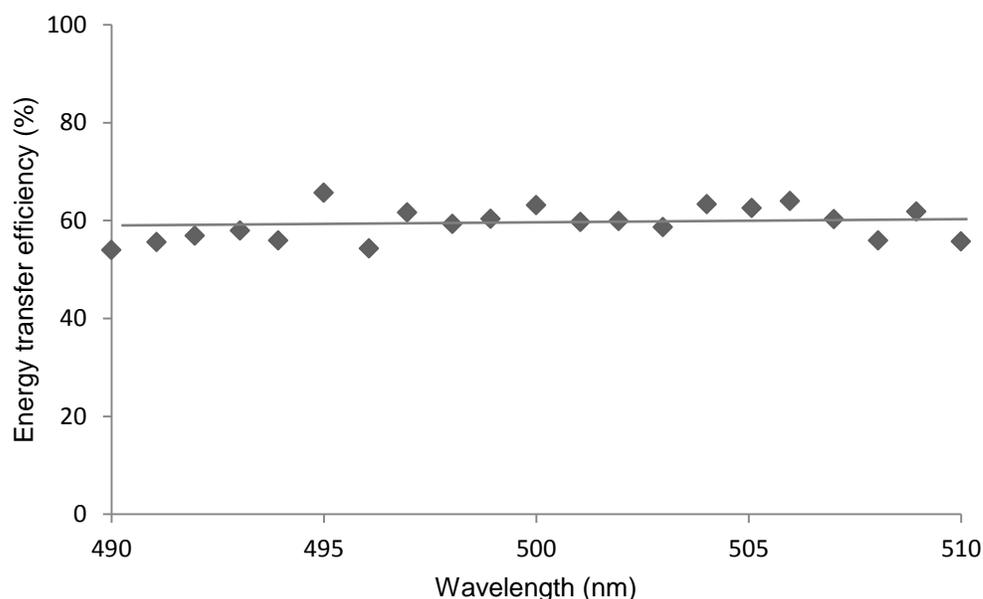


Figure 3.8: Energy transfer efficiency (59 ± 4 percent across 490-510 nm wavelength range) between donor (mTFP) and acceptor (6-FAM) in the hetero-FRET system produced after hybridization of mTFP-PNA to DNA beacon obtained by intensity measurements.

3.2.2.2 Lifetime measurements

The mTFP-PNA:Beacon:PNA2 system exhibited changes in frequency domain lifetime characteristics upon assembly (Fig 3.5). In the absence of DNA Beacon, mTFP-PNA exhibited a single component lifetime of 3.3 ns. Under conditions of high excess Beacon:PNA2, the system was dominated by free Beacon:PNA2 (2.4 ns). At near stoichiometric conditions, the acceptor 6-FAM exhibited lifetime heterogeneity with τ_{ϕ} (3.1 ns) greater than τ_m (2.5 ns). This behaviour is characteristic of FRET and results in the appearances of points outside the semi-circle of single component lifetimes in a polar plot (Fig 3.9).^{54, 60} These data are consistent with partial unfolding of the Beacon in the presence of PNA2 and with the production of an assembled system of mTFP-PNA:Beacon:PNA2 in which there is energy transfer from mTFP to 6-FAM.

[Pick the date]

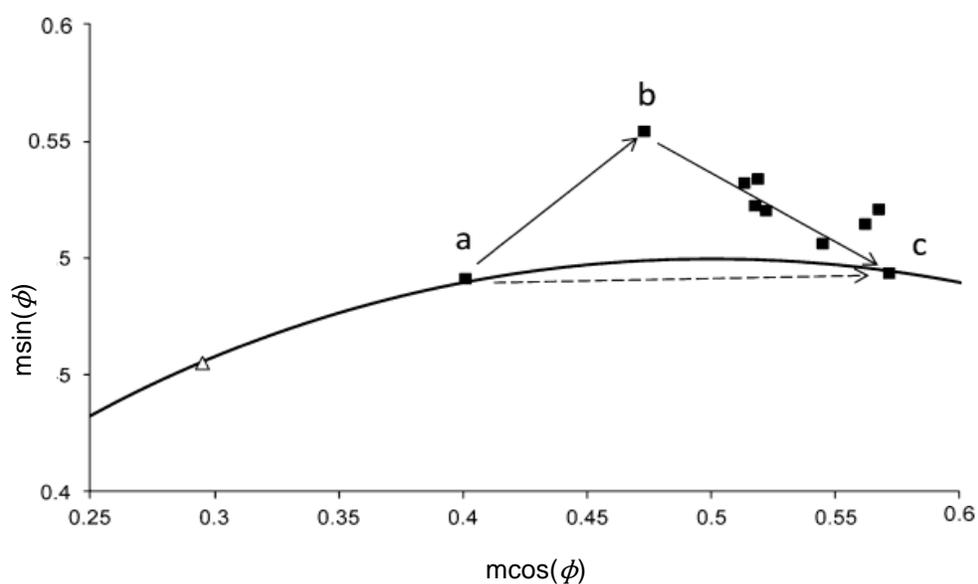


Figure 3.9: Polar co-ordinate presentation of frequency domain lifetime measurements from mTFP-PNA: Beacon: PNA2 complexes. The data correspond to measurements dominated by 6-FAM (■). The position of the rhodamine 6G standard is also shown (△). Position (a) indicates the position of mTFP-PNA in the absence of added Beacon:PNA2 and corresponds 3.3 ns. Position (b) represents the position of a mixture of mTFP-PNA (2.5 μM) and Beacon:PNA2 (2.7 μM). As the concentration of Beacon:PNA2 increases relative to mTFP-PNA the system follows the trajectory from (b) to (c). At point (c) the signal is dominated by Beacon:PNA2 (2.4 ns). The dotted arrow represents the expected path of a non-interacting mixture of mTFP-PNA and Beacon:PNA2.

3.2.2.3 Anisotropy measurement of mTFP-PNA: DNA beacon assembly

In the developed hetero-FRET system, the modification and assembly of mTFP-PNA resulted in a slight increase in fluorescence anisotropy of mTFP. The initial anisotropy of mTFP (0.32) increased to 0.33 after ligation to PNA. A further increase (0.35) was observed after subsequent hybridization to the DNA beacon. The correlation times (eqn 3.8, Experimental part) calculated for mTFP-PNA1 (θ_1) and mTFP-PNA1:DNA (27 μM DNA) (θ_2) were 18 and 25 ns respectively, using experimentally determined average lifetimes and $r_0=0.39$ from previous studies of fluorescent proteins. These estimated rotational correlation times are in the typical range of monomeric fluorescent FPs in buffer^{33, 52}.

The 36% increase of θ which was observed when mTFP-PNA (29567 Da) formed the mTFP-PNA:Beacon assembly (37368 Da) was close to the 26% increase in mass. These results were interpreted as supporting the previous intensity and polar plot analyses which together proved the success of ligation and assembly strategy and the accessibility of assemblies to photophysical analysis.

3.2.3 Homo-FRET system demonstrating dimer and oligomer assembly

3.2.3.1 Anisotropy measurements of homo-FRET system

Anisotropy changes of the system were studied by adding a range of DNA1 and DNA2 concentration (0–2.5 μM) to 2.5 μM of mTFP-PNA. At a 1:1 ratio of mTFP-PNA to DNA1 and DNA2 (monomer:DNA), maximum anisotropy values of 0.34 (DNA1) and 0.36 (DNA2) were observed. In comparison with the observed anisotropy for free mTFP-PNA (0.32), these increases were consistent with a slight increase in the rotational correlation time of the larger assemblies and similar to the values obtained in the presence of DNA beacon.

Adding DNA1 and DNA2 scaffolds to the mTFP-PNA system resulted in a decrease of measured anisotropy indicative of homo-FRET. Titration of mTFP-PNA with DNA1 showed a gradual decrease of anisotropy in comparison with mTFP-PNA:DNA1 monomer. At a 2:1 ratio of mTFP-PNA to DNA1, a maximum decrease of 45% was observed which was close to the expected 50% reduction (eqn 3.7, Experimental part). Addition of more mTFP-PNA resulted in a gradual recovery of the anisotropy (Fig 3.10).

[Pick the date]

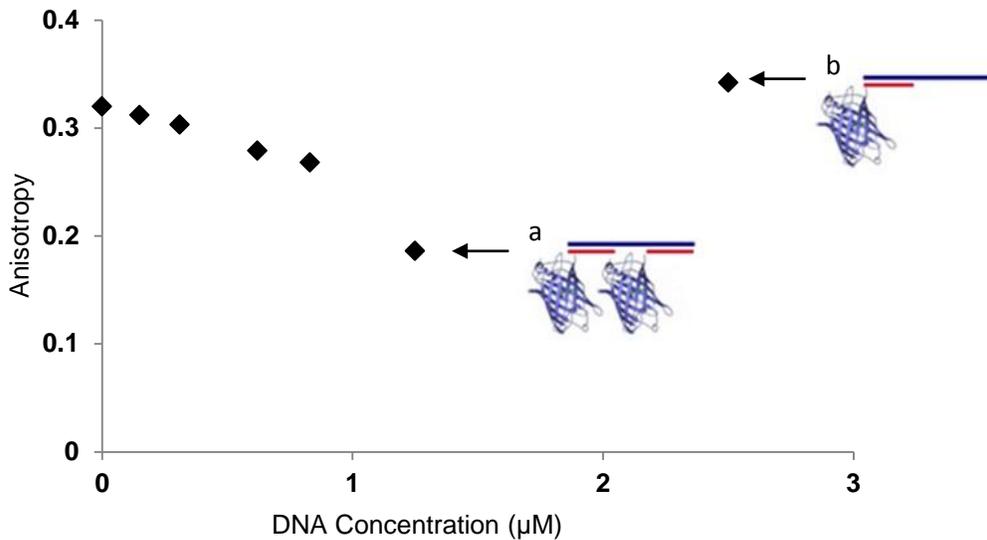


Figure 3. 10: Anisotropy changes of the DNA1-mTFP-PNA system after adding different concentration of DNA1 (0-2.5 μM) to 2.5 μM mTFP-PNA (a) 1.25 μM : DNA1-dimer and b) 2.5 μM: DNA1-monomer).

On the DNA2 scaffold (Fig 3.10), a 15% decrease in anisotropy was observed with a 2:1 ratio of mTFP-PNA:DNA2. This value is significantly less than predicted by eqn 3.7, this could be because of the expected distance range (from 3.3-10 nm) between monomer units on four complementary binding locations on DNA2 scaffold and distribution of different species in solution. The larger distances are outside $0.8 R_0^{42, 61}$ resulting in less change in anisotropy than seen in the DNA1 assembly. Similarly, at a 3:1 ratio of mTFP-PNA to DNA2 a 44% decrease was observed in the system. The lowest observed anisotropy (0.082 showing 77% decrease of anisotropy) with a 4:1 ratio mTFP-PNA to DNA2 was consistent with a tetramer. The resulting DNA directed tetramer assembly was in reasonable agreement with the predictions of eqn 3.7 in Experimental part (Fig 3.11).

[Pick the date]

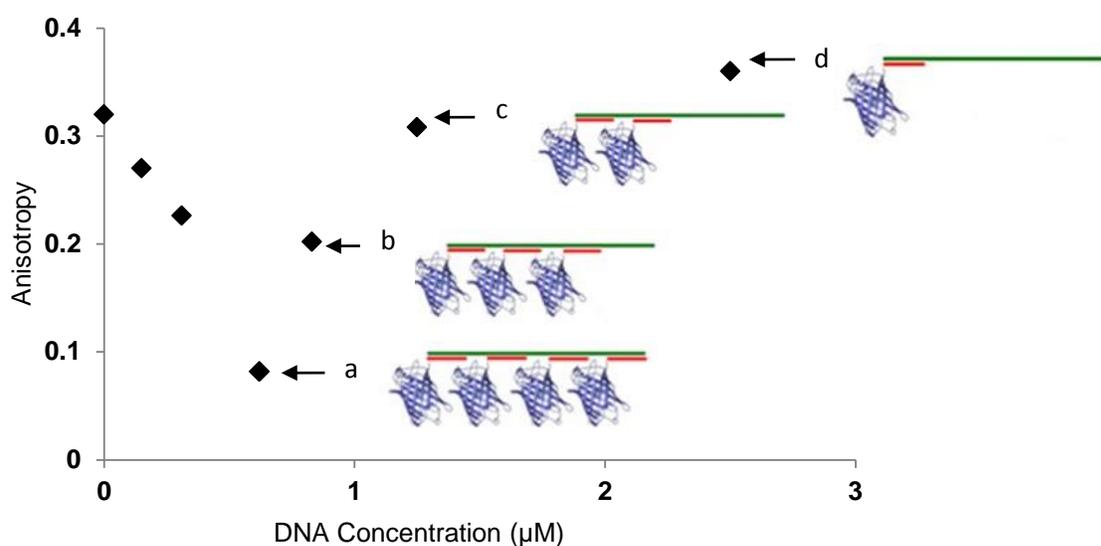


Figure 3.11: Anisotropy changes of the DNA2-mTFP-PNA system after adding different concentration of DNA2(0-2.5 μM) to 2.5 μM mTFP-PNA (a) 0.62 μM : DNA2-tetramer, b) 0.83 μM : DNA2-trimer, c) 1.25 μM : DNA2-dimer and d) 2.5 μM : DNA2-monomer). Assemblies illustrated at points a, b, and c represents the stoichiometry. The mixtures of assembled species at these locations are more complex. For example, the 2:1 species can exist in 3 forms and a randomly assembled system will include some amount of 1:1, 2:1, 3:1, and 4:1 species.

The PNA-directed assemblies developed in this chapter can provide a well-defined scaffold for adjusting the number, distance and distribution of proteins in cluster forms. The current model system can be used to induce aggregation of proteins of interest such as membrane proteins by modifying them with fluorescent proteins as protein tags through cloning methods. The signature of protein assemblies could be traced by fluorescence measurements. As described in section 3.3, hetero- and homo-FRET permitted noninvasive detection and quantification of the protein clusters by measuring the intensity, lifetime, and anisotropy of the systems. The clear and readily interpreted photophysical properties of the assembled mTFP-PNA model system can provide a practical tool to investigate the behavior of protein clusters. This may introduce a novel approach to control the nano scale spatial organization of cell-membrane proteins in living systems and to study the relevant consequences on cellular signaling mechanisms.

3.3 Conclusion and outlook

The FP-PNA units generated by NCL (Chapter 2) proved to be a good system for directed protein self-assembly demonstrating a new way to controllably induce and monitor protein assembly. The characteristics of PNA and the capability of studying this model system by FRET techniques provides a promising way to study protein assemblies in controlled conditions. The convenient formation of dimers and oligomers on DNA scaffolds was performed by the precise recognition capability of PNA to complementary DNA. The assemblies gave clear and readily interpreted photophysical signatures providing a practical tool to investigate the behavior of protein clusters. By fusing EPL enabled FPs to protein of interest, such as cell membrane proteins, control of both assembly and visualization is provided following addition of template to the PNA ligated proteins. Further work could readily extend this approach from homo dimers and oligomers to binary, ternary, and higher oligomer systems containing any number of different dyes or fluorescent proteins in precisely engineered arrangements.

3.4 Experimental part

3.4.1 Hybridization and assembly of mTFP-PNA with DNA

A DNA beacon (6FAM5`ACAGCTGCATGGTCAGTGCTGT3`Dabcyl) (The Midland Certified Reagent Company, Inc, USA) was used to assemble mTFP-PNA. In the same way, two DNA sequences (DNA1:5`TGCATGGATCTGCATG3`) and (DNA2:5`TGCATGGATCTGCATGGATCTGCATGGATCTGCATG3`) (Life technologies corporation, USA) were used to make assembled mTFP-PNA dimers and tetramers, respectively. The titration of 1.25 μ M mTFP-PNA with (1.25-2.5 μ M) of DNA1, and 0.62 μ M mTFP-PNA with 0.62-2.5 μ M DNA2 was carried out in phosphate buffer (100mM, NaCl 200 mM, pH 7) with 2 hrs incubation at room temperature. The results were assessed by SDS-PAGE electrophoresis and SEC-HPLC at (214 nm, 462 and 492 nm using phosphate buffer (150 mM, pH 7) as mobile phase). The SEC-HPLC was carried out on a SRT SEC-150, 5 μ m, 4.6x300mm; Chromex Scientific, UK) calibrated with molecular weight protein marker kit 12-200 kDa (sigma).

3.4.2 Photophysical measurements

A range of a DNA beacon concentrations (0-27 μM) were added to 2.5 μM mTFP-PNA solution and PNA2 (CAGTCA) (Advanced Peptide Inc., USA). Similarly, varying concentration of DNA1 and DNA2 (0-2.5 μM) were added to 2.5 μM mTFP-PNA and the emission intensity and anisotropy recorded between 480-530 nm (excitation at 462 nm) using a fluorimeter equipped with polarisers (Cary Eclipse; Varian). Energy transfer efficiency (E) can be measured based on eqn 3.2. R_0 is the Förster distance in which the transfer efficiency E is 50% and r is the distance between fluorophores.³³

$$E = \frac{R_0^6}{R_0^6 + r^6} \quad [3.2]$$

In a hetero-FRET system, the energy transfer efficiency can be measured by lifetimes or intensity,

$$E = 1 - \frac{\tau_2}{\tau_1}, \quad E = 1 - \frac{I_2}{I_1} \quad [3.3]$$

Where τ_1 and τ_2 are respectively the lifetimes of donor in the absence and presence of the acceptor. Similarly, I_1 and I_2 are the intensities of the donor in the absence and presence of acceptor. Fluorescence lifetimes measurement were performed using a spectroscopic fluorescence lifetime imaging system operating in the frequency domain⁶² with rhodamine 6G as a lifetime reference.

In case of frequency domain lifetime measurements, the measured phase, ϕ_i , and modulation, m_i , may be presented in polar co-ordinates.

[Pick the date]

$$x_i = m_i \cos(\phi_i) \quad [3.4]$$

$$y_i = m_i \sin(\phi_i) \quad [3.5]$$

In polar co-ordinate analysis, the following equation gives the estimated fraction of mTFP-PNA in the hybridized form (α):

$$\alpha = \sqrt{\frac{(x_1 - x_i)^2 + (y_1 - y_i)^2}{(x_1 - x_2)^2 + (y_1 - y_2)^2}} \quad [3.6]$$

Where (x_1, y_1) , (x_2, y_2) , and (x_i, y_i) are the coordinates in the absence of acceptor, the fully hybridized form, and mixture of free and hybridized forms, respectively.

In homo-energy transfer, lifetime does not in general change therefore the fluorescence anisotropy can be used as a useful indicator of assembly. To a first approximation, the anisotropy can be estimated based on eqn 3.7 in which $r_{monomer}$ and $r_{oligomer}$ are the anisotropy of a fluorophore monomer and the anisotropy of oligomer of fluorophores undergoing homo-FRET, respectively, and n is the number of fluorophores in the oligomer.

$$r_{oligomer} \cong \frac{r_{monomer}}{n} \quad [3.7]$$

This equation can be used when the inter-fluorophore distance is $< 0.8 R_0$. In more complicated systems, treatment of homo-FRET is more difficult since it can be affected by the cluster size, the orientation between fluorophores, and inter-fluorophore distances.^{61, 63}

An increase of molecular mass competes with homo-FRET induced depolarisation. This phenomenon can be estimated based on the Perrin equation:^{33, 64}

[Pick the date]

$$r = \frac{r_0}{1 + \frac{\tau}{\theta}} \quad [3.8]$$

Where r , r_0 , τ , and θ are measured and initial anisotropy of the system, fluorescence lifetime, and the rotational correlation time, respectively. The following equation indicates the dependency of rotational correlation time on viscosity, η , the molecular weight, M , the specific volume and hydration, $\bar{v} + h$, the gas constant, R , and the temperature, T .

$$\theta = \frac{\eta M}{RT} (\bar{v} + h) \quad [3.9]$$

Assuming all other parameters are equal (temperature, viscosity, etc.), the rotational correlation time scales with the mass showing that as mass increases, the measured r will approach r_0 .

$$\frac{\theta_1}{\theta_2} \cong \frac{M_1}{M_2} \quad [3.10]$$

3.5 References

1. R. J. Conrado, G. C. Wu, J. T. Boock, H. Xu, S. Y. Chen, T. Lebar, J. Turnsek, N. Tomsic, M. Avbelj, R. Gaber, T. Koprivnjak, J. Mori, V. Glavnik, I. Vovk, M. Benčina, V. Hodnik, G. Anderluh, J. E. Dueber, R. Jerala and M. P. DeLisa, *Nucleic Acids Research*, 2012, **40**, 1879-1889.
2. S. Le Gac, E. Schwartz, M. Koepf, J. J. L. M. Cornelissen, A. E. Rowan and R. J. M. Nolte, *Chemistry – A European Journal*, 2010, **16**, 6176-6186.
3. B. R. Griffith, B. L. Allen, A. C. Rapraeger and L. L. Kiessling, *Journal of the American Chemical Society*, 2004, **126**, 1608-1609.
4. C. J. Delebecque, A. B. Lindner, P. A. Silver and F. A. Aldaye, *Science*, **333**, 470-474.
5. D. A. Uhlenheuer, L.-G. Milroy, P. Neiryck and L. Brunsveld, *Journal of Materials Chemistry*, 2011, **21**, 18919-18922.
6. D. A. Uhlenheuer, D. Wasserberg, C. Haase, H. D. Nguyen, J. H. Schenkel, J. Huskens, B. J. Ravoo, P. Jonkheijm and L. Brunsveld, *Chemistry - A European Journal*, 2012, **18**, 6788-6794.
7. H. Li, S. H. Park, J. H. Reif, T. H. LaBean and H. Yan, *Journal of the American Chemical Society*, 2003, **126**, 418-419.
8. A. Gangar, A. Fegan, S. C. Kumarapperuma and C. R. Wagner, *Journal of the American Chemical Society*, 2012, **134**, 2895-2897.
9. R. Sharma, A. G. Davies and C. Walti, *Nanotechnology*, 2012, **23**, 365301.
10. S. H. Park, P. Yin, Y. Liu, J. H. Reif, T. H. LaBean and H. Yan, *Nano Letters*, 2005, **5**, 729-733.
11. M. Bouvier, *Nature reviews. Neuroscience*, 2001, **2**, 274-286.

12. J. C. Sacchettini, L. G. Baum and C. F. Brewer, *Biochemistry*, 2001, **40**, 3009-3015.
13. S. Li, X. Zhang and W. Wang, *Biophysical Journal*, 2010, **98**, 2554-2563.
14. M. Ma and D. Bong, *Organic & Biomolecular Chemistry*, 2011, **9**, 7296-7299.
15. Z. L. Pianowski and N. Winssinger, *Chemical Society Reviews*, 2008, **37**, 1330-1336.
16. K. K. Sadhu, M. Röthlingshöfer and N. Winssinger, *Israel Journal of Chemistry*, 2013, **53**, 75-86.
17. J. L. Harris and N. Winssinger, *Chemistry – A European Journal*, 2005, **11**, 6792-6801.
18. J. P. Dagher, M. Ciobanu, S. Alvarez, S. Barluenga and N. Winssinger, *Chemical Science*, 2011, **2**, 625-632.
19. K.-T. Huang, K. Gorska, S. Alvarez, S. Barluenga and N. Winssinger, *ChemBioChem*, 2011, **12**, 56-60.
20. H. D. Urbina, F. Debaene, B. Jost, C. Bole-Feysot, D. E. Mason, P. Kuzmic, J. L. Harris and N. Winssinger, *ChemBioChem*, 2006, **7**, 1790-1797.
21. M. Ciobanu, K.-T. Huang, J.-P. Dagher, S. Barluenga, O. Chaloin, E. Schaeffer, C. G. Mueller, D. A. Mitchell and N. Winssinger, *Chemical Communications*, 2011, **47**, 9321-9323.
22. S. A. Kazane, J. Y. Axup, C. H. Kim, M. Ciobanu, E. D. Wold, S. Barluenga, B. A. Hutchins, P. G. Schultz, N. Winssinger and V. V. Smider, *Journal of the American Chemical Society*, 2013, **135**, 340-346.
23. K. Gorska, J. Beyrath, S. Fournel, G. Guichard and N. Winssinger, *Chemical Communications*, 2010, **46**, 7742-7744.
24. K. Gorska, K.-T. Huang, O. Chaloin and N. Winssinger, *Angewandte Chemie International Edition*, 2009, **48**, 7695-7700.
25. C. Scheibe, S. Wedepohl, S. B. Riese, J. Dervedde and O. Seitz, *ChemBioChem*, 2013, **14**, 236-250.
26. A. S. Harding and J. F. Hancock, *Trends in Cell Biology*, 2008, **18**, 364-371.
27. J. Ichinose, M. Murata, T. Yanagida and Y. Sako, *Biochemical and Biophysical Research Communications*, 2004, **324**, 1143-1149.
28. T. J. Drake and W. Tan, *Applied Spectroscopy*, 2004, **58**, 269A-280A.
29. A. Marx and O. Seitz, *Molecular Beacons: Signalling Nucleic Acid Probes, Methods, and Protocols*, Humana Press Inc., USA, 2008.
30. S. Jockusch, A. A. Marti, N. J. Turro, Z. Li, X. Li, J. Ju, N. Stevens and D. L. Akins, *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology*, 2006, **5**, 493-498.
31. H. Kuhn, V. V. Demidov, J. M. Coull, M. J. Fiandaca, B. D. Gildea and M. D. Frank-Kamenetskii, *Journal of the American Chemical Society*, 2002, **124**, 1097-1103.
32. H. Kuhn, V. V. Demidov, B. D. Gildea, M. J. Fiandaca, J. C. Coull and M. D. Frank-Kamenetskii, *Antisense & Nucleic Acid Drug Development*, 2001, **11**, 265-270.
33. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, USA, 2006.
34. M. J. Roberti, L. Giordano, T. M. Jovin and E. A. Jares-Erijman, *ChemPhysChem : a European Journal of Chemical Physics and Physical Chemistry*, 2011, **12**, 563-566.
35. E. A. Jares-Erijman and T. M. Jovin, *Nature Biotechnology*, 2003, **21**, 1387-1395.
36. A. E. Johnson, *Traffic (Copenhagen, Denmark)*, 2005, **6**, 1078-1092.
37. T. Förster, *Annalen der Physik*, 1948, **437**, 55-75.
38. J. R. Oppenheimer, *Physic.Rev.*, 1941, **60**.
39. Y. Sun, H. Wallrabe, S. A. Seo and A. Periasamy, *ChemPhysChem : a European Journal of Chemical Physics and Physical Chemistry*, 2011, **12**, 462-474.
40. D. W. Piston and G. J. Kremers, *Trends in Biochemical Sciences*, 2007, **32**, 407-414.
41. E. A. Jares-Erijman and T. M. Jovin, *Analytical Techniques / Mechanisms*, 2006, **10**, 409-416.
42. A. N. Bader, E. G. Hofman, J. Voortman, P. M. en Henegouwen and H. C. Gerritsen, *Biophysical Journal*, 2009, **97**, 2613-2622.
43. Y. Yan and G. Marriott, *Current Opinion in Chemical Biology*, 2003, **7**, 635-640.
44. A. H. Clayton, Q. S. Hanley, D. J. Arndt-Jovin, V. Subramaniam and T. M. Jovin, *Biophysical Journal*, 2002, **83**, 1631-1649.
45. Y. Chen, J. P. Mauldin, R. N. Day and A. Periasamy, *Journal of Microscopy*, 2007, **228**, 139-152.
46. B. A. Pollok and R. Heim, *Trends in Cell Biology*, 1999, **9**, 57-60.
47. M. Heidecker, Y. Yan-Marriott and G. Marriott, *Biochemistry*, 1995, **34**, 11017-11025.
48. F. T. S. Chan, C. F. Kaminski and G. S. Kaminski Schierle, *ChemPhysChem*, 2011, **12**, 500-509.

[Pick the date]

49. A. N. Bader, S. Hoetzl, E. G. Hofman, J. Voortman, P. M. P. van Bergen en Henegouwen, G. van Meer and H. C. Gerritsen, *ChemPhysChem*, 2011, **12**, 475-483.
50. R. B. Sekar and A. Periasamy, *The Journal of Cell Biology*, 2003, **160**, 629-633.
51. A. Leray, F. B. Riquet, E. Richard, C. Spriet, D. Trinel and L. Heliot, *Microscopy Research and Technique*, 2009, **72**, 371-379.
52. V. Subramaniam, Q. S. Hanley, A. H. Clayton and T. M. Jovin, *Methods in Enzymology*, 2003, **360**, 178-201.
53. Q. S. Hanley and A. H. Clayton, *Journal of Microscopy*, 2005, **218**, 62-67.
54. T. S. Forde and Q. S. Hanley, *Photochemical & Photobiological Sciences*, 2005, **4**, 609-616.
55. A. H. Clayton, Q. S. Hanley and P. J. Verveer, *Journal of Microscopy*, 2004, **213**, 1-5.
56. Q. S. Hanley, *Journal of The Royal Society Interface*, 2009, **6**, S83-S92.
57. G. Redford and R. Clegg, *Journal of Fluorescence*, 2005, **15**, 805-815.
58. R. A. Colyer, C. Lee and E. Gratton, *Microscopy Research and Technique*, 2008, **71**, 201-213.
59. M. A. Digman, V. R. Caiolfa, M. Zamai and E. Gratton, *Biophysical Journal*, 2008, **94**, L14-L16.
60. M. Štefl, N. G. James, J. A. Ross and D. M. Jameson, *Anal. Biochem.*, 2011, **410**, 62-69.
61. L. W. Runnels and S. F. Scarlata, *Biophysical Journal*, 1995, **69**, 1569-1583.
62. Y. Zhou, J. M. Dickenson and Q. S. Hanley, *Journal of Microscopy*, 2009, **234**, 80-88.
63. I. Gautier, M. Tramier, C. Durieux, J. Coppey, R. B. Pansu, J. C. Nicolas, K. Kemnitz and M. Coppey-Moisan, *Biophysical Journal*, 2001, **80**, 3000-3008.
64. F. Perrin, *J. Phys. Radium*, 1926, **7**, 390-401.

Chapter 4

Coupling of Purified SNAP protein with fluorescent PNA

Gaining control over protein assembly holds promise for the investigation of protein aggregation. The useful properties of PNA and its sequence-specific recognition for DNA templates have been applied in previous chapters to create assembly units. Assembled units were produced via expressed protein ligation of fluorescent proteins with PNA and subsequent oligomerization induced using DNA templates to create FRET systems. Although fluorescent proteins have many unique characteristics which make them a powerful tool as imaging tags when fused to proteins, they have some limitations such as slow maturation and a tendency to form aggregates. A self-labeling moiety such as the SNAP-tag is another approach used widely to fuse to proteins of interest and study their behavior. This method has been adopted here based on ligation of PNA to SNAP protein. The ligation requires modification of PNA with O⁶-benzyl guanine which is a specific substrate for SNAP. To create FRET systems, the modified PNA has been labeled with organic fluorophore. Successful modification of PNA with O⁶-benzyl guanine and fluorophore followed by rapid and selective coupling to SNAP protein has been demonstrated in this chapter for use in later studies of assembly.

4.1 Introduction

During the last century, fundamental studies of proteins have been influenced significantly by chemical site specific labeling of protein with organic fluorophores and other biophysical probes.¹ The labeling of purified proteins based on selective reactivity with cysteine and lysine residues in living cells is not as effective as *in vitro* due to the presence of various proteins and reactive species in the cell environment.²

As mentioned in previous chapters, the application of FPs as selective, genetically encoded tags to study protein-protein interaction, protein localization and oligomerization has had major impact on live cell imaging progress.³⁻⁹ In spite of the significant role of FPs in studies of cellular processes, they have limitations. The size of FPs (238 amino acid, ~30 kDa) can interfere with the assembly, function, and dynamic behavior of proteins.⁹⁻¹¹ Although FPs with a broad range of absorption and emission spectra are available providing multicolour imaging of several proteins in a single cell^{12, 13} they are not comparable with the wide variety of organic fluorophores in terms of photostability and the brightness.¹⁰ The slow maturation of FPs and their tendency to form aggregates can restrict applications in more sophisticated biophysical studies in living cells.^{14, 15}

Recently, application of tag-mediated labeling methods has increased widely for studying protein dynamics, localization and trafficking in live cells.¹⁶⁻¹⁹ Therefore, more alternative approaches to label proteins of interest are via self-labeling and assemblies protein tags. One such approach could be via conjugation of PNA.

A 15 amino-acid tetracysteine (TC) with the consensus sequence CCXXCC (X could be any amino acid but not cysteine) was the first peptide tag for specific protein labeling.²⁰ TC binds specifically to membrane permeable bis-arsenical fluorescent compound such as FIAsh, a fluorescein derivative, or ReAsH, a resorufin derivative. The fluorescence of these dyes is switched on by binding to the TC tag.^{2, 21, 22} In similar approach, metal ion chelating

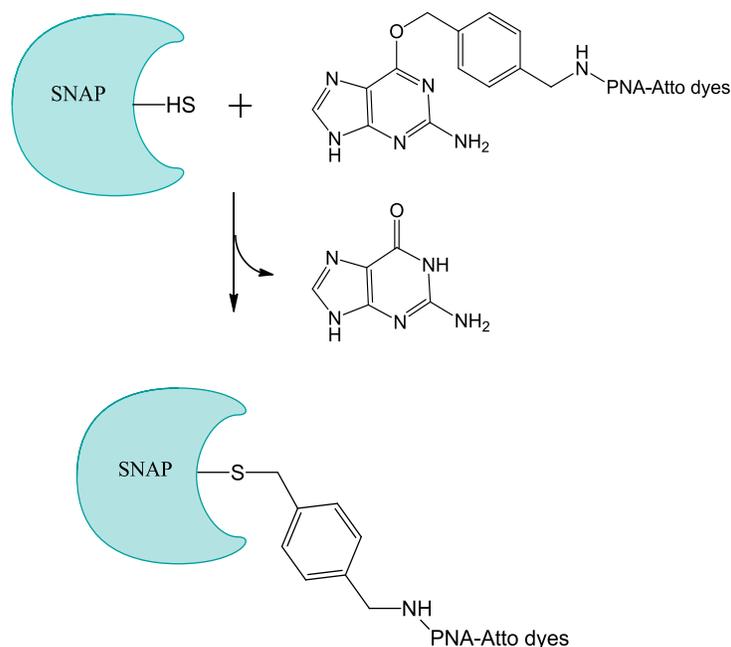
nitrilotriacetate (NTA) derivatives bind reversibly to oligohistidine tags on fusion proteins.²³ This approach is also reported based on Zn^{2+} and lanthanide fluorophore complexes.^{24, 25} The limitations of these strategies are the cytotoxicity of high dose metal complexes and non-specific binding to similar, non-target sequences.² To overcome the specificity limitation of short peptide tags, the introduction of protein-based tags provides a simple, selective strategy to label target proteins with a protein receptor or enzyme followed by subsequent labeling to substrate-probe.¹⁰

Self-labeling protein tags including Halo-tag,²⁶ interacting with Halo ligand, SNAP tag and Clip tag^{2, 27} interacting with benzylguanine/benzylcytosine derivatives are smaller than FPs.¹⁰ Moreover, subsequent modifications with organic fluorophores provide a vast array of colors with better photophysical properties and functionalities.²⁸⁻³⁰

A 19-20 kDa SNAP-tag is available commercially based on mutation of the DNA repair protein O⁶-alkylguanine-DNA alkyltransferase (AGT). The labeling step involves the specific reaction of benzylguanine (BG) derivatives with a reactive cysteine residue of AGT leading to irreversible covalent linkage.^{2, 29-36} *In vitro* labeling of the SNAP-tag proceeds with a relatively fast reaction and 100% efficiency regardless of the fused protein attached.^{22, 27, 37} The large variety of BG derivatives available to use in different conjugation procedures and the specificity of their labeling are considerable advantages with this approach. Since the nature of the ligand attached to BG does not influence the rate of the SNAP-tag reaction with BG derivatives it allows labeling to a wide variety of synthetic probes.^{16, 30, 38, 39}

The versatility and efficiency of the SNAP-tag labeling makes it a promising method to localize and study fusion protein behaviors in living systems. A number of *in vitro* and *in vivo* applications have been reported using SNAP-tag labeling including, localization and trafficking of the fusion protein in cell membranes, labeling of antibody fragments, designing a fluorescent biosensor, control of yeast transcription and visualization of metabolite signaling.⁴⁰⁻⁴⁷

In the last stage of this project, purified SNAP protein was used to ligate a BG-PNA-dye conjugate (Scheme 4.1). Then a DNA template was used to assemble these units in dimer and oligomer form to generate an assembled FRET system (Chapter 5). The chemistry allowing the SNAP proteins to be modified with PNAs will be discussed in this chapter.



Scheme 4.1: Coupling of BG-PNA-Atto dyes to SNAP protein by producing a thioether bond between the cysteine of SNAP and the benzyl linker of a modified construct.

4.2 Results and Discussion

To prepare a fluorescent SNAP-PNA conjugate, the first step is coupling PNA to O⁶-benzylguanine (BG) derivatives for use as SNAP substrates. The resulting PNA-BG conjugate can be labeled with fluorescent dyes.

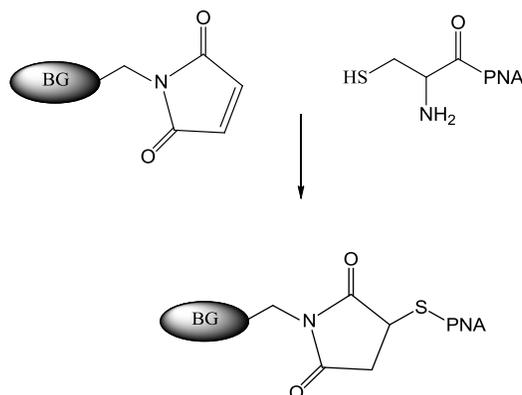
Three PNA sequences with N-terminal cysteine and C-terminal lysine or aspartic acid residues (PNA1: Cys-O-ACGTAC-Lys, PNA2: Cys-O-ACGTAC-Asp and PNA3: Cys-O-CAATGA-Lys) were labeled firstly with BG derivatives and then with fluorescent dyes containing thiol-reactive and amine reactive groups. The Atto and Alexa dyes were selected and BG-PNA conjugates were labeled with Atto488, Atto532 and Alexa fluor488. This allows a SNAP protein to be labeled specifically with BG attached to a fluorescent PNA. After transferring the benzyl group of the BG

to a cysteine residue in the active site of the SNAP, fluorescent PNA is covalently conjugated to SNAP. This gives a handle for directed assembly and allows visualization of the protein.

4.2.1 Labeling of PNA to BG substrates

Two crosslinking approaches based on the reactivity of BG derivatives toward thiol and amine groups of PNAs were used. In the first approach (Scheme 4.2), a maleimide derivative of BG underwent an alkylation reaction with the thiol group of cysteine at the N-terminus of PNA1, PNA2 and PNA3 sequences (Cys-O-ACGTAC-Lys, Cys-O-ACGTAC-Asp and Cys-O-CAATGA-Lys) forming stable thioether bonds⁴⁸ in the presence of a reducing agent (TCEP) at 4°C overnight at pH 7.2.

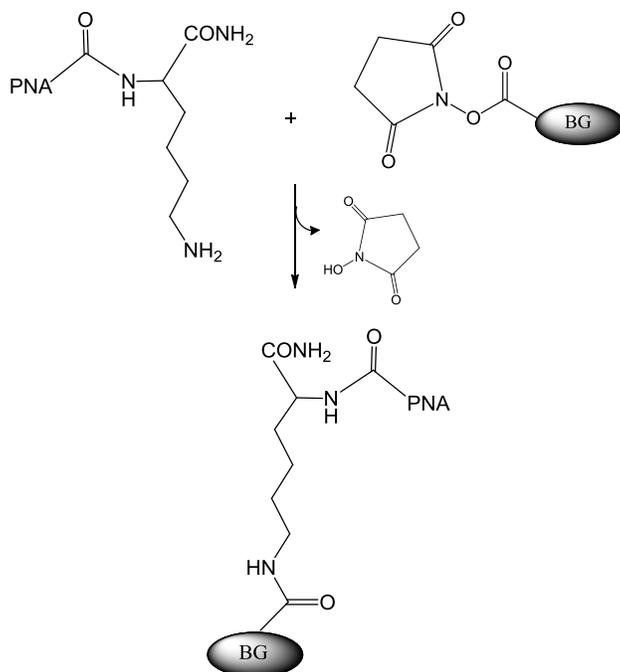
The specific reactivity of maleimide with thiol groups is optimal in the pH ranges 6.5-7.5. Above pH 7.5, maleimide begins to undergo hydrolysis and react with amine groups. Thiols tend to form disulfide bonds in the presence of oxygen. Therefore, a reducing agent such as DTT or TCEP restricts the formation of disulfide bonds.⁴⁸



Scheme 4.2: Conjugation of thiol group of cysteine residue of PNA to BG-maleimide.

In the second strategy (Scheme 4.3), an N-hydroxysuccinimidyl (NHS-ester) derivative of BG was conjugated to the amine group of the C-terminal lysine residues of PNA1 (Cys-O-ACGTAC-Lys) and PNA3 (Cys-O -CAATGA-Lys). The carbonyl group of NHS ester undergoes a nucleophilic attack by the amine group of lysine followed by rapid leaving of the NHS group generating an

amide bond between BG and PNA. This reaction was run at pH 7.2 as both amine reactivity and hydrolysis of NHS ester increase above physiological pH⁴⁸.



Scheme 4.3: Conjugation of amine group of lysine residue of PNA to BG-NHS ester derivative.

The production of BG-PNA conjugates based on these strategies was investigated by MALDI-TOF-MS and reverse phase HPLC (RP-HPLC). The data indicated that the reaction of BG with the thiol group of the PNA cysteine residue was more efficient than coupling using NHS-ester reactivity. The RP-HPLC showed a single peak for PNA1 at 24.32 min which is completely disappeared after conjugation to BG- maleimide, indicating complete conversion to PNA1-BG conjugated form in comparison to the 74% conversion using NHS-ester coupling.

Mass spectrometry of conjugates showed peaks at 2492 Da (calculated: 2494 Da), 2479 Da (calculated: 2481 Da) and 2514 Da (calculated: 2516) related to BG-PNA1 (Fig 4.1a), BG-PNA2 (Fig 4.1b) and BG-PNA3 (Fig 4.1c), respectively.

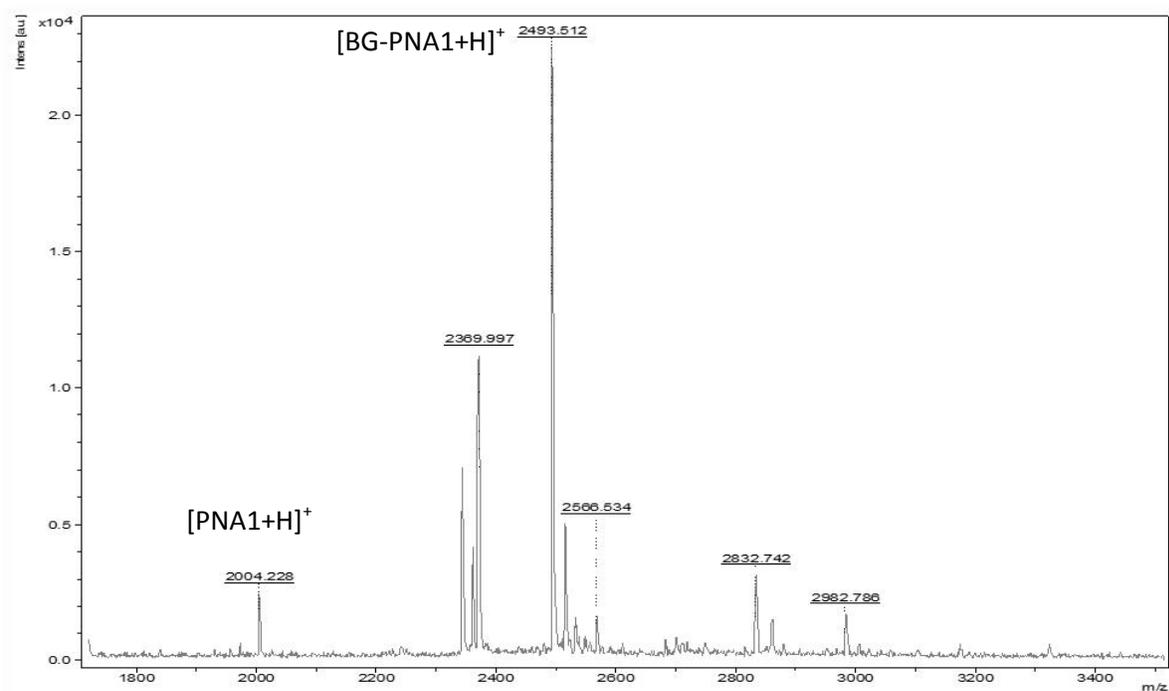


Figure 4.1a: Mass spectrometry data of BG-PNA1 conjugate at 2492 Da (calculated: 2494 Da) .

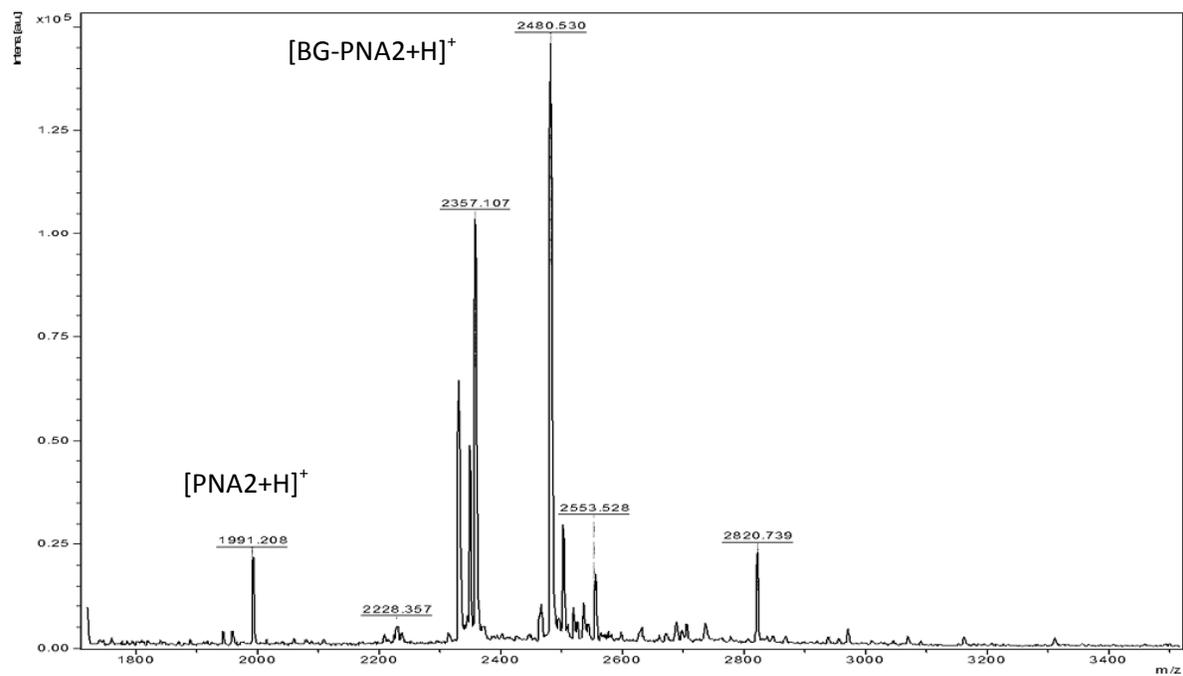


Figure 4.1b: Mass spectrometry data of BG-PNA2 at 2479 Da (calculated: 2481 Da).

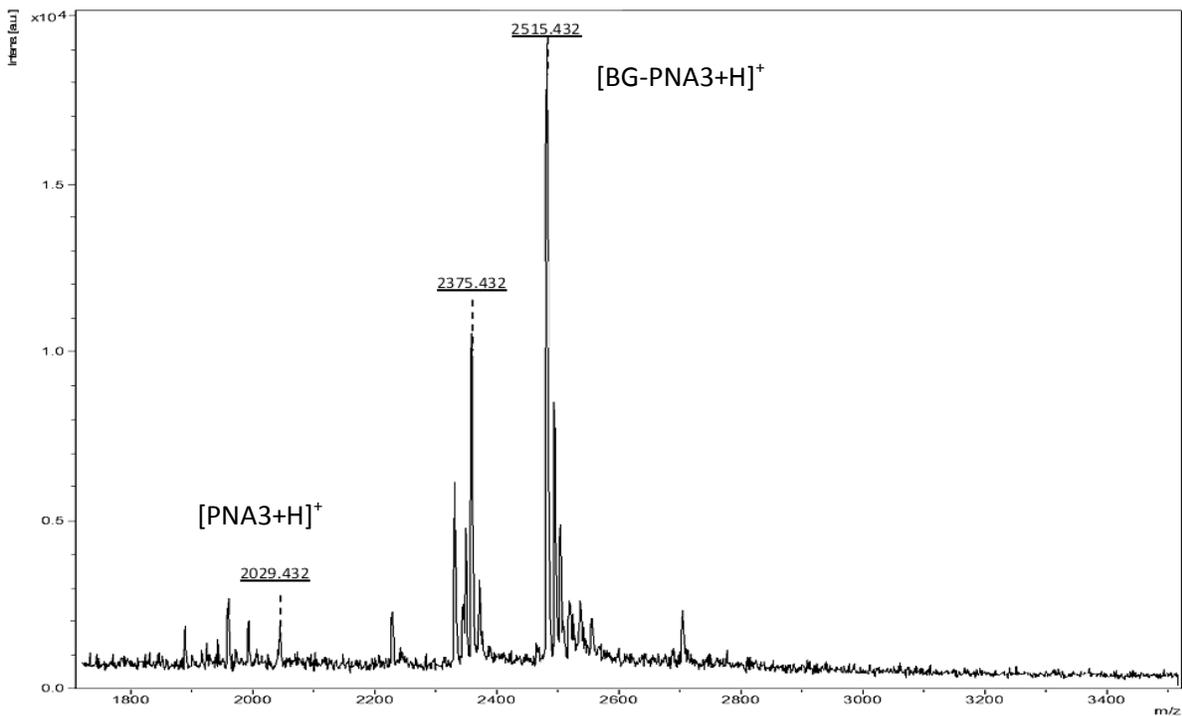


Figure 4.1c: Mass spectrometry data of BG-PNA3 at 2514 Da (calculated: 2516).

The RP-HPLC results showed a distinct peak for each product of BG-PNAs (BG-PNA1 at 23.83 min, BG-PNA2 at 23.74 min and BG-PNA3 at 23.8 min) and confirmed the purity of final BG-PNA1 conjugate. The chromatograms of BG-PNA2 and BG-PNA3 showed that BG-PNA conjugates are dominant products and a small amount of contamination was observed (Fig 4. 2).

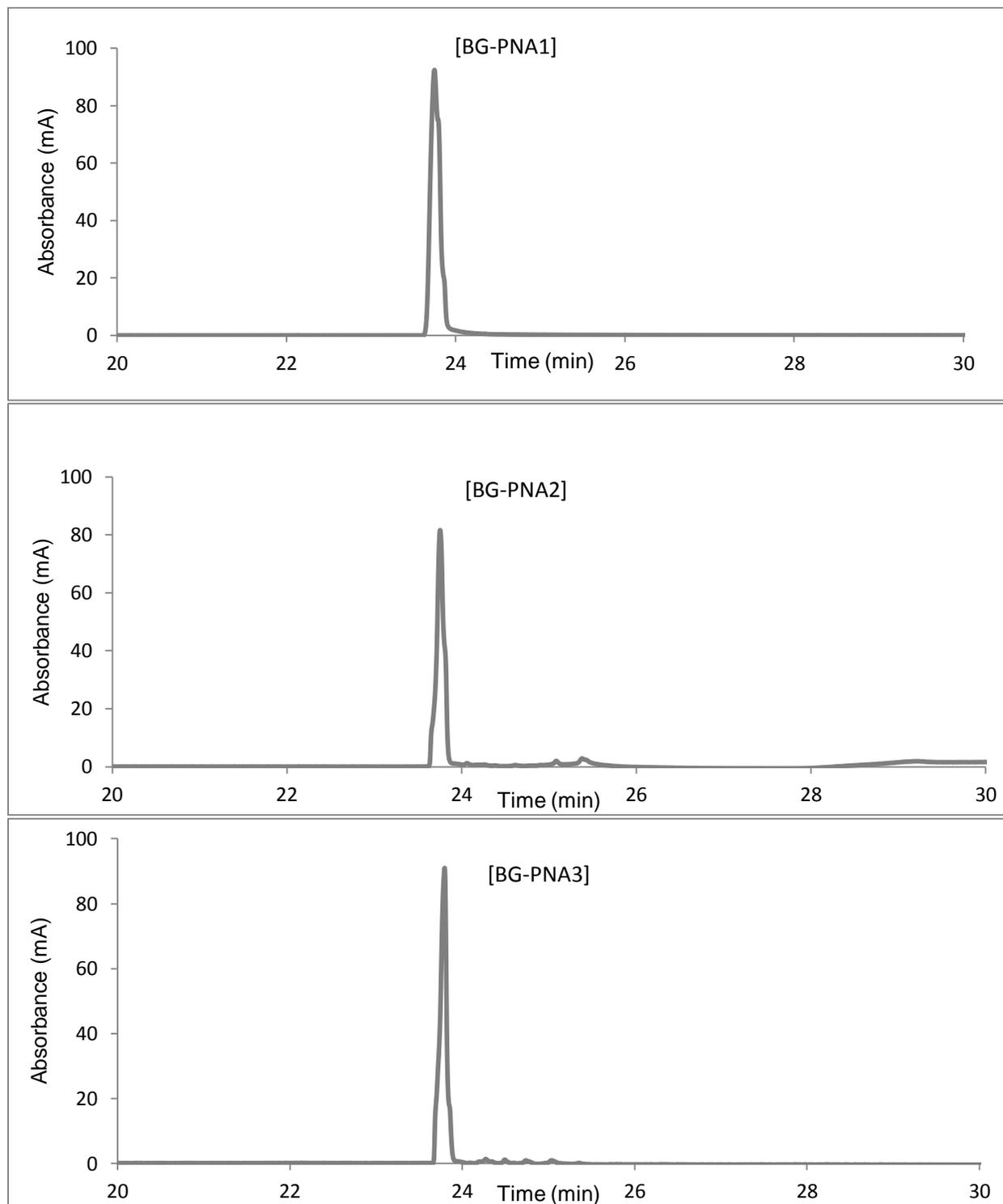


Figure 4.2: RP-HPLC chromatographs of BG-PNA1 at 23.83 min, BG-PNA2 at 23.74 min and BG-PNA3 at 23.8 min (from top to bottom).

4.2.2 Labeling of BG -PNA conjugates with fluorescent dyes

The BG-PNAs produced in the previous section were subsequently labeled with fluorescent dyes. Due to their high photostability and brightness two Atto dyes (Atto488 λ_{ex} : 501nm and λ_{em} : 523nm and Atto532 λ_{ex} : 532nm and λ_{em} : 553nm) and Alexa fluor488 (λ_{ex} : 493nm and λ_{em} : 517nm) were chosen to label BG–PNA conjugates. Labeling was carried out via NHS-esters (Atto dyes) and carboxyl reactive groups (Alexa fluor488)(Fig 4.3).

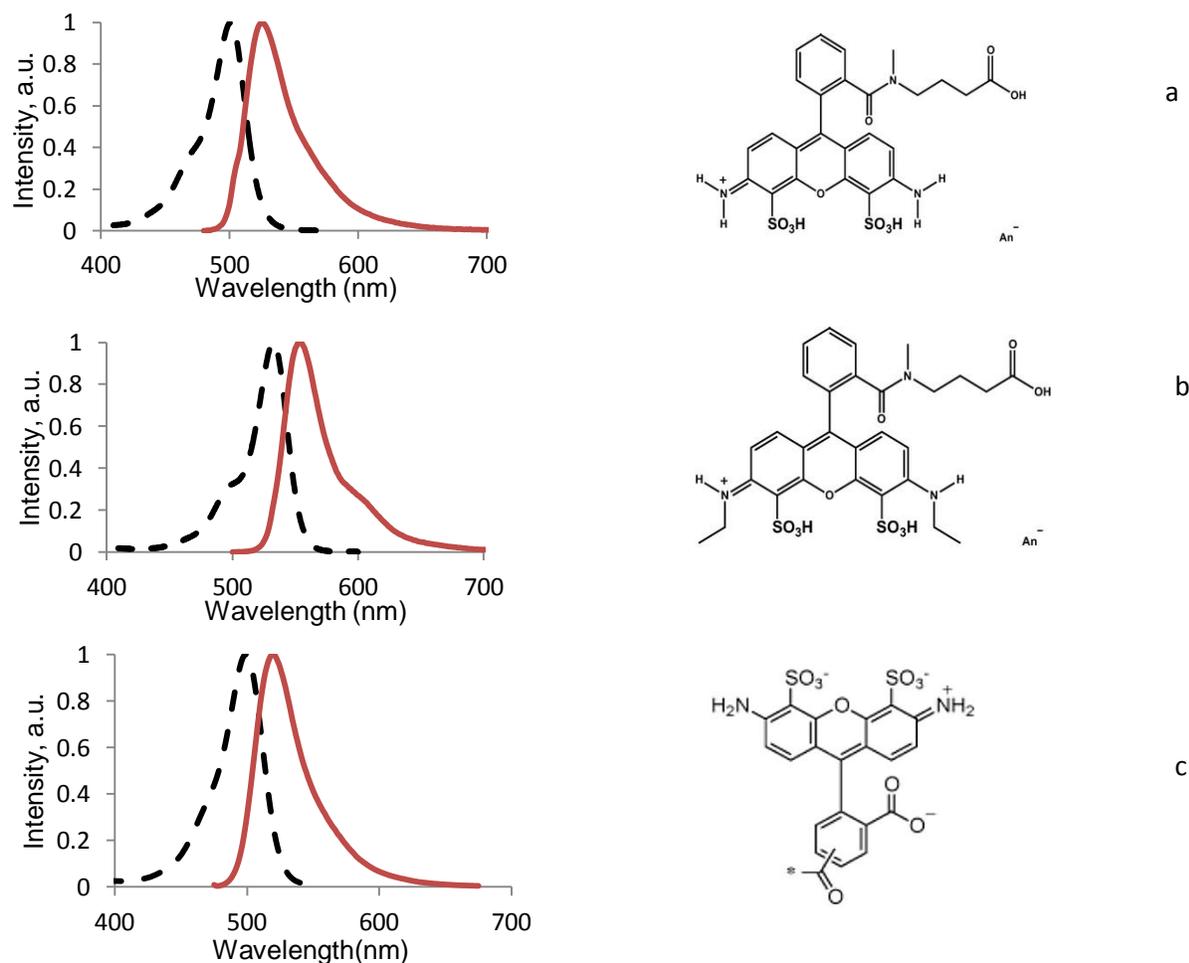
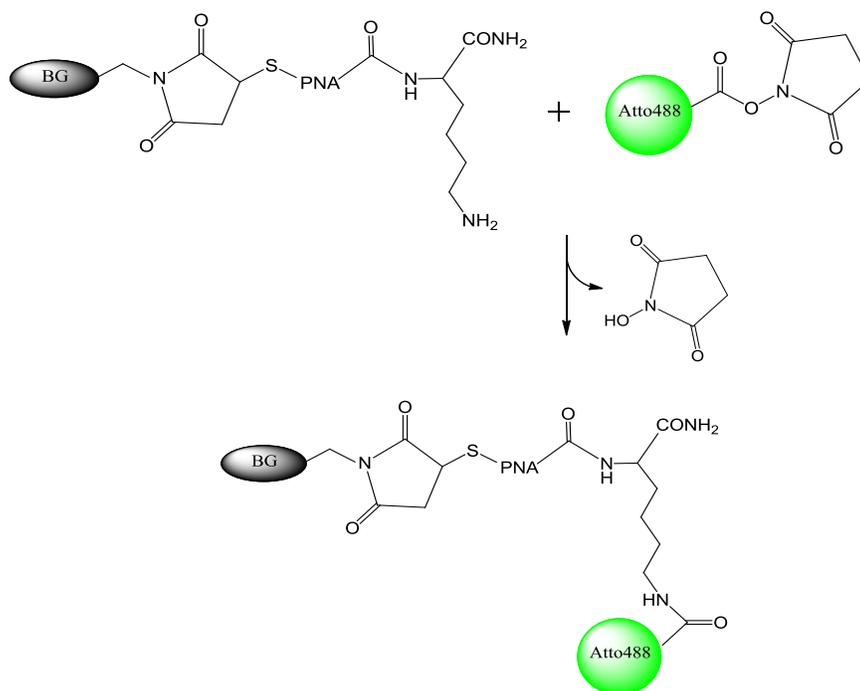


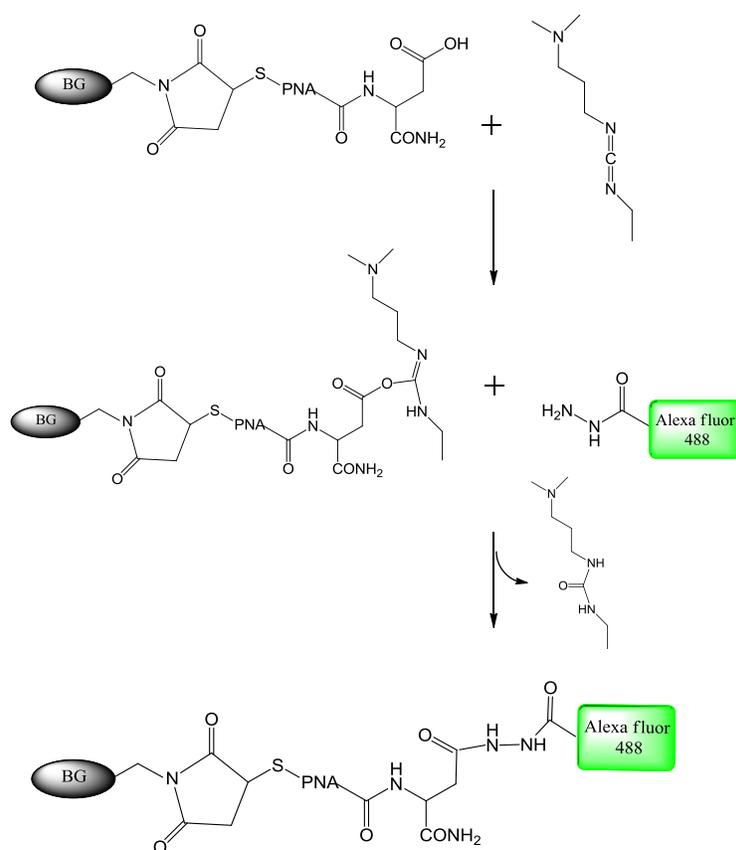
Figure 4.3: Excitation (dashed lines) and emission (solid lines) Spectra and chemical structures of a) Atto488 (λ_{ex} : 501nm and λ_{em} : 523nm, Q(Quantum yield):80%, ϵ_{max} (Max molar absorptivity):90000M⁻¹cm⁻¹, b) Atto532 (λ_{ex} : 532nm and λ_{em} : 553nm, Q: 90%, ϵ_{max} :115000 M⁻¹cm⁻¹) and c) Alexa fluor488 (λ_{ex} : 493nm and λ_{em} : 517nm, Quantum yield:92%, ϵ_{max} :73000 M⁻¹cm⁻¹)(Data from ATTOTECH GmbH, and Life Technologies) .

Atto488 and Atto532 were coupled to BG-PNA via reaction of the NHS group on the Atto dyes with the primary amine group of the C-terminal lysine of BG-PNA1 and BG-PNA3 conjugates forming a stable amide bond (Scheme 4.4). The reaction with NHS-ester derivative was performed in pH 7.2 at room temperature with 1 hour incubation in phosphate buffer in dark.



Scheme 4.4: Conjugation of the amine group of the lysine residue at the C-terminus of BG-PNA1 to an Atto488-NHS ester derivative.

Alexa flour488 hydrazide was used to label BG-PNA2 possessing an aspartic acid at the C-terminus. The carboxylate group of aspartic acid was coupled using Alexa fluor488 hydrazide via a carbodiimide (EDC)–mediated reaction forming an amide linkage (Scheme 4.5).⁴⁸



Scheme 4.5: Conjugation of the carboxyl group of aspartic acid residue at C-terminal of BG-PNA2 to Alexa fluor488 hydrazide mediated by 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC).

Mass spectral analysis revealed peaks at 3066 Da (calculated: 3065 Da) and 3147 Da (calculated: 3148 Da) confirming the production of BG-PNA1-Atto488 and BG-PNA3-Atto532 (Fig 4.4a and b) .

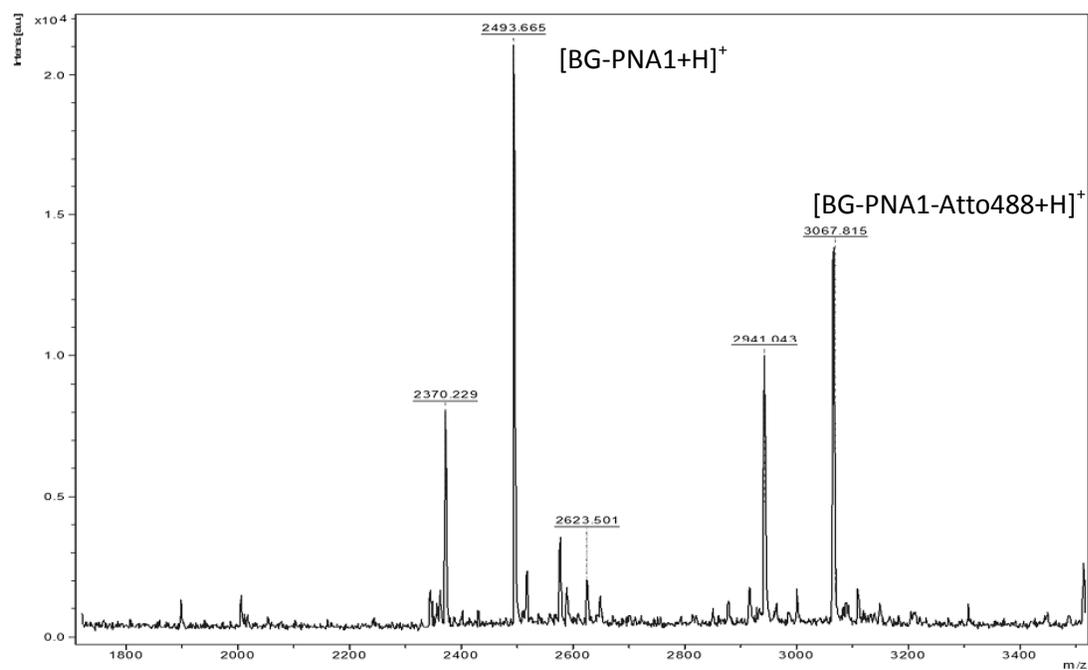


Figure 4.4a: Mass spectrometry data of BG-PNA1-Atto488 at 3066 Da (calculated: 3065 Da).

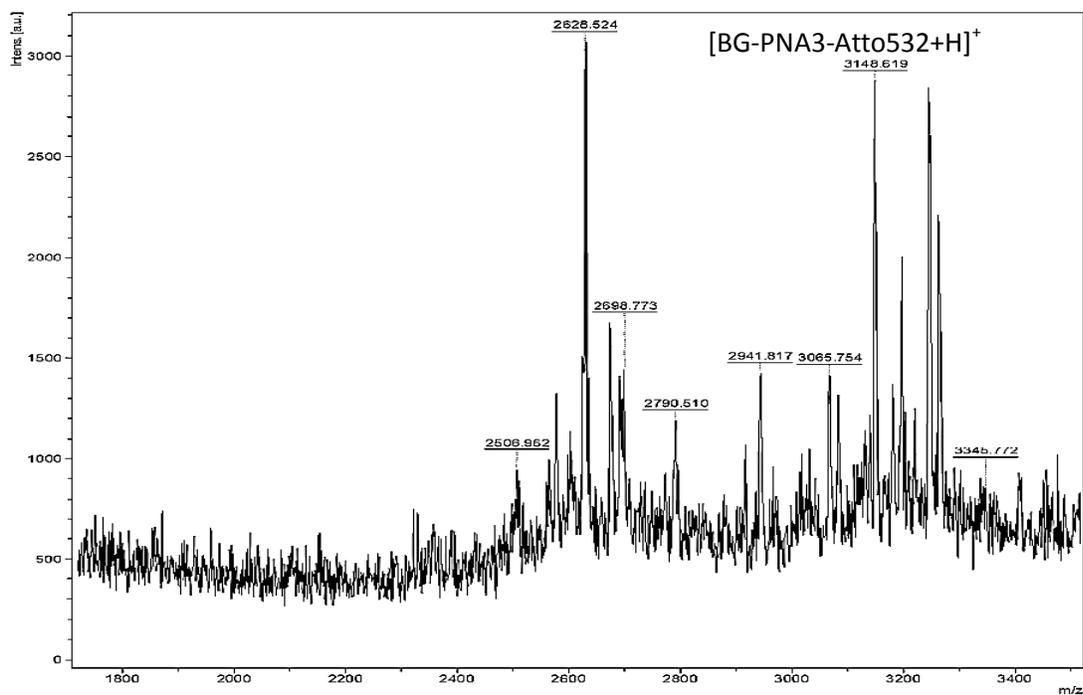


Figure 4.4b: Mass spectrometry data of BG-PNA3-Atto532 at 3147 Da (calculated: 3148 Da).

The RP-HPLC data showed complete conversion to BG-PNA- Atto conjugates by showing new distinct peaks at 23.78 and 23.74 min related to BG-PNA1-Atto488 and BG-PNA3-Atto532 conjugates, respectively, in comparison with BG-PNA1 and BG-PNA3 at 501 nm and 532 nm (Fig 4.5).

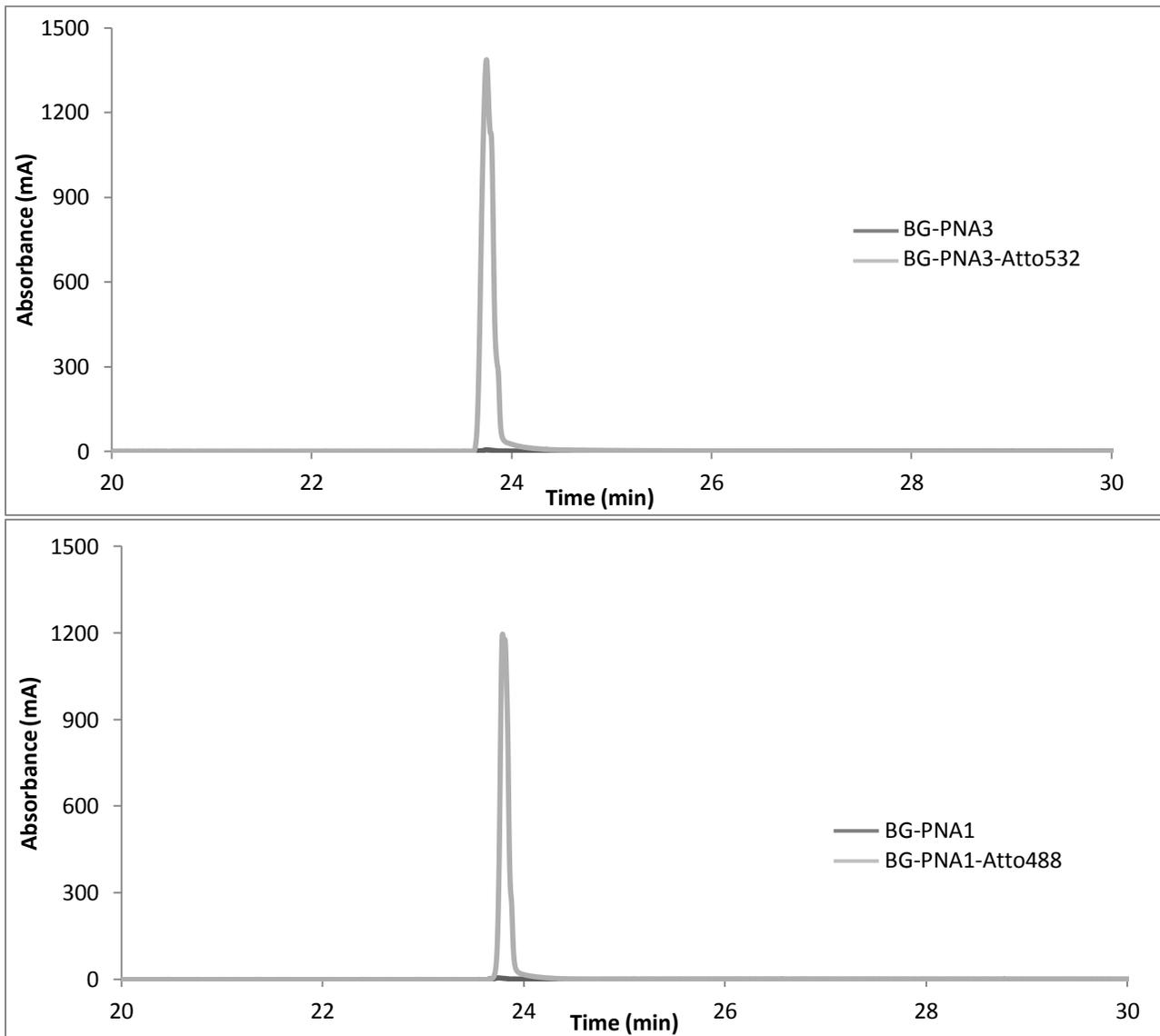


Figure 4.5: RP-HPLC chromatograms of BG-PNA3-Atto532 (grey line) at 23.74 min compare to the same concentration of BG-PNA3 (black line) at 532 nm (top).BG-PNA1-Atto488 (grey line) at 23.78 min compared to the same concentration of BG-PNA1 (black line) at 501 nm (bottom)

Neither MALDI-TOF nor RP-HPLC indicated the production of BG-PNA2-Alexa fluor488. Therefore, BG-PNA1-Atto488 and BG-PNA3-Atto532 constructs were chosen to couple with SNAP protein.

4.2.3 Coupling of BG-PNA-fluorophore to SNAP protein

The coupling of BG-PNA-Atto constructs prepared in the previous section to SNAP was performed by mixing two equivalents of modified PNA1-BG-Atto488 and PNA3-BG-Atto532 constructs with one equivalent SNAP in phosphate buffer pH 7.4 and incubated for 2hrs at 37°C in dark. The reaction proceeded successfully with about 100% yield by producing a thioether bond between cysteine of SNAP and benzyl linker of modified constructs (Scheme 4.1). The final products were purified subsequently by dialysis to remove unreacted compounds.

The fluorescent intensity measurements showed 1:1 ratio of PNA-Atto dyes was attached to SNAP with no significant changes or shift of Atto dye's spectra (Fig 4.6).

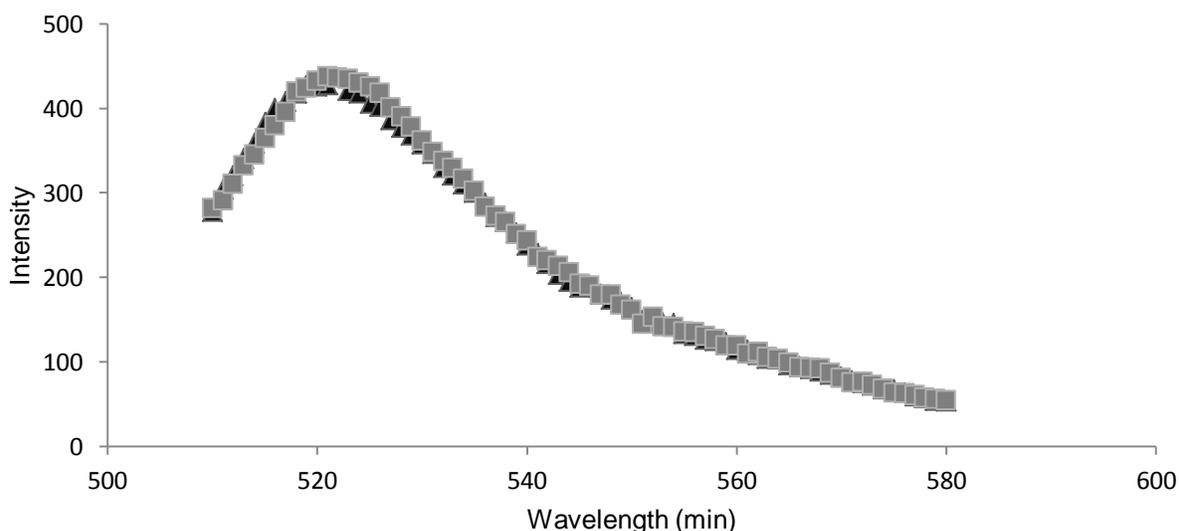


Fig 4.6: The normalized emission peak related to free Atto488 (■) and SNAP-PNA-Atto 488 (5µM)(▲) after excitation at 501 nm (maximum emission at 521 nm).

MALDI-TOF confirmed the production of SNAP-PNA-Atto488 and SNAP-PNA-Atto532 conjugates at 22900 kDa (calculated: 22896 kDa) and 22980 kDa (calculated: 22976 kDa), respectively. However, because of the fragmentation behaviour of MALDI-TOF system, both peaks related to unconjugated and conjugated SNAPS were observed (Fig 4.7 a and b).

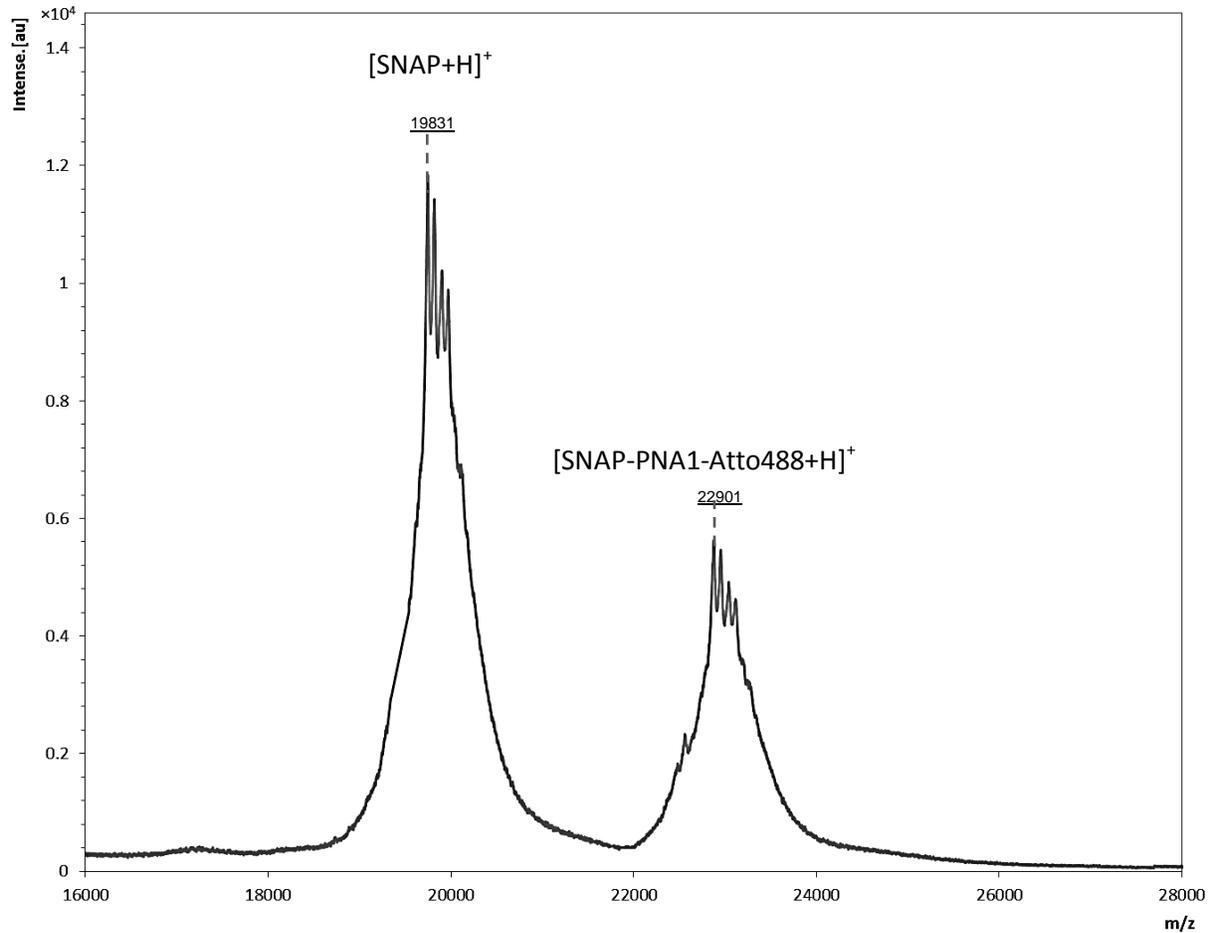


Figure 4.7 a: Mass spectrometry data related to pure SNAP protein at 19830 kDa (calculated 19838 kDa) and SNAP-PNA1-Atto488 at 22900 kDa (calculated: 22895 kDa).

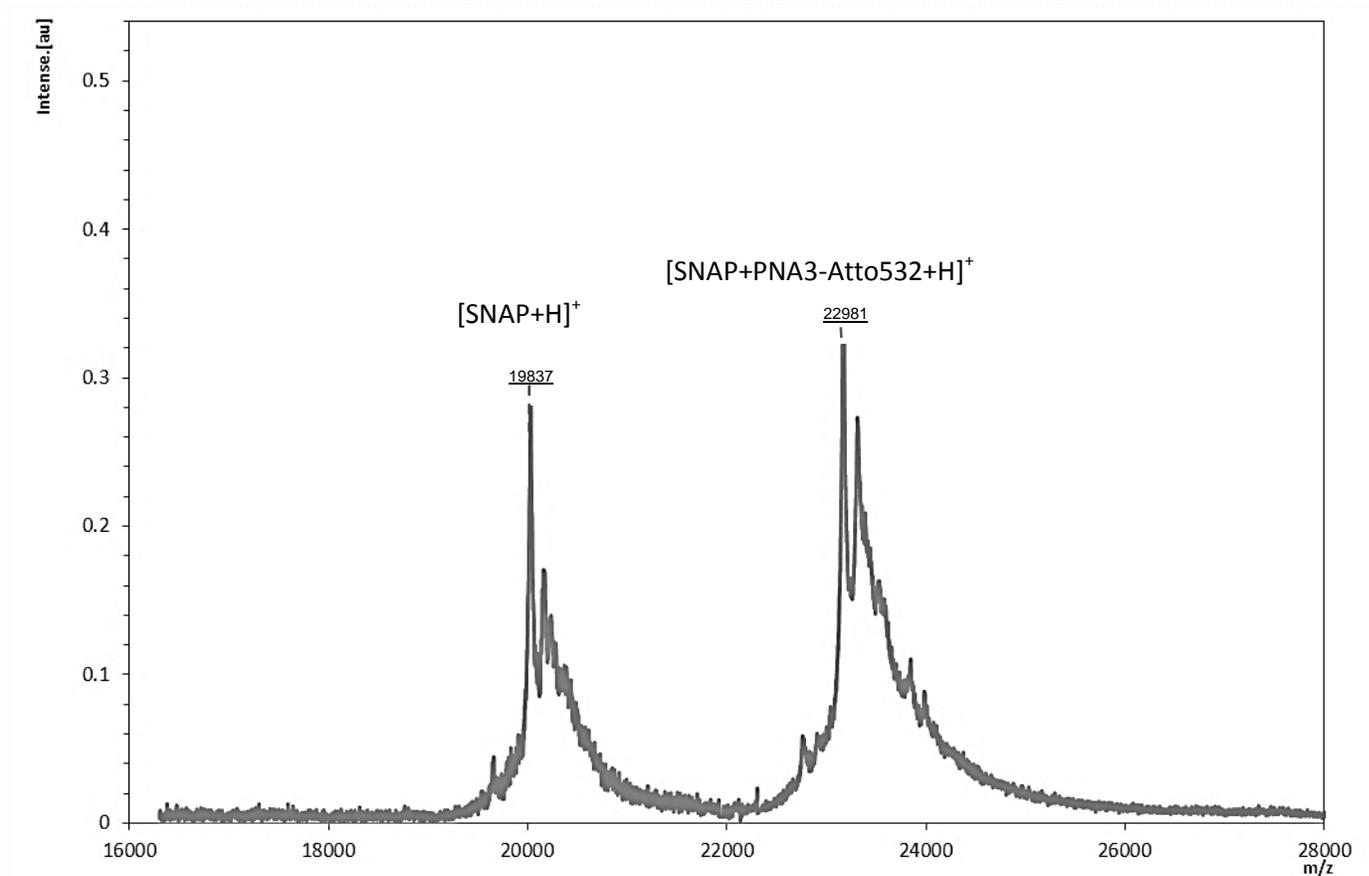


Figure 4.7b: Mass spectrometry data related to pure SNAP protein at 19836 kDa (calculated 19838 kDa) and SNAP-PNA3-Atto532 at 22980 kDa (calculated: 22975 kDa).

The regular SDS-PAGE of SNAP-PNA-Atto dyes conjugates indicated almost complete conversion of SNAP protein to conjugated forms (Fig 4.8).

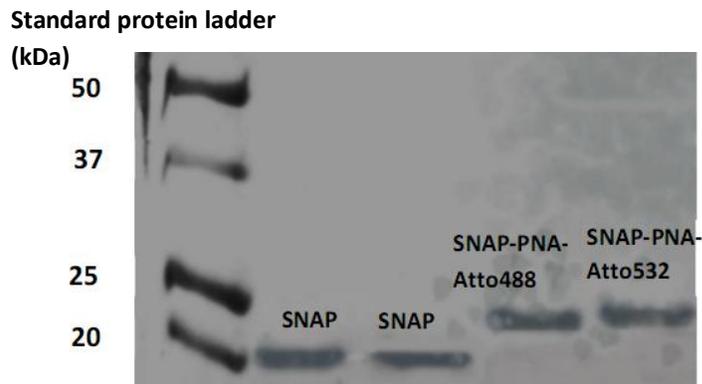


Figure 4.8: SDS-PAGE results for SNAP-PNA-Atto conjugates. The bands related to purified SNAP, SNAP-PNA-Atto488 and SNAP-PNA-Atto532conjugate at 20, 23 kDa, and 23 kDa, respectively.

SEC-HPLC data indicated a single peak related to the final products when observed at the maximum absorption wavelength of Atto 488 (λ_{ex} 501 nm) for SNAP-PNA-Atto488 at 9.4 min and Atto 532 (λ_{ex} 532nm) for SNAP-PNA-Atto532 at 9.44 min with similar elution time with purified SNAP at 9.37 nm (Fig 4.9).

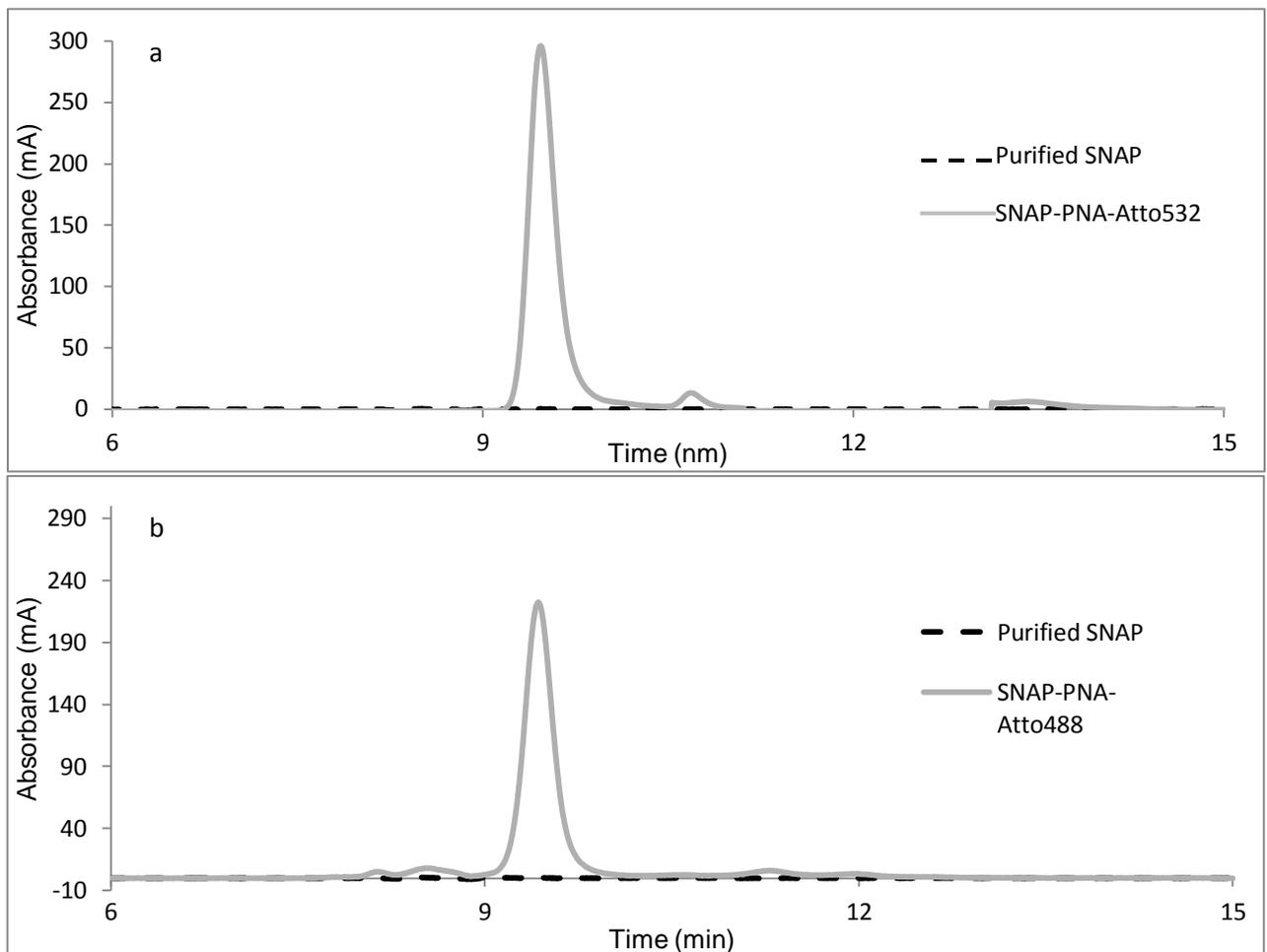


Figure 4.9: SEC-HPLC chromatograms of: a) SNAP-PNA-Atto532 conjugate (grey line) at 532 nm, peak at 9.4 min, insert: SDS-PAGE bands related to purified SNAP and SNAP-PNA-Atto532 conjugate at 20 and 23 kDa. b) SNAP-PNA-Atto488 conjugate (grey line) at 501 nm , peak at 9.44 min.

The SEC-HPLC data showed that the higher ratio of construct to SNAP and longer incubation time did not have a strong effect on conjugation. However, performing the reaction at room temperature dramatically decreased the coupling efficiency (Fig 4.10 a and b).

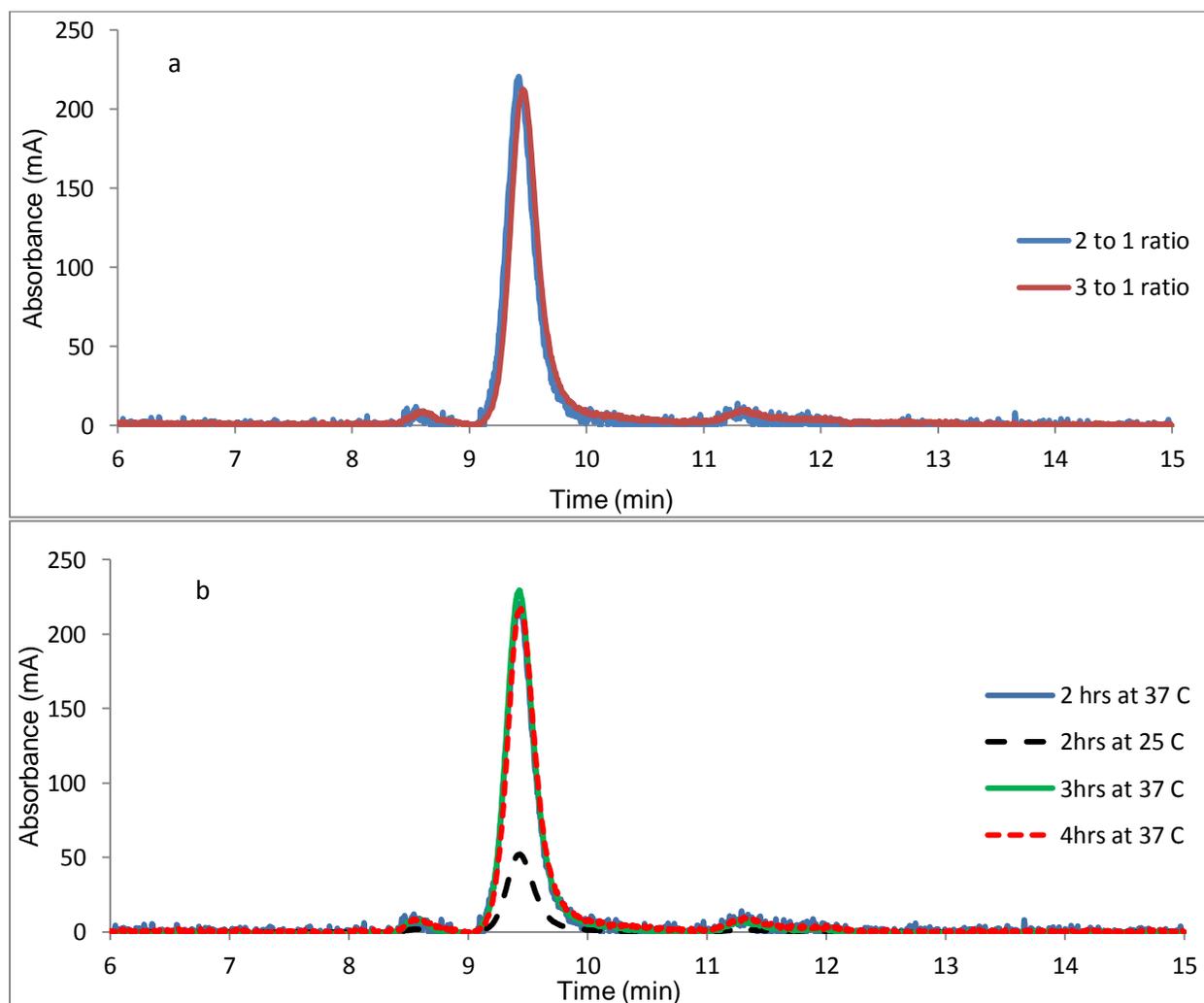


Figure 4.10: SEC-HPLC chromatograms of the SNAP-PNA-Atto488 conjugate at 501 nm: a) with different ratio of BG-PNA-Atto 488 concentration to SNAP, b) with different incubation time at 37 and 25 °C (room temperature).

The specificity of ligation and the speed of reaction in both FP-PNA (Chapter 2) and SNAP-PNA-fluorophore (SNAP-PNA-F) preparations were comparable. An advantage of SNAP-PNA-F is the labeled fluorophore can be easily replaced with a broad range of available organic fluorophores with unique photophysical characteristics to provide different FRET systems. Also, the final

SNAP-PNA-fluorophore (~23kDa) is smaller than FP-PNA (~29kDa) which may interfere less with SNAP-PNA-fluorophore conjugates allowing introduction of assembly directing moieties on any protein expressing the SNAP sequence. Such proteins can be subjected to template directed oligomerization similar to those described in Chapter 3.

4.3 Conclusion and outlook

In this chapter, the successful synthesis of modified PNAs with benzyl guanine (BG-PNA) and fluorescent dyes (BG-PNA-Atto) was shown. After applying different labeling approaches, BG was conjugated to PNA through the reaction of a BG maleimide derivative with a cysteine thiol residue on PNA with high efficiency. The resulting BG-PNA was labeled with NHS-EDC derivatives of Atto dyes (488 and 532). The mass spectrometry and RP-HPLC Data confirmed the final product.

The couplings of prepared constructs with SNAP were performed efficiently with high specificity and nearly complete conversion of SNAP to the conjugated form. SDS PAGE, SEC-HPLC, mass spectrometry and fluorescence anisotropy measurements (Chapter 5) all verified generation of the final conjugates.

The SNAP-PNA conjugate was designed as a unit to assemble into dimer and higher oligomers via hybridization to a complementary DNA or PNA as a framework to provide homo- and hetero-FRET systems (Chapter 5).

4.4 Experimental Part

4.4.1 Conjugation of BG to PNA

4.4.1.1 Conjugation of BG maleimide to cysteine residue at the N-terminus of PNA

The conjugation was run on a 500 μ l scale using 1.5-fold excess of BG-maleimide (150 μ M) (New England Biolabs, USA). The conjugation solution was 100 μ M of PNA (PANAGENE, USA) in 500 μ l PBS buffer pH 7.2, 150 μ M freshly prepared BG-maleimide, and 1mM of TCEP (sigma-Aldrich) as a reducing agent. The mixture incubated overnight at 4°C.

4.4.1.2 Conjugation of BG-NHS ester to lysine residue at the C-terminus of PNA

Ten equivalents of freshly prepared BG-NHS ester (New England Biolabs, USA) (1mM) were added to 100 μ M of PNA in 500 μ l PBS buffer pH 7.2. The reaction solution was incubated 2hrs at room temperature with shaking.

4.4.2 Conjugation of BG-PNA to fluorophore

4.4.2.1 Coupling of Atto dyes (488 and 532) -NHS ester derivative to lysine residue at the C-terminus of BG-PNA

Coupling reactions were carried out by adding approximately 10 equivalents of freshly prepared solutions of dye NHS-ester (1mM) (ATTO-TEC GmbH, Germany) to 200 μ l of prepared BG-PNA conjugate in section 4.5.1.1. The reaction mixture then incubated at room temperature for 2hrs with shaking in the dark.

4.4.2.2 Coupling of Alexa-fluor 488 hydrazide derivative to aspartic acid residue at the C-terminus of BG-PNA

Approximately 5 equivalents of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (Sigma-Aldrich) were combined with BG-PNA prepared previously in section 4.5.1.1 to a final volume

of 200 μ l. The mixture was incubated for 30 min at room temperature to activate the free carboxyl group of aspartic acid at the C-terminus. Thereafter, 10 equivalents of freshly prepared Alexa-flour488 hydrazide (1mM) (Life technologies, UK) were added to the reaction solution and incubated for 2hrs at room temperature with shaking in the dark.

4.4.3 Coupling of purified SNAP protein with BG-PNA-Atto 488 and 532 constructs

Coupling was run on a 200 μ l scale using 2-fold excess of BG-PNA-Atto. The conjugation solution was 5 μ M of purified SNAP protein (New England Biolab, USA) in 200 μ l phosphate buffer pH 7.4, 10 μ M of the prepared BG-PNA-Atto 488 and 532, and 1 mM of TCEP as a reducing agent. The reaction mixture incubated for 2hrs at 37°C in the dark. Afterwards, the final solution was purified by dialysis to remove unreacted compounds. These conditions were selected after evaluating the effects of temperature (37°C and room temperature (25°C)), 2 and 3-fold excess of BG-PNA-Atto (10 μ M and 15 μ M) and different incubation times (2, 3, and 4 hrs) on the conjugation reaction while maintaining other conditions constant.

4.4.4 Dialysis purification of SNAP-PNA-Atto constructs

The dialysis tube (Avg. flat width 35 mm (1.4 in), 12000 Da MW) (Sigma-Aldrich) was boiled in 10 mM bicarbonate buffer (pH 9) and rinsed with distilled deionized water. The dialysis was carried out twice at room temperature for 2-3 hrs each time and finally overnight at 4°C. The final product was stored at -20°C in the dark.

4.4.5 SDS-PAGE electrophoresis of SNAP-PNA-Atto 488 and 532 conjugates

SNAP-PNA-Atto488 and SNAP-PNA-Atto 532 conjugates samples (10 μ M) were heated at 90°C for 5 min and mixed with the same equivalent of SDS sample loading buffer (Biorad) on precast acrylamide gels (Biorad) (12% resolving polyacrylamide gel and 5% stacking gel). The electrophoresis ran for 70 min at 100 V. To stain the protein bands Brilliant coomassie blue (Sigma-Alderich) was used.

4.4.6 MALDI-TOF mass spectrometry analysis

MALDI-TOF mass spectrometry analysis was carried out as described in Chapter 2 after C18 zip-tip preparation of samples.

4.4.7 RP-HPLC of BG-PNA and BG-PNA-Atto 488 and 532 conjugates

The RP-HPLC was performed using C18 reverse phase column eluted with flow rate 1ml/min by 40 min linear gradient of Acetonitrile/0.1% TFA in H₂O started at 10-90% . Eluted compounds were detected by absorbance.

4.4.8 SEC-HPLC of SNAP-PNA-Atto 488 and SNAP-PNA-Atto 532 conjugates

The SEC-HPLC analysis was carried out as described in Chapter 3. The detection wavelengths were 214, 501 (maximum absorption wavelength of Atto 488) and 532 nm (maximum absorption wavelength of Atto 532).

4.5 References

1. H. M. O'Hare, K. Johnsson and A. Gautier, *Current Opinion in Structural Biology*, 2007, **17**, 488-494.
2. D. Jung, K. Min, J. Jung, W. Jang and Y. Kwon, *Molecular BioSystems*, 2013, **9**, 862-872.
3. M. Chalfie, *Photochemistry and Photobiology*, 1995, **62**, 651-656.
4. M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher, *Science (New York, N.Y.)*, 1994, **263**, 802-805.
5. R. A. Cinelli, A. Ferrari, V. Pellegrini, M. Tyagi, M. Giacca and F. Beltram, *Photochemistry and Photobiology*, 2000, **71**, 771-776.
6. N. Hayes, E. Howard-Cofield and W. Gullick, *Cancer Letters*, 2004, **206**, 129-135.
7. P. Ray, H. Pimenta, R. Paulmurugan, F. Berger, M. E. Phelps, M. Iyer and S. S. Gambhir, *Proceedings of the National Academy of Sciences*, 2002, **99**, 3105-3110.
8. S. H. D. Haddock, N. Mastroianni and L. M. Christianson, *Proceedings of the Royal Society B: Biological Sciences*, 2010, **277**, 1155-1160.
9. G. Crivat and J. W. Taraska, *Trends in Biotechnology*, 2012, **30**, 8-16.
10. C. Jing and V. W. Cornish, *Accounts of Chemical Research*, 2011, **44**, 784-792.
11. Y. Yano and K. Matsuzaki, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 2009, **1788**, 2124-2131.
12. N. C. Shaner, P. A. Steinbach and R. Y. Tsien, *Nature Methods*, 2005, **2**, 905-909.
13. N. Yamamoto, H. Tsuchiya and R. Hoffman, *International Journal of Clinical Oncology*, 2011, **16**, 84-91.
14. C. S. Lisenbee, S. K. Karnik and R. N. Trelease, *Traffic*, 2003, **4**, 491-501.
15. E. Snapp, in *Current Protocols in Cell Biology*, John Wiley and Sons, Inc., Editon edn., 2005.
16. T. Komatsu, K. Johnsson, H. Okuno, H. Bito, T. Inoue, T. Nagano and Y. Urano, *Journal of the American Chemical Society*, 2011, **133**, 6745-6751.
17. L. E. T. Jansen, B. E. Black, D. R. Foltz and D. W. Cleveland, *The Journal of Cell Biology*, 2007, **176**, 795-805.
18. J. G. Cole Nb Fau - Donaldson and J. G. Donaldson, *ACS Chem. Biol.*, 2012, **7**, 464-469.
19. G. A. Farr, M. Hull, I. Mellman and M. J. Caplan, *The Journal of Cell Biology*, 2009, **186**, 269-282.

20. S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis and R. Y. Tsien, *Journal of the American Chemical Society*, 2002, **124**, 6063-6076.
21. T. Gronemeyer, G. Godin and K. Johnsson, *Protein Technologies and Commercial Enzymes*, 2005, **16**, 453-458.
22. M. J. Hinner and K. Johnsson, *Chemical Biotechnology & Pharmaceutical Biotechnology*, 2010, **21**, 766-776.
23. R. Guignet, E. Fau - Hovius, H. Hovius, R. Fau - Vogel and H. Vogel, *Nat Biotechnol.*, 2004, **22**, 440-444.
24. R. Y. Hauser, C. Fau - Tsien and R. Y. Tsien, *Proc Natl Acad Sci U S A.*, 2007, **104**, 3693-3697.
25. N. Soh, *Sensors*, 2008, **8**, 1004-1024.
26. G. V. Los, L. P. Encell, M. G. McDougall, D. D. Hartzell, N. Karassina, C. Zimprich, M. G. Wood, R. Learish, R. F. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. H. Klaubert, R. F. Bulleit and K. V. Wood, *ACS Chemical Biology*, 2008, **3**, 373-382.
27. X. Sun, A. Zhang, B. Baker, L. Sun, A. Howard, J. Buswell, D. Maurel, A. Masharina, K. Johnsson, C. J. Noren, M.-Q. Xu and I. R. Correa, *ChemBioChem*, 2011, **12**, 2217-2226.
28. A. Juillerat, C. Heinis, I. Sielaff, J. Barnikow, H. Jaccard, B. Kunz, A. Terskikh and K. Johnsson, *ChemBioChem*, 2005, **6**, 1263-1269.
29. N. Johnsson and K. Johnsson, *ACS Chemical Biology*, 2007, **2**, 31-38.
30. K. Johnsson, *Nature Chemical Biology*, 2009, **5**, 63-65.
31. A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel and K. Johnsson, *Nat Biotech*, 2003, **21**, 86-89.
32. A. Keppler, H. Pick, C. Arrivoli, H. Vogel and K. Johnsson, *Proceedings of the National Academy of Sciences of the United States of America*, 2004, **101**, 9955-9959.
33. C. Keppler, A. Fau - Arrivoli, L. Arrivoli, C. Fau - Sironi, J. Sironi, L. Fau - Ellenberg and J. Ellenberg, *Biotechniques*, 2006, **41**, 167-170, 172, 174-165.
34. M. Keppler, A. Fau - Kindermann, S. Kindermann, M. Fau - Gendreizig, H. Gendreizig, S. Fau - Pick, H. Pick, H. Fau - Vogel, K. Vogel, H. Fau - Johnsson and K. Johnsson, *Methods*, 2004, **32**, 437-444.
35. M. Z. Lin and L. Wang, *Physiology*, 2008, **23**, 131-141.
36. T. Gronemeyer, C. Chidley, A. Juillerat, C. Heinis and K. Johnsson, *Protein Engineering Design and Selection*, 2006, **19**, 309-316.
37. M. A. Gauthier and H. A. Klok, *Chemical Communications* 2008, **23**, 2591-2611.
38. D. Maurel, S. Banala, T. Laroche and K. Johnsson, *ACS Chemical Biology*, 2010, **5**, 507-516.
39. C. Campos, M. Kamiya, S. Banala, K. Johnsson and M. González-Gaitán, *Developmental Dynamics*, 2011, **240**, 820-827.
40. M. A. Brun, K.-T. Tan, E. Nakata, M. J. Hinner and K. Johnsson, *Journal of the American Chemical Society*, 2009, **131**, 5873-5884.
41. C. Donovan and M. Bramkamp, *Microbiology*, 2009, **155**, 1786-1799.
42. A. F. Hussain, F. Kampmeier, V. von Felbert, H. F. Merk, M. K. Tur and S. Barth, *Bioconjugate Chemistry*, 2011, **22**, 2487-2495.
43. F. Kampmeier, M. Ribbert, T. Nachreiner, S. Dembski, F. Beaufils, A. Brecht and S. Barth, *Bioconjugate Chemistry*, 2009, **20**, 1010-1015.
44. M. Bannwarth, I. R. Correa, M. Sztretye, S. Pouvreau, C. Fellay, A. Aebischer, L. Royer, E. Riñós and K. Johnsson, *ACS Chemical Biology*, 2009, **4**, 179-190.
45. M. Kamiya and K. Johnsson, *Analytical Chemistry*, 2010, **82**, 6472-6479.
46. H. Haruki, M. R. Gonzalez and K. Johnsson, *PLoS ONE*, 2012, **7**, e37598.
47. J. P. Pin, L. Comps-Agrar, D. Maurel, C. Monnier, M. L. Rives, E. Trinquet, J. Kniazeff, P. Rondard and L. Prézeau, *The Journal of Physiology*, 2009, **587**, 5337-5344.
48. G. T. Hermanson, *Bioconjugate Techniques*, 2nd edn., Academic Press, London, 2008.

Chapter 5

Hetero- and Homo-FRET systems created by assembly of fluorescent SNAP-PNA units in dimer and higher oligomer forms

The SNAP protein is a widely used self-labelling tag that can be used as an alternative or orthogonal group to FPs for tracking protein localization and trafficking in living systems. A model system providing controlled alignment of SNAP-tag units can provide a new way to study clustering of fusion proteins. In chapter 3, the precise recognition characteristics of PNA for complementary DNA templates led to successful assembly of FP-PNA units in dimer and tetramer forms. In this chapter, the fluorescent SNAP-PNA constructs prepared in chapter 4 were assembled to create dimer, trimer, and tetramer forms on different DNA scaffolds. Dimer forms were prepared based on template directed assembly of both homo- and hetero-dimers of SNAP-PNA constructs. Two longer DNA scaffolds were applied to provide inducible assembly of fluorescent SNAP-PNA constructs into higher oligomers. Size exclusion chromatography, SDS-PAGE and fluorescence techniques confirmed the assembly formation. The results demonstrated the production of the assembled dimers and higher oligomers exhibiting hetero- and homo-FRET which can be used to study clustering behavior of other proteins attached to the SNAP-tag. Variations of this system are capable of directed assembly of non-fluorescent molecules which may allow study of downstream processes in cells.

5.1 Introduction

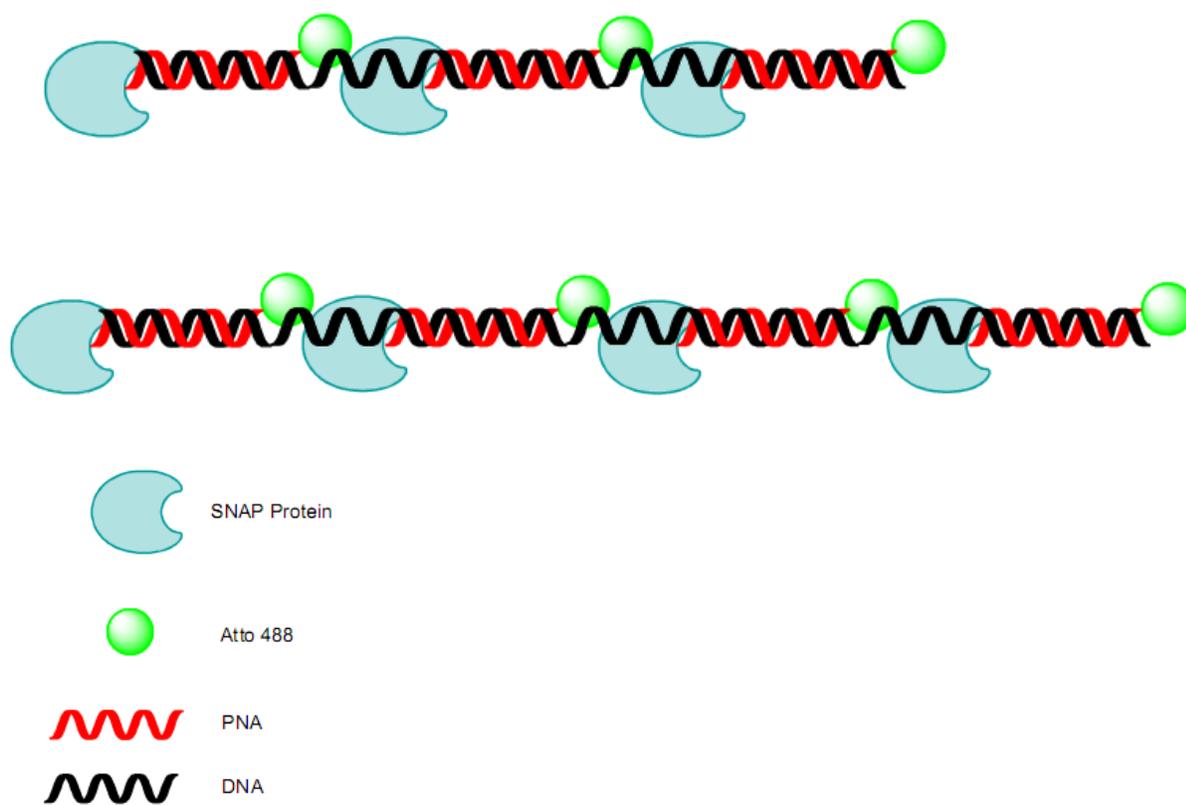
The role of dimer and higher oligomer formation in the functionality of proteins remains an important question, especially in the case of cell membrane proteins.¹ Biophysical tools helping to address this relationship would bring better understanding of the role of malfunctioning proteins in diseases and potentially aid in fighting those diseases.²

In Chapters 2 and 3, the combination of protein bioconjugation techniques to produce protein-PNA with identical programmability of DNA provided a controllable self-assembled FRET-system which was used successfully to investigate fluorescent proteins oligomerization.

As stated in chapter 4, the self-labelling SNAP protein technology has been widely used to visualize and study protein localization, dynamics and trafficking in living cells.³⁻¹⁵ For example, the SNAP-tag approach has been recently used to address the oligomerization behavior of GPCRs. It showed the predominant assembly of mGlu receptors as a homodimer and the existence of homomers of the heteromeric GABA_B receptors.^{1, 16, 17} The importance of these assemblies on the function of receptors is still under investigation.

The high efficiency and specificity of this labelling technology combined with ease of use and speed of application makes it a promising platform to study proteins of interest. The SNAP technology makes accessible a wide range of chemical probes and unlike fluorescent proteins does not show detrimental effects on fused proteins.¹⁸⁻²¹

FRET has been widely employed to detect SNAP-tag fusion protein localization and oligomerization.^{6, 17, 22-28} Specific conjugation of ligands to SNAP tag fusion proteins allowed the analysis of ligand-receptor interaction using time resolved fluorescence energy transfer (TR-FRET).²⁹ The combination of SNAP technology and TR-FRET was also applied for investigation of GPCRs oligomerization behaviour.^{1, 17}



Scheme 5.2: Assembly of SNAP-PNA1-Atto488 in higher oligomer forms. The top one is related to assembly of three SNAP-PNA1-Atto488 constructs into trimer (DNA3-trimer) and likewise the lower assembly consists of four SNAP-PNA1-Atto488 constructs in a tetrameric form (DNA4-tetramer).

Table 5.1: Oligonucleotide sequences used for making assembled SNAP dimer and oligomer forms.

Name	Oligonucleotide sequence
PNA1	5'- ACGTAC-3'
PNA3	5'- CAATGA-3'
DNA1	5'- TGCATGGATCTGCATG-3'
DNA2	5'- TGCATGGATCGTTACT-3'
DNA3	5'-TGCATGGATCTGCATGGATCTGCATG3'
DNA4	5'- TGCATGGATCTGCATGGATCTGCATGGATCTGCATG-3'

After assembly SDS-PAGE analysis was carried out after 5 min heating of each sample at 90°C . Adding 2:1 ratio of SNAP-PNA1-Atto 488: DNA1 resulted in three distinct bands at 23, 27 and 50 kDa related to SNAP-PNA-Atto488 (monomer), monomer:DNA1 and dimer:DNA1, respectively. Similarly, mixing 1:1:1 ratio of SNAP-PNA1-Atto488: SNAP-PNA3-Atto532:DNA2 indicated the same three bands at 23, 27 and 50 kDa related to SNAP-PNA1-Atto488 and SNAP-PNA3-Atto532 (both as monomer), monomer:DNA2 and dimer:DNA2 consists of both SNAP-PNA1-Atto488 and SNAP-PNA3-Atto532 hybridized to DNA2 .Under similar conditions, adding a 3:1 ratio of SNAP-PNA1-Atto 488: DNA3 produced four distinct bands at 23, 31, 54 and 77 kDa related to SNAP-PNA1-Atto 488 (monomer), monomer:DNA3, dimer:DNA3 and trimer:DNA3 , respectively. Also, 4:1 ratio of SNAP-PNA1-Atto 488:DNA4 resulted in five distinct band at 23, 34, 57, 80,103 kDa related to SNAP-PNA1-Atto488 (monomer), monomer:DNA4, dimer:DNA4, trimer:DNA4 and tetramer:DNA4 , respectively (Fig 5.1).

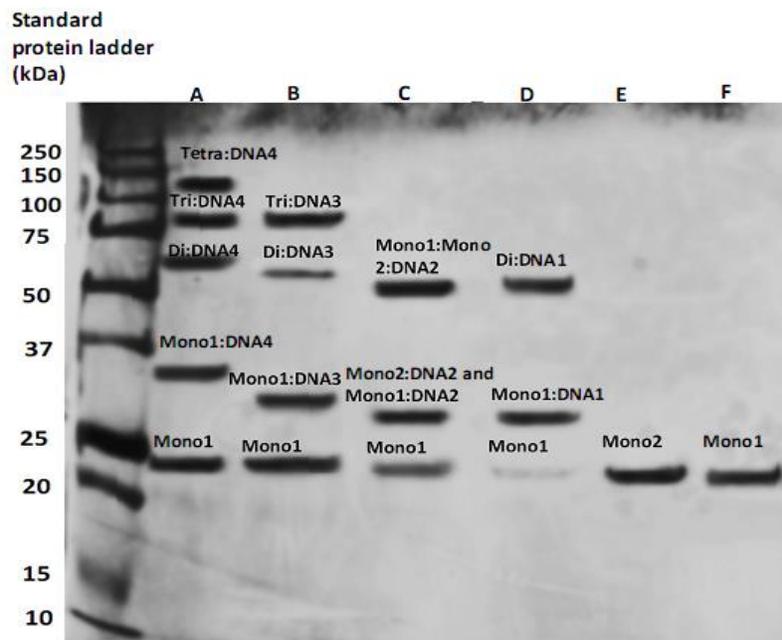


Figure 5.1: SDS-PAGE analysis of dimers (homo-dimer: column D and hetero-dimer: column C with bands at 23, 27, and 50 kDa) and oligomers (trimer : column B with bands at 23, 31, 54, and 77kDa and tetramer:column A with bands at 23, 34, 57, 80, and 103 kDa). Mono1 and Mono2 are related to SNAP-PNA-Atto488 (column F with bands at 23 kDa) and SNAP-PNA3-Atto532 (Column E with bands at 23 kDa) (both as monomers).

SEC-HPLC data was calibrated against a molecular weight standard (Chapter3, Fig 3.3). The chromatograms of SNAP-PNA1-Atto488 and SNAP-PNA3-Atto532 monomers showed single peaks at 9.4 and 9.44 min at the wavelengths of maximum absorption of Atto488 (501 nm) and Atto532 (532 nm), respectively. The solution containing 2:1 ratio of SNAP-PNA1-Atto 488:DNA1 indicated a new peak at 8.59 min (estimated to be 8.64 min) at 214 and 501 nm attributed to the efficient assembly of homo-dimer on the DNA1 scaffold (Fig 5.2). The same result was observed when-PNA1-Atto 488 SNAP and SNAP-PNA3-Atto 532 assembled as a hetero-dimer on the DNA2 scaffold by showing a new peak at 8.58 min (estimated to be 8.64 min) at the maximum wavelength of Atto 488 and Atto 532 and at 214 nm (Fig 5.3). A titration of DNA1 with an increasing amount of SNAP-PNA1-Atto488 indicated a gradual decrease of peak related to free SNAP-PNA1-Atto488 monomer which reached zero at 2:1 ratio of SNAP-PNA1-Atto488 to DNA1 while the peak related to DNA1:dimer reached its highest, showing the complete conversion of monomer to dimer assembly.

The solutions featuring 3:1 SNAP-PNA1-Atto488:DNA3 and 4:1 ratio of SNAP-PNA1-Atto488:DNA4 showed additional peaks at 8.16 min (estimated to be 8.14 min) and 7.79 min (estimated to be 7.65 min) at 501 and 214 nm respectively related to assembly in trimer form on DNA3 scaffold (Fig 5.4) and tetramer form on DNA4 scaffold (Fig 5.5). Titrations of these solutions with an increasing amount of SNAP-PNA1-Atto488 showed the same gradual decrease of monomer peak which eventually reached zero at 3:1 and 4:1 ratio of SNAP-PNA1-Atto488:DNA3 and SNAP-PNA1-Atto488:DNA4, respectively, while both peaks of trimer:DNA3 and tetramer:DNA4 reached their maximum which were attributed to the efficient formation of trimer and tetramer assemblies (Fig 5.6).

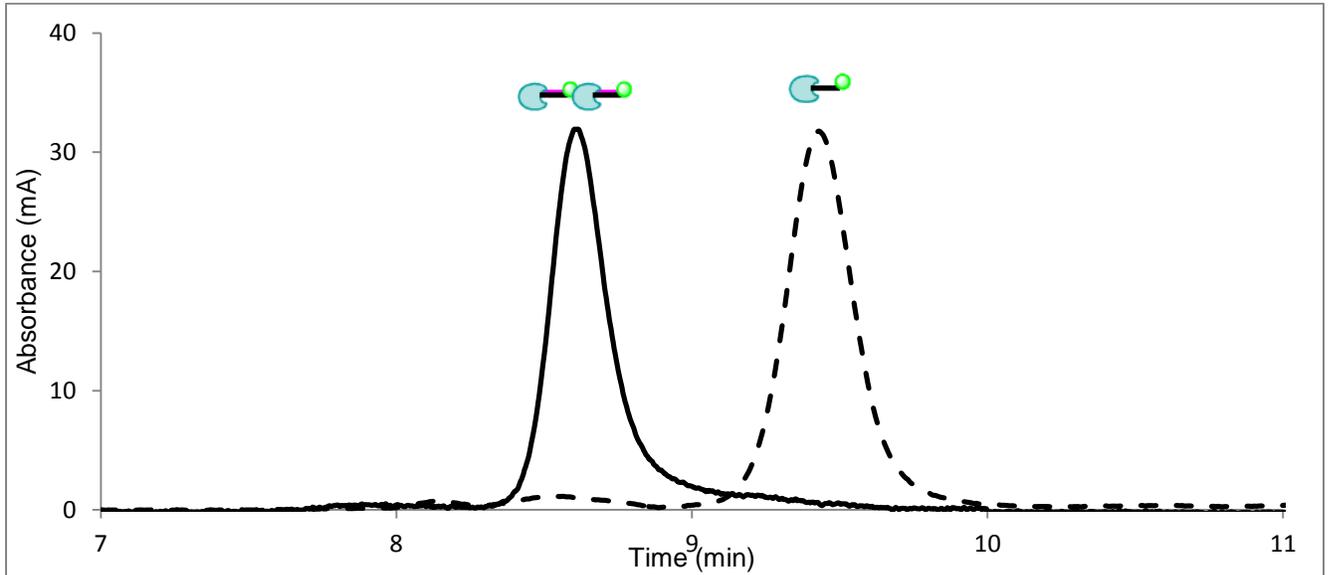


Figure 5.2: Normalized SEC-HPLC chromatograms of SNAP-PNA1-Atto488 (dashed line, peak at 9.4 min) at 214nm and SNAP-PNA1-Atto488-DNA1(solid line, peak at 8.59 min related to DNA1:dimer) at 501 nm.

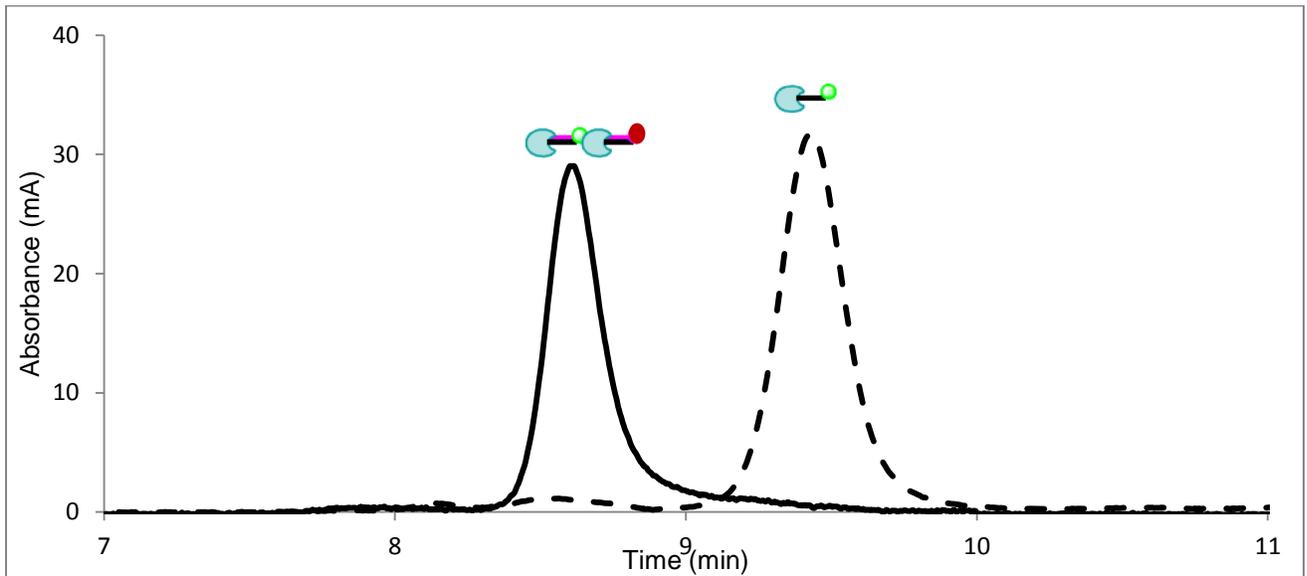


Figure 5.3: Normalized SEC-HPLC chromatograms of SNAP-PNA1-Atto488 (dashed line, peak at 9.4 min) and SNAP-PNA1-Atto488- SNAP-PNA3-Atto532- DNA2 (solid line, peak at 8.58min related to DNA2:dimer) at 532 nm

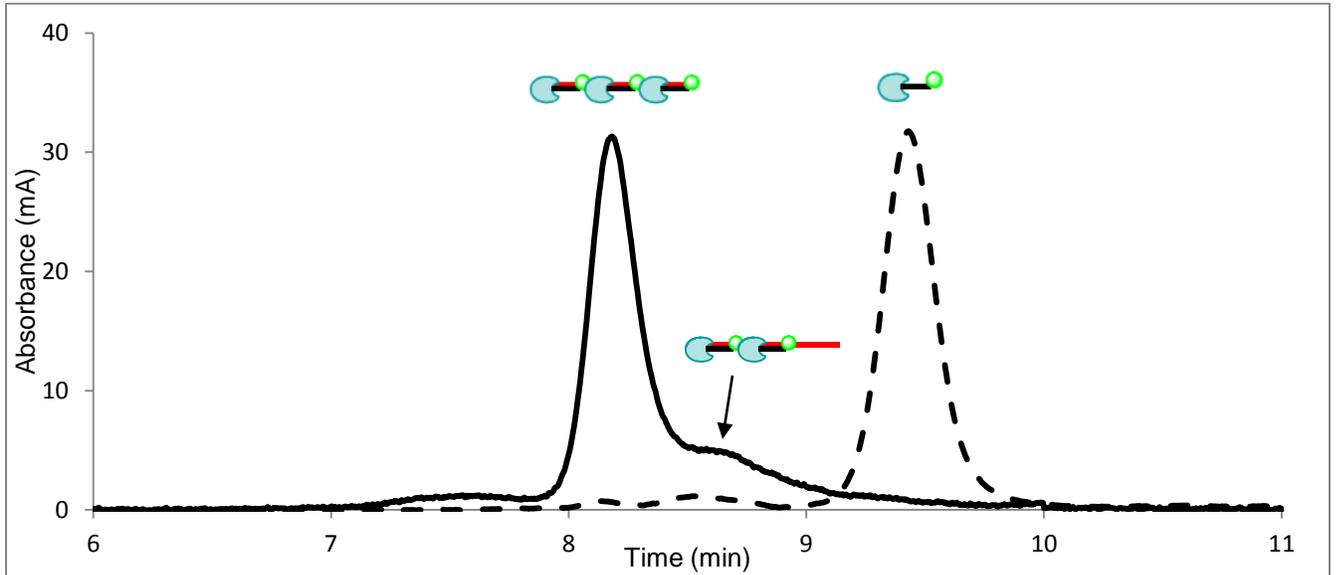


Figure 5.4: Normalized SEC–HPLC chromatograms of SNAP-PNA1-Atto488 (dashed line, peak at 9.4 min) and SNAP-PNA1-Atto488- DNA3 (solid line, peak at 8.16 min related to DNA3:trimer) at 501 nm.

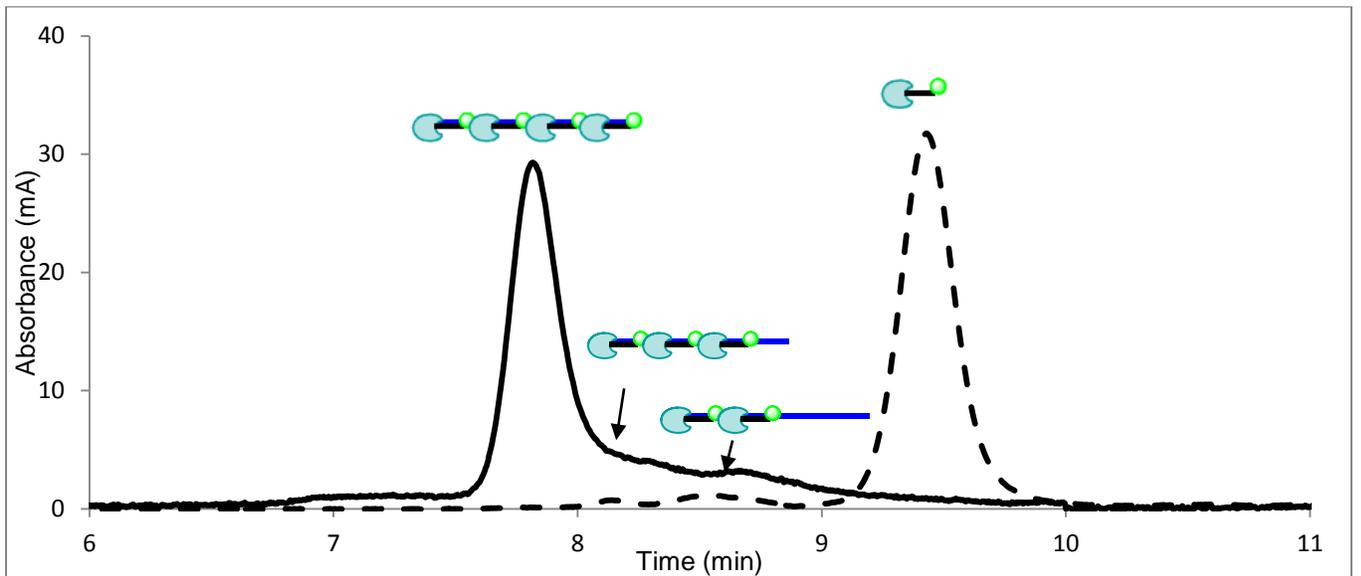


Figure 5.5: Normalized SEC–HPLC chromatograms of SNAP-PNA1-Atto488 (dashed line, peak at 9.4 min) and SNAP-PNA1-Atto488- DNA4 (solid line, peak at 7.79 min related to DNA4:tetramer) at 501 nm.

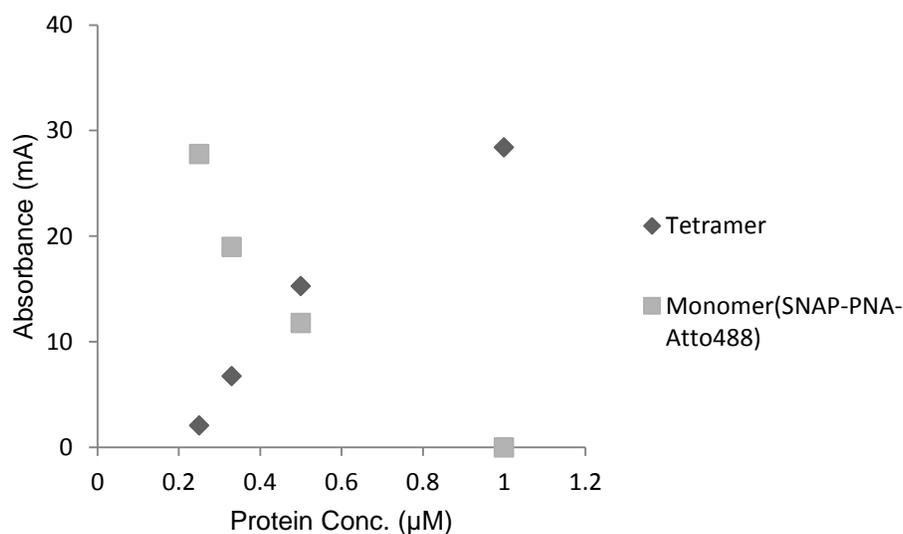


Figure 5.6: SEC analysis of showing the increase of the tetramer peak (7.79 min) and decrease of the monomer peak (9.4 min) monitored at 501 nm while titrating DNA4 (0.25 µM) with SNAP-PNA-Atto488 (0.25-1 µM).

The characterization of the assembled homo- and hetero-dimers and oligomers by SDS-PAGE electrophoresis and SEC-HPLC confirmed the assembly of SNAP-PNA-Atto488 and SNAP-PNA-Atto532 monomers using the specific recognition capability of conjugated PNAs to DNA templates.

5.2.2 Hetero-FRET system

To study the ability of fluorescent SNAP-PNA to assemble, DNA2 template (5'-TGCATGGATCGTTACT-3') was used as a framework to make hetero-dimer. The DNA template had two complementary parts for SNAP-PNA1-Atto488 and SNAP-PNA2-Atto532 with the distance (3.7 nm) below the Forster distance of Atto488 and Atto532 as a FRET pair (6.4 nm) (Scheme 5.1- bottom). Due to the spectral overlap of the dyes, the excitation energy of donor fluorophore (Atto488) transfers to acceptor fluorophore (Atto532) after assembly indicating the formation of dimer.

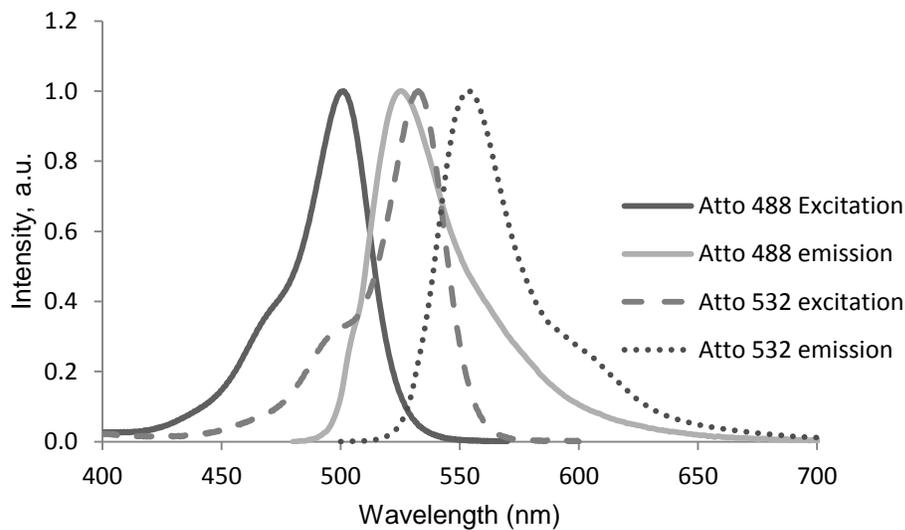


Figure 5.7: Spectral overlap between Atto488 emission and Atto532 excitation as a FRET pair: solid black and grey peaks are related to normalized excitation and emission of Atto488, dashed and dotted peaks is related to normalized excitation and emission of Atto532, respectively (data from ATTo-TEC GmbH).

To study the assembly of SNAP-PNA constructs on DNA2, the emission intensity of the solution containing 1:1 ratio of SNAP-PNA1-Atto488:DNA2 which plays the role of donor in expected hetero-FRET systems was measured with excitation at 501 nm. Simultaneous decrease of donor emission intensity was observed by adding the same ratio of acceptor, SNAP-PNA-Atto532, while a new peak related to sensitized emission intensity of acceptor at 553 nm was appeared (Fig 5.8).

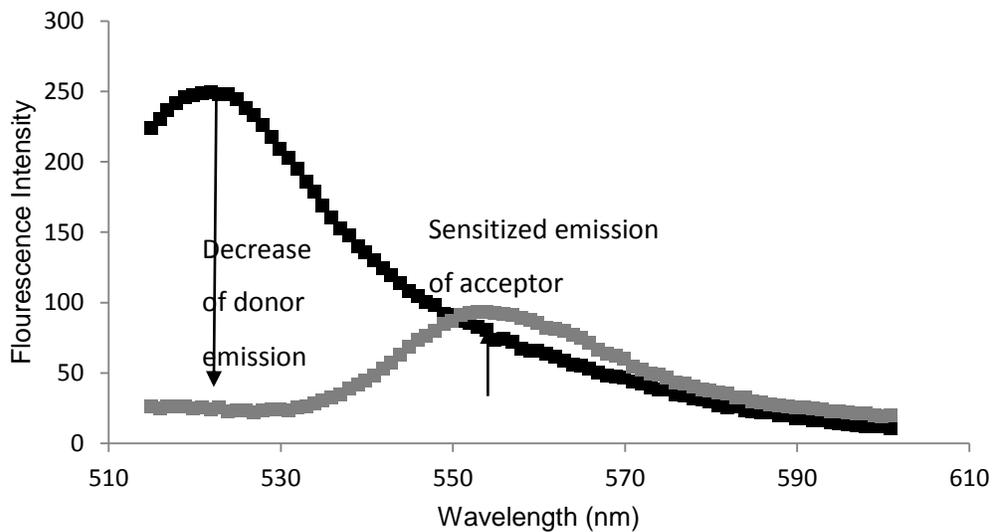


Figure 5.8: (■) Emission intensity spectrum of SNAP-PNA1-Atto 488 was excited at 501 nm. (■) decreasing the emission intensity of SNAP-PNA1-Atto 488 after addition the same ratio of acceptor (SNAP-PNA3-Atto532) which was also excited at 501 nm.

Four replicate measurements (Fig 5.9) demonstrated similar sensitized emission intensity. The donor region (514- 525 nm) appeared to be nearly completely quenched in all but one of the four samples with any residual fluorescence close to instrument offset. The dashed line (Fig 5.9) was the only one which showed the structure of donor fluorescence and exhibited $90\pm 1\%$ energy transfer over the range of 514-525 nm which was less than the other three, perhaps due to an excess of Atto488. The average FRET efficiency of the system based on all 4 graphs of Fig 5.7 was estimated to be about 93 ± 2 percent which confirms the assembly of two monomers separated by approximately 4.1 ± 0.2 nm ($=0.64 R_0$) and efficient energy transfer between donor and acceptor. This is in reasonable agreement with homo-FRET data in section 5.3.2 which would predict $\sim 94\%$ energy transfer between two fluorophores when they are located in $< 0.8 R_0 = 4$ nm distance from each other in dimer assembly.

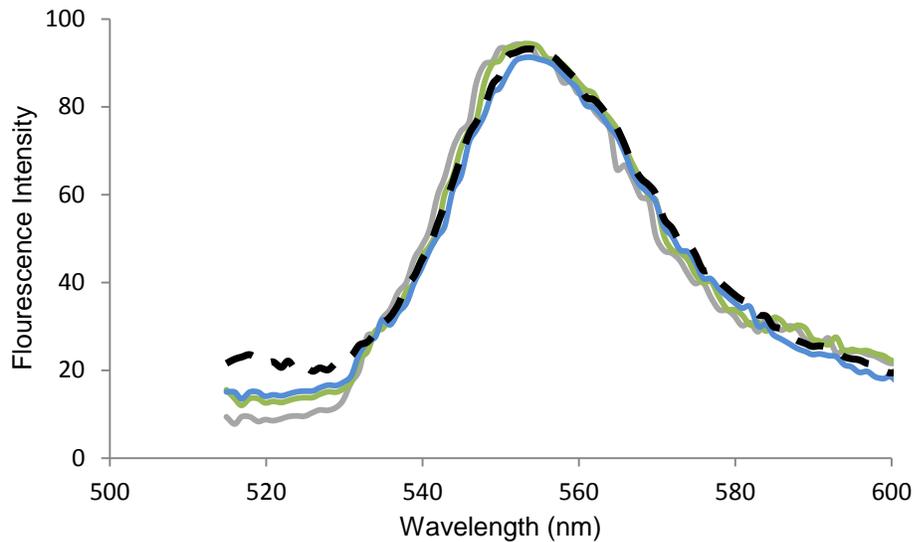


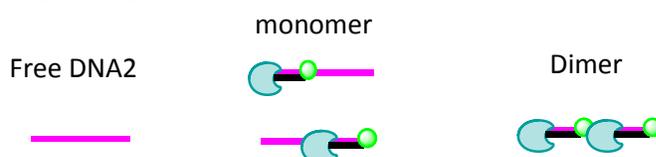
Figure 5.9: Replicates observed for sensitized emission intensity (■ in Fig 5.6), only the dashed graph shows the structure of donor spectrum over the range of 514-525 nm an. The rest look like offsets over that range.

A PNA-DNA double helix (16 base pairs per turn and 28° twist and 3.3 angstrom rise³⁰) demonstrates different conformation compared to regular B-DNA. A four base gap (GATC) was engineered between each monomer location on the DNA scaffold in our system which makes the helix conformation more complex. The estimated distance between Atto488 and Atto532 in the dimer assembly was about 3.7 nm based on summation of the lengths of 4 bases (GATC) gap, 6 PNA-DNA base pairs and a lysine residue (~ 4 Angstrom³¹). This distance assumes two fluorophores in the PNA-DNA duplex (3.7 nm) are on the same side. The estimated distance based on FRET (4.1 nm) suggests a twist in PNA-DNA helix introducing a slight increase in the spatial position of fluorophores with respect to each other in dimer assembly.

5.2.3 Homo-FRET system

5.2.3.1 Anisotropy measurement of Homo-FRET system indicating the assembly in dimer and higher oligomer forms

A range of DNA1 and DNA3 and DNA4 concentrations (0–1 μM) were added to 1 μM of SNAP-PNA1-Atto488 to study anisotropy changes of the system. In comparison to the measured anisotropy for SNAP-PNA1-Atto488 (monomer) (0.237 ± 0.002), the maximum anisotropy observed for each system was at 1:1 ratio DNA:monomer. The measured anisotropies of DNA1:monomer, DNA3:monomer and DNA4:monomer were 0.255 ± 0.002 , 0.275 ± 0.002 , and 0.264 ± 0.002 respectively, which is in accord with the slight increase of rotational correlation time of the larger size of the assemblies. It should be noted that 1:1 ratio represents the stoichiometry and the mixture of randomly assembled species existed in this ratio is more complex. For example, at 1:1 stoichiometry, DNA2:monomer can exist in four forms and a randomly assembled systems will include some amount of free DNA2, monomer, and dimer species as shown in Scheme 5.3.



Scheme 5.3: a distribution of randomly assembled species at 1:1 stoichiometry of DNA2:monomer.

As mentioned in previous chapters, fluorescence anisotropy measurement is an indicator of homo-energy transfer generated after assembly. The following equation gives an approximation of the expected anisotropy of an oligomer undergoing homo-FRET when the inter-fluorophore distance is $< 0.8 R_0$. r_{oligomer} , r_{monomer} , are the anisotropy in oligomer and monomer forms, and n is the number of fluorophores in the oligomer.³²

$$r_{\text{oligomer}} \cong \frac{r_{\text{monomer}}}{n} \quad [5.1]$$

Depending on the orientation, the distance between fluorophores, and the cluster size, this approximation may not hold rigorously and more detailed treatment required.³²⁻³⁴

A decrease of measured anisotropy after adding more monomer to the complementary DNA templates confirmed production of homo-FRET systems. Titration of SNAP-PNA-Atto488 with DNA1 resulted in a gradual decrease of anisotropy compared to DNA1:monomer which reaches maximum decrease of 52% at 2:1 ratio of SNAP-PNA-Atto488:DNA1. Since the distance between two complementary parts for SNAP-PNA-Atto488 constructs on DNA1 scaffold is less than the $0.8 R_0$ for Atto488 in homo-transfer ($R_0=5\text{nm}$) eqn 5.1 predicts 50% depolarization which is very close to the measured decrease. Upon further addition of SNAP-PNA-Atto488 the anisotropy gradually increased (Fig 5.10).

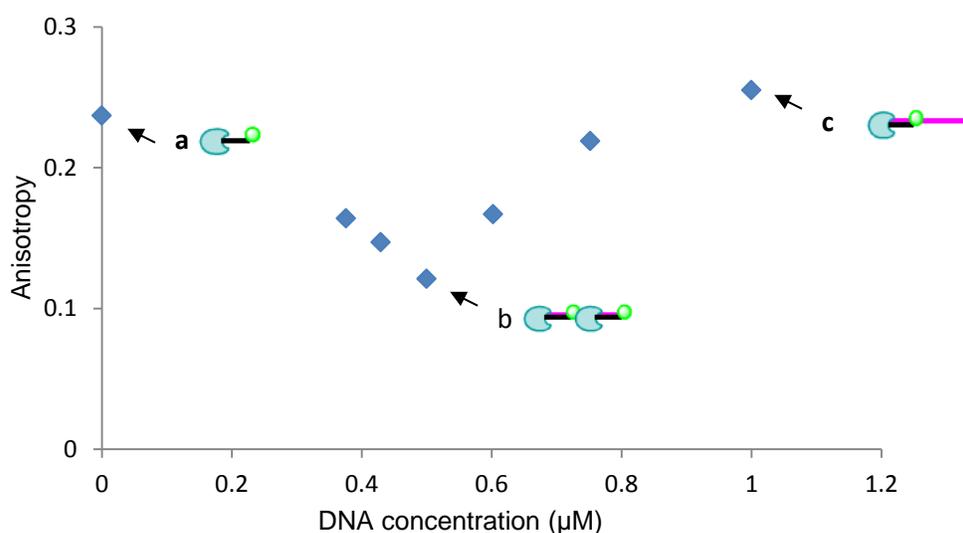
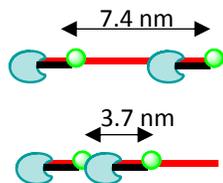


Figure 5.10: Anisotropy changes of the SNAP-PNA-Atto488:DNA1 homo FRET system after adding different concentration of DNA1 (0-1 μM) to 1 μM SNAP-PNA-Atto488.:a) monomer, b) 0.5 μM : DNA1:dimer, and c) 1 μM : DNA1:monomer. The assembly illustrated at point c represents the stoichiometry. The mixture of randomly assembled species in these locations is more complex (see Scheme 5.4).

In template directed oligomer formation via DNA3, a 48% decrease of anisotropy was obtained with a 2:1 ratio of SNAP-PNA-Atto488:DNA3. The maximum decrease of 48% was observed

when a 2:1 ratio of SNAP-PNA-Atto488 to DNA3 was applied to make DNA3:dimer (Fig 5.11). This value is close to the approximately 50% decrease predicted by eqn 5.1 but less than DNA2:dimer. As mentioned in chapter 3, this might be due to the distribution of different probable assemblies with different distances between SNAP-PNA-Atto488 monomers on DNA3 scaffold (the probable inter-fluorophore distances could be between 3.7-7.4 nm, Scheme 5.4).



Scheme 5.4: an example of different Inter-fluorophore distances (3.7-7.4 nm) between two monomer on DNA3 which has 3 locations to bind (at 2:1 stoichiometry of SNAP-PNA-Atto488:DNA3, the expected randomly assembled species is more than the above two species).

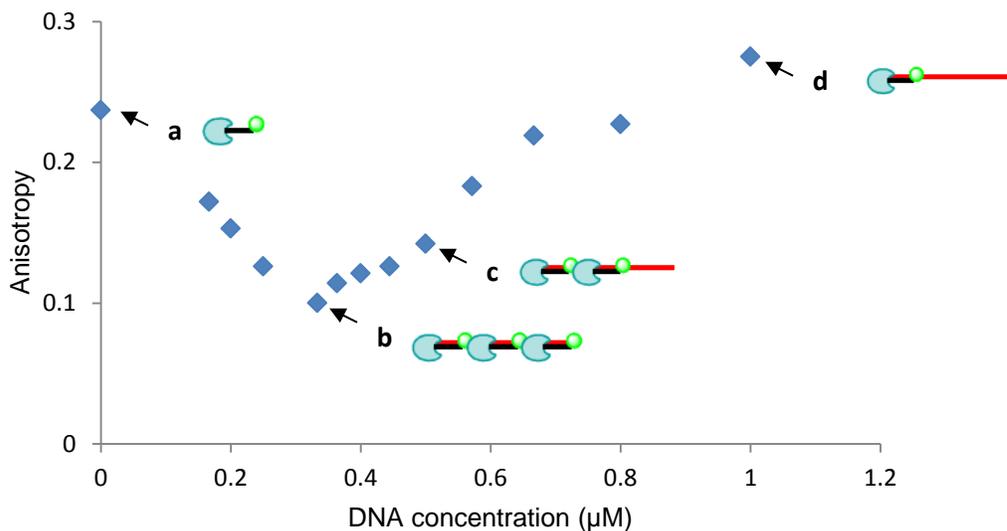


Figure 5.11: Anisotropy changes of the DNA3:trimer homo FRET system after adding (0-1 µM) concentration of DNA3 to 1 µM SNAP-PNA-Atto488: a) monomer b) 0.33 µM: DNA3: trimer , c) 0.5 µM: DNA3: dimer , and d) 1 µM: DNA3: monomer. The assemblies illustrated at points c and d represent the stoichiometry. The mixture of randomly assembled species in these locations is more complex.

For 3:1 ratio of SNAP-PNA-Atto488 to DNA3 64% decrease was observed which is close to the predicted decrease given by eqn 5.1 (Fig 5.11).

Similarly, maximum 37% anisotropy decrease was observed when 2:1 ratio of SNAP-PNA-Atto488 to DNA4 was applied to make DNA4:dimer (Fig 5.12) which is less than predicted by eqn 5.1. However, DNA4 has 4 locations to bind, and with two SNAP-PNA-Atto488 attached, the inter-fluorophore distances between monomer units is expected to range from 3.7-11 nm. The larger distances are outside $0.8 R_0$ resulting in less change in anisotropy than seen in DNA2 and DNA3 assemblies. Also 57% decrease for 3:1 ratio of SNAP-PNA-Atto488 to DNA4 was observed. The lowest anisotropy measured for tetramer assembly on DNA scaffolds was 0.088 ± 0.002 in the solution featuring 4:1 ratio of SNAP-PNA-Atto488 monomer to DNA4 showing 67% anisotropy decrease which is slightly less than the predicted 75% decrease of anisotropy by eqn 5.1 due to the mixture of randomly assembled species in system.

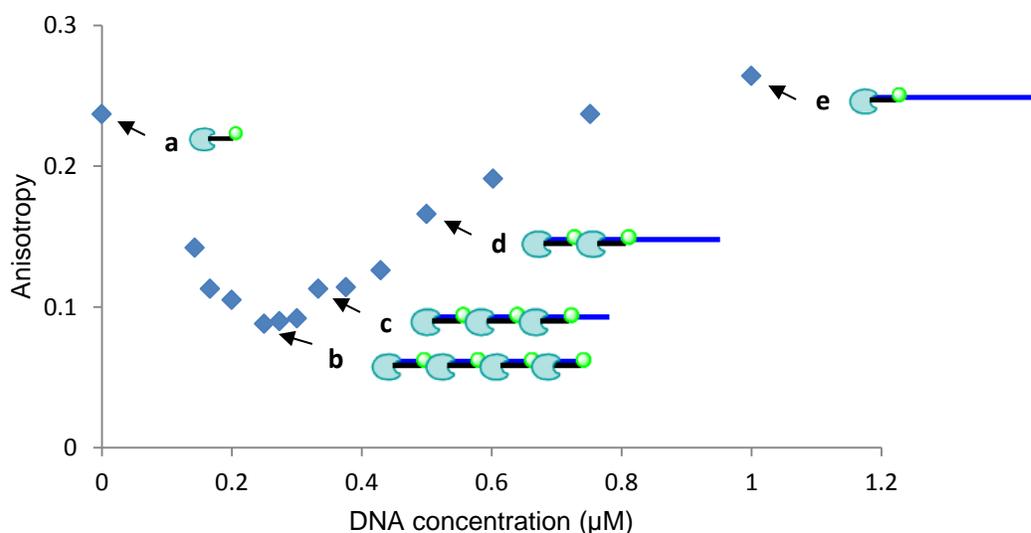


Figure 5.12: Anisotropy changes of the DNA4:tetramer homo FRET system after adding different concentration of DNA4 (0-1 μM) to 1 μM SNAP-PNA-Atto488: a) monomer, b) 0.25 μM : DNA4: tetramer, c) 0.33 μM : DNA4: trimer, d) 0.5 μM : DNA4: dimer, and e) 1 μM : DNA3: monomer. The assemblies illustrated at points c, d and e represent the stoichiometry. The mixture of randomly assembled species in these locations is more complex.

As Described in chapter3, the PNA-directed fluorescent SNAP protein assembled system can be used to induce aggregation of other proteins of interest by fusion to SNAP-tag through molecular cloning techniques. It offers site-selective modification of SNAP-tagged fused protein with a vast variety of organic flourophores and other biophysical probes. The straightforward photophysical interpretation of the assembled systems confirmed the benefit of homo- and hetero-FRET measurements to study protein clustering behaviour.

5.3 Conclusion and outlook

In conclusion, model systems incorporating controllable assembly of SNAP-PNA conjugates on DNA frameworks were demonstrated in this chapter. Taking advantage of the unique recognition capability of PNA for complementary DNA scaffolds, fluorescent SNAP-PNA units were precisely assembled in dimer and higher oligomer forms. The assemblies were characterised by HPLC, SDS-PAGE and fluorescent techniques. Two dimer forms were created in this chapter using DNA scaffolds. The first one was a homo dimer composed of two SNAP-PNA1-Atto488 constructs. The second one was a hetero-dimer of SNAP-PNA1-Atto488 and SNAP-PNA3-Atto532. SEC-HPLC and SDS-PAGE showed distinct bands and peaks related to each dimer species. Photophysical studies of both systems confirmed FRET occurred between the fluorophores. Assembly of SNAP-PNA1-Atto488 into trimer and tetramer forms was also confirmed with HPLC and SDS-PAGE. The anisotropy measurement of homo-dimer and other homo-oligomers showed a clear evidence of homo-FRET among the flourophores. The induced hetero- and homo-FRET systems based on fluorescent SNAP-PNA will provide a useful tool to study hetero- and homo oligomerization of interest proteins which can be fused to a SNAP tag *in vitro* or *in vivo*.

5.4 Experimental part

5.4.1 Hybridization and assembly of fluorescent SNAP-PNA with DNA

Four DNA scaffolds (Life Technologies, USA) were used to hybridize to fluorescent SNAP-PNA constructs prepared in chapter 4. DNA1 (5'- TGCATGGATCTGCATG-3') was used as template to assemble two SNAP-PNA1-Atto488 constructs in dimer form. In the same way, DNA2 (5'- TGCATGGATCGTTACT-3') was applied to make hetero-dimer of SNAP-PNA1-Atto488 and SNAP-PNA3-Atto532. DNA3 (5'-TGCATGGATCTGCATGGATCTGCATG3') and DNA4 (5'- TGCATGGATCTGCATGGATCTGCATGGATCTGCATG-3') scaffolds were both used to assemble SNAP-PNA1-Atto488 constructs in trimer and tetramer forms, respectively. The titration of 0.5 μM DNA1, 0.33 μM DNA3, and 0.25 μM DNA4 with different concentrations of SNAP-PNA-Atto488 (0-1 μM) was carried out in phosphate buffer (100 mM, NaCl 200 mM, pH 7) with 2 hrs incubation at room temperature. The results were assessed by SDS-PAGE electrophoresis and SEC-HPLC at (214 nm and 501 and 532 nm) as described in previous chapters. The SEC-HPLC calibrated with molecular weight protein marker kit 12-200 kDa (MWGF200-sigma).

5.4.2 SDS-PAGE electrophoresis analysis

SDS-PAGE electrophoresis of SNAP-PNA-Atto 488 and 532 conjugates SNAP-PNA-Atto488 and SNAP-PNA-Atto 532 conjugates samples (10 μM) were heated at 90°C for 5 min and mixed with the same equivalent of SDS sample loading buffer (Biorad) on precast acrylamide gels (Biorad) (12% resolving polyacrylamide gel and 5% stacking gel). The electrophoresis ran for 70 min at 100 V. To stain the protein bands Brilliant coomassie blue (Sigma-Aldrich) was used.

5.4.3 SEC HPLC analysis

SEC-HPLC of SNAP-PNA-Atto 488 and SNAP-PNA-Atto 532 conjugates The SEC-HPLC analysis was carried out as described in Chapter3. The detection wavelengths were 214, 501 (maximum absorption wavelength of Atto 488) and 532 nm (maximum absorption wavelength of Atto 532).

5.4.4 Photophysical measurements

Different concentration of DNA1, DNA3 and DNA4 (0-1 μ M) were added to 1 μ M SNAP-PNA1-Atto488 in phosphate buffer (100 mM, NaCl 200 mM, pH 7) with 2 hrs incubation at room temperature. The anisotropy was recorded using a multimode microplate reader (Infinite F200 PRO, Tecan Group Ltd.). The excitation and emission wavelengths were 485 and 535nm, respectively.

Similarly, 1 μ M of SNAP-PNA1-Atto 488 and SNAP-PNA3-Atto 532 were added to the same concentration of DNA2. The intensity measurement was carried out using fluorimeter (Cary Eclipse; Varian) over the range of 515-600 nm when the solution was excited at 501 nm which is the maximum absorption wavelength of Atto 488 .

5.5 References

1. A. Goddard and A. Watts, *Biophysical Reviews*, 2012, **4**, 291-298.
2. Y. D. Paila, M. Kombrabail, G. Krishnamoorthy and A. Chattopadhyay, *The Journal of Physical Chemistry B*, 2011, **115**, 11439-11447.
3. G. A. Farr, M. Hull, I. Mellman and M. J. Caplan, *The Journal of Cell Biology*, 2009, **186**, 269-282.
4. L. E. T. Jansen, B. E. Black, D. R. Foltz and D. W. Cleveland, *The Journal of Cell Biology*, 2007, **176**, 795-805.
5. T. Komatsu, K. Johnsson, H. Okuno, H. Bito, T. Inoue, T. Nagano and Y. Urano, *Journal of the American Chemical Society*, 2011, **133**, 6745-6751.
6. X. Sun, A. Zhang, B. Baker, L. Sun, A. Howard, J. Buswell, D. Maurel, A. Masharina, K. Johnsson, C. J. Noren, M.-Q. Xu and I. R. Correa, *ChemBioChem*, 2011, **12**, 2217-2226.
7. M. A. Brun, K.-T. Tan, E. Nakata, M. J. Hinner and K. Johnsson, *Journal of the American Chemical Society*, 2009, **131**, 5873-5884.
8. C. Donovan and M. Bramkamp, *Microbiology*, 2009, **155**, 1786-1799.
9. A. F. Hussain, F. Kampmeier, V. von Felbert, H. F. Merk, M. K. Tur and S. Barth, *Bioconjugate Chemistry*, 2011, **22**, 2487-2495.
10. D. Srikun, A. E. Albers, C. I. Nam, A. T. Iavarone and C. J. Chang, *Journal of the American Chemical Society*, 2010, **132**, 4455-4465.
11. A. Regoes and A. B. Hehl, *Biotechniques*, 2005, **39**, 809-812.
12. C.-J. Zhang, L. Li, G. Y. J. Chen, Q.-H. Xu and S. Q. Yao, *Organic Letters*, 2011, **13**, 4160-4163.
13. S. Chattopadhyaya, R. Srinivasan, D. S. Y. Yeo, G. Y. J. Chen and S. Q. Yao, *Bioorganic & Medicinal Chemistry*, 2009, **17**, 981-989.
14. C. Campos, M. Kamiya, S. Banala, K. Johnsson and M. González-Gaitán, *Developmental Dynamics*, 2011, **240**, 820-827.
15. T. Gronemeyer, G. Godin and K. Johnsson, *Protein Technologies and Commercial Enzymes*, 2005, **16**, 453-458.
16. P. A. Pellett, X. Sun, T. J. Gould, J. E. Rothman, M.-Q. Xu, I. R. Corrêa and J. Bewersdorf, *Biomed. Opt. Express*, 2011, **2**, 2364-2371.

17. D. Maurel, L. Comps-Agrar, C. Brock, M. L. Rives, E. Bourrier, M. A. Ayoub, H. Bazin, N. Tinel, T. Durroux, L. Prezeau, E. Trinquet and J. P. Pin, *Nature Methods*, 2008, **5**, 561-567.
18. A. Tirat, F. Freuler, T. Stettler, L. M. Mayr and L. Leder, *International Journal of Biological Macromolecules*, 2006, **39**, 66-76.
19. C. S. Lisenbee, S. K. Karnik and R. N. Trelease, *Traffic*, 2003, **4**, 491-501.
20. E. Snapp, in *Current Protocols in Cell Biology*, John Wiley and Sons, Inc., Editon edn., 2005.
21. A. Gautier, A. Juillerat, C. Heinis, I. R. Corrãa Jr, M. Kindermann, F. Beaufils and K. Johnsson, *Chemistry & Biology*, 2008, **15**, 128-136.
22. F. Ciruela, J. P. Vilardaga and V. Fernandez-Duenas, *Trends in Biotechnology*, 2010, **28**, 407-415.
23. M.-L. Prezeau L Fau - Rives, L. Rives MI Fau - Comps-Agrar, D. Comps-Agrar L Fau - Maurel, J. Maurel D Fau - Kniazeff, J.-P. Kniazeff J Fau - Pin and J. P. Pin, *Current Opinion in Pharmacology*, 2010, **10**, 6-13
24. T.-R. Xu, R. J. Ward, J. D. Pediani and G. Milligan, *Biochemical Journal*, 2011, **439**, 171-183.
25. E. Doumazane, P. Scholler, J. M. Zwier, E. Trinquet, P. Rondard and J.-P. Pin, *The FASEB Journal*, 2010, **25**, 66-77.
26. J. P. Pin, L. Comps-Agrar, D. Maurel, C. Monnier, M. L. Rives, E. Trinquet, J. Kniazeff, P. Rondard and L. Prézeau, *The Journal of Physiology*, 2009, **587**, 5337-5344.
27. A. Masharina, L. Raymond, D. Maurel, K. Umezawa and K. Johnsson, *Journal of the American Chemical Society*, 2012, **134**, 19026-19034.
28. O. Cottet M Fau - Faklaris, D. Faklaris O Fau - Maurel, P. Maurel D Fau - Scholler, E. Scholler P Fau - Doumazane, E. Doumazane E Fau - Trinquet, J.-P. Trinquet E Fau - Pin, T. Pin Jp Fau - Durroux and T. Durroux, *Frontiers in Endocrinology*, 2012, **3**, 1664-2392.
29. H. Haruki, M. R. Gonzalez and K. Johnsson, *PLoS ONE*, 2012, **7**, e37598.
30. V. Menchise, G. De Simone, T. Tedeschi, R. Corradini, S. Sforza, R. Marchelli, D. Capasso, M. Saviano and C. Pedone, *Proceedings of the National Academy of Sciences of the United States of America*, 2003, **100**, 12021-12026.
31. M. Carrion-Vazquez, A. F. Oberhauser, T. E. Fisher, P. E. Marszalek, H. Li and J. M. Fernandez, *Progress in Biophysics and Molecular Biology*, 2000, **74**, 63-91.
32. L. W. Runnels and S. F. Scarlata, *Biophysical Journal*, 1995, **69**, 1569-1583.
33. I. Gautier, M. Tramier, C. Durieux, J. Coppey, R. B. Pansu, J. C. Nicolas, K. Kemnitz and M. Coppey-Moisan, *Biophysical Journal*, 2001, **80**, 3000-3008.
34. A. N. Bader, S. Hoetzel, E. G. Hofman, J. Voortman, P. M. P. van Bergen en Henegouwen, G. van Meer and H. C. Gerritsen, *ChemPhysChem*, 2011, **12**, 475-483.

Summary

The functional DNA mimic peptide nucleic acid (PNA) provides superb control properties over the dynamics of assembled systems. Precise recognition of PNA for complementary DNA or PNA templates enables self-assembly of PNA-protein conjugates in a well-defined programmed manner. The PNA-guided assemblies developed in this work provide scaffolding for adjusting the number, distance and distribution of proteins using chemical and molecular biology. This model system can be used to induce aggregation of proteins of interest which are modified with tags by molecular or chemical biology in living systems. Such a programmed self-assembly model system may bridge the gap between studies on isolated proteins that cannot account for protein clustering in the native environment and whole cell studies, which do not allow currently the controlled actuation of the nano-clusters.

Plasma membrane-resident signaling proteins such as glycosylphosphatidylinositol (GPI) anchored proteins and EGFR show complex dynamic ordering, such as clustering, on a nanoscale level which typically changes after external stimulation.¹⁻¹⁰ Controlling the distribution of plasma membrane-resident signaling proteins will allow the study of relevant cellular signaling mechanisms that depend on complex nanoscale molecular ordering. Inducing nano-scaffolds that would bind multiple membrane proteins in a defined way into a cluster has been employed previously but they have some limitations including incomplete control over the location and number of proteins in clusters.¹¹⁻¹³ The PNA-directed nano-clustering of proteins described here may provide a well-defined model system to feature controlled levels of aggregation and surface density at the plasma membrane which has not, to our knowledge, been reported so far.

In summary, the aim of this research was to use the unique characteristics of PNA to create controllable protein assemblies directed by precise PNA-DNA hybridization, studying the photo-

physical signatures of this self-assembled model system via FRET and providing a powerful tool for live system imaging. Fluorescent protein (FP) and SNAP tag groups were chosen as two popular protein tags to conjugate to PNA and assemble in dimeric and higher order oligomeric forms.

In Chapter 2, expressed protein ligation (EPL) provided a site-selective and facile conjugation of fluorescent proteins to PNA to create units which can be assembled through specific recognition of PNA for DNA frameworks. In EPL, an intein-mediated protein expression and purification procedure was employed to create fluorescent proteins bearing a C-terminal thioester. The modified FPs were able to conjugate to the thiol group of a cysteine residue at the N-terminus of a PNA using mercaptophenylacetic acid (MPAA) as a catalyst. Different modified fluorescent proteins with C-terminal thioesters were expressed and purified. Mass spectrometry and SDS-PAGE analysis confirmed the final proteins purity. Among those, the monomeric teal fluorescent protein (mTFP) was chosen for conjugation to PNA and subsequent assembly studies due to the greater brightness and photostability. The conjugation was assessed by mass spectrometry and spectrophotometry and showed complete conversion of the FP to a ligated form.

Assembly of mTFP-PNA units on DNA frameworks can provide a model system mimicking protein aggregation which can be studied by FRET techniques. Such a model system was described in Chapter 3. To prove the principle of PNA directed inducible assembly, a DNA beacon with 6-FAM and Dabcyl at its ends was used to assemble a hetero-FRET system with the mTFP-PNA conjugate. Using fluorescence techniques such as intensity, frequency domain lifetime and anisotropy measurements, the assembled system exhibited decreased donor intensity, changes in frequency domain lifetime, and increased anisotropy as an indicator of hetero-FRET. Different DNA scaffolds allowing the alignment of multiple mTFP-PNA in a controllable manner provided model systems exhibiting homo-FRET. Efficient assembly of protein in dimeric and oligomeric forms on the DNA-PNA frameworks was confirmed with size exclusion chromatography (SEC) and SDS-PAGE. The assemblies gave clear and readily

interpreted photophysical signatures providing a practical tool to investigate the behaviour of protein clusters. The well-defined characteristics of this model system can be extended to any other protein through fusion to FPs.

Self-labeling tags such as the SNAP-tag are smaller than FPs and can be easily modified with organic fluorophores providing a broad range of colours with an extended range of photophysical properties and functionalities. The versatility and efficiency of SNAP-tag labeling makes it a promising method to localize and study fusion protein behaviors in living systems. In Chapter 4, the SNAP-PNA conjugate was created as an alternative assembly unit. The ligation of SNAP-PNA required modification of a PNA with O⁶-benzyl guanine which is a specific substrate for SNAP. To create FRET systems, the modified PNA was labeled with Atto dyes. Successful modification of PNA with O⁶-benzyl guanine and Atto dyes was confirmed by mass spectrometry and reverse phase HPLC. The efficient coupling of a BG-PNA-Atto dye construct with SNAP was performed with high specificity and almost complete conversion of SNAP to the conjugated form. SDS-PAGE, SEC-HPLC, mass spectrometry and fluorescent measurements all verified generation of the final conjugates.

In Chapter 5, the controllable assembly of fluorescent SNAP-PNA units on DNA scaffolds to create dimeric and higher order oligomeric forms was discussed. The assemblies were characterized by SEC-HPLC, SDS-PAGE and fluorescent techniques. Different DNA scaffolds were used to create hetero-dimer, homo-dimers and homo-oligomers. Photophysical studies of assembled systems confirmed FRET occurred between the different fluorophores in the hetero-dimer (hetero-FRET) and among similar fluorophores in homo-dimers and homo-oligomers (homo-FRET). This model system can be used to study clustering behavior of other proteins expressing the SNAP sequence.

Further work could readily extend these systems from homodimers and oligomers to binary, ternary, and higher oligomer systems containing any number of different dyes or fluorescent proteins in precisely engineered arrangements.

Fluorescent proteins and fluorescently labeled SNAP-tags have been widely used as markers for visualization of cell membrane protein clusters using FRET combined with fluorescent microscopy.¹⁴⁻¹⁷ To investigate and address unresolved questions in regard to the role of cell membrane proteins (e.g. EGFR cluster formation and activation), gaining external control on clustering across cell membrane could be informative and helpful.¹⁸ Variations of the model systems developed in this research are capable of directed assembly of proteins which may provide a new insights into studying the mechanism of cell membrane proteins clustering, leading to new methods for regulation of signaling pathways which rely on biophysical and chemical inputs.

To address this goal, in collaboration with Max Planck Institute of Molecular Physiology in Dortmund, the developed SNAP-PNA:DNA assembled model system in Chapter 4 and 5 were applied to controllably assemble dimeric forms of EGFR *in vivo*. Therefore, the BG-PNA-Atto488 constructs (Chapter 4) were used to modify EGFR –or EGFR transmembrane domain (TMD)-SNAP tag fusion proteins (SNAP-EGFR-mCherry, SNAP-TMD-mCherry and SNAP-TMD-mCitrine) which were expressed in transfected MCF-7 cell line (breast cancer cell line). Thereafter, DNA1 template (Chapter 5) was added to the transfected cells to assemble the modified PNA-SNAP-EGFR in dimer forms. The fluorescent anisotropy imaging and total internal reflection fluorescence (TIRF) microscopy were used to analyse labelling and assembly efficiency. The initial result was not that conclusive and more control experiments will be needed to optimize the labelling and assembly. Unfortunately, lack of time did not allow further progress within the timeline of this research but due to the important goals behind approach, it is highly suggested as a future work.

6.1 References

1. A. H. Clayton, F. Walker, S. G. Orchard, C. Henderson, D. Fuchs, J. Rothacker, E. C. Nice and A. W. Burgess, *Journal of Biological Chemistry*, 2005, **280**, 30392-30399.
2. L. t. Comps-Agrar, J. Kniazeff, C. Brock, E. Trinquet and J.-P. Pin, *FASEB Journal*, 2012, **26**, 3430-3439.
3. D. Calebiro, F. Rieken, J. Wagner, T. Sungkaworn, U. Zabel, A. Borzi, E. Cocucci, A. Zurn and M. J. Lohse, *Proceedings of the National Academy of Sciences*, 2013, **110**, 743-748.
4. D. Maurel, L. Comps-Agrar, C. Brock, M. L. Rives, E. Bourrier, M. A. Ayoub, H. Bazin, N. Tinel, T. Durrourx, L. Prezeau, E. Trinquet and J. P. Pin, *Nature Methods*, 2008, **5**, 561-567.

5. T. W. Gadella, Jr. and T. M. Jovin, *Journal of Cell Biology*, 1995, **129**, 1543-1558.
6. J. Ichinose, M. Murata, T. Yanagida and Y. Sako, *Biochemical and Biophysical Research Communications*, 2004, **324**, 1143-1149.
7. L. Albizu, M. Cottet, M. Kralikova, S. Stoev, R. Seyer, I. Brabet, T. Roux, H. Bazin, E. Bourrier, L. Lamarque, C. Breton, M.-L. Rives, A. Newman, J. Javitch, E. Trinquet, M. Manning, J.-P. Pin, B. Mouillac and T. Durroux, *Nature Chemical Biology*, 2010, **6**, 587-594.
8. G. Milligan, *Molecular Pharmacology*, 2013, **84**, 158-169.
9. L. Pin Jp Fau-Comps-Agrar, D. Comps-Agrar L Fau-Maurel, C. Maurel D Fau - Monnier, M. L. Monnier C Fau - Rives, E. Rives Ml Fau-Trinquet, J. Trinquet E Fau-Kniazeff, P. Kniazeff J Fau-Rondard, L. Rondard P Fau - Prezeau and L. Prezeau, *Journal of Physiology*, 2009, **587**, 5337-5344.
10. M.-L. Prezeau L Fau - Rives, L. Rives Ml Fau - Comps-Agrar, D. Comps-Agrar L Fau - Maurel, J. Maurel D Fau - Kniazeff, J.-P. Kniazeff J Fau - Pin and J. P. Pin, *Current Opinion in Pharmacology*, 2010, **10**, 6-13
11. K. G. N. Suzuki, T. K. Fujiwara, F. Sanematsu, R. Iino, M. Edidin and A. Kusumi, *Journal of Cell Biology*, 2007, **177**, 717-730.
12. A. T. Reynolds, C ; Verveer, PJ ; Rocks, O ; Bastiaens, PIH, *Nature Cell Biology*, 2003, **5**, 447-453.
13. P. J. W. Verveer, F S ; Reynolds, A R ; Bastiaens, P I, *Science (New York, N.Y.)*, 2000, **290**, 1567-1570.
14. A. N. Bader, E. G. Hofman, J. Voortman, P. M. en Henegouwen and H. C. Gerritsen, *Biophysical Journal*, 2009, **97**, 2613-2622.
15. N. Kozer, D. Barua, S. Orchard, E. C. Nice, A. W. Burgess, W. S. Hlavacek and A. H. A. Clayton, *Molecular BioSystems*, 2013, **9**, 1849-1863.
16. K. Noga, H. Christine, T. J. Jacob, C. N. Edouard, W. B. Antony and H. A. C. Andrew, *Physical Biology*, 2011, **8**, 066002.
17. A. Sorkin, M. McClure, F. Huang and R. Carter, *Current Biology : CB*, 2000, **10**, 1395-1398.
18. D. Stabley, S. Retterer, S. Marshall and K. Salaita, *Integrative Biology*, 2013, **5**, 659-668.