

Templated self-assembly of the bacterial flagellar motor torque ring *in vitro*

Samuel E Tusk

St. Hilda's College
University of Oxford

*A thesis submitted for the degree of
Doctor of Philosophy*

Trinity 2016

Abstract

The Bacterial Flagellar Motor is a self-assembled, ion-driven rotary motor, capable of rotating at >1000Hz, switching direction within ms, and continually rebuilding itself to tune performance and sensitivity.

Motor self-assembly begins with a membrane-embedded ring of the protein FliF, which templates the sequential assembly of FliG and dynamic FliM₁:FliN₃ subunits. FliG, FliM and FliN comprise the C ring: site of torque generation and motor switching. Averaged cryo-EM structures of purified motors feature an unexplained symmetry mismatch between the FliF ring (~26-fold) and C-ring (~34-fold), and stoichiometries of both FliG and FliM₁:FliN₃ subunits are contested. A recent domain-swap polymerization model for FliG may explain the symmetry mismatch.

This thesis describes attempts to template short (1-mer to 5-mer) FliG oligomers *in vitro* by substituting the FliF ring with a variety of rationally-designed DNA templates. Variation of FliG stoichiometry with template design could be used to test the domain-swap polymerization model.

Template design and characterization are described, along with purification of FliG-fluorophore conjugates and measurement of labelling fraction. Single-molecule fluorescence assays are developed for clear counting of FliG stoichiometry via bleaching steps, and counting with native PAGE is also demonstrated. Multiple DNA-FliG conjugation strategies are described. 5-NTA DNA oligos bind 10xHis-FliG more stably than previously described 3-NTA oligos, but do not reliably template proximal FliGs. Covalent DNA-FliG conjugation rectifies this, allowing controlled assembly of FliG on templates.

This opens the door to testing domain-swap-polymerization, and is the first step to templating a fully-functional C-ring *in vitro*, both as a platform for *in vitro* study and a tool to study self assembly from the bottom-up. This represents a new approach to the study of large protein complexes.

Templated self-assembly of the bacterial flagellar motor torque ring *in vitro*



Samuel E Tusk
St. Hilda's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2016

Acknowledgements

Of the many people who contributed to this thesis, I'd first like to thank my supervisors Dr. Richard Berry and Prof. Andrew Turberfield, for their support, guidance, patience and rigour. In particular, thanks to Richard for allowing me so much freedom and to Andrew for providing a counter-balance, without which nothing would have gotten done. Thanks also to Andrew for reading the draft of this thesis.

I'd also like to thank the Berry and Turberfield groups for their help, support and friendship; in particular, Dr. Jon Bath who has been a constant source of invaluable advice, and a role model both professionally and personally. I'd also especially like to thank Joel Spratt, who has contributed more than anyone to the content of this thesis, and has been a joy to work with so closely.

I'm also grateful to the Physics support staff, and those from elsewhere in the Biophysics department who have assisted with various protocols and techniques. Likewise, the groups of Lawrence Lee and Keiichi Namba, who have been a pleasure to work with.

Thanks also to the EPSRC for funding me, the HFSP for funding the project, and also to the LSI DTC programme for their support and training.

Thanks to innumerable friends who have enriched the last three years of my life, including my assorted musical partners-in-crime, house-mates at 17 Warneford, and in particular the irreplaceable trio of Emma, Adam and Michelle.

Finally, thanks to my partner Bethan and my family; especially my parents, who have been so supportive since the beginning.

Abstract

The Bacterial Flagellar Motor is a self-assembled, ion-driven rotary motor, capable of rotating at $>1000\text{Hz}$, switching direction within ms, and continually rebuilding itself to tune performance and sensitivity.

Motor self-assembly begins with a membrane-embedded ring of the protein FliF, which templates the sequential assembly of FliG and dynamic FliM₁:FliN₃ subunits. FliG, FliM and FliN comprise the C ring: site of torque generation and motor switching. Averaged cryo-EM structures of purified motors feature an unexplained symmetry mismatch between the FliF ring (~ 26 -fold) and C-ring (~ 34 -fold), and stoichiometries of both FliG and FliM₁:FliN₃ subunits are contested. A recent domain-swap polymerization model for FliG may explain the symmetry mismatch.

This thesis describes attempts to template short (1-mer to 5-mer) FliG oligomers *in vitro* by substituting the FliF ring with a variety of rationally-designed DNA templates. Variation of FliG stoichiometry with template design could be used to test the domain-swap polymerization model.

Template design and characterization are described, along with purification of FliG-fluorophore conjugates and measurement of labelling fraction. Single-molecule fluorescence assays are developed for clear counting of FliG stoichiometry via bleaching steps, and counting with native PAGE is also demonstrated. Multiple DNA-FliG conjugation strategies are described. 5-NTA DNA oligos bind 10xHis-FliG more stably than previously described 3-NTA oligos, but do not reliably template proximal FliGs. Covalent DNA-FliG conjugation rectifies this, allowing controlled assembly of FliG on templates.

This opens the door to testing domain-swap-polymerization, and is the first step to templating a fully-functional C-ring *in vitro*, both as a platform for *in vitro* study and a tool to study self assembly from the bottom-up. This represents a new approach to the study of large protein complexes.

Contents

List of Figures	xi
List of Abbreviations	xiii
1 Background	1
1.1 The Bacterial Flagellar Motor	2
1.1.1 Bacterial Motility	2
1.1.1.1 Bacteria in liquids	2
1.1.1.2 Bacteria on surfaces	4
1.1.2 Why study the Bacterial Flagellar Motor?	6
1.1.3 Bacterial Flagellar Motor Overview	8
1.1.4 FliF ring	10
1.1.5 Aside: EM microscopy and the BFM	11
1.1.6 MS-ring C-ring interface	12
1.1.7 C-ring	14
1.1.8 Protein Composition	14
1.1.8.1 Symmetry Mismatch	15
1.1.8.2 FliG	17
1.1.8.3 FliG domain-swap polymerization model	23
1.1.8.4 FliM and FliN	23
1.1.8.5 C ring Turnover	30
1.1.9 Export Apparatus	33
1.1.10 Rod, Hook and Filament	35
1.1.11 Stators	36
1.1.12 Accessory proteins	39
1.2 Why template a C ring in vitro?	41
1.2.0.1 It avoids purification damage and allows observation of assembly	42
1.2.0.2 Superior techniques are available in vitro	42
1.2.0.3 The benefits of a controllable template	44
1.2.0.4 Drawbacks	46
1.3 DNA Nanotechnology	47

1.3.1	Overview	47
1.3.2	DNA nanostructure templates	49
1.4	Fluorescence techniques	54
1.5	Project Overview	58
1.5.1	Initial Plan	60
1.5.2	Chronological overview	63
2	Building Blocks	67
2.1	DNA templates	67
2.1.1	10nt Linear Templates	67
2.1.1.1	Original sequences	68
2.1.1.2	Redesign 1	71
2.1.1.3	10nt templates - second redesign	72
2.1.1.4	Summary	75
2.1.2	20nt Linear Templates	75
2.1.2.1	Utility	75
2.1.2.2	Assembly and suspected dimerization	75
2.1.3	DNA Tiles	76
2.2	Protein	78
2.2.1	Constructs	78
2.2.2	Expression, Purification and dye labeling	79
2.2.2.1	Expression	79
2.2.2.2	Purification and Dye labelling	80
2.2.3	Quantification of dye labelling	80
2.2.3.1	Canonical method	82
2.2.3.2	Cysteine accessibility - a dead end	85
2.2.3.3	Surface binding kinetics - speculative and complex	86
2.2.3.4	Tag co-localization - hopefully definitive	88
2.2.4	Stability and function	88
2.2.4.1	Protein Thermal Shift Assay	88
2.2.4.2	Recreating <i>in vivo</i> behavior <i>in vitro</i>	90
2.3	Protein-DNA conjugation	90
2.3.1	Overview	90
2.3.2	NTA-DNA	92
2.3.2.1	tris-NTA	92
2.3.2.2	Synthesis	92
2.3.2.3	Purification	94
2.3.2.4	Observations of binding	96
2.3.2.5	Pentakis-NTA	96

2.3.2.6	Competition with histag-histag binding	99
2.3.2.7	Conclusion and motivation for covalent methods . .	100
2.3.3	Covalent Methods	102
2.3.3.1	Guided Conjugation	102
2.3.3.2	Maleimide-cysteine	104
2.3.3.3	Purification	105
3	Single Molecule Microscopy	113
3.1	Hardware	114
3.1.1	Microscope	114
3.1.2	Flow chambers	114
3.2	Surfaces and experimental approach	117
3.2.1	Surface Development	119
3.2.1.1	BSA,BSA-biotin	119
3.2.1.2	PEG	119
3.2.1.3	Future approaches	120
3.2.2	Template distribution on surfaces	122
3.3	Software and counting pipeline	122
3.3.1	Version 1 - Intensity in ROI	123
3.3.1.1	Identifying Spots	123
3.3.1.2	Drift Correction	124
3.3.1.3	Generation of intensity traces	125
3.3.1.4	Counting	126
3.3.2	Version 2 - Gaussian Fitting	126
3.3.2.1	Identifying and Fitting Spots	127
3.3.2.2	Generation of intensity traces	127
3.3.2.3	Step Counting	128
3.4	Anti-bleaching system	128
3.5	Dye behaviour and counting rules	129
3.5.1	Counting large numbers of dyes	130
3.5.1.1	12-dye DNA structures	130
3.5.1.2	DNA tiles	131
3.6	Conclusion	132
3.6.1	Gel Electrophoresis: an alternative?	133

4	Results & Future Directions	135
4.1	Templated assembly experiments	136
4.1.1	FliG on 10nt templates tested by fluorescence counting . . .	136
4.1.2	10nt DNA templates with fluorescent linkers	137
4.1.3	FliG on 20nt templates tested by fluorescence counting . . .	139
4.1.4	FliG on 20nt templates tested by native PAGE	141
4.1.5	FliG on 10nt templates tested by native PAGE	142
4.1.6	Conclusions	144
4.2	Coda	144
Appendices		
A	Appendix	151
A.1	Buffers used	151
A.2	DNA sequences	152
A.3	Plasmid amplification	153
A.4	Protein expression, purification and dye labelling	154
A.4.1	Expression and lysis	154
A.4.2	Purification and labelling	155
A.5	BCA assay	156
A.6	Proteinase K digestion to measure dye labelling	156
A.7	Surface Modification	157
A.7.1	BSA surface	157
A.7.2	PLL-PEG surface	157
A.7.3	PEG vectabond surface	157
A.8	Anti-bleaching system	159
A.9	FliG assembly Gels	160
	References	161

List of Figures

1.1	Bacterial swimming strategies	5
1.2	BFM overview	8
1.3	BFM components	10
1.4	FliF-FliG fusion motors	14
1.5	Structural diversity of BFMs	15
1.6	FliG full-length crystal structure	18
1.7	FliG assorted cystal structures	20
1.8	FliG domain-swap polymerization	24
1.9	FliM and FliN assembly	28
1.10	Flagellar filament	37
1.11	Early history of DNA nanotechnology	48
1.12	DNA origami	50
1.13	DNA templating optically active particles, part 1	52
1.14	DNA templating optically active particles, part 2	53
1.15	DNA templating proteins, part 1	55
1.16	DNA templating proteins, part 2	56
1.17	Fluorescence techniques	59
1.18	Project summary - essential background	61
1.19	Testing domain-swap polymerization with templating	62
2.1	FliG spacing on DNA templates	68
2.2	10nt linear template designs	69
2.3	5 x 10nt template assembly	70
2.4	Poor assembly of gapped 10nt templates	71
2.5	10nt tempates - first redesign	72
2.6	SYBR green assay for template assembly	74
2.7	Assembly of 3 x 20nt templates and weak dimerization	76
2.8	DNA tile	77
2.9	FliG constructs	79
2.10	FliG purification	81
2.11	FliG concentration measurement with BCA assay	83
2.12	Proteinase K assay to test dye labelling	84

2.13	Measuring dye labelling efficiency through transient binding . . .	87
2.14	Protein Thermal Shift Assay	89
2.15	Tris-NTA modification of DNA	93
2.16	NTA-DNA purification	95
2.17	Observations of Tris-NTA binding	97
2.18	Pentakis-NTA DNA	98
2.19	FliG histag-histag interactions	101
2.20	Guided conjugation strategy	103
2.21	Limited proteolysis of FliG	104
2.22	DNA affinity column	107
2.23	PAGE purification	109
2.24	Ion exchange chromatography	111
2.25	Conjugate purification with Size Exclusion Chromatography	112
3.1	TIRF Microscope optical arrangement	115
3.2	Flow chamber construction	116
3.3	Experimental approach for fluorescence experiments	118
3.4	Surface development	121
3.5	Surface distribution of templates	123
3.6	Fluorescence counting pipeline - version 1	125
3.7	Fluorescence counting pipeline - version 2 (Gaussian fitting)	126
3.8	Anti-bleaching system	129
3.9	Dye behavior during bleaching	131
3.10	Bleaching traces with large numbers of dyes	132
4.1	FliG on 5x10nt template counted with fluorescence	137
4.2	DNA linkers on 5x10nt template counted with fluorescence	138
4.3	FliG on 3x20nt templates counted with fluorescence	140
4.4	FliG on 3x20nt templates counted via native PAGE	143
4.5	FliG on 5x10nt templates counted by native PAGE	145
4.6	Coda: Covalent FliG-DNA conjugates on 5x10nt templates counted by native PAGE	146

List of Abbreviations

PMF	Photon Motive Force
BFM	Bacterial Flagellar Motor
CW	Clockwise
CCW	Counter-Clockwise
T3SS	Type 3 secretion system
EM	Electron Micrography
NMR	Nuclear Magnetic Resonance
DLS	Dynamic Light Scattering
SAXS	Small Angle X-ray Scattering
GFP	Green Fluorescent Protein
SMF	Sodium Motive Force
FRET	Fluorescence Resonance Energy Transfer
bp	Base pair
nt	Nucleotide
AF488	Alexa Fluor 488 (dye)
AF647	Alexa Fluor 647 (dye)
SEC	Size Exclusion Chromatography
TIRF	Total Internal Reflection Fluorescence
PEG	Polyethylene glycol
MB	Buffer used for FliG experiments. (Not to be confused with motility buffer!) See A.1
EPPS RB	Buffer used for FliG experiments; See A.1

1

Background

Contents

1.1	The Bacterial Flagellar Motor	2
1.1.1	Bacterial Motility	2
1.1.2	Why study the Bacterial Flagellar Motor?	6
1.1.3	Bacterial Flagellar Motor Overview	8
1.1.4	FliF ring	10
1.1.5	Aside: EM microscopy and the BFM	11
1.1.6	MS-ring C-ring interface	12
1.1.7	C-ring	14
1.1.8	Protein Composition	14
1.1.9	Export Apparatus	33
1.1.10	Rod, Hook and Filament	35
1.1.11	Stators	36
1.1.12	Accessory proteins	39
1.2	Why template a C ring in vitro?	41
1.3	DNA Nanotechnology	47
1.3.1	Overview	47
1.3.2	DNA nanostructure templates	49
1.4	Fluorescence techniques	54
1.5	Project Overview	58
1.5.1	Initial Plan	60
1.5.2	Chronological overview	63

In this chapter, I will briefly describe the biological context of the Bacterial Flagellar Motor (BFM) (section 1.1.1) and elaborate on its worth as a subject of study (section 1.1.2), before reviewing the BFM literature with a comprehensive

focus on C-ring structure. (The most essential points will be summarized later, in section 1.5.) I will go on to make an argument for templating C-ring construction *in vitro* (section 1.2), before briefly reviewing relevant aspects of DNA nanotechnology (section 1.3) and single-molecule fluorescence (section 1.4), which are logical tools for templating and measuring such constructs, respectively. Finally (section 1.5), I will overview the work in this thesis, and its place in a larger C-ring-templating project.

1.1 The Bacterial Flagellar Motor

1.1.1 Bacterial Motility

Most species of bacteria, in their natural habitats, are capable of directed motion in response to sensory input. Most commonly this is used to navigate gradients in the environment, and relocate to areas optimal for growth. This may include gradients in chemical concentrations, temperature and pH [1]. Internal measurement of cellular energy generation (e.g. through the cellular redox state or proton motive force (PMF)) also allows indirect navigation of gradients in light intensity, oxygen availability, metabolite concentration, or anything else which might affect cellular metabolism [2]. While much of the older bacterial motility literature focuses on “chemotaxis” (i.e. navigation based on the sensing of chemical gradients), results also apply to navigation based on the other inputs described above, which use essentially the same signal transduction pathways [2]. Likewise, descriptions of “chemotaxis” system later in this text also apply to other inputs.

Let us quickly survey the variety of mechanisms that underlie bacterial motion in different contexts.

1.1.1.1 Bacteria in liquids

With a few known exceptions (e.g. the marine cyanobacterium *Synechococcus* which swims by generating travelling waves on its cell body [3], and pathogenic bacteria which move by controlling actin machinery in their eukaryotic host cells [4]), directed bacterial motility in liquid environments is mediated by the bacterial flagellar motor (BFM); a rotary electric motor driving a long ($\sim 10\mu\text{m}$) thin ($\sim 20\text{nm}$) semi-rigid

helical filament (“flagellum”) [5]. A universal joint (“hook”) transmits torque from motor to filament while allowing flexibility in the filament orientation[5].

In the model species, *Escherichia Coli* (*E. coli*), multiple filaments attached to counter clockwise (CCW¹) rotating motors on a cell body form a bundle which propels the cell smoothly. However, if a motor switches stoichastically to clockwise (CW) rotation, a conformational change is induced in the filament which expels it from the bundle and causes a “tumble”: a random change in swimming direction [5] (Figure 1.1a).

Bias for CW rotation over CCW is controlled by the phosphorylation of the signalling protein CheY, regulated in turn by the cell’s chemotactic system. When the environment is becoming more favourable (e.g. when toxin levels are lower than they were some moments previously), phosphorylation of CheY is reduced, favouring CW rotation and smooth propulsion. Alternatively, when the environment is becoming less favourable, CheY phosphorylation is increased, favouring increased switching rates and CW rotation bias, leading to more frequent direction changes. Thus the cell can navigate up or down environmental gradients, as required.

Slow adaption processes in both the sensory systems[6] and the motor itself[7] work to restore switching rates to some steady state value, thus adapting the sensitivity of the system to match local conditions. This allows exquisite sensitivity to ligand concentration to be maintained, in some cases, over at least five orders of magnitude[6].

While many well studied bacteria such as *Salmonella Typhimurium* (*S. typhimurium*) swim similarly to *E. coli*, others strategies exist. For example:

In *Sinorhizobium meliloti* (*S. meliloti*) (Figure 1.1b), the multiple motors of a cell rotate only CW, and individual filaments are locked into one conformation. Instead of modulating motor direction, the chemotaxis system modulates motor speed. With all motors at full speed, a bundle is formed as in *E. coli*. When the speeds of individual flagella decline at different rates, however, the bundle is broken apart and the cell tumbles or turns[8, 9].

¹Defined looking from the filament towards the motor.

Rhodobacter Sphaeroides (*R. Sphaeroides*) (Figure 1.1c) also has a unidirectional motor, but only a single motor per cell. Though there is some small degree of speed modulation, the main response to sensory signals is through a molecular brake, which can stop motor rotation entirely. The absence of motor torque allows the filament to relax to a compact coiled form, increasing the cell's rotational diffusion constant to permit reorientation[10–13].

Vibrio alginolyticus *V. alginolyticus*) in liquid (Figure 1.1d). has a single polar motor which reverses direction like an *E. coli* motor. In the absence of a bundle, CW rotation pulls the cell backwards rather than inducing tumbling. However, the switch from CW back to CCW causes a "flick" of the filament which re-orientes the cell [14].

Finally, long corckscrew-shaped “spirochaetes” such as *Borrelia burgdorferi* (*B. burgdorferi*;Figure 1.1e) have multiple flagella filaments trapped between the inner and outer membrane. When the motors rotate in the same direction, they propagate waves of movement along the cell which drive smooth swimming. When a motor switches direction however, and filaments counter-rotate, it flexes the cell in such a way as to change its orientation [15].

Despite the diversity of these strategies, all rely on some kind of stochastic switching or modulation in motor behaviour which induces reorientation of the cell. These switching mechanisms, even those where switching is not obviously binary (e.g. the speed-changing motor of *S. meliloti*[9]), are all mechanistically related, and all controlled by the phosphorylation of CheY (or homologues). For example, while *V. alginolyticus* in liquid expresses a direction-switching BFM, on a surface (see next section) it expresses a genetically distinct speed-changing BFM, yet a single genomic CheY controls both[16].

1.1.1.2 Bacteria on surfaces

Recently, there has been an increasing shift from the study of bacteria in liquid to bacteria on surfaces. In particular, there has been a growing understanding that biofilms (surface-bound communities of bacteria) are associated with a wide spectrum of bacteria in a diverse range on environments, and are central to many

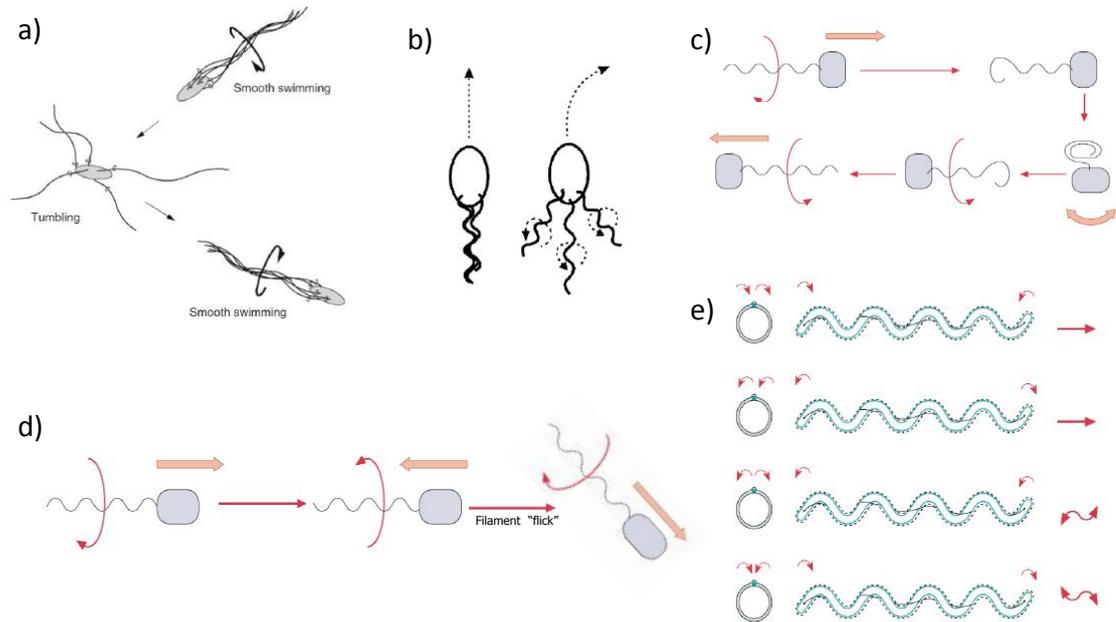


Figure 1.1: Examples of bacterial swimming strategies. a) *E. coli*, *S. typhimurium*, b) *S. meliloti*, c) *R. Sphaeroides*, d) *Vibrio alginolyticus*, e) spirochaetes e.g. *B. burgdorferi*. Figures adapted from [1, 13, 17]

problems in industry and medicine (biofilms are involved in up to 80% of human bacterial infections, for example)[18–20].

There is a rapidly growing body of evidence that the flagellar motor plays a crucial role both in accelerating surface adhesion (by overcoming repulsive forces) and in sensing the presence of a surface, priming the bacteria for biofilm formation[18]. Flagella can also form part of the meshwork holding biofilms together[21].

While cells in biofilms tend to be static, there are also a range of strategies that different bacteria adopt for movement on surfaces. The first, defined as “swarming”, is flagella-driven, and associated with the dynamic “rafts” of bacteria moving cooperatively side-by-side, possibly with flagella bundling between cells[22]. It is also associated with multi-flagellate bacteria. Indeed, some bacteria such as *V. alginolyticus* have a single polar flagella for swimming, but for swarming express a genetically distinct set of lateral flagella [16]. While there is strong evidence that swarming ability is connected to the chemotaxis system and motor switching, is not obvious that swarms actually chemotax (i.e. move up/down environmental

gradients) [22]. It may be that swarming motility is primarily an undirected method to expand bacterial colonies on a surface.

Other surface motility strategies are independent of flagella. “Twitching motility” is driven by Type IV pili: proteinaceous extrusions which cyclically extend, bind to surface, and then retract, pulling the cell with them as they do so. As with swarming, it is typically associated with rafts of co-operating cells, and sometimes important in biofilm formation.

Strategies involving neither flagella or type IV pili certainly exist, but are even less well understood. For example, the individual gliding motility of *Myxococcus xanthus* is prospectively mediated by membrane-spanning surface-binding protein complexes which are driven along the cell’s cytoskeleton [23]. Similarly, the individual gliding motility of *Flavobacterium johnsoniae* relies on filamentous surface-binding proteins, which are apparently propelled by proton-powered rotary motors (distinct from the BFM) along an extracellular track[24].

1.1.2 Why study the Bacterial Flagellar Motor?

There are a number of good reasons to study the BFM. Firstly, as described above, the BFM is central not just to liquid motility but also to surface motility and biofilm formation. The latter in particular is crucial not just to understanding how bacteria function in a range of habitats, but also to understanding and manipulating bacteria (both pathogenic and beneficial) in the human body and in industrial processes (e.g. food production). Indeed, the BFM is critical for virulence in a wide range of pathogenic bacteria, serving a wide range of roles: navigating to infection sites, sensing arrival, manipulation of immune responses (which are often sensitive to flagellar proteins), and escape from host cells post-replication[21]. It is not an inconceivable target for the development of new antibiotics. This is not, however, the only reason to study it.

Secondly, the motor is a close genetic relative of the less well studied bacterial Type 3 Secretion System (“T3SS”). As the BFM must export filament components through the membrane in an ordered manner, the T3SS exports protein substrates

to a eukaryotic host cell[25]. The genetic relationship is sufficiently close that in the BFM-free *chlamydia* family, a few conserved BFM orthologue proteins apparently interact with the T3SS[26], and in a number of species it is actually the BFM, and not the T3SS, that excretes virulence factors[21]. As the T3SS is essential for pathogenicity in many bacteria, and has a similar structure across a broad range of species, it is an attractive drug target. There is an argument that, as opposed to antibiotics which kill bacteria or inhibit their growth, drugs targeting the T3SS should infer less selection pressure for drug-resistant strains, and additionally be more selective for pathogenic bacteria over host-beneficial bacteria[27]. There is a close interplay between the fields of BFM and T3SS study, and it is not uncommon for discoveries in one to direct discoveries in the latter[28, 29].

Thirdly, the motor is arguably the best studied of all large protein complexes. By thoroughly understanding this one model complex, we can hope to discover design principles and conceptual tools that help us understand the plethora of other large protein complexes in nature. (This carries the caveat, of course, that we have no special reason to believe the motor is representative of protein complexes in general.)

Similarly, the entire chemotactic pathway, from the chemical receptors (input) through to the motors (output) is the best studied signalling system in nature. As an almost self-contained model system, we can aspire to a complete understanding of the interaction between its constituent parts, learning general principles which might apply to other signalling systems, and systems biology in general[6]. It is increasingly clear that the motor is a highly reactive and adaptive[7] part of this network, and thus understanding its detailed operation is crucial to a full understanding of the entire chemotactic system.

If the motor and chemotactic pathway can serve as models for understanding the principles of biological systems, the same principles may also inform biomimetic nanotechnology and reaction networks. It goes without saying that the self assembly of the BFM and its performance as a motor far eclipses any rationally designed system to date; likewise the incredible sensitivity and adaptivity built into the chemotaxis system. The BFM (and likewise the T3SS) may also serve as useful

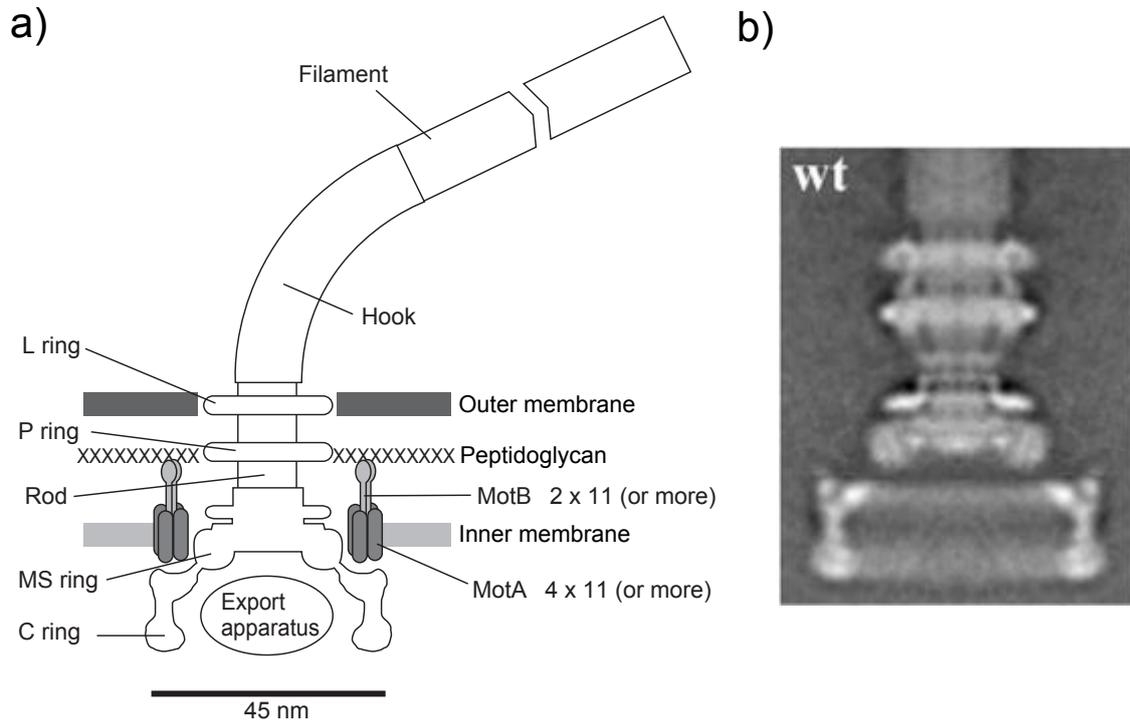


Figure 1.2: a) Schematic of *E. coli* or *S. typhimurium* Flagellar Motor (adapted from [5]). b) EM micrograph of isolated *S. typhimurium* BFM basal bodies, averaged over many motors [30].

components in genetically re-purposed bacteria; for example, both are capable of exporting recombinant proteins, and could prospectively be re-purposed for drug and vaccine delivery[25].

In the following sections, I will go into deeper detail regarding the motor's structure, assembly and operation. I hope the reader can be convinced in the process that, in addition to the reasons given above, the BFM is a fascinating and beautiful machine worthy of study in its own right.

1.1.3 Bacterial Flagellar Motor Overview

The Bacterial Flagellar Motor (Figure 1.2) is a large (~ 11 MDa[5]) self-assembled protein complex spanning both membranes. In the motors characteristic of *E. coli* and *S. typhimurium*, the rotor spans both membranes, reaching from the C-ring in the cytoplasm to the hook and filament outside the cell, all of which rotate as one unit[31]. One or more stators, peptidoglycan-anchored complexes made of proteins MotA and MotB, create H^+ -selective ion channels through the inner

membrane. Apparatus elsewhere in the cell creates a proton motive force across the membrane, meaning that gradients in the electric field and H^+ concentration favour H^+ movement from the periplasm into the cytoplasm, thus driving protons through the ion channels. This in turn drives a conformational rearrangement in the stator which interacts with the top of the C-ring, in such a way as to drive its rotation[5]. (Thus, the C-ring is also referred to as the “torque ring”). Speeds in some species can reach a phenomenal 1700Hz[32]. Rotation is propagated to the filament via the hook, which acts as a universal joint. Stochastic motor switching events (directional switching in the case of *E. coli* and *S. typhimurium*) are related to a rapid (1-100ms [33]) conformational change in the C ring, which presumably re-orientes the top of the C-ring where interactions with the stators occur. Switching dynamics are consistent with a conformational spread model, i.e. the individual units of the C ring may be in CCW-favoring or CW-favoring configurations, but prefer to be in the same configuration as their neighbours, such that the ring is usually found with all units in the same switch state[33, 34]. In *E. coli* and *S. typhimurium*, a switch from CCW to CW rotation also induces a conformational transformation in the filament, reversing the handedness of its helical screw pattern. This switching is regulated by the binding of phosphorylated CheY (CheY-P) to the C ring, which increases the chance that a motor will be found in the CW state. The variation of rotational bias with cytoplasmic CheY-P concentration is phenomenally sensitive; the measured Hill coefficient of ~ 20 is the highest of any known allosteric protein complex[35]. Although implied by the conformational spread model and not conclusively disproven, cooperativity in CheY-P binding has not been observed, leading to suggestions that cooperativity must be embedded in the switching mechanism[36, 37].

Having given this rough overview, I will cover all aspects of the motor in more detail, with a focus on the MS and C rings, which will be the main subject of study later in this thesis. Along the way I will overview some techniques which have been central to the study of the motor.

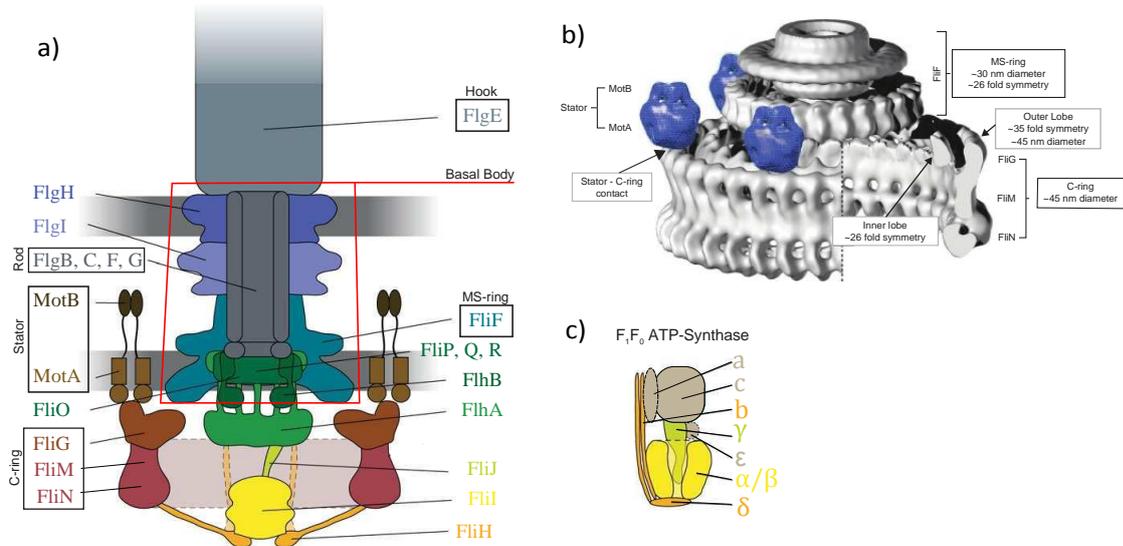


Figure 1.3: a) Protein composition of *E. coli* or *S. typhimurium* Flagellar Motor (adapted from [25]). b) Averaged EM structure of isolated MS and C rings from *S. typhimurium* [38]. EM structures of stator complexes from *V. alginolyticus*[39] are shown approximating stator placement observed in situ[40]. Figure from [41]. c) Structure of F_1F_0 ATP-Synthase. Colour indicates homology with BFM components (a). Figure adapted from [25].

1.1.4 FliF ring

The assembly of FliF monomers into the MS-ring (Figure 1.3) is one of the first steps in motor assembly. Over-expression of *S. typhimurium* FliF in *E. coli* lacking other motor components is sufficient to produce membrane-bound MS rings. When purified and imaged with EM microscopy (see next section), these FliF-only rings account for most of the MS ring density seen in fully assembled motors[42–45] (Figure 1.4a,d). However, at native expression levels, fluorescence assays suggest that complete FliF rings only form in the presence of FliG, the next assembly component. This points to a model of co-operative assembly which we will see repeated in later sections, for other motor components. In *E. coli*[46], but not *S. typhimurium*[47], the membrane-bound export protein FlhA is also required for complete FliF rings, and the self-assembly of FlhA into a ring seems to precede FliF assembly[46].

The full behaviour of FliF during later stages of assembly is poorly understood. The MS ring must have a hole through which hook and filament proteins are exported. Yet when FliF is overexpressed, *E. coli* easily survives having 50% of its inner

membrane surface occupied by MS rings, suggesting that any pore must be initially tightly sealed[42], and undergo some conformational change later in assembly [44].

Biochemical evidence implies ~ 26 FliF units in purified basal bodies [48], consistent with the 26-fold, 25-fold and 24-fold MS symmetry classes observed via EM[38] (see next section). As the FliF ring seems stable[46] and FliF mutations affect assembly but otherwise do not affect rotation [49], MS ring is typically considered to be a rigid structural part of the rotor which serves as a platform for FliG assembly, but not otherwise involved in motor function.

1.1.5 **Aside: EM microscopy and the BFM**

Electron microscopy (EM) exploits the relatively small de Broglie wavelength of electrons to perform imaging analogous to a light microscope, but with a much higher resolution. The low penetration depth of electrons necessitates thin samples and careful use of staining to overcome low contrast of biological samples. Therefore, the first EM images of the BFM (45 years ago!) were of purified motors[50]. In the purification process, most of the motor was stripped away, leaving only the MS ring through to the hook (the “hook-basal body”). In the 1990s, gentler purification protocols allowed the retention of the C ring, and the identity of the protein components was elucidated through antibody labelling and comparison of different mutants[51–53]. More recently, isolated stators have also been imaged, reconstituted in liposomes[39, 54].

Single particle imaging techniques have been used to align and average many noisy single motor images into high resolution structures. Commonly, images are rotationally averaged, guided by rotational power spectra which infer the rotational symmetry of individual motors. This has produced structures of high resolution (e.g. figures 1.2b and 1.3b), sufficient for useful comparison with protein crystal structures[38, 55–57].

However, there are caveats to this averaging process. Disordered components (e.g. freely dangling protein domains) will be lost in the averaging, as will any features in the ring which break perfect rotational symmetry (e.g. gaps). The EM

symmetry mismatch between MS/C rings (Figure 1.3b)[30, 58–60] and dynamic remodelling of FliM/N in the C-ring[7, 61–64] make it almost impossible to envision a motor without one or both of these features, as reflected by their appearance in all leading models[41, 59, 63, 65–67]. The dynamic remodelling should also make us question how similar a purified motor looks to a motor at work in the living cell.

Fortuitously, advances in Cryo Electron Tomography (Cryo-ET) now permit us to image the BFM in situ, unstained, in flash frozen (but intact) bacteria[68]. These motors include visible features usually lost in purification such as the export apparatus and stators (although the later are not clear in all species[69].) As with the purified motors, contrast is low, so high resolution structures require multi-particle averaging with all its caveats. Even with averaging, only a few features (e.g. stator rings in spirochaetes[15, 69–72] and export apparatus of *Leptospira interrogans* [40]) show clear rotational symmetries. Nevertheless, the technology has already greatly expanded our knowledge of motor structural diversity across species[69, 72] and is developing rapidly; further advances should be expected.

1.1.6 MS-ring C-ring interface

Detailed biochemical studies show that only the 38 C-terminal amino acids of FliF[73–75], located on the cytoplasmic side of the MS ring[76], are required for FliG binding. Likewise, only the 46 N-terminal amino acids of FliG are involved in FliF binding[73, 77, 78]. These studies were preceded by the accidental 1992 discovery of two mutants where the genes encoding FliF and FliG had been fused together. The first (the fusion mutant) contains essentially full-length FliF and FliG, whereas the second (the deletion-fusion mutant) lost 56 and 94 amino acids from the C-terminal of FliF and N-terminus of FliG, respectively. Remarkably, both fusions resulted in functional (but somewhat impaired) motors[58].

When the FliF-FliG fusions are over-expressed, they form rings which can be purified and imaged with EM, for comparison with wild type FliF rings (figure 1.4). The FliF-FliG fusion rings (Figure 1.4h) lack anything resembling a partial C ring: only a MS ring with some small additional density compared to the FliF-only case.

This density is not nearly enough to account for the full mass of the FliG[44]. It is also not enough to account for the full mass of the MS ring as seen in complete purified basal bodies[44](Figure 1.4g,e). When complete motors (both wild-type and fusion mutants) are purified in harsh conditions which remove FliM and FliN (but not FliG) from the motor, the MS ring has more mass and is comparable to MS rings from basal bodies (Figure 1.4d-f), but the additional mass is still not sufficient to account for the full mass of FliG[30, 44, 58]. In less harsh purification conditions where FliM and FliN are retained (figure 1.4a-c), the fusion mutant makes a motor similar to the wild type, whereas the fusion-deletion mutant forms a motor with a significantly smaller C-ring, and a missing lobe of density[30, 79].

This all suggests a model whereby part of FliG resides in the MS ring, another part serves as a permanent linker between MS and C rings, FliM and FliN don't interact directly with the MS ring at all, and parts of FliG usually present in the C-ring are disordered (and thus not visible in the averaged EM structures) when FliM and FliN are absent. The extra density in complete motors with FliM/N removed, as compared to over-expressed FliF/FliF-FliG rings, could be parts of FliG residing in the MS ring which are ordered or disordered depending on assembly history, or may be other proteins ([44] suggests a mix of both).

In all EM images of purified motors, the EM density linking the MS and C rings is very low, suggesting that the corresponding portion of FliG is flexible. This is consistent with recent in-situ Cryo-ET images from *B. Burgdorferi* where the link between MS and C rings is pliable enough to permit the C-ring to flex, following the curvature of the cell membrane[70]. Tilt of the C ring relative to the MS ring can also be seen in purified basal bodies ².

Given that the motor is thought to have ~ 26 FliFs, the fusion mutant also points to a model where the motor has only 26 FliGs. In section 1.1.7, we will see that this contradicts other evidence; making the stoichiometry of FliG in the motor controversial. Accessory proteins proposed to later the MS ring - C ring interface will be covered in section 1.1.12.

²Personal communication, Keiichi Namba

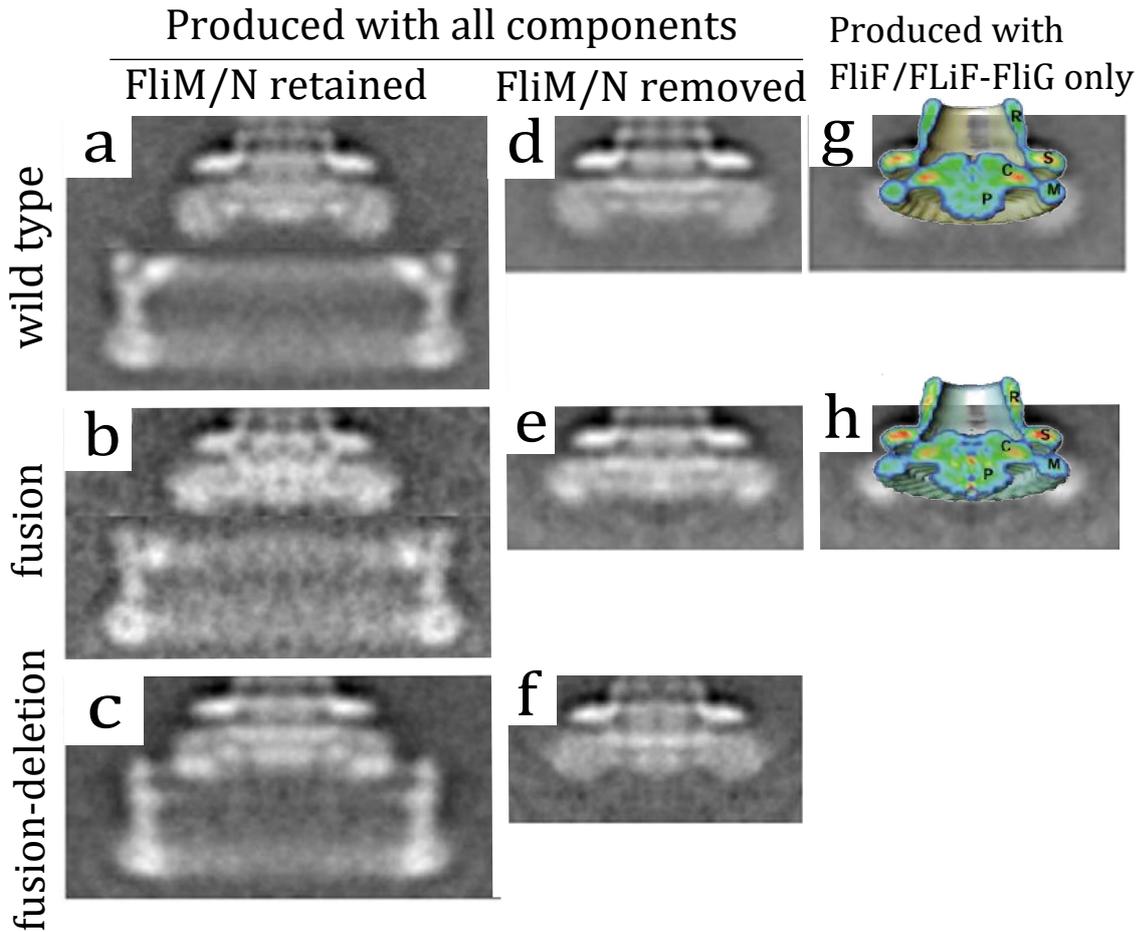


Figure 1.4: Comparative EM of wild type, FliF-FliG fusion and FliF-FliG deletion fusion mutants. a,b,c) Motors produced in presence of all components, purified retaining FliM/FliN [30] d,e,f) Motors produced in presence of all components, purified under conditions where FliM/FliN are lost [30] g,h) In colour: density maps for MS rings produced with over-expressed FliF or FliF-FliG fusion, respectively[44]. These have been overlaid onto the structures from figures d and e (grayscale) for comparison.

1.1.7 C-ring

1.1.8 Protein Composition

Mutation studies from the 1960s through to the 1980s firmly established that FliG, FliM and FliN were all involved in both torque generation and switching. This was followed by intergenic suppression evidence that all 3 operate together in a “switch complex” which interacts with CheY and CheZ; cytoplasmic components of the chemosensory system[49, 80]. Later, it was found that retention of FliG, FliM and FliN during motor purification was associated with the appearance of a cytoplasmic “C-ring”[53] (synonymous with “switch complex”[65] and “torque

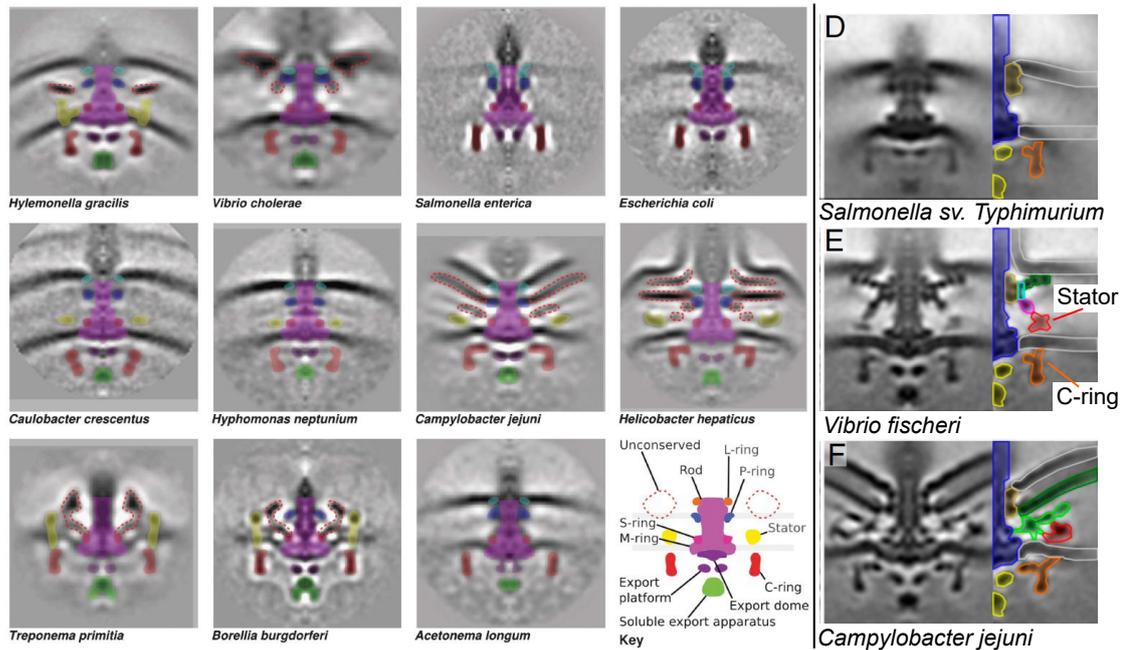


Figure 1.5: Subtomogram averages of BFM in various species, imaged *in situ* with CryoET. Images on the left and right are from [69] and [72] respectively.

ring”[57]). Putative stator complexes had already been (correctly) identified *in-situ* by this time[81], and it was recognised that the C-ring was at an appropriate radius to interact with them[53]. Since then, *in-situ* CryoET structures have consistently shown stators, when visible, near the top of the outer lobe of C-ring density (Figure 1.5) [15, 40, 69–72, 82].

Early models put FliG with FliF in the MS ring, and FliM with FliN in the C ring[53]. Later, more detailed information about the binding relations between FliG/M/N, the observed structural effects of the fusion-deletion mutant and evidence that FliG was the site of direct stator interaction led to the model of figure 1.3, with FliG occupying the top of the C ring (and some small density on the MS ring), with FliM then FliN beneath it. This broad arrangement is widely accepted, but the details are contested. In sections 1.1.8.2 and 1.1.8.4 we will look more closely at the biochemical data on these three proteins, to better understand the contested details.

1.1.8.1 Symmetry Mismatch

The structure of the C ring as seen in EM of purified motors has some interesting features, particularly relating to symmetry. The top of the C ring is split into inner

and outer lobes of density (Figure 1.3b). Consistently, the outer lobe of the C ring (presumably containing FliM, FliN and some of FliG) is seen with ~ 34 -fold symmetry, whereas the inner lobe (presumably containing part of FliG) is seen with ~ 26 -fold symmetry, identical to the MS ring. This “symmetry mismatch” has been studied most closely in purified motors from a clockwise-locked mutant of *S. typhimurium*[38]. In that study, MS ring symmetry was found to vary from 23-fold to 26-fold, and always matched the symmetry of the inner lobe of the C ring. The outer lobe C ring symmetry varied from 32-fold to 36-fold, and was matched by changes in ring size, consistent with a fixed subunit spacing. There was no clear correspondence between MS ring and C ring symmetries; i.e. all possible combinations were observed. Similar results were found in wild type *S. typhimurium* motors, where C ring symmetries of 31-fold to 38-fold were observed[83]. In the fusion-deletion mutant, the reduced C-ring diameter implies ~ 31 -fold symmetry[79]. Furthermore, although in-situ CryoET structures can not yet resolve C ring symmetries, C ring sizes are seen to vary widely between species (from 34nm to 57nm diameter) [69, 72]; we can easily imagine a corresponding variation in symmetries. Notably, the size of the MS ring is conserved between species[69] (Figure 1.5). This all suggests that the stoichiometry of C ring proteins is mismatched to the stoichiometry of MS ring proteins, and free to vary both within and between species. Although this mismatch was first observed 15 years ago[79], there is still no consensus as to what function it serves, or how it is mediated.

Some early hypothesis as to function were that non-commensurate symmetries were part of a mechanism in which C ring (FliM and FliN) and MS rings (FliF and FliG) rotate relative to each other [79], but this was never very well supported biochemically, and has been recently disproven by the fluorescence observation of FliN rotating at similar frequencies to the hook[31]. That leaves more structural explanations, one of which might be that it allows the C ring size to vary without a change in MS ring stoichiometry. For example, perhaps the C ring can expand indefinitely through the incorporation of additional subunits, until the linker portion of FliG connecting the MS and C rings is fully extended. This might explain how

variation in C ring size between species is mediated[69, 72], or even be part of a dynamic remodelling process, in accordance with growing evidence that FliM and FliN numbers in the motor are dynamic (see section 1.1.8.4).

Whatever the function, a mechanical explanation is difficult. In particular, it is difficult to explain how FliG fits into this story, given that most leading models include FliG_c (the site of torque generation) in the ~34-fold outer lobe of the C ring (section 1.1.8.2). One possibility is that 26 FliG's bind to the MS ring, and that 26 FliG's are present in the outer C ring, with 8 gaps. 26-fold symmetric stepping during rotation is often taken to imply 26 FliGs[84, 85]; a flawed interpretation, given that the outer lobe of the C ring has 34-symmetry regardless of the FliG stoichiometry. Equally, given that the motor can run with a single stator (section 1.1.11), it is hard to explain how gaps in the FliG_c ring would not halt rotation. A second possibility is that 34 FliG's are present in the outer C ring, 26 of which bind to FliF and the rest are unbound, leaving FliG_n unstructured and dangling inside the C ring; averaged away in EM structures. For the fusion mutant, these dangling regions would include entire copies of FliF. A 1:1 FliG:FliM stoichiometry is loosely supported by one biochemical study[86], but not at all conclusively. A final possibility is that FliG (or at least FliG_c) does not extend to the outer C ring at all, and is present only in the inner C ring. However, this is inconsistent both with stator localization in CryoET, and the structure of the fusion-deletion mutant (Figure 1.4c), which lacks an inner lobe yet includes most of FliG, and remains functional.

1.1.8.2 FliG

Crystal structures of partial and full-length FliG from *Aquifex aeolicus*, *Thermotoga maritima* and *Helicobacter pylori* show 3 domains; FliG_n, FliG_m and FliG_c (Figure 1.6) [56, 57, 65, 87–89]. These are separated by flexible linkers, mutations in which are linked to motor switching bias, and which vary in conformation between different crystal structures[57, 65, 88, 90].

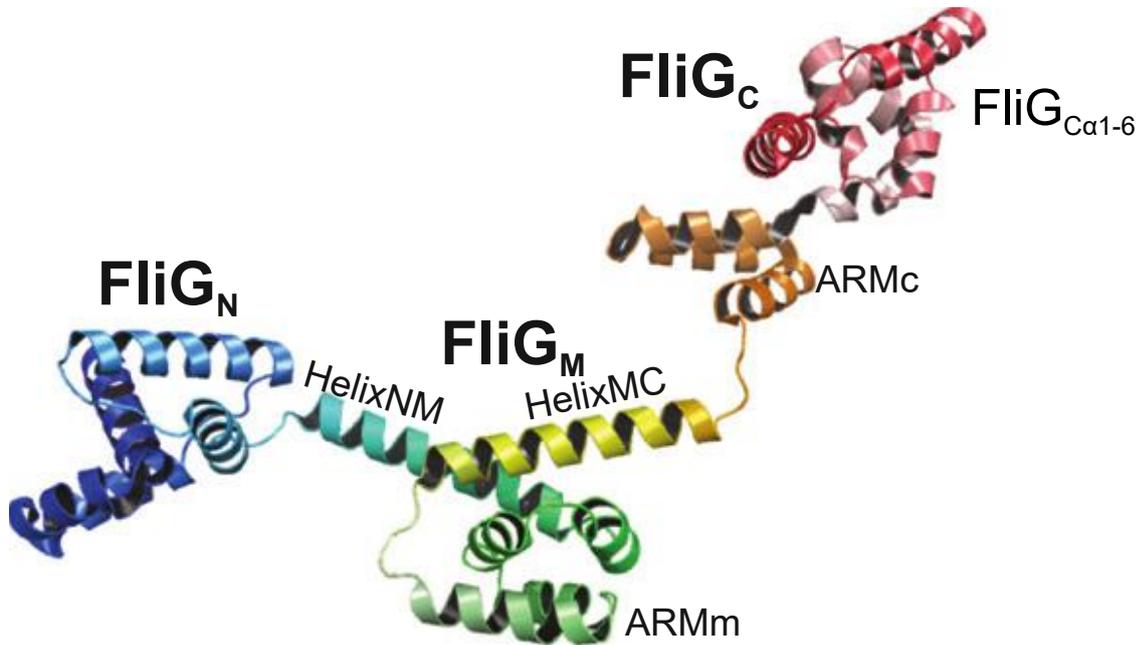


Figure 1.6: Crystal structure of full length FliG from *Aquifex aeolicus*[57]

FliG_n As mentioned in section 1.1.4, the N-terminal domain of FliG, FliG_n, is exclusively responsible for FliF binding, and requires only the last 46 amino acids[30, 58, 73, 74, 77, 91]. There is limited evidence that, in some species, this same FliF-binding sequence might promote FliG homodimerization in the cytoplasm prior to binding FliF[74]. Some cross-linking evidence suggests FliG_n is proximal to FliG_m in the motor [59], but this is highly controversial [41]. In the deletion-fusion mutant, most of FliG_n is lost (along with the inner lobe of C-ring density), and yet motors are still functional. This implies FliG_n corresponds to the inner lobe of the C ring (separate from FliG_m), and has no important role in switching or torque generation, consistent with a lack of motility or switching mutants[58].

FliG_c The C-terminal domain, FliG_c, can be conceptualized as two subdomains: ARM_c and FliG_{cα1-6} (Figure 1.6).

FliG_{cα1-6} contains the “torque helix”; the site of direct C-ring interaction with the stator, and thus torque generation[92–94]. Charged residues along this helix interact electrostatically with charged residues on MotA, as demonstrated by targeted intergenic suppression mutants [94]. This interaction is sufficiently

general that FliG_c and MotA from different species make functional motors[95, 96], as do charge-reversing pair mutations[94]. There is some biochemical evidence that separate groups of charges are responsible for steering stators into position above FliG and transmitting the subsequent stator power-stroke[97]. One model, supported by simulation *in silico* suggests the charged residues are responsible only for steering, and the power-stroke is transmitted sterically[98]. This steering hypothesis is consistent with in-situ CryoTM observations that stators held at a slightly larger or smaller radius than the C ring can still drive rotation efficiently[72]. Motor switching is widely thought to involve rotation of the torque helix relative to the stator [56, 57, 85, 88, 89, 98, 99], and may involve rotation of FliG_{ca1-6} relative to ARM_c via a highly conserved[100] linker, as implied by crosslinking and mutational studies[56, 93, 101–103] and consistent with variation between crystal structures(Figure 1.7).

The role of ARM_c is disputed. Unlike FliG_{ca1-6}, it is required for motor assembly[92], and mutational studies *in vivo* and *in vitro* (using full-length FliG) implicate a hydrophobic patch on ARM_c in binding to FliM_m [59, 67, 87, 91]. There are also reports of *in-vivo* crosslinking (albeit low-yield) to support this interaction[59]. However, studies of FliG_c-FliM binding *in vitro* have been mixed; one report of clear binding via ARM_c[59] contradicts reports of no binding[104, 105], very weak binding[65] or weak binding via a region near the FliG_{ca1-6} torque helix rather than the ARM_c hydrophobic patch (inferred by NMR)[106]. One possible explanation is that an ARM_c-FliM interaction may occur only in an assembled motor and not in isolated proteins [104]. Alternatively, perhaps the limited reports of crosslinking and *in-vitro* binding[59] are erroneous, while the mutational studies might be explained by the binding of FliM via an interface elsewhere on FliG that is somehow remotely disrupted by the mutation of the ARM_c hydrophobic patch[57]. While this seems far-fetched, it is supported by the observation that the deletion of FliG_c from *E. coli* FliG weakens *in-vitro* binding to FliM and significantly reduces the entropic (i.e. potentially hydrophobicity-related) component of the interaction, yet FliG_c alone shows no detectable binding[105]. (Interestingly, the deletion of

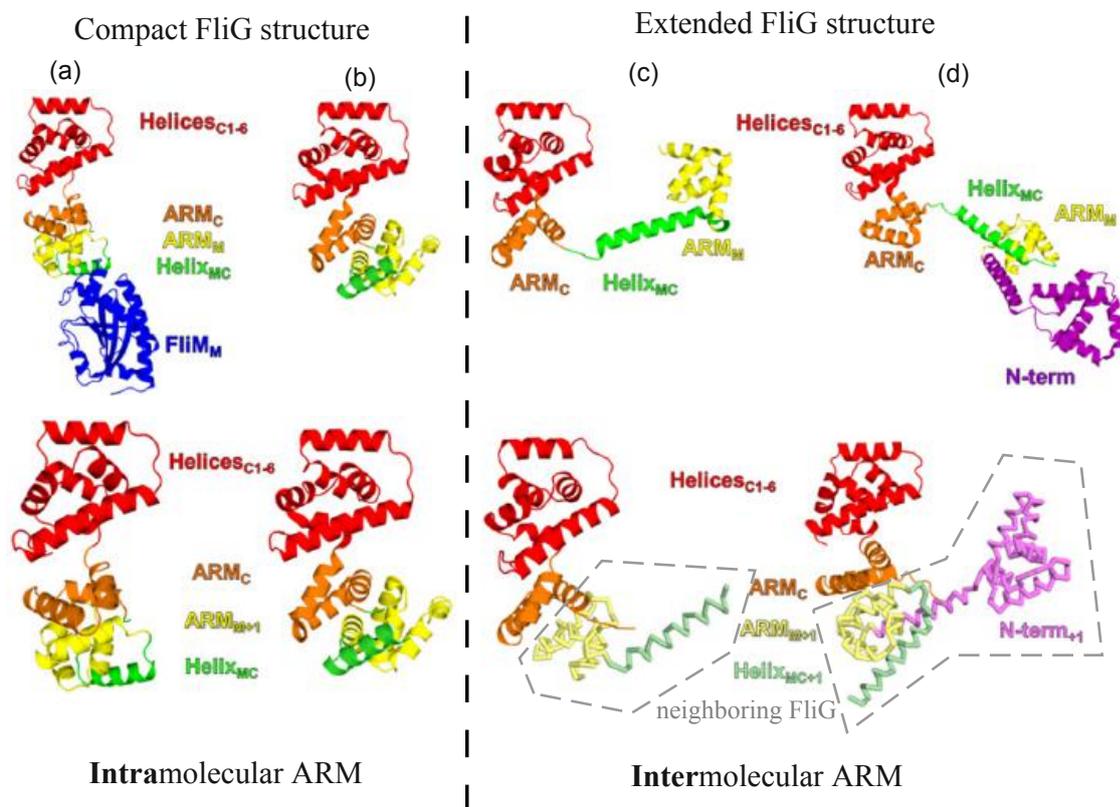


Figure 1.7: Various crystal structures of FliG, demonstrating rotation of FliG _{α 1–6} relative to ARM_c and intramolecular and intermolecular ARM interaction. a) FliG_{mc} bound to FliM_m from *T. maritima*[65], b) FliG_{mc} from *T. maritima* [88], c) FliG_{mc} from *T. maritima*[87], d) FliG full length from *A. aeolicus*[57]. Figure adapted from [65].

FliG_c in *H. pylori* did not have the same effect, pointing to differences between species which might confuse this issue.) We will discuss an alternate plausible role for ARM_c, before returning to this FliM-binding hypothesis.

FliG Armadillo motifs ARM_c and ARM_m in FliG_m are so-called because they resemble armadillo (ARM) motifs. These 3-helix motifs are found in a broad range of eukaryotic proteins, where ARMs stack together via hydrophobic interfaces to form a superhelix[107–109]. Genomic evidence that ARM motifs may be common in prokaryotes and archaea is just beginning to emerge[110], but has not been verified: ARM domains remain poorly characterized outside of eukaryotes. The role of ARM_c and ARM_m in FliG are controversial.

Unlike the well-characterized eukaryotic ARMs, ARM_c and ARM_m have a linker between them. However, they both have hydrophobic patches as expected[57],

and every FliG_{mc} crystal structure so far shows ARM-ARM stacking via the hydrophobic patches, consistent with the eukaryotic model[56, 57, 65, 87, 88, 104, 105] (Figure 1.7). Intriguingly, in one of these structures[65] (and in another with some ambiguity[88]) the linker between the two ARMs, HelixMC, is partially unfolded which allows intramolecular ARM stacking (i.e. ARM_c and ARM_m from within the same protein are bound to each other)(Figure 1.7a-b). However, in other crystal structures, HelixMC is longer, making an intramolecular ARM stack impossible, and instead an intermolecular ARM stack is seen (i.e. ARM_c is bound to the ARM_m of a *neighbouring* FliG)[56, 57, 87](Figure 1.7c-d). The nature of the ARM stack in the working motor is a key point of contention, and will be a central component of this thesis.

One argument proposes that ARM stacking is a crystallographic artefact, consistent with a failure to detect inter-FliG cross-linking *in vivo*[59]. However, a more recent study has shown intermolecular ARM crosslinking in membrane-bound (i.e. motor-bound) fractions of FliG, explaining the previous null result as an over-expression artefact[111]. The growing body of evidence for physiologically relevant ARM stacking also includes evolutionary covariance[111], solution SAXS structures[111], MALLS[104] and pulsed dipolar ESR spectroscopy (which measures distances between labelled sites *in vitro*)[104].

This leaves us with two models for the role of ARM_c. In one, the hydrophobic patch of ARM_c directly binds to FliM and the two ARM domains do not interact *in vivo*. In another, ARM_c is bound to ARM_m, and ARM_c only binds FliM indirectly through ARM_m (see below). The balance of evidence is shifting towards the latter model, or perhaps a mix of the two.

FliG_m FliG_m consists of ARM_m (see above) sandwiched by two linker helices; HelixMC and HelixMN. Crosslinking evidence suggests that neighbouring FliG_m are in close contact in the motor [101]. Uncontested biochemical[87, 89, 91, 100, 102, 105, 112] and crystallographic [59, 65, 87, 104, 105] evidence shows that FliM_m binds to FliG_m via a conserved EHPQ motif and a hydrophobic pocket formed by

HelixMC and the bottom of ARM_m (e.g. Figure 1.7a) (not to be confused with the hydrophobic patch on the top of ARM_m which mediates the interaction with ARM_c, or the proposed FliM binding site on FliG_c).

The physiological arrangement of HelixMC is not entirely clear. A conserved Gly-Gly motif joining HelixMC with ARM_c is important for switching [90] and expected biochemically to be a flexible linker[87] consistent with crystal structures[57, 65, 87] and limited proteolysis experiments[88]. Variation between crystal structures also shows that the linker between HelixMC and ARM_m is flexible[88]. Limited proteolysis and switching mutations suggest HelixMC itself is flexible and heavily involved in switching[88, 90].

In some crystal structures[57, 87, 88] (e.g. Figure1.7c,d), HelixMC is in an extended form which prevents intramolecular ARM stacking and promotes intermolecular ARM stacking. In this extended form, part of HelixMC may stack against ARM_m(e.g. Figure1.7d) or point away from it(e.g. Figure1.7c), maybe stacking against a neighbouring ARM_m in the process[88]. Mutations expected to disfavour the former configuration are associated with CW switching mutants[57, 88]. In other structures (Figure1.7a,b), including all FliM_m co-crystals (e.g. Figure 1.7a), the ARM_m-proximal half of HelixMC structured and stacked against ARM_m, while the rest is unstructured and often not resolved at all, allowing an intramolecular ARM interaction. In the FliM_m co-crystals, the structured part of HelixMC forms part of the FliM binding interface, consistent with all the biochemical evidence mentioned above[59, 65, 104, 105].

Some argue that the extended form is unphysiological, and in some structures may be stabilized artificially by crystal contacts[65, 87]. Others argue that these crystal contacts are physiological, and that the extended form of HelixMC is not only physiological but (consistent with HelixMC switching mutants) may mediate interaction with neighbouring FliG units, either directly [88] or by allowing intermolecular ARM stacking [57, 111]. As FliM binding to the motor is dynamic (see below), it is also conceivable that transitions between a FliM-stabilized compact form and an extended unstacked form might be related to the switching mechanism.

1.1.8.3 FliG domain-swap polymerization model

A recent study by our collaborators[111] uses small angle X-ray scattering (a very low-resolution measure of protein conformation in solution) to argue that monomeric FliG switches between compact and extended forms (Figure 1.8a). Their *in vivo* crosslinking additionally implies intermolecular ARM stacking (and therefore an extended form) in the motor, consistent their earlier model of motor structure [57]. Thus, they propose a novel assembly model where FliG is largely in compact form in the cytoplasm, preventing aggregation (consistent with an observed lack of cytoplasmic FliG-FliG interaction[78]). Then templated at high effective concentrations by the MS ring however, rare transitions to the extended form would be locked in place by intermolecular ARM stacking (Figure 1.8b). This assembly model is not obviously consistent with the observation that crystal structures favour the compact form in the presence of FliM and that the extended HelixMC is usually seen stabilized by crystal contacts. Momentary unstacking of the ARM motifs should also carry quite an energetic penalty. Therefore evidence for the model is not overwhelming. However, its appeal is that it explains how FliGs could be tightly bound in the motor (facilitating switching co-operativity), while avoiding self-interactions in the cytoplasm. It also provides a possible explanation for the symmetry mismatch, as templated polymers of extended FliG display ARM binding sites that could recruit cytoplasmic FliG even in the absence of a spare template site (Figure 1.8c). In addition, although they are evolutionarily distinct, recent studies on FliM and FliN (section 1.1.8.4) hint that they might assemble by a similar mechanism, implying this could be a common feature in the assembly of large protein complexes. **Testing the FliG domain-swap polymerization model will be a primary aim of this thesis.**

1.1.8.4 FliM and FliN

Early EM studies on isolated motors established that FliM and FliN make up most of the mass of the C ring[53]. Biochemical studies consistently show that FliM interacts with itself, FliN and FliG [58, 89, 102, 112–116] whereas FliN interacts

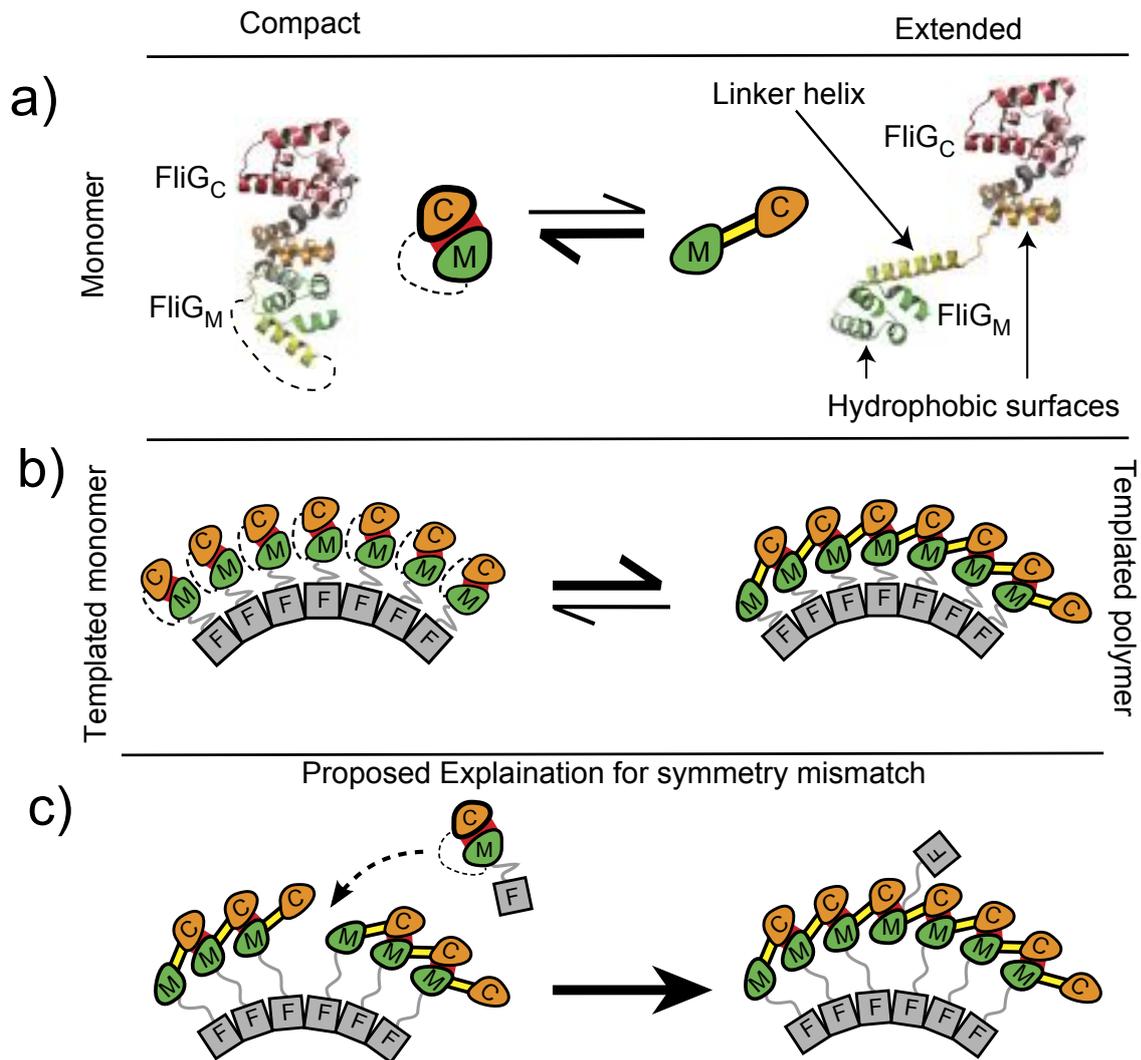


Figure 1.8: a) Compact and open forms of FliG. Red = ARM-ARM stack. Yellow = HelixMC. b) Templated domain-swap polymerization. c) Domain-swap polymerization as an explanation for gap filling. C,M,F represent FliG_c, FliG_m and FliG_f, respectively. Figures adapted from [111]

with FliM but no other C ring proteins [46, 58, 61, 114, 116]. Combined with crystallographic evidence that the FliG-binding domain of FliM is opposite the FliN binding domain[89], this indicates FliM is in the middle of the C ring and FliN is at the bottom (Figure 1.3). We will also see strong evidence that FliM and FliN pre-associate before assembly into the motor, and act as as one unit: thus, we will discuss them together. As with FliG, we will look at the subunits on FliM and FliN in turn, before focusing on the details of CheY binding.

FliM_n FliM_n can be considered in two parts. At the N-terminus, a highly conserved 16 residue sequence has long been established as the primary site for CheY-FliM binding[116–118]. This CheY-binding peptide has been co-crystallized with CheY[119–121], and remains at least partially structured in CheY’s absence[106].

The remaining ~35 residues of FliM_n have never been crystallized, consistent with size exclusion chromatography[115] and NMR[106] which both point to a disordered, flexible linker. The sequence of this linker is poorly conserved, except for a highly conserved but poorly understood Asp/Tyr pair which may have some role in switching[106, 112, 115]. A five-residue extension of the linker has minimal effect on motor function[112], whereas 10 residue deletions disrupt CheY binding[116] and therefore switching[118], but nevertheless do not inhibit rotation or motor assembly[118]. Indeed, some species have motors without FliM_n[16]. Thus, FliM_n serves only as a CheY-binding domain on a flexible tether.

FliM_m FliM_m, a structured globular domain, has been structurally well-characterized, both alone [115] and in association with FliG_m[89, 104, 105].

The region surrounding a very well-conserved[100] GGXG motif is the site of binding to FliG_m, consistent with a range of biochemical, mutational and crystallographic evidence[59, 87, 89, 104–106, 112, 116]. This is also the same site implicated in the contested binding to FliG_c (section 1.1.8.2), and is opposite FliM_m’s N and C termini, where we expect CheY and FliN to bind via FliM_n and FliM_c respectively (e.g. Figure 1.9a).

While strong FliM_m-FliM_m association outside the motor requires at least one FliM_m to have FliM_c attached[78, 118], *in vivo* crosslinking in the motor has mapped out extensive surfaces of FliM_m-FliM_m interaction, consistent with a ring of parallel FliM_m[89, 101, 102, 104, 115]. Variation of crosslinking yields with motor switch state and differential clustering of switch-biasing mutations along the interaction surfaces suggest that relative movement of adjacent FliMs is part of the switching mechanism[89, 104, 115].

Although FliM_n is the main site of CheY binding, there is increasing evidence that FliM_m also has a CheY-binding role. FliM itself has distant sequence and structure homology to a class of CheY phosphatases which includes CheC, CheX and FliY[106, 115] (the last of which is an additional C ring protein in some species (see below). While FliM_m lacks the conserved features associated with phosphatase activity, there is mounting evidence that it weakly binds to the phosphorylation site of CheY[106]. Some evidence suggests this disrupts FliG_c-FliM_m binding, but not FliG_m-FliM_m binding[106].

FliM_c and FliN interaction FliM_c is established as the sole site of binding to FliN[78, 116], consistent with a FliM-FliN fusion protein that produces functional motors (albeit enhanced by the addition of extra free FliN)[122, 123]. A truncated FliN has been crystalized as a homodimer, and in solution (without FliM) forms dimers (*T. maritima*) or tetramers (*E. coli*)[114]. Contacts elucidated by crosslinking are consistent with donut-shaped tetramers (dimers of dimers) which fit the EM density at the bottom of the C ring[55] (figure 1.9a). FliM and FliN exist as complexes in solution, and indeed FliM cannot be stably purified without FliN[113, 114, 124, 125]. Early reports estimated the stoichiometry of these complexes at 1:4 FliM:FliN[63, 64, 114], and models based on crosslinking and mutations proposed that FliM_c alternates with FliN tetramers along the bottom of the C ring[55, 126] (Figure 1.9a).

More recently however, an alternate model has gained traction, inspired by study of FliM and FliN homologues in the T3SS, a new crystal structure of the FliM-FliN fusion and the use of analytical mass spectrometry[123, 127]. In brief: FliM_c and FliN contain SpoA domains, which have remarkable structural similarity. In much the same way as FliN forms a homodimer, FliM_c and FliN may form a heterodimer (as crystalized in the form of the FliM-FliN fusion[123]). And in much the same way as two FliN dimers are expected to come together and form a doughnut tetramer, one FliM-FliN dimer and one FliN-FliN dimer are thought to bind and make a 1:3

FliM:FliN complex, consistent with some older biochemical measurements of FliM and FliN stoichiometry in purified motors (~ 35 and ~ 111 respectively[128]).

The T3SS equivalent of this unit (comprising 2 FliN homologues and a FliM_c-FliN fusion homologue) shows oligomerization in solution consistent with an open lock-washer conformation. Mutational studies suggest this oligomerization is necessary for T3SS function, and when this arrangement is extrapolated to FliM-FliN (Figure 1.9b) the resulting structure is both consistent with crosslinking and mutation mapping, and a convincing fit for the EM density seen in purified C rings (figure 1.9c)[127].

Although the FliM:FliN units are less prone to oligomerization in solution than their T3SS equivalents, it is tempting to think they might be in conformational equilibrium between open and closed forms (figure 1.9d), and assemble into the motor by the same domain swap mechanism proposed for FliG. This remains untested, but if this were true despite the non-homology of FliG and FliM/FliN, it would indicate convergent evolution that may be replicated in many other protein assemblies.

FliN function FliN is known to have two main functions, both associated with the same conserved hydrophobic patch on a FliN-FliN dimer[114, 126, 129]. Firstly, this patch is implicated in switching dynamics[114, 126, 129]. CW-biasing mutations here can often be reversed by overexpression of CheY[129], but direct evidence of CheY binding has been observed only recently. *In vitro*, it turns out that CheY binds well to FliN, but only once CheY has been activated by phosphorylation *and* binding of the FliM_n CheY-binding peptide in the relevant CheY binding pocket[130]. Additionally, this same patch is essential for the binding of FliH, through which FliN templates all of the cytoplasmic export apparatus. FliN is thus necessary for flagellation[125, 129, 131, 132]. This interaction has been measured *in vitro*[129, 132] and the ~ 10 residue peptide at the end of FliH responsible for binding[133] has recently been crystalized bound to the FliM-FliN fusion[123]. Given that differential crosslinking yields imply relative movement of FliN and FliM_c during switching, it is likely that FliN moves during switching[126], perhaps altering the accessibility of the CheY and FliH binding sites.

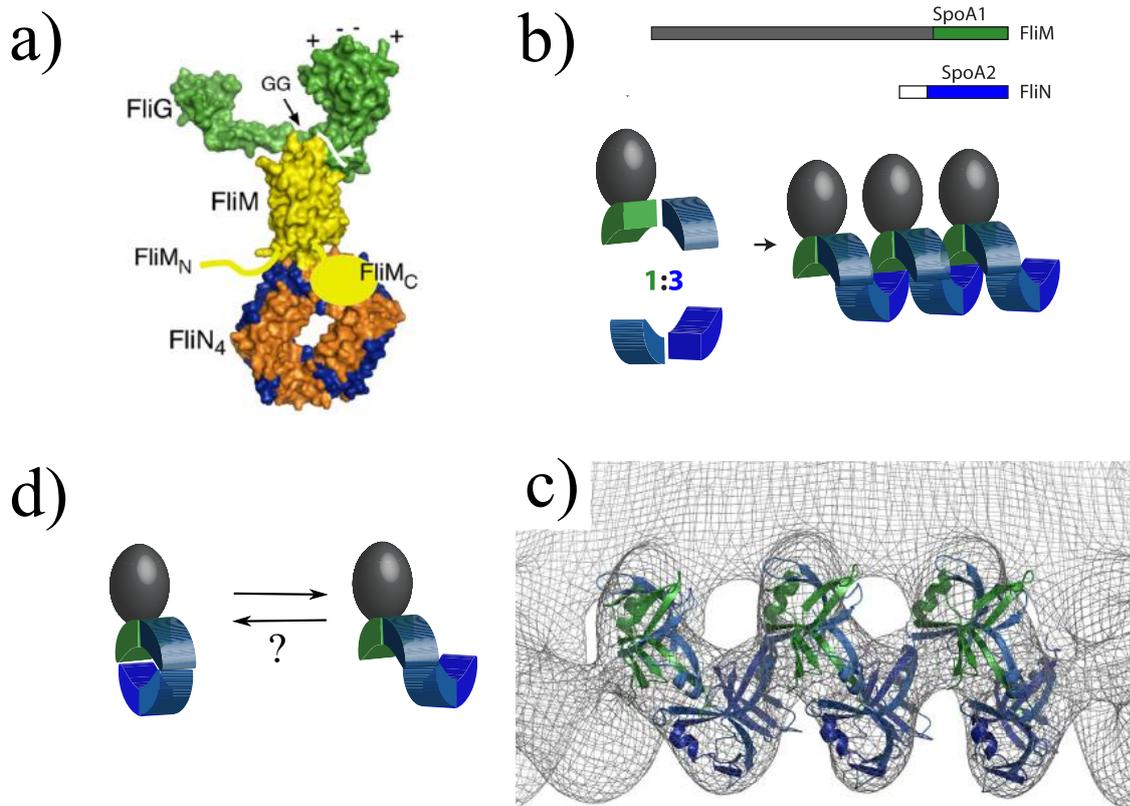


Figure 1.9: a) Older model of FliN tetramers in the C ring[126] b) recent model of 1:3 FliM:FliN complexes[127] c) Docking of the above into EM structure of the lower C ring [38] d) Prospective FliM/N domain swap model

While the C terminal 2/3 of FliN includes the conserved binding patch and the SpoA domains responsible for FliN-FliN and FliN-FliM interactions, the N terminal $\sim 1/3$ of FliN is poorly conserved and not essential for motor function[125, 134]. Its deletion does however cause some limited inhibition of the motor[125], and it has been proposed to modulate CheY binding by competitively blocking the binding patch [129, 130]. Presumably, this could also block FliH binding.

CheY binding We have painted a rather complicated picture of CheY binding. CheY has a phosphorylation site, and when phosphorylated, the surrounding surface binds weakly to FliM_m[106]. Opposite the phosphorylation site is a binding pocket for the FliM_n CheY-binding peptide, the affinity of which again is much higher when CheY is phosphorylated[117, 119]. Finally, CheY binds to FliN, with some mutational evidence that this is mediated by the peptide-filled binding pocket on

CheY[130]. The elucidation of these 3 interactions has led to a “tethered bait” model where phosphorylated CheY binds first to the FliM_n peptide, which then increases the local concentration of CheY, promoting subsequent binding to FliM_m and/or FliN, either of which could plausibly be linked to conformational switching of the motor[106, 126, 130]. This could perhaps reconcile the co-operativity of switching with the failure to observe co-operativity in CheY-P binding[135].

Switching Besides the details of CheY binding, there are other aspects of switching which remain unclear. For example, in section 1.1.3, I claimed that switching was stochastic, and the concentration of cytoplasmic CheY-P alters the bias of that switching. While this is the canonical view[5, 35], observations of synchronised switching between motors has fuelled a minority view that switching events are not stochastic, but always the result of a change in CheY-P concentration[136–139].

Additionally, there are numerous observations that frequency of switching is dependent on motor load and speed [140–143]. Mechanisms to explain this have been proposed[144, 145], but not tested.

Switch sensitivity to temperature has also been observed; indeed, motors in the absence of CheY usually rotate CCW, but can switch to CW if the temperature is dropped[146]. This may indicate that hydrophobic surfaces are exposed in the CW state, given that the free-energetic penalty for exposing such surfaces would decrease at lower temperature[89].

FliY An additional C-ring protein, FliY, is expressed in a number of species, both with (*H. pylori*) and without (*B. subtilis*, *T.maritima*) separate expression of FliN[147]. The sequence mirrors the three-domain structure of FliM; N domain with a conserved CheY-binding peptide, a middle domain related to a family of CheY phosphatases, and a C domain homologous to FliN[147]. However, whereas FliM_m binds CheY-P only weakly (requiring assistance from FliM_n) and has no phosphatase activity, FliY_m binds CheY-P strongly (without FliY_n assistance) and retains strong phosphatase activity[147, 148]. Recent evidence also suggests that FliY may be expressed both as a full length product and a truncated product containing only

the FliN-homologous part; analagous to the expression of the FliM-FliN fusion homologue of the T3SS[127]. Cryo-ET images of C rings in FliY-containing species do not look obviously different to those of FliY-lacking species[147]. However, as FliY shows no FliG-binding activity[147], its function in the ring is likely distinct from FliM. While it is likely that FliY_c incorporates into the bottom of the C ring as FliN does, it is unclear whether FliY_m incorporates into FliM_m ring (perhaps leaving gaps in the FliG ring above), or is disordered. It may even be possible that only truncated (FliN-homologous) FliY incorporates into the motor; this will only be resolved with further experiments.

1.1.8.5 C ring Turnover

Turnover, i.e. the continual dissociation and association of components in the working motor, was first observed in GFP-tagged stators in live *E. coli*[149]. Stators recruited to the motor were seen to last only ~ 30 s before diffusing away. Later studies showed that the recruitment of stators to the motor is sensitive to the stator's conjugate ion motive force[150–152] and the external load on the motor[153–155], indicating dynamic adaption of the motor to changing conditions, alongside a load-sensing role.

This inspired the discovery that FliM and FliN are also undergoing turnover. Multiple studies have painted a reasonably consistent, but complex, picture. Firstly, all observations are consistent with the notion that FliM and FliN join and leave the motor as pre-assembled FliM₁:FliN₄ or FliM₁:FliN₃ units (see above)[62, 64]. Secondly, some sub-population of FliM:FliN units in the motor are tightly bound (not exchanging), whereas another population is weakly bound (exchanging on a ~ 30 s timescale). Thirdly, increased CW rotation bias (whether induced by FliG mutation or CheY-P concentration) results in fewer tightly bound FliM:FliN units and fewer FliM:FliN units in total[62–64], consistent with a model where the C ring presents a constant number of FliM:FliN binding sites, and rotation direction changes the proportion of those which are weak-binding (and thus only partially occupied)[62]. This has been explained as a mechanism to tune motor switching

sensitivity to the average CheY-P concentration over a long time scale (\sim minutes)[7, 35, 156]. Thus, as with the stators, turnover serves to adapt the motor to changing conditions. A mechanistic description of the mechanism has not been elucidated.

The dynamic nature of FliM and FliN stoichiometry also confuses the symmetry mismatch issue, and there is disagreement as to what the upper (CCW) and lower (CW) limits of FliM₁:FliN₃ stoichiometry are. Purified CW and CCW motors display the same C ring symmetry in EM[30, 38], offering little guidance. One school of thought is that the lower limit of stoichiometry corresponds to a complete 34-fold ring, which can be expanded by additional FliMs[62]. Another is that weak and strong binding sites correspond to the (disputed) separate FliM binding sites on FliG_m and FliG_c[64], with the two sites perhaps creating two concentric FliM/FliN rings with gaps, or perhaps a single ring (with or without gaps) plus additional bound but unstructured FliMs/FliNs, dangling inside or outside the ring. Alternatively, perhaps the *upper* limit of FliM stoichiometry corresponds to a full 34-fold ring as seen in cryo-EM, and lower stoichiometries correspond to a ring with gaps[61, 63]. This corresponds with some estimations based on fluorescence intensities[61, 63], and perhaps relates to C-ring gaps seen in early cryo-EM images of single purified motors[157] and more recently in *in situ* with cryo-ET[15]. Presumably conformational spread requires at least one protein to form a gapless ring, in which case models with gappy FliM/FliN rings require FliG to mediate conformational spread, and be gapless (i.e. present in 34-fold stoichiometry).

This story is also potentially complemented by recent observations of turnover in the T3SS analogue of FliM₁:FliN₃ [28]. While the total number of units is approximately constant, they turn over, with a timescale dependant on whether the T3SS is actively secreting and whether the export ATPase (see below) is catalytically functional. In addition, the large number of units (\sim 22) inferred by fluorescence [28] and the biochemical evidence for inter-unit bonds[127] point to the T3SS having a BFM-like C ring, whereas most *in situ* Cryo-ET structures show no density where a C-ring ought to be[158, 159], or distinct “pods” of density connected to the export ATPase, but not joined into a ring[160]. The latter result

has driven speculation that some FliM₁:FliN₃ homologues are weakly bound and not visualised at all, while others are more strongly bound through anchoring to the export ATPase[127]. While the turnover studies in the T3SS did not explicitly quantify strongly and weakly bound populations[28], it is tempting to speculate that the export apparatus may be involved in BFM FliM₁:FliN₃ turnover. A role in constraining C ring size is also conceivable[41].

Turnover of FliG has been hard to probe until recently due to the difficulty of obtaining a functional fluorescent protein fusion, but recent work in our lab (manuscript in preparation) suggests a complete absence of FliG turnover.

Assembly In addition to the above, there are also open questions regarding assembly. While the sequential assembly of FliF, FliG, FliM, FliN and the cytoplasmic export apparatus is well-established, it is unclear if rings form strictly sequentially (i.e. FliF ring forms, then FliG ring forms, then FliM ring forms, etc) or if the process is more co-operative (i.e. FliG ring does not fully assemble without help from FliM, etc.) I would argue that available evidence points to the latter: recall that in purified basal bodies of the FliF-FliG fusion, ordered FliG density in the C ring is only observed if FliM and FliN are present[44], suggesting disorder in the absence of FliM and FliN (or, alternatively, damage during the purification process that removes FliM and FliN). Similarly, mutation of the FliM_m FliG-binding motif can reduce FliG-FliG crosslinking yields[89]. Finally, an *in vivo* fluorescence study of part-assembled motors using GFP-tagged FliF, FliG and FliM in *E. coli* showed less FliF when FliG was deleted, less FliG when FliM was deleted and less FliM when FliN was deleted, hinting strongly at a co-operative assembly process, with the caveat that the bulky fluorescent proteins might be destabilizing native interactions[46]. Some studies even show that, for example, the FliF ring (at native FliF concentration) does not form at all without FliG[47]. Likewise, in some species, FliM does not obviously incorporate into the motor without FliN [161]

The above fluorescence study along with others (e.g. FliM and FliN turnover studies) also note three well-defined populations of fluorescent foci: fixed bright

spots at the cell pole (plausibly aggregates), fixed medium-intensity spots always seen at the site of working motors, and $\sim 45\%$ dimmer spots which are mobile in the cell membrane and do not exhibit FliM/FliN turnover[46, 61, 63, 162]. The dimmer spots, assumed to be assembly intermediates (perhaps mobile because their hook/rod has not yet punctured the cell wall) make up a substantial fraction of observed spots; more than half in some reports[46]. The exact nature of these intermediates and their abundance remains to be explained. The bright spots at the cell pole also exhibit turnover[46], perhaps a hint that they are more than just aggregates.

We have already mentioned gappy C rings occasionally seen *in situ* with cryo-ET[15]. This includes motors which seem to lack a C ring entirely, despite having fully formed flagella (which we know requires a C ring to assemble). Again, these structures are poorly understood.

1.1.9 Export Apparatus

We have covered the C ring in comprehensive detail, as it will be central to the aims of this thesis. The remaining parts of the motor are less central, and we will cover them in less detail.

The sequential construction of the rod, hook and flagellar filament (refer to fig 1.2a) requires the delivery of protein components down a narrow channel inside the growing rod/hook/filament. This is orchestrated by the export apparatus, which must identify the current stage of construction, select the appropriate components from the cytoplasm, potentially unfold them (to fit through the narrow channel), and feed them into the pore[25]. This is non-trivial.

The export apparatus is assembled from many different protein families, which we can broadly divide into membrane-embedded and cytoplasmic parts. The membrane-embedded parts (green in figure 1.3a) include FliP/FliQ/FliR/FliO, about which not much is known, and FlhA/FlhB. All are associated with the MS ring. FlhA has a cytoplasmic domain (FlhA_c), expected (by comparison with cryo-ET of its T3SS analogue) to form a nonamer ring inside the C ring[47, 160, 163]. Fluorescence studies suggest that in *E. coli* FlhA assembles before FliF

with ~ 20 fold stoichiometry[46], whereas in *S. typhimurium* it assembles after FliF with ~ 10 fold stoichiometry and cooperatively with the other membrane-embedded export components[47]. FlhB also has a cytoplasmic domain (FlhB_c) which interacts with FlhA_c[164], together forming the “export gate”[165], thought to be the site of binding for export substrates[166].

The cytoplasmic components, FliI/FliH/FliJ are notable for their homology to the F₁ part of F₁F₀ ATP-synthase[167]; another rotary motor (Figure 1.3 c). Briefly, FliI is homologous to the α and β subunits of F1 which form a 3-fold symmetric barrel that undergoes conformational changes driven by ATP binding and hydrolysis, FliJ is homologous to the barrel-embedded rotor (γ subunit) which rotates as a result of those changes, and FliH is homologous to the stator which holds the barrel in place.

FliH and FliI form stable FliH₂FliI₁ complexes, and FliH has a short peptide tag on the end which can bind either to the hydrophobic patch of FliN or to FlhA_c. Current understanding is that these subunits assemble onto the bottom of the C ring, which (together with ATP-binding and the addition of FliJ) promotes the formation of a FliI₆ barrel, highly similar to its ATP-synthase homologue[165, 167]. The FliJ in this barrel then docks into the FlhA_c ring[168].

FliH₂FliI₁ complexes also play a second role, which is binding to export substrates, usually accompanied by a substrate-specific chaperone protein (which may possibly compete with FliH for binding to FliI[169]). As FliH and FliI both have FlhA_c-binding activity, they presumably help shuttle substrates to the export gate, consistent with observations of FliI turnover (albeit not fast enough to account for typical export rates)[165]. FliI is implicated in ATP-aided cheparone-removal and substrate unfolding, in preparation for export[170].

Despite all this, FliI is not essential for export; its absence can be overcome somewhat by mutations in FlhA/FlhB, overexpression of other export proteins or an increase in PMF, which is the real driving force between substrate export, rather than ATP hydrolysis[166, 171]. Biochemical evidence points to FlhA forming the H⁺ pore which must facilitate this. In some species, export can additionally utilize a sodium motive force (SMF)[166]. The role of the FliI₆FliJ₁ ATPase seems to be that

of a rotary ignition key, which somehow switches the export gate into a more efficient mode of export[166, 172]. This requires ATP hydrolysis, but only at a low rate[173].

FlhA_c has been crystallized as 4 separate flexibly-linked domains[174, 175], and isolated monomers weakly bind all substrate-chaperone complexes with varying affinity[176]. It is likely that a conformational change in the ring of FlhA (and probably FlhB) switches the affinity for rod/hook/filament proteins, co-ordinating assembly[174]. It is already known that a molecular ruler, FliK, senses the length of the hook and triggers a switch in substrate-selectivity through an interaction with FlhB[25] which involves FlhB self-cleavage[177]. Slow turnover has been observed in GFP-tagged FlhA[46, 47] (though not the T3SS analogue [28]), and is unexplained.

The actual tagging of proteins for export is poorly understood: unstructured N-terminal sequences are typically seen, but are not sufficiently similar that it is easy to identify export-tagged proteins. Some proteins, but not others, require additional tag sequences[25]. Information from untranslated regions of mRNA also seems to be used for substrate recognition[178].

The details of protein export through the growing rod/hook/filament are also hotly contested; some propose that single-file diffusion of substrates down the channel is sufficient, while others suggest that export substrates are bound together in a long chain, and folding of proteins once they exit the channel helps pull through subsequent proteins[25].

1.1.10 Rod, Hook and Filament

Rod and Bushing The Rod is complex; joined to the FliF ring with FliE and comprising of FlgB, FlgC, FlgF and FlgG, all of which are similar and excreted by the export apparatus[17, 179]. A FlgJ cap on the growing rod is involved in rod formation, and hydrolyses peptidoglycan, allowing the rod to pierce the cell wall[180]. A variety of accessory proteins in the periplasm facilitate the construction of the P and L rings (FlgI and FlgH respectively; Figure 1.3a) around the rod, presumably acting as a motor bushing[17]. In addition, the P-ring either scaffolds

stators[181] or scaffolds additional rings which themselves scaffold stators, depending on the species (section 1.1.11)[72].

Hook and filament The hook (FlgE) adjoins the rod, constructed with the aid of a FlgD cap which dissociates after construction. Hook length is regulated at $\sim 55\text{nm}$, ~ 120 copies of FlgE. Helical rows of FlgE (Figure 1.10a) can compress and expand as the hook rotates, allowing it to act as a universal joint; flexibly translating rotation from the motor axis to the axis of the filament. This enables flagella bundle formation[17, 182].

The adapter proteins FlgK and FlgL join the flexible hook to the more rigid filament, constructed of FliC (flagellin) with the aid of a FliD cap. The filament may grow up to $15\mu\text{m}$ or longer, incorporating upwards of 20,000FliCs. FliC may switch between two forms corresponding to right handed or left handed protofilaments; the helical shape of the filament comprises a mixture of these forms (Figure 1.10b,c). A change in motor direction exerts a force which switches the handedness of the filament through a change in the ratio of FliC conformations[5, 17, 183]

1.1.11 Stators

Structure The stators, the “fixed” elements of the motor (relative to the cell) are stable 4:2 complexes of the integral membrane proteins MotA and MotB[185], which together form selective ion channels in the inner membrane[186] that are responsible for transducing electrochemical energy into mechanical energy.

MotA has large cytoplasmic loops containing the charged residues essential for torque generation through an electrostatic interaction with FliG, as discussed in section 1.1.8.2 [93, 94, 187]. MotB on the other hand is mostly located in the periplasm. While stator crystal structures are limited to some periplasmic parts of MotB[188], crosslinking studies have mapped the rough relation of transmembrane helices[189], which show two trans-membrane channels formed at MotA:MotB interfaces. Protonation of a conserved aspartate residue projecting into these

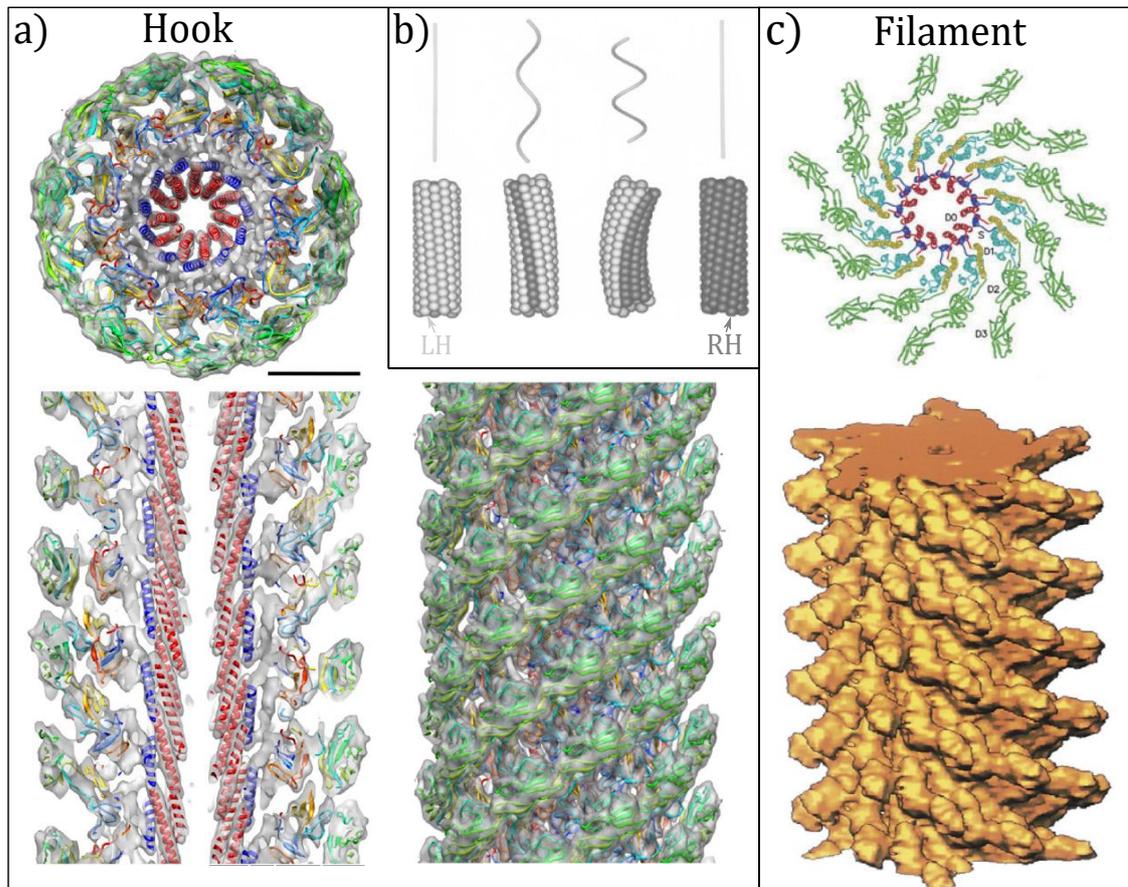


Figure 1.10: a) Crystal structure of FlgE docked into EM structure of the hook[184] b) Filament structure determined by mixture of L-type and R-type conformations of FliC[17] c) FliC arrangement and EM structure of the filament [5]

channels from MotB is almost certainly linked to a conformational change which propagates down to the cytoplasmic loops of MotA[94, 189, 190].

The stator channels only conduct when incorporated into the motor[188, 191]. Outside the motor, a “plug” in the periplasmic part of MotB blocks the channel[188, 192]. The removal of this plug (speculatively triggered by MotA-FliG interaction[193]) is probably associated with an extension of the periplasmic part of MotB, allowing peptidoglycan-binding and periplasmic-disc-binding motifs (see below) to reach their respective targets[188, 194]. The periplasmic part of MotB also seems to be involved in the load-dependency of stator incorporation (see below)[195].

Diversity of conjugate ions While the stators of *E. coli* and *S. typhimurium* are powered by H^+ transit, a great many other species have stators powered by Na^+ .

Recently, reports of a K^+/Rb^+ -powered stator in *Bacillus alcalophilus* (associated with alkaline K^+ -rich environments)[196] and a Ca^{2+}/Mg^{2+} -powered stator in a *Paenibacillus sp.* strain (isolated from a Ca^{2+} -rich hot spring)[197] suggest a wide range of possible coupling ions reflecting a variety of bacterial habitats.

Excepting a few small differences[95, 198], H^+ and Na^+ stators show remarkable similarity in their structure and interaction with the motor, to the extent that functional hybrid motors (rotors and stators from different species) and hybrid stators (mixed MotA and MotB sub-domains from different species) are easily engineered[95, 199]. Even the Na^+ stator of *T. maritima*, an ancient hypertherophile genetically very distant from *E. coli*, is functional in the *E. coli* motor if the periplasmic MotB domain is replaced with its *E. coli* equivalent[200]. A number of species have now also been found in which the same rotor is powered simultaneously by H^+ and Na^+ stators[151, 201]. Although the new K^+/Rb^+ and Ca^{2+}/Mg^{2+} stators have been less well studied, the former is functional in *E. coli* motors[196], and the latter is both functional in *B. subtilis* motors and naturally co-powers a single rotor alongside a H^+ stator[197], suggesting compatibility and similarity amongst all known stator types.

Assembly into the motor Expression of stators in a stator-deletion background leads to a stepwise recovery of motor speed, suggesting that stators are independent torque-generating modules[202]. Fluorescent studies in *E. coli* and a number of other species show continual assembly and disassembly of stators in the motor, with a typical half life of $\sim 30s$ [149–151]. Rates of turnover are responsive to ion motive force, which is required for stator localization to the motor in some cases, and responsible in dual-ion motors for tuning the ratio of stator types[150, 151]. Turnover is also responsive to torque on the motor[153–155], recruiting extra stators to compensate for increased load, and plausibly mediating the surface-sensing role of the BFM in biofilm formation[18].

In *E. coli*, the periplasmic part of MotB interacts with the periplasmic P ring (FlgI, Figure 1.3a)[181], whereas in some other species the P ring may template

a variety of related periplasmic structures which themselves interact with the periplasmic part of MotB[69, 72, 203] (Figure 1.5). These various structures may serve to stabilize the motor at high rotation speeds (e.g. in the high speed motor of *V. fischeri*)[188, 204] and / or increase the motor torque output by increasing the number and radius of stators[72]. Freeze-fracture EM images, steps in speed recovery and fluorescence studies all suggest a maximum of 11 stators in the *E. coli* motor, whereas Cryo-ET studies show 13-fold[72], 16-fold[69], and 17-fold[72] symmetry of stator placement in various species with extended periplasmic structures. Interestingly, motors with these extended structures have stators much more detectable in Cryo-ET than those of *E. coli* type motors, possibly hinting that the stators may be more stably incorporated[69, 72]. Interaction with these periplasmic structures has been shown to be necessary for stator incorporation into the motor in a number of species[72, 188, 205, 206].

Additionally, it has been shown in *E. coli*/*S. typhimurium*-type motors that the FliG_c-MotA interaction is necessary (at wild-type expression levels) for efficient stator incorporation into the motor[97, 207, 208]. A recent report also suggests that in a number of species, peptidoglycan maturation enzymes (and therefore stator-peptidoglycan interactions) are also necessary[209]. Thus, it seems likely that stators rely on multiple interactions to localize at the motor, presumably coupled to an unblocking of the ion channel.

1.1.12 Accessory proteins

In addition to the core motor proteins outlined above, I will briefly outline some of the additional proteins which apparently interact with the motor. Note that many of these are postulated to interact with FliG, and thus could be relevant to templating of FliG *in vitro*.

YcgR YcgR appears to downregulate motor activity in response to high levels of Cyclic di-GMP; a signalling molecule associated with a switch to surface adhesion and biofilm formation[210, 211]. It both acts as a brake on the motor [212] and

introduces a strong CCW motor bias[210, 213]. It is thought to operate by adjusting the FliG_{ca1-6}-MotA interface[56, 212, 214, 215], but there is great disagreement about whether this is mediated by binding to MotA[46, 211, 212] or FliG/FliM[56, 213, 214]. There is even some suggestion that YcgR could be a permanent part of the motor[214], though this is inconsistent with fluorescence studies in some species[211]. In *Pseudomonas aeruginosa*, a YcgR homologue inhibits surface motility by tuning competition between two distinct sets of H⁺ stators (only one of which is swarming-competent). Strains lacking this homologue additionally have altered cyclic-di-GMP levels, hinting that perhaps YcgR it is responsible for communicating the output of the torque sensor.

FRD Fumarate Reductase (FRD) binds to FliG[46, 216] in response to increased levels of fumarate, increasing CW rotation bias[217–219]. It essential for chemotaxis in some bacteria, while completely absent in others[216]. As fumarate is involved with cell metabolism, this presumably modulates motor switching in response to cell metabolism, perhaps contributing to energy taxis.

H-NS H-NS is a DNA-binding protein involved in the regulation of BFM pathways. However, a body of evidence says that it additionally binds to the motor directly via FliG[46, 91, 220], explaining non-motile (but flagellated)[221] and speed-increasing[222] mutants. Recently, a strong argument has been made that these mutants can be explained via indirect regulatory effects, and that much of the evidence for direct H-NS-FliG interaction can not be reproduced or is weak[223]. Measurements of in-vivo FRET binding[46], however, have yet to be explained. Thus the case for direct H-NS interaction with the motor, along with the possible role of such binding, is unclear.

FliL FliL is an inner-membrane protein with both periplasmic and cytoplasmic components[224], observed in *B. burgdorferi* Cryo-ET to sit between the motor and stators in the periplasm[225]. Its role varies between species, but always seems to involve stator-rotor interactions. In *S. typhimurium* it is essential for swarming

motility, associating with the motor[226] via FliF, FliG and the stators[46, 227] and stabilizing it at high load (i.e. during swarming), preventing the ejection of the filament[228] and increasing torque generation[227]. In *C. crescentus* on the other hand, it *aids* ejection of the filament; part of a regulated transition during surface adhesion[229]. In *V. alginolyticus* it aids stator incorporation into the motor, especially at high loads[224], and in *R. sphaeroides*[230] and *B. burgdorferi*[225], it is essential even for swimming motility. In both *E. coli*[231] and *P. mirabilis*[231–234] it is not essential for swimming, but clearly plays some role in surface sensing, and triggering the transition from swimming to surface-swarming morphology.

EpseE EpseE is another protein involved in biofilm formation, halting flagella rotation[235] in *B. subtilis* through a clutch mechanism; binding to FliG and somehow disengaging it from the stators[236].

Hsp90 Heat shock protein 90 (Hsp90) is a chaperone known to aid construction of oligomeric protein complexes in eukaryotes. An *in-vivo* FRET observation of direct interaction between Hsp90 and FliN/FliI in *E. coli* suggests that it could play a similar role in BFM construction[46].

F₁F_O ATP-Synthase A single report has been made of F₁F_O ATP-synthase association with the BFM in *E. coli*; the F₁ β -subunit interacting with FliG, and even affecting switching dynamics. F₁F_O is also postulated to maintain the PMF in the vicinity of the motor through ATP hydrolysis. The same report also claims NADH-ubiquinone oxidoreductase association with the motor[237].

1.2 Why template a C ring *in vitro*?

This thesis concerns the first steps towards templating a C ring *in vitro*, replacing the MS ring with an artificial FliG-binding mimic. The short-term aim of this (elaborated in section 1.5) is to test the FliG domain-swap polymerization model (sec 1.1.8.3). However, this would be a stepping stone on the way to a complete C ring *in vitro*, and we should take a moment to justify why this is worthwhile.

1.2.0.1 It avoids purification damage and allows observation of assembly

Many of the pressing structural questions about the C ring concern stoichiometries. Are there 26 or 34 FliGs? Does a full ring of FliM₁FliN₃ subunits correspond to the maximum or a minimum stoichiometry? As we have already discussed (section 1.1.5), EM images of purified motors cannot resolve gaps or disordered components. Apparent inconsistencies with *in-vivo* fluorescence studies also give us good reason to question how much the motor is altered during purification, especially given the dynamic nature of the motor and the harsh purification conditions required. We already know that no reported purification procedure retains the cytoplasmic export apparatus, despite direct connection to the C ring and a plausible role in motor turnover dynamics. This problem is shared by all biochemical assays which might be performed on a purified motor: chromatography, binding assays, and so on. So if we want to study a motor *in vitro*, it may be sensible to also assemble it *in vitro*. This not only avoids possible purification damage, but allows us to study the kinetics of C ring assembly, resolving (for example) the degree to which subunit assembly is cooperative.

1.2.0.2 Superior techniques are available *in vitro*

Given the issues with purification, and the potential difficulty of constructing the C ring *in vitro*, are *in vitro* experiments really worthwhile as compared to *in vivo*? Consider EM imaging: as compared with EM of purified motors *in vivo*, Cryo-ET of motors *in situ* lacks sufficient resolution to count subunits, and many biochemical assays (e.g. various chromatography methods) are impossible or very difficult, while those that are possible and useful (e.g. targeted crosslinking) are nevertheless prone to overexpression artefacts, crosstalk with other proteins in the cell, and similar problems which could be overcome *in vitro*.

Likewise, consider fluorescence studies, which are of particular interest for observing stoichiometry and dynamics. *In vivo* fluorescence can assess *relative* changes in stoichiometry, and can theoretically quantify numbers through stepwise

photo-bleaching, as has been attempted for FliM[61] (see section 1.4). However, fluorescent proteins are bulky; questions always remain over how much they affect function, and for some proteins such as FliG functional fusions have been very difficult to obtain, and plagued by questions about fluorescence maturation and unwanted cleavage. Furthermore, poor fluorophore performance and complicated background signal makes accurate counting very difficult[238]. FliN has proved impossible to count directly in this way[63], and while work is ongoing in our lab to count FliG, interpretation is still obscured by background issues. Organic dyes on the other hand are much smaller and thus less perturbing, giving much better performance and potentially making for very clear counting when combined with the low background of an *in vitro* experiment. Methods to label proteins with organic dyes *in vivo* are limited in yield and difficult, involving artificial amino acids or bulky protein tags. Furthermore, while it is possible to label proteins *in vitro* and deliver them into a bacteria via electroporation, yields are highly variable, and the process perturbing to the cell[239]. Therefore, if we want to observe a motor where every protein of a given family (or a well-known fraction) is labelled with an organic dye (as is necessary for stoichiometry experiments), only construction *in vitro* seems plausible.

The superior performance of organic dyes is also desirable for single-molecule tracking, and high-resolution measurements of distances and structural conformations using Fluorescence Resonance Energy Transfer (FRET)[240]. If FRET could be employed as a readout for the structural changes involved in motor switching, and correlated with the stoichiometry of bound CheY, a vast array of experiments could be performed. It would provide, for example, an opportunity to convincingly disprove the cooperativity of CheY binding to the motor, and (with the aid of various mutants) probe the tethered-bait binding model for CheY. The absolute control of CheY-P levels, impossible *in vivo*, would allow a definitive test of the controversial claim[136] that motor switching is not random, but always associated with change in CheY-P levels. Likewise, control over solvents could probe the involvement of hydrophobic interactions in switching. The details of CheY binding

and ring switch state could be correlated with FliM turnover dynamics. This is not an exhaustive list of target experiments, but hopefully makes the point that the assays available *in vitro* could be very powerful. It is also possible to envisage using an *in vivo* C ring as a base to build the external parts of the export apparatus, both to investigate its role in C ring dynamics, and to study in its own right.

1.2.0.3 The benefits of a controllable template

The most novel part of this project is the idea of substituting a physical template (the MS ring) with a controllable synthetic template. One reason for this is practical; as a membrane protein, FliF is difficult to purify and assemble into a ring *in vitro*. I am only aware of one attempt to construct a C-ring on a purified MS-ring, which (comparing their findings with later *in vivo* results) was clearly unsuccessful[113] (although it is not obvious that a new attempt would also be unsuccessful, given the benefit of the last 20 years' research.)

However, beyond these practical reasons is the concept that a changeable template can be used as an experimental tool. In section 1.5.1 we describe a scheme to correlate changes in a templated FliG structure to different configurations of template, in such a way as to test the domain-swap polymerization model. A range of similar experiments could be imagined, for example

- Perform a comparable experiment to probe the FliM domain-swap polymerization model.
- Change FliG-template linker length to test the hypothesis that ring size will grow until limited by FliG extension.
- Change template curvature and size to establish if FliG and other C ring proteins have curvature built in, or if they are flexible and controlled by the template.
- Change the strength of template-FliG binding and observe assembly kinetics to establish the importance of FliG-FliG interactions. The template could even be removed post-ring-formation to test this.

- Compare FliM binding to proximally-templated or distally-templated FliGs, to infer the importance of FliM-FliM interactions in assembly.

With a fully functional templated (switching) C ring, further possibilities arise, e.g.

- Control patterning of FliG switch mutants around the ring, and observe effects on whole-ring switching dynamics to parametrize the energetics of conformational spread
- Make FRET assays much easier by alternating FliGs with differently coloured double-labels, rather than having to label each FliG with two dyes.
- Template physiological and non-physiological mimics of the export apparatus, binding to FliN via the FliH FliN-binding region, to probe the hypothesis that the export apparatus is involved in turnover dynamics.

Additionally, patterning of templates on a surface could be used to facilitate cryo-EM data collection, creating dense but non-overlapping arrays of proteins or structures not otherwise achievable[241]. Templating helical filaments of FliG, for example, could facilitate imaging of the subdomain arrangements. Again, this is not meant to be an exhaustive list, but some examples of how the technique could be powerful.

Finally, consider that the templating approach could be useful for the study and control of other protein complexes. Even within the narrow field of the BFM, we can imagine multiple applications. For example, an artificial template could replace the role of the C ring in templating parts of the cytoplasmic export apparatus (FliH,FliI,FliJ,etc). Likewise, FlhB dimerizes via hard-to-reconstitute transmembrane domains, but has weak interactions between the cytoplasmic domains (FlhB_c) which plausibly become stable when templated in proximity by the transmembrane dimerization[164]. Similarly, its companion FlhA has a cytoplasmic part (FlhA_c) tethered flexibly to a transmembrane part. As the nonomer FlhA_c ring does not readily form from purified FlhA_c, we can easily imagine that it

requires templating via its transmembrane parts. It is widely supposed that the arrangement of the four domains of FlhA_c in the ring can change co-operatively, switching the specificity of export-substrate binding[174]. So for the cytoplasmic parts of FlhA and (also plausibly for FlhB), physiological interactions can only be measured *in vitro* if templated into their physiological structures. Given the difficulties of working with membrane proteins, and the unknown role of additional components in the membrane, this may be achieved most easily with an artificial template. Thus, templating a C ring *in vitro* may serve as a proof-of-concept experiment to inspire utility in other systems.

1.2.0.4 Drawbacks

There are, of course, caveats to this approach. The most obvious is that we have little way to know if our structure is physiological. *In vivo*, it is possible to focus our attention on demonstrably functional motors. *In vitro*, construction of a fully functional motor would be extremely challenging, not least because of the requirement for an energised membrane and the large list of required components. Thus, in the foreseeable future, we can only assess our constructs by looking for behaviour seen *in vivo*, and forming hypothesis based on *in vitro* measurements which can then be tested *in vivo*. The most convincing readouts of physiological behavior (e.g. a C ring with bistable conformational switching, responsive to CheY-P binding, conforming to measured FliM/N turnover dynamics) are years in the future, and certainly beyond the work in this thesis.

Furthermore, there are good reasons we might expect *in vitro* assembly not to work. It is not implausible that accessory proteins such as H-NS or Hsp90 may be required for assembly (see section 1.1.12), and if assembly is very sensitive to solution composition or precise details of the template, it may be difficult to find the right conditions by trial and error (See section 2.2.4 for elaboration). It may even be that assembly requires interactions with the membrane, or other cellular components which have yet to be identified (e.g. the associated ATPase complexes proposed by [237]). Looking further ahead, perhaps the dynamics of the

C ring; switching, turnover, etc, all rely on incorporation into an energised motor. We already know that switching dynamics are responsive to applied torque[140, 143], and motor switching statistics in the absence of CheY imply coupling to a non-equilibrium process (i.e. torque generation)[141, 142].

To conclude; a C ring *in vitro* would be amenable to powerful techniques not available *in vivo*, potentially capable of solving some of the most pressing questions about C ring structure, function and dynamics. Construction of this ring *in vitro* avoids the realistic prospect of damage during purification and allows us to investigate assembly dynamics. Assembly on an artificial template may both be the easiest way to achieve this, and opens up a wide range of additional investigative possibilities. However, a close interplay with *in vivo* experiments will be required to verify the utility of any templated constructs, and there is a possibility that our understanding of assembly is not comprehensive enough for assembly *in vitro* to be successful.

1.3 DNA Nanotechnology

A MS-ring-mimicking template needs to be robust, addressable with nanometre precision, producible *en masse*, and easily modified. Realistically, there is only one technology currently capable of achieving this: DNA nanotechnology.

1.3.1 Overview

The specificity of Watson-Crick base pairing and the ease of DNA synthesis make DNA a versatile material for the construction of self-assembled nanostructures and machines. DNA sequences may be rationally generated that, when annealed, associate with their designed partner sequences, but not with other unintended partners. This forms the basis of almost all DNA nanotechnology, and allows remarkably robust self-assembly of structures and operation of molecular machines.

The first designed DNA structures involved simple 4-way junctions, self assembled into a 2D lattice[242] (Figure 1.11a). This was followed by the design of small (~5-10nm) well-defined structures comprising a number of unique short (10s of nucleotide)

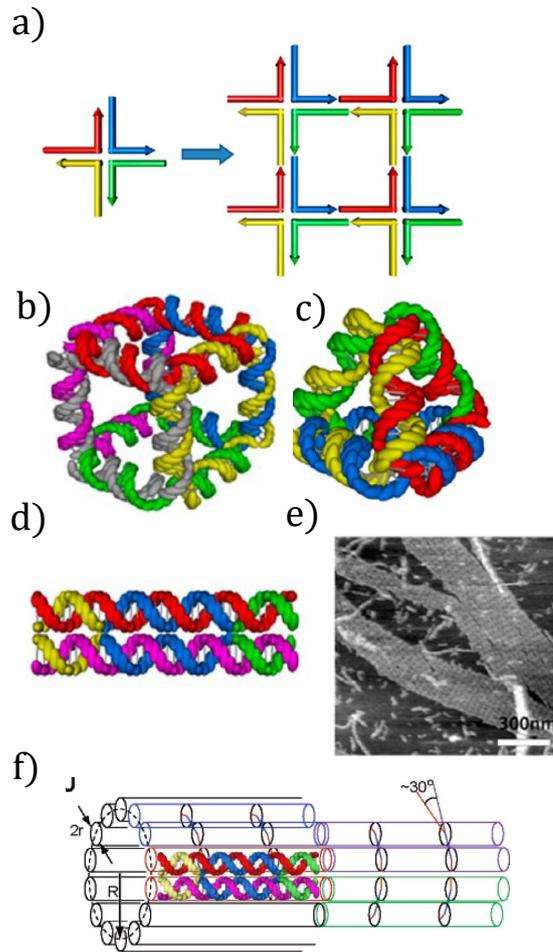


Figure 1.11: a) Design of 2D arrays b) DNA cube of [243] c) DNA tetrahedron of [244] d) Double crossover (DX) tile [245] e) 2D array formed from DX tiles (AFM) [245] f) Tubes formed from DX tiles [246]. Figures a-e adapted from [248].

oligos; early designs requiring enzymatic ligation and complex assembly with poor yields [243] (Figure 1.11b), but later designs achieving one-pot assembly with high yields, reliant only on Watson-Crick base pairing [244] (Figure 1.11c). Simultaneously, rigid double crossover (DX) tiles (featuring oligos which cross between neighboring duplexes, holding them in close proximity; Figure 1.11d) were being established as a building block for periodic 2D arrays [245] (Figure 1.11e), tubes [246, 247] (Figure 1.11f) and other structures [248].

A major breakthrough came with the development of DNA origami in 2006 [249]. Whereas previous structures involved small numbers of unique short oligos, origami comprises one very long “scaffold” oligo (1000s of nucleotides; typically phage DNA) folded into a designed shape by hundreds of unique short “staple”

oligos, which from crossovers between different sections of template (figure 1.12a). This makes 2D shapes $\sim 100\text{nm}$ in dimension (figure 1.12b) in remarkably high yield, with individually addressable staples sites for surface patterning (Figure 1.12c). In the 10 years since, construction of such 2D shapes has become routine in nanotechnology labs, and there has been an explosion of more sophisticated 3D origami techniques, including controlled curvature of parallel helices (figure 1.12d) and the rendering of arbitrary polyhedral surfaces (figure 1.12e). Recent years have also seen structures constructed from hundreds of small DNA “brick” oligos (figure 1.12f), alongside many other novel strategies conferring incredible variety in the structures that can be constructed. The full depth of DNA structural technology, and the wide array of DNA machinery and computation networks is far beyond the scope of this thesis; we will limit ourselves to a brief overview of the application of DNA nanostructures to templating various molecules. We will cover details of various design strategies as they become relevant in later chapters.

1.3.2 DNA nanostructure templates

Optically active particles Plasmonic nanoparticles, organic dyes, quantum dots and other optically active molecules may have intermolecular interactions that depend precisely on their distances. DNA has been used to both understand and exploit these effects.

One strand of research has involved using programmed pairwise interactions between DNA-labelled nanoparticles to self assemble lattice structures[253](Figure 1.13a). This approach has been applied to wide range of nanoparticles and lattice structures[254, 255], with dynamic control of the DNA enabling lattice reshaping[256], or fragmentation into small nanoclusters[257]. Similarly, self-interacting DNA structures can be decorated with nanoparticles to form 2D arrays[258, 259] (Figure 1.13b) or helically-labelled tubes[260] (Figure 1.13c). Most recently, assembly of rigid addressable DNA origami disks around nanoparticles has fused the approaches, allowing angular control of interactions to form different 2D lattices[261](Figure 1.13d).

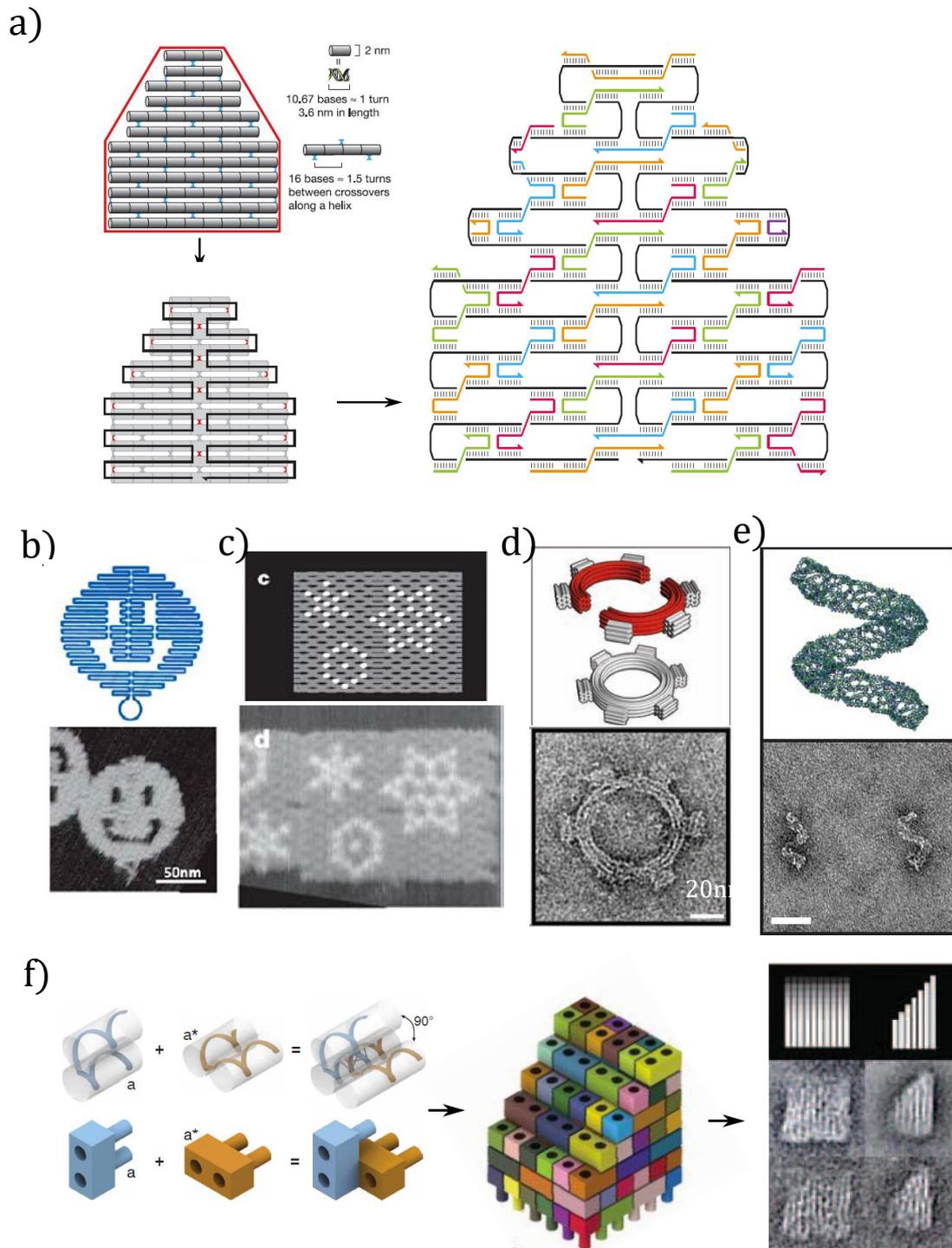


Figure 1.12: a) DNA origami[249] b) 2D DNA origami shapes[249] c) Addressability of DNA origami attachment points[249] d) Complex 3D designs of DNA origami[250] e) DNA origami polyhedra[251] f) Assembly with DNA bricks[252]

Another research approach involves discrete DNA structures, which are addressed specifically with nano-particle labelling sites. Early examples used simple short linear DNA templates[262] (Figure 1.13e), which may be extended into long polymers through rolling-circle polymerization of the DNA [263]. More recent examples involve DNA origami or helical bundle templates. For example, 2D origami tiles have been used for precise patterning of self-similar nano-particle chains[264] (Figure 1.14a), gold nanorods with controlled alignment[265] (figure 1.14b), and differential spacing of gold nanoparticles with organic dyes to measure the variation of dye-quenching with distance[266](figure 1.14c). 3D origami structures have been used to control the spacing of gold nanoparticles to within $\sim 1\text{nm}$ accuracy, creating localized light-enhanced regions for surface-enhanced Raman spectroscopy[267] (Figure 1.14d) and single-molecule fluorescence[268] (Figure 1.14e). Origami tubes have been used to make nano-particle helices[269] (Figure 1.14f), and 7-helix bundles and 2D tiles have both been used to template light-harvesting chains[270, 271]. Simple control of dye placement, stoichiometry and distance have also made DNA origami nanostructures popular calibration standards fluorescence microscopy techniques[272–276].

Proteins DNA has also been used, in a number of ways, to manipulate proteins and their activity. Details of protein-DNA conjugation will be covered in section 2.3.1; here we will just summarize applications.

DNA can be used to bring together separate DNA-conjugated parts of a split protein such as GFP[277] (Figure 1.15a) and the enzyme murine dihydrofolate reductase[278] (Figure 1.15b), controllably restoring natural function. Likewise, an enzyme which relies on electron transfer between redox centres in two adjacent domains can have the isolated domains brought close via a DNA linker, controlling enzymatic activity[279] (Figure 1.15c).

DNA devices may also control the interaction of enzymes with tethered co-factors[280] (Figure 1.15d) or inhibitors[281] (Figure 1.15e), and even inactivate enzymes through a hybridization-driven application of mechanical force[282].

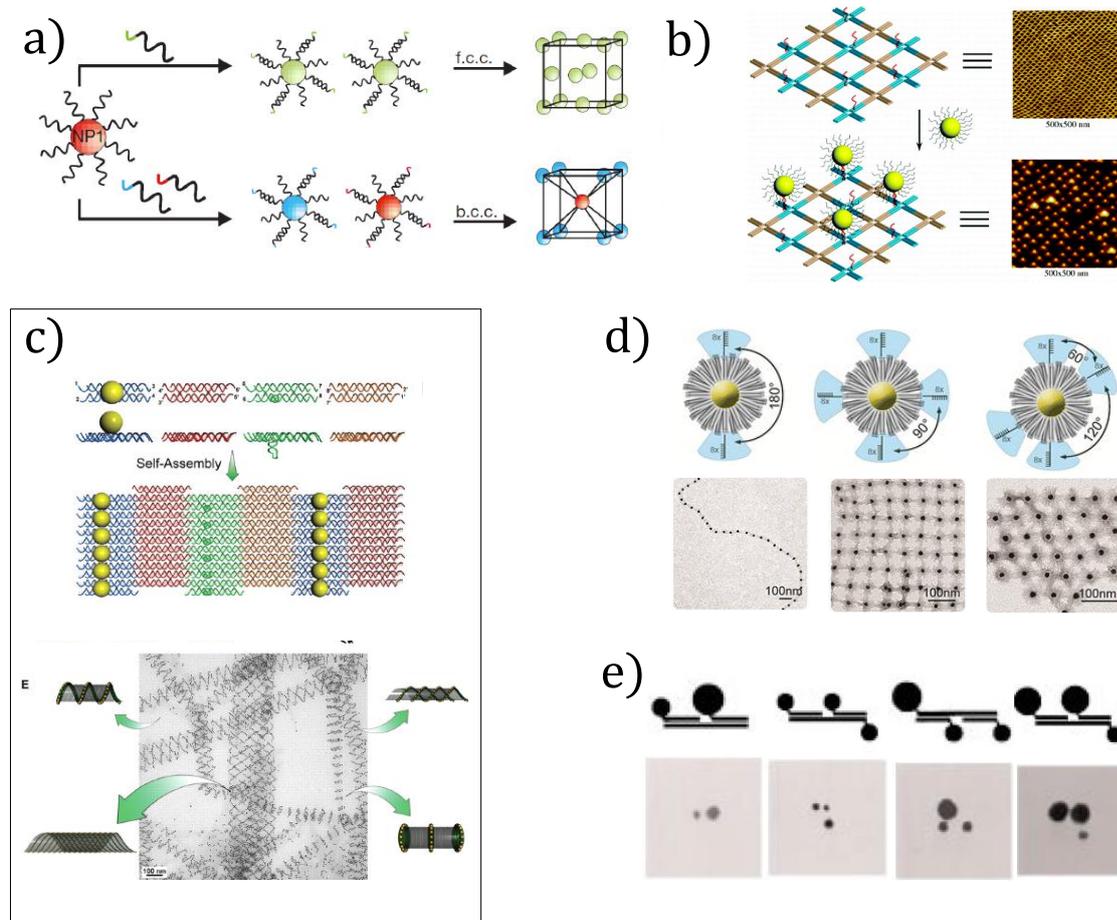


Figure 1.13: DNA templating of plasmonic nanoparticles. a) Nanoparticle lattices mediated by pairwise DNA interactions[253] b) 2D nanoparticle arrays[258] c) Helically-labelled polymer tubes[260] d) DNA “nanoflower”[261] e) Nanoparticle assemblies on short linear DNA templates[262]

DNA templates may also tether together protein-binding ligands at different distances, either to co-operatively combine binding properties for vastly increased affinity[283, 284] (Figure 1.15f), or to (through optimization of the ligand-ligand spacing) probe the distance of different binding sites on the protein surface[285, 286] (Figure 1.15g).

Many studies recently have involved enzyme cascades, templating enzymes in proximity to control rates of diffusive substrate transfer. For example, glucose oxidase and horseradise peroxidase have been templated on 2D DNA lattices[287], origami tiles[288], and the interior of origami nanotubes[289, 290]. Rates of substrate transfer can be increased further by using a noncatalytic protein “bridge” to connect

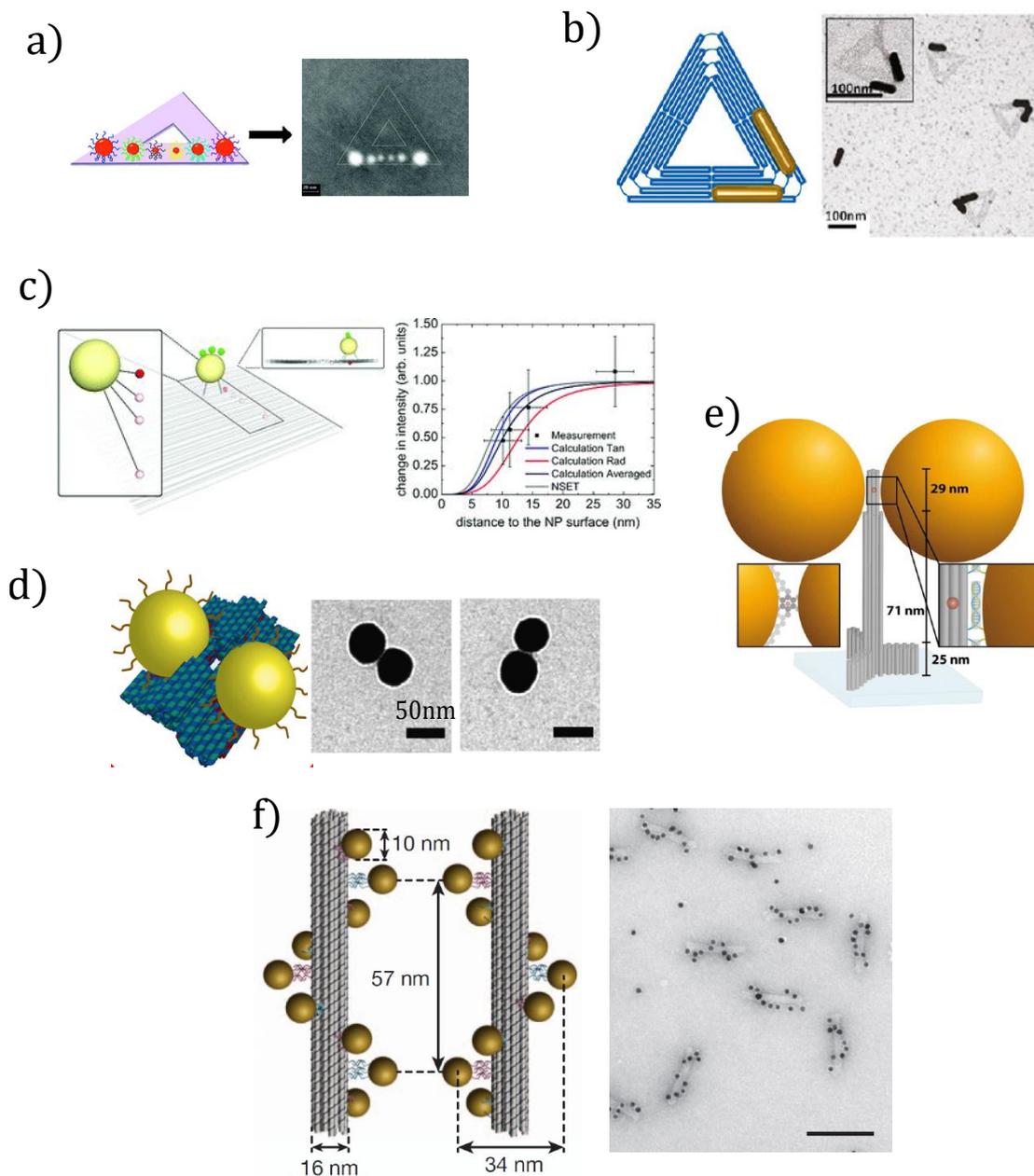


Figure 1.14: DNA templating of plasmonic nanoparticles. a) Nanoparticle chains on 2D DNA origami[264] b) Gold nanorods on 2D DNA origami[265] c) Variable nanoparticle - dye distances on DNA origami to investigate quenching[266] d,e) Control of nanoparticle positioning for surface-enhanced Raman spectroscopy[267] and single-molecule fluorescence[268] respectively f) Nanoparticle helices templated on origami tubes[269]

the hydration shells of the two enzymes[288] (Figure 1.16a), or tethering of the substrate to a swinging arm[291] (Figure 1.16b). In a photonic equivalent, DNA has been used to attach a FRET cascade of dyes to cytochrome *c*, forming an artificial light harvesting system[292] (Figure 1.16c).

DNA templating has also been used to template larger groupings of proteins. For instance a DNA origami tube was used as a platform for a controllable number of dyneins and kinesins; the detailed movement of this structure along microtubules was used to make inferences about the cooperative and antagonistic interactions between the motor proteins[293] (Figure 1.16d). A similar study templated different varieties of myosin on a 2D origami tile, and compared collective trajectories on actin networks infer differences between the different myosins[294]. DNA-conjugated motor proteins have also facilitated artificial microtubule networks, assembled or disassembled through crosslinking by DNA structures templating multiple kinesins. A second class of kinesin-labelled DNA template can controllably move cargo on this network[295] (Figure 1.16e). Polymerization of amyloid fibrils has also been directed inside DNA tubes by a nucleating peptide, with the direction of the tube (and thus the fibril) controlled by tube templating on an origami superstructure[296]. Finally, periodic DNA structures have been used to template dense 2D arrays of protein to facilitate single-particle cryo-EM data collection[241] (Figure 1.16f).

To my knowledge, however, nobody has yet templated the constituent parts of a protein complex on a DNA scaffold.

1.4 Fluorescence techniques

Many of the unresolved questions about the C-ring concern subunit stoichiometries, which may be dynamic and heterogeneous between motors. Single-molecule fluorescence is the only easy way to measure the subunit stoichiometry of individual complexes with some degree of time resolution, and (as discussed in section 1.1.8.5) has been widely used to study the BFM *in vivo*, along with a vast array of other biological structures both *in vivo* and *in vitro*[297, 298].

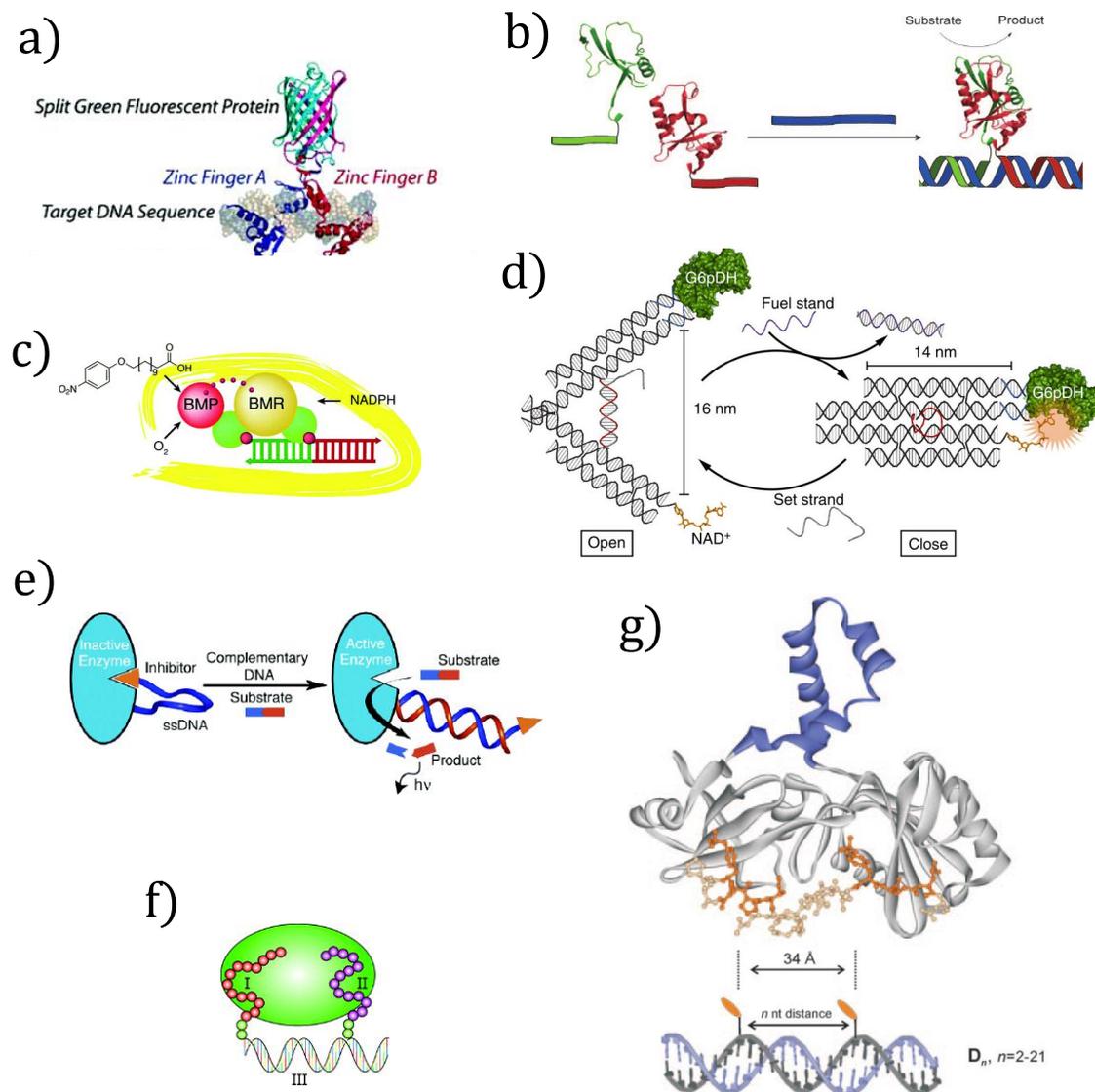


Figure 1.15: DNA templating of various proteins. a) Split GFP[277] b) split murine dihydrofolate reductase[278] c) Interacting redox domains of Cytochrome P450 BM3[279] d) G6pDH and cofactor[280] e) Enzyme and inhibitor[281] f) Protein-binding peptides[283] g) Templating of peptides to measure separation of their conjugate binding pockets[285]

There are a number of ways to measure stoichiometries with fluorescence. In the most popular *in vivo* methods, all fluorescent dyes associated with a structure will contribute to the intensity of a single diffraction-limited fluorescent foci (Figure 1.17 a). However, cellular autofluorescence and freely-diffusing dye-labelled proteins and will contribute to a noisy background beneath this spot. If structures of interest are near to the coverslip, Total Internal Reflection Fluorescence (TIRF) may be used to reduce the background by limiting illumination to within $\sim 100\text{nm}$

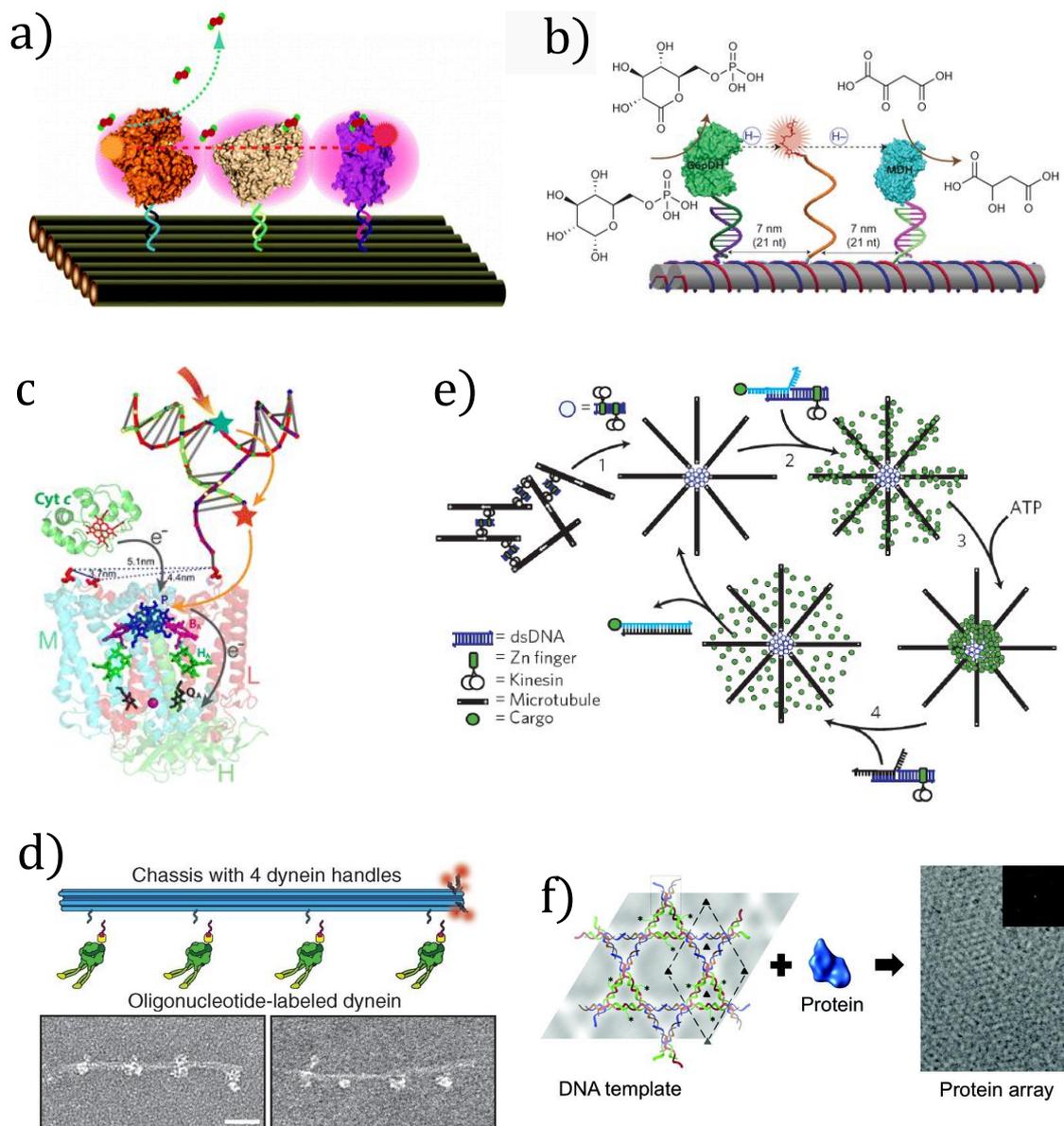


Figure 1.16: DNA templating of various proteins. a) Enzyme cascade using a noncatalytic bridge protein[288] b) Enzyme cascade with substrate tethered to swinging arm c) Artificial light-harvesting system[292] d) Templated Dyneins e) DNA controlled microtubule networks[295] f) Templated protein array for efficient cryo-EM data collection.

of the coverslip[149, 297]. A number of protocols exist for the subtraction of this background; none particularly accurate, especially when the spot intensity is low[299]. Nevertheless, background subtraction is required to form an estimate of the spot intensity, which can then be converted to a stoichiometry estimate if the average intensity per fluorophore is known[298]. This is usually found in one of two ways: the first uses a control structure with known subunit stoichiometry, labelled with the same dye and measured in maximally similar conditions[63, 298]. Accuracy is limited by the difficulty of imaging control structures in an identical environment. The second approach avoids this problem by calculating the single-fluorophore brightness from the individual spot of interest itself, rather than a control structure. This relies on the stochastic de-activation (bleaching) of individual fluorophores during sustained illumination. As fluorophores bleach, the intensity of the spot should decrease in a stepwise fashion (Figure 1.17 a). *In vivo*, noise usually obscures these steps, and Chung-Kennedy filtering[300] is applied to try and recover them. In some cases, filtered traces are clear enough to count every step[301]. More typically, however, only a subset of steps are well resolved, and these are used to measure the single-fluorophore brightness, assuming that it is relatively uniform between fluorophores[149].

In vitro, superior dyes and lower background signals enable measurements so clear that steps can be counted individually, even without filtering (Figure 1.17 b). This is reassuringly direct, and does not rely on control structures or any assumptions about dye brightness. It is, however, confused by dye photophysics (e.g. blinking) and near-simultaneous bleaching events which cannot be resolved due to limited time resolution. There is also often ambiguity in the detection of bleaching events. This can be avoided by analysing total bleach times rather than individual bleaching events, but only for homogenous populations[302].

Other methods are available. For example, the photoactivatable fluorophores used for super-resolution imaging can also be used for counting; statistics of blinking can be used to infer the number of dyes present. However, accuracy depends on meticulous characterization and modelling of dye behaviour to correct

for inefficient photo-conversion, over-counting due to repeated activations, and localization uncertainties[303, 304]. A similar counting approach has also been successful *in vitro* using DNA PAINT[305], where blinking is caused not by dye photophysics, but by transient binding of dye-labelled DNA oligos to binding targets on a DNA origami structure[273]. However, this required long acquisition times, and counting precision was poor unless averaged over many structures in a homogeneous sample.

Another family of techniques uses pulses of confocal laser illumination too short for a single dye to undergo more than one excitation/emission cycle. Statistical analysis of the photons emitted per illumination pulse can be used to infer dye stoichiometry very quickly, with minimal photobleaching and minimal sensitivity to heterogeneous dye brightness[274] (Figure 1.17 c). The same approach can be applied to stimulated emission depletion (STED) microscopy, combining counting and super-resolution imaging[275]. However, while the time resolution is certainly appealing, very specialized equipment is required, and the precision is not obviously greater than counting of bleach steps.

1.5 Project Overview

In section 1.1 we introduced the BFM (Figure 1.18a), its self assembly (Figure 1.18b) and the mismatched symmetries of MS ring and C ring components (Figure 1.18c). We also, in section 1.1.8.2, introduced the domain-swap polymerization model for FliG, which could explain that symmetry mismatch (Figure 1.18d). Later (section 1.2), we went on to make the argument for templating the C ring (FliG, FliM and FliN) *in vitro* with an artificial template replacing the MS ring (FliF). The obvious material for such a template is DNA (section 1.3). This would be not only useful for understanding of the BFM, but demonstrate a new approach to the study of large protein complexes.

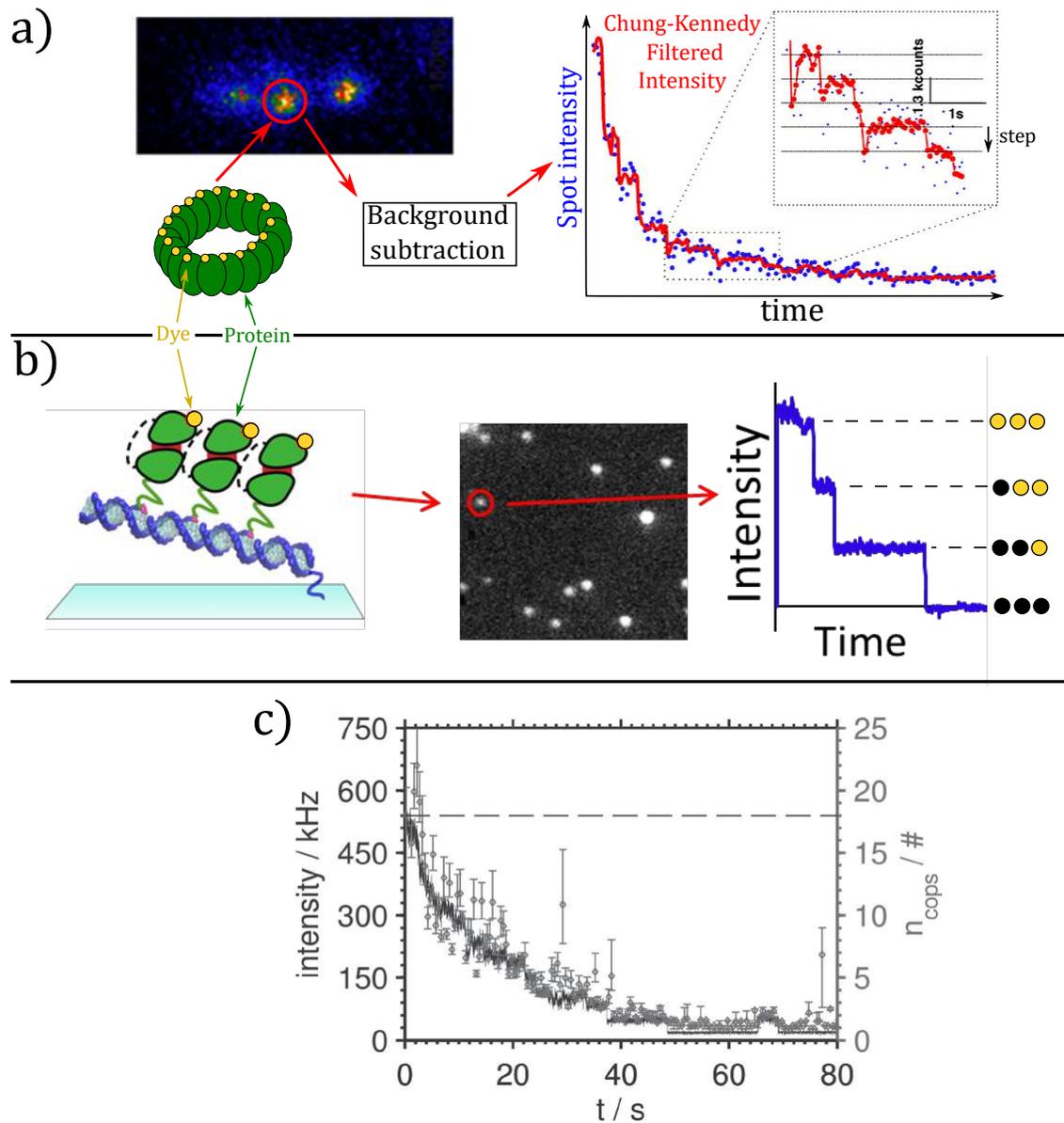


Figure 1.17: a) FliM-YPet imaged in E-coli with TIRF, counted via photo-bleaching steps[61]. b) Templated FliG structures imaged *in vitro* and counted via photo-bleaching steps (section 4.1.3). c) DNA tile with 18 fluorophores; simultaneously counted by photobleaching (black) and photon statistics (grey) [274].

1.5.1 Initial Plan

The long term goal of this project, then, is to template a working C ring on a DNA scaffold. This is a large project, in partnership with the groups of Lawrence Lee in Sydney and Keiichi Namba in Osaka; experts in protein crystallography and cryo-EM respectively. The original plan for the project was as follows: here in Oxford, the Turberfield group would design short DNA templates mimicking MS-ring fragments. The Lee group would produce FliG constructs, and various FliG mutants of interest. Together, these would be the basis for DNA-templated FliG polymers (with the addition of FliM and FliN, if necessary). A number of techniques would be used to study these polymers: single-molecule-photobleaching in the Berry group to count the stoichiometry of FliG on individual templates and Biolayer Interferometry (BLI)³ in the Lee group to measure bulk kinetics. The stoichiometry and binding dynamics of FliG on differently designed linear templates would be used to test the FliG domain swap model (Figure 1.19). After confirmed polymerization, results could be verified by targeted crosslinking *in vitro* and comparison of different FliG mutants. Namba group could then perform cryo-EM imaging of templated polymers, perhaps on long helically-templated filaments designed for efficient data collection. All this would pave the way for construction of complete C rings on more elaborate templates.

In the months prior to my involvement in the project, Jon Bath (Turberfield group, Oxford) had designed origami structures to template rings of a variety of sizes. The Lee group had produced a number of his-tagged FliG constructs, which were to be dye labelled through an engineered cysteine, and assembled on short linear DNA templates presenting NTA-modified (histag-binding) DNA. At this point, I began working on the project. My focus was to be on the single-molecule fluorescence aspect of the project, but there were a number of building blocks required before informative domain-swap experiments could be performed; design and assembly of linear templates (section 2.1), FliG expression and labelling (section 2.2), DNA-FliG conjugation (section 2.3), and development of hardware, software

³A technique similar to surface plasmon resonance

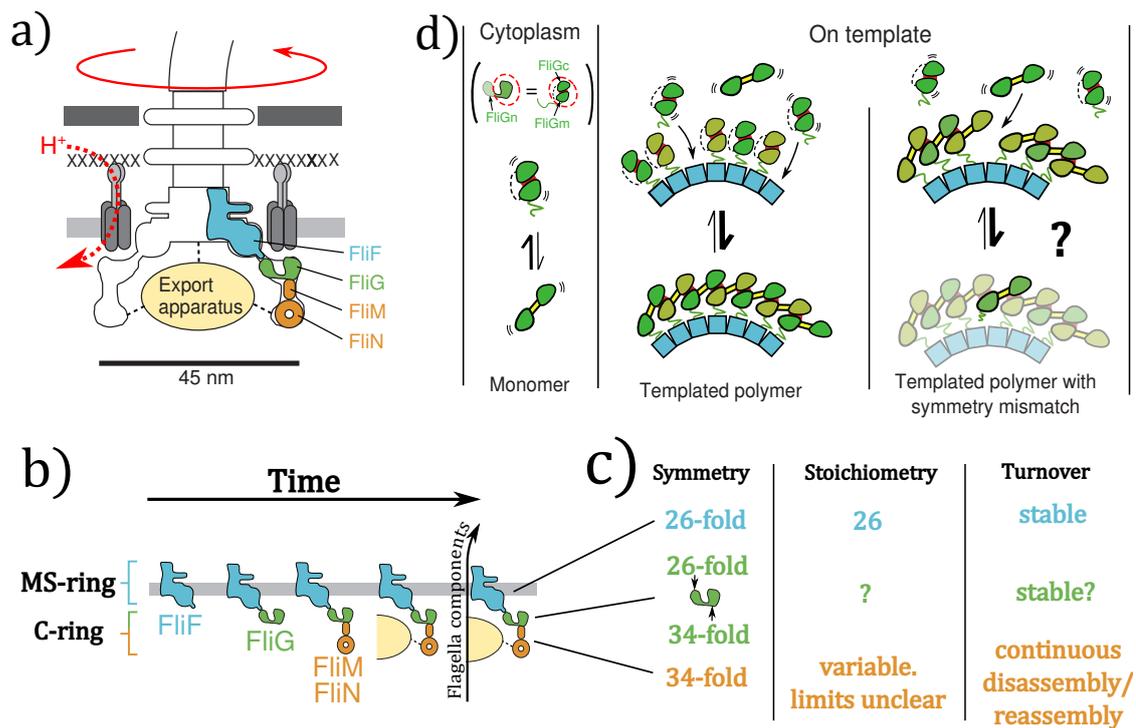


Figure 1.18: (a) The BFM: a rotary motor powered by ion flux across a membrane, essential for swimming, swarming and surface sensing; a canonical large protein complex. (b) The sequential self-assembly of the MS and C rings. (c) Symmetries observed in averaged EM structures of purified motors, and protein turnover observed with *in vivo* fluorescence. (d) Speculative domain-swap polymerization model for FliG assembly, which could explain the symmetry mismatch.

and protocols for the microscopy assays (section 3). Reliable measurement of N FliG's on N-site templates would be the proof that all these components were working as expected. As this goal has only just been reached at the time of writing, this thesis is a story of the development and testing of those components.

As these components were developed in parallel, informing each other in various ways, and shared in parts with my colleague Joel Spratt, it is difficult to form a linear narrative. In the interests of clarity, I will separately describe the development of each of the biophysical building blocks of the system (Chapter 2) and the development of microscopy techniques (Chapter 3). In the final chapter (Chapter 4), I will describe some key microscopy and PAGE experiments, in the loose context of a chronological narrative, bringing together elements described in the earlier chapters. Though I hope each section is clear in isolation, I will additionally give a detailed chronological overview here.

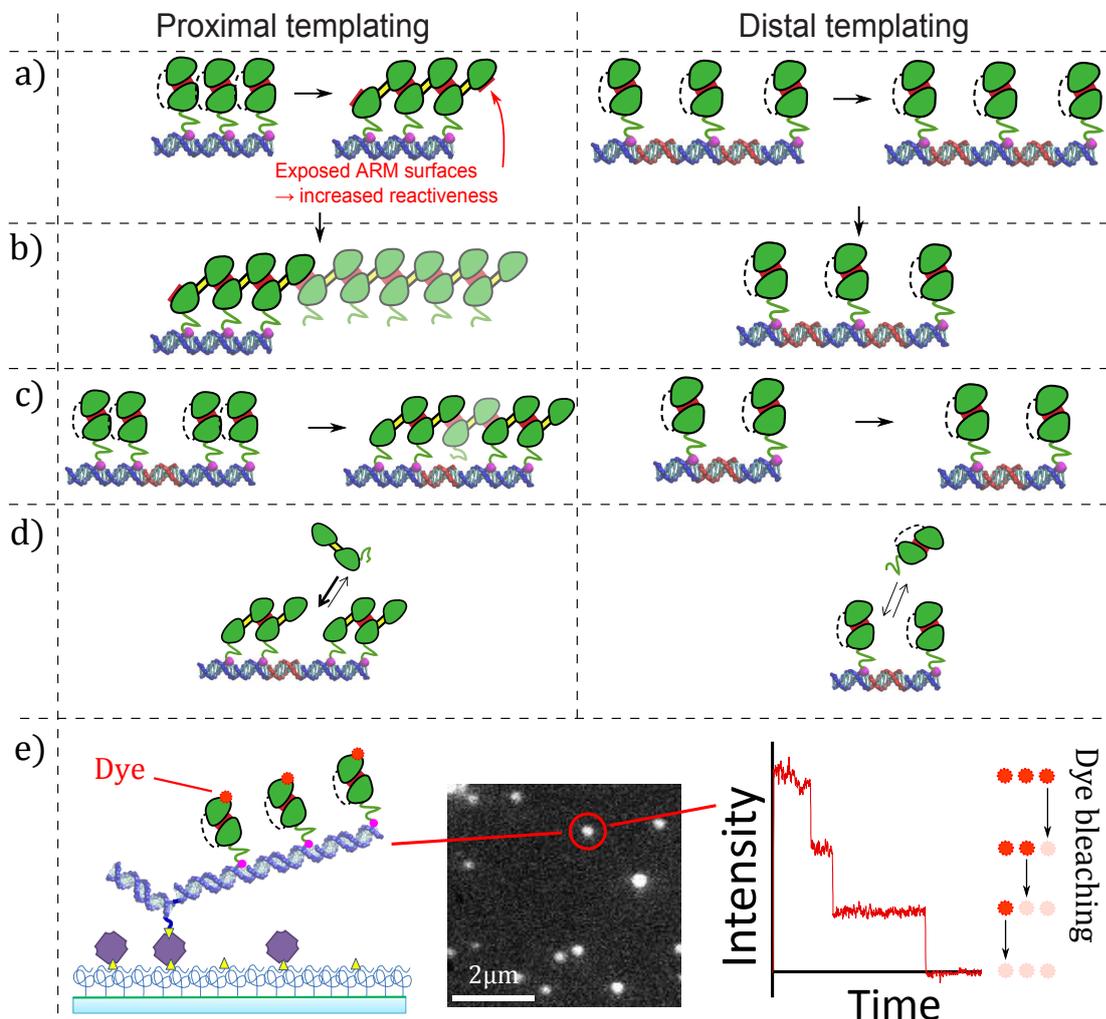


Figure 1.19: (a) We want to test a central hypothesis of domain-swap polymerization (see section 1.1.8.3), which is that proximally-templated FliGs will adopt a domain-swapped configuration whereby neighbours form intermolecular ARM-ARM interactions. At the ends of domain swapped polymers, exposed ARM surfaces increase reactivity to FliG present in solution. We are looking for signatures of this increased reactivity. (b) **Runaway Polymerization** In the strongest case, the exposed ARM surfaces seed a growing polymer of domain-swapped FliG. (c) **Gap filling** A slightly weaker possibility is that free FliG will fill only gaps flanked by exposed ARM surfaces (providing, in the process, a clear mechanism for a FliF-FliG symmetry mismatch). (d) **Gap filling kinetics** Weaker still is the possibility that FliG ultimately fills gaps regardless of exposed ARM surfaces, but that the exposed ARM surfaces have a measurable effect on the binding/unbinding kinetics. (e) **Dye counting via photobleaching steps** might be an ideal way to count FliGs to infer all the above effects.

1.5.2 Chronological overview

For the first few terms I was working (with Jon Bath's assistance) on NTA-modification of DNA, expression, dye labeling and purification of a FliG construct, design of linear DNA templates, and development of microscopy protocols and software. FliG expression/labelling/purification was reasonably straightforward and successful. NTA modification of DNA, however, was repeatedly unsuccessful, in that size exclusion chromatography showed no sign of FliG binding to the DNA. Simultaneously, I was testing different surface modification approaches for the microscopy, and struggling to get surfaces with low levels of non-specific FliG binding; essential for single-molecule experiments. In the following terms, NTA-modification of DNA was still unsuccessful, as was an alternate covalent conjugation strategy using maleimide-DNA. Concurrently, I established a surface modification strategy that was sufficient in terms of performance, but had some intrinsic fluorescence that was an unavoidable feature of the surface chemistry. This was not a problem at longer wavelengths, and inspired the relabelling of FliG with an alternative dye. I also at this point discovered and overcame some flaws in the canonical method of measuring dye labelling stoichiometry, eventually measuring labelling efficiencies of $\sim 100\%$. The microscope was part-rebuilt to allow a larger illumination and imaging area.

Later, I started to see signs of NTA-DNA binding to FliG with size exclusion chromatography, but not with gel electrophoresis (consistent with an understanding that electrophoresis breaks the NTA-histag link, as reported in previous work). Binding was also not visible with single-molecule microscopy, possibly due to problems with the surface modification. At the same time, I began experimenting with anti-bleaching systems for fluorescence, which were so effective that the microscope was part-rebuilt again, to maximise illumination intensity and reducing bleaching times. I also started characterizing assembly of the DNA templates. The following term, I could finally count FliG molecules on templates with clear bleaching steps. However, I found only $\sim 1/3$ of designed template sites appeared to be filled, indicating a problem with DNA template assembly, DNA-FliG conjugation, or dye

labelling. This coincided with the identification of uncertainties in the purification of NTA DNA. Attempts at alternate NTA modification strategies were unsuccessful.

At this time, a project student Joel Spratt, started working very closely with me on the project. After some work together trying to improve NTA-DNA purification, he spent most of the 10 week project trying to implement a guided covalent conjugation strategy. By the end of his project, that appeared to be successful, and I started work on developing a purification strategy for the covalent conjugates, and also on scaling up his protocols. The following term, Joel returned as a DPhil student working on the project. We expanded our stocks of dye-labelled proteins, including FliG from *salmonella* for compatibility with FliM/FliN (which can be purified from *salmonella* [125] but not *E. coli*), and both worked on scaling up DNA modification protocols, developing purification strategies, and characterizing the covalent conjugates. At the microscope, I counted the designed stoichiometry on fluorescent-DNA control templates, validating template assembly and the dye-counting method.

In the final phase of work, some problems with assembly of gappy DNA templates were identified and resolved, and I started provisional work on counting structures with >5 dyes. Unfortunately, the guided covalent conjugation permanently stopped working at this point, for reasons unknown. However, Joel discovered (contrary to previous understanding) that NTA-histag binding could be visualised by native PAGE. This helped us establish that the 5-NTA DNA modifications improved binding strength over the previous 3-NTA modifications. This was confirmed with single-molecule microscopy, but reliable stoichiometric assembly on templates was still not observed. With native PAGE, however, stoichiometric binding of FliG to DNA templates was achievable, but only on templates with non-proximal FliGs. This and a number of other factors suggested that covalent FliG-DNA conjugates would still be preferable; Joel successfully resurrected the maleimide-DNA covalent conjugation strategy, and at the time of writing has just managed to demonstrate controlled stoichiometric assembly on both proximal and non-proximal templates, with native PAGE. While some work is needed to make these conjugates viable

for fluorescence experiments, we are finally at the point where we can perform the experiments outlined in Figure 1.19.

2

Building Blocks

2.1 DNA templates

2.1.1 10nt Linear Templates

If a DNA template is to mimic the role of the MS ring, it ought to have a comparable spacing of attachment points. EM structures[38] imply a $\sim 3.5\text{nm}$ inter-subunit spacing along the maximal circumference of the MS ring, and a $\sim 3.8\text{nm}$ spacing along the outer C ring; $\sim 3.2\text{nm}$ at the edge of the inner lobe of density. Fortunately, the DNA double helix rotates $\sim 34^\circ$ per basepair (bp) with a rise of $\sim 0.34\text{nm}$ [306], such that periodic binding sites with a 10bp spacing ($\sim 3.4\text{nm}$) should be projected approximately in the same plane; ideal for FliG templating (Figure 2.1).

However, such a short sequence would usually have a low melting temperature. A 10 nucleotide (10nt) sequence “N10” with very high GC content (i.e. high stability) was designed prior to my involvement in the project. Fixing this as the FliG-linking strand, I used NUPACK[307] to design a series of short linear template sequences to test for FliG gap-filling (Figure 2.2)(see section 1.5.1 for gap-filling). For templates incorporating gaps, a spacer sequence “dN10” (dummy N10) is designed to keep the template rigid, maintaining the desired spacing. A tag sequence on the end of the template is designed to bind a biotinylated anchor strand “BA” (for attachment to a surface), or to an attachment point on a DNA

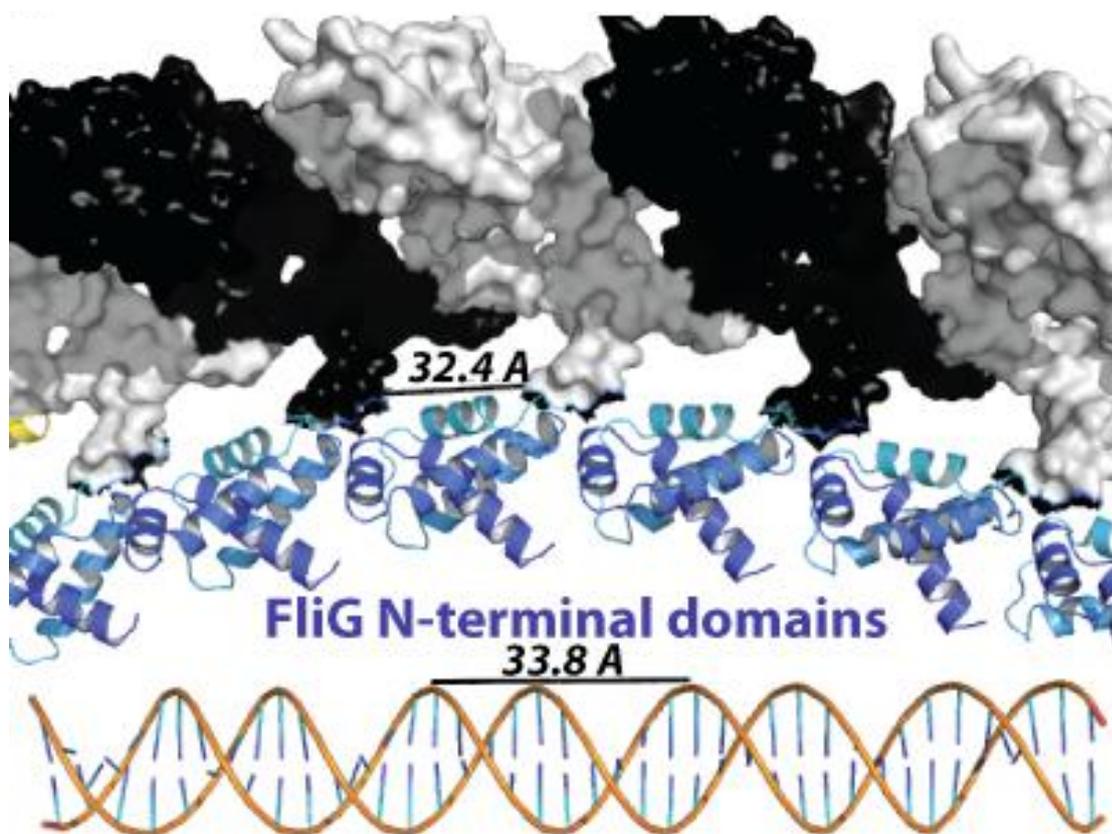


Figure 2.1: Figure from Lee lab, comparing 10nt repeat of DNA with inter-FliG spacing in their molecular model of the C ring[57].

tile (Figure 2.2 and section 2.1.3). Poly-T linkers were used to ensure flexibility between templating and surface-attachment regions.

The sequences went through a number of redesigns, listed in appendix A.2. Oligos were all synthesized by IDT, with HPLC purification for sequences over ~ 30 nt.

2.1.1.1 Original sequences

Room temperature assembly of NTA-modified N10 (section 2.3.2.1) on a gapless template (T5N10) was confirmed with native poly-acrylamide gel electrophoresis (PAGE); titration of NTA-N10 against T5N10 produces stepwise mobility shifts corresponding to increasing N10 stoichiometry (Figure 2.3a). 12.5mM MgCl_2 was required for full assembly, in addition to the 400mM NaCl present in our experimental buffer (refer to section A.1). Blurring of bands indicates slight

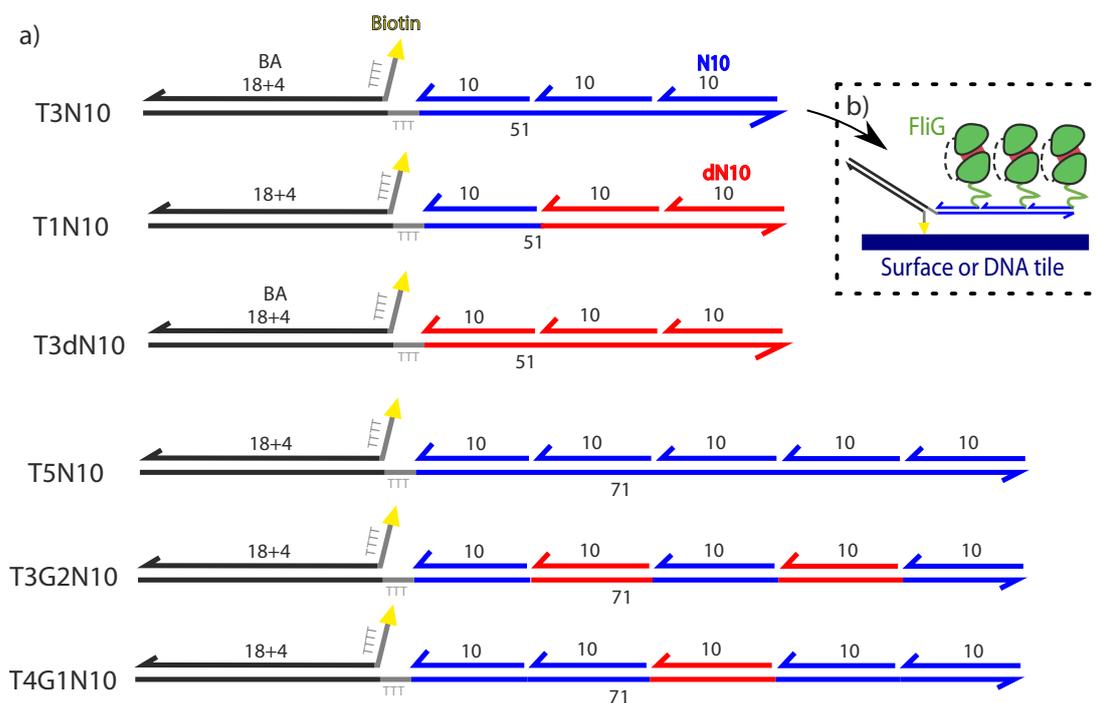


Figure 2.2: a) 10nt linear templates. Numbers indicate oligo length in nucleotides. Colour indicates sequence complementarity. Oligo sequence names are indicated. 3' end indicated by arrow. b) Example of use.

N10 dissociation during the running of the gel, probably due to the lack of salt during electrophoresis. Repeat attempts to clarify this by running gels containing 10mM $MgCl_2$ were unsuccessful.

The correct assembly of N10 conjugated to Alexa Fluor 647 (AF647) via a TTT linker onto T5N10 was also confirmed by single-molecule photo-bleaching. The majority of spots showed 5 bleach steps (Figure 2.3b) (see section 4.1.2 for more detail). Spots were imaged in constant wash conditions, and a small degree of disassembly was observed over the 120m of the experiment (Figure 2.3c). However, the imaging buffer contained only 150mM NaCl and no $MgCl$; with increased salt, greater stability would be expected. A similar experiment counting FliG bound to 10nt NTA-DNA linkers on T5N10 likewise showed minimal dissociation over 200m, suggesting that linkers were stably (if not necessarily stoichiometrically) bound (section 4.1.1).

Despite successful assembly of NTA-modified N10 onto templates, gels showed no sign of dN10 (the gap-bracing oligo) binding (figure 2.4), consistent over a wide

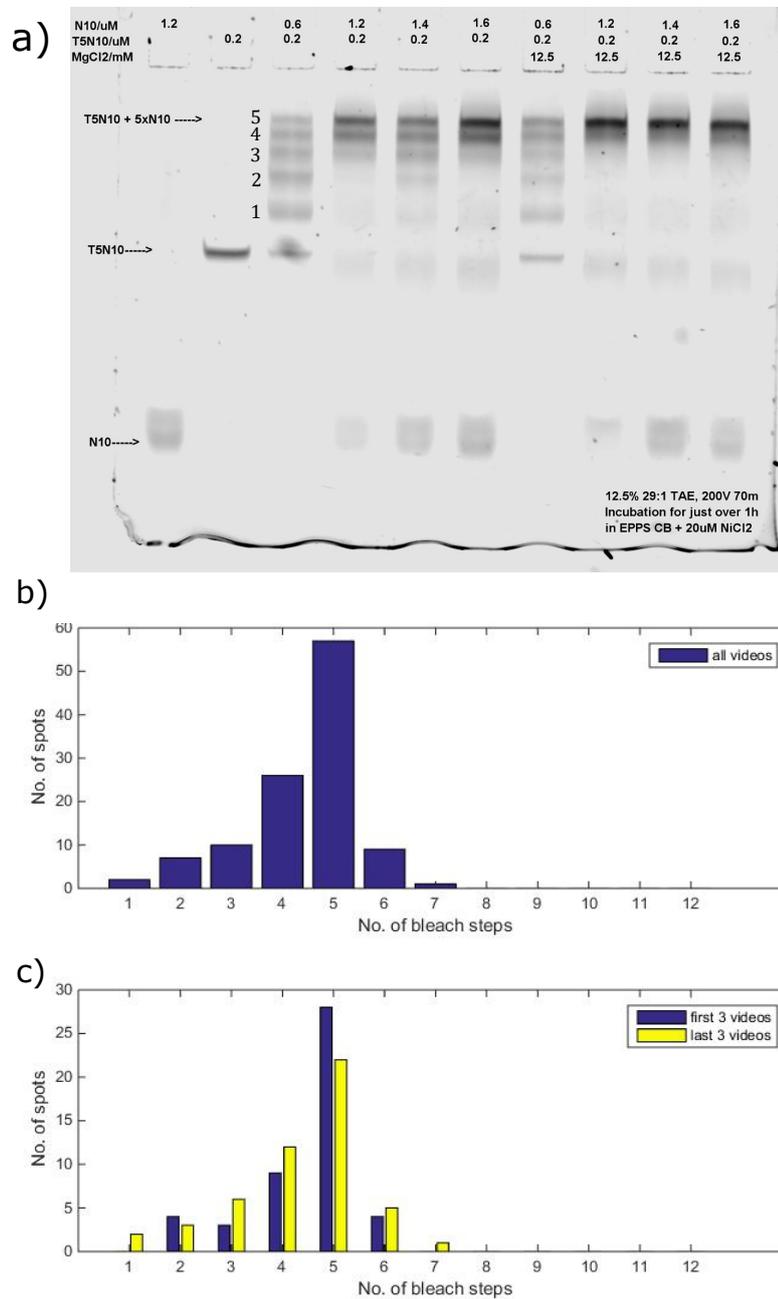


Figure 2.3: a) Assembly of NTA-modified N10 assembled on a 5-site template (T5N10). b,c) Fluorescence counting of fluorescent N10 on T5N10. Data comprised 12 videos spread over \sim 120m.

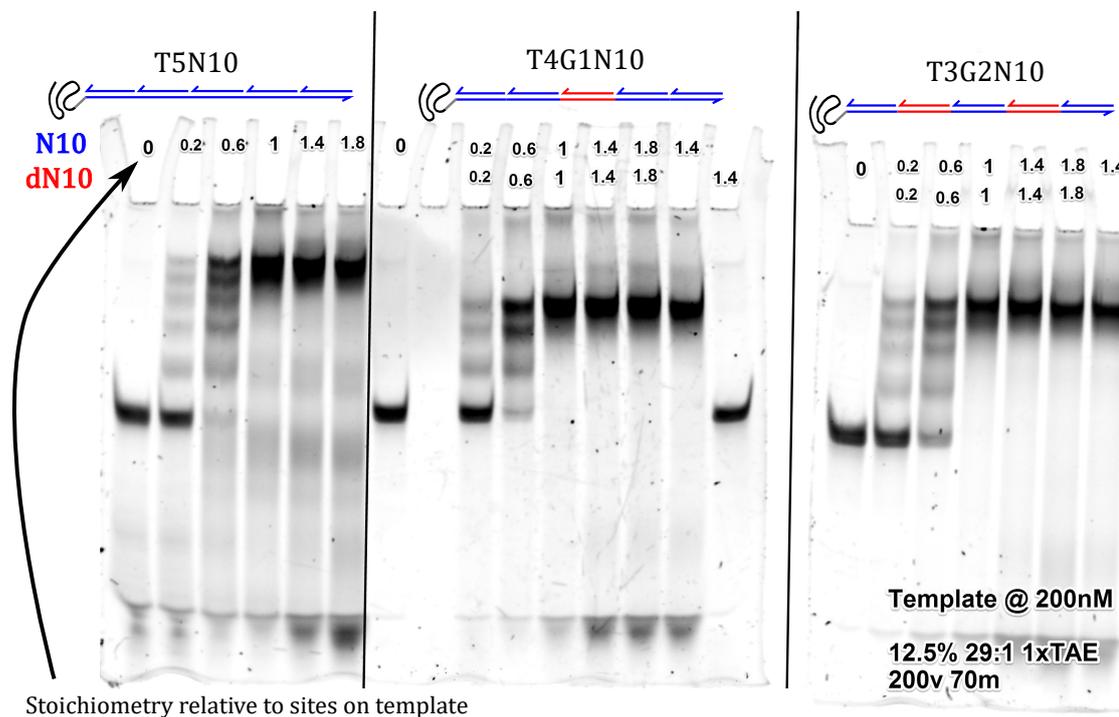


Figure 2.4: Titration of N10 and dN10 against various templates shows binding of N10 but not dN10.

range of conditions. Re-evaluating the sequences, the GC content (i.e. binding stability) of dN10 was much lower than N10; dN10 sequences were redesigned for increased stability.

2.1.1.2 Redesign 1

The first redesign (annotated “v3”) showed perplexing behaviour. A gel shift indicating dN10 binding to the template was seen, but only in the presence of low salt (Figure 2.5a); counter-intuitive, given that salt should stabilize binding. We noted that this dN10 sequence contained 3 Gs in a row, unlike the previous sequence, or the N10 sequence which had 2 at most. Guanines from 4 separate oligos may stack into a square planar structure, and stretches of repeated guanines may form a stable ladder of of such structures; a G quadruplex (Figure 2.5b). A variety of salts including both $MgCl^2$ and $NaCl$ may stabilize G quadruplexes, and we hypothesized that these could be competing with dN10 binding to the template, although this would be somewhat surprising for such a short stretch of Gs[308, 309].

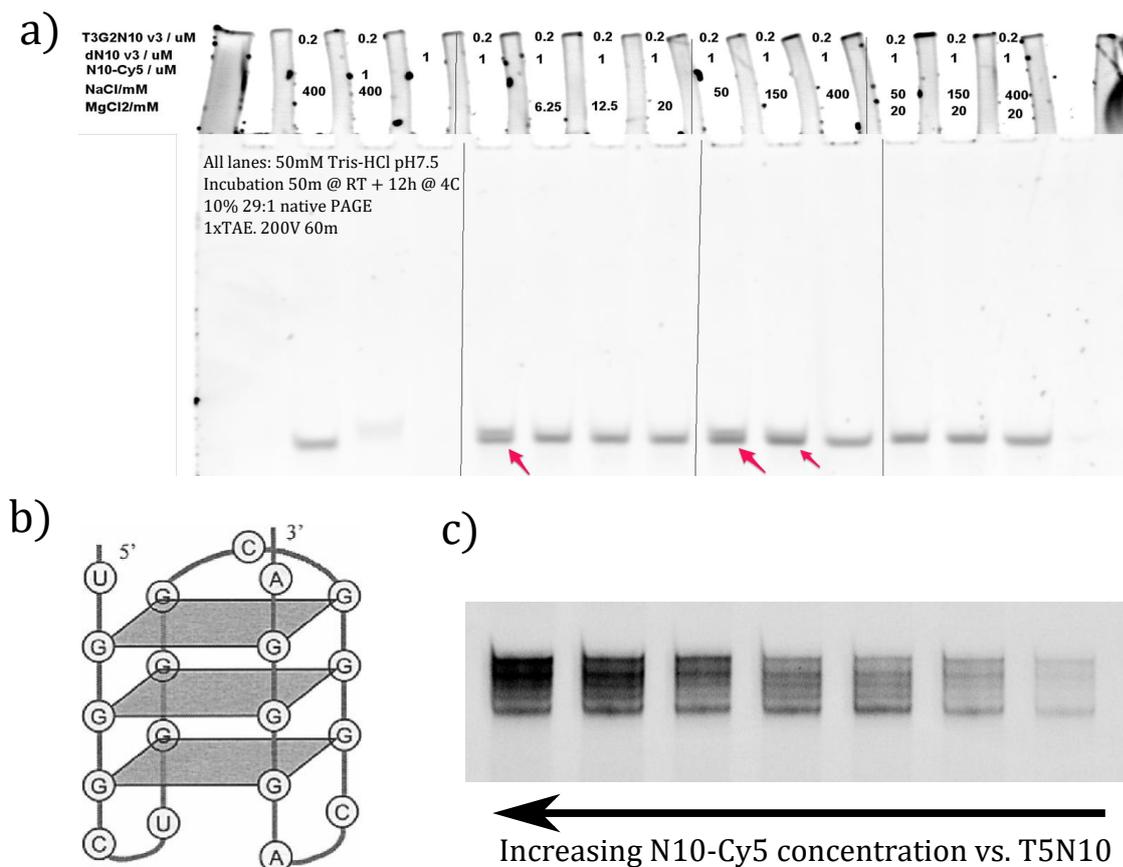


Figure 2.5: Analysis of first redesign of template sequences. a) dN10 binding in different salt conditions. Arrows indicate binding at low salt. b) Schematic of a G quadruplex, from [310] c) Titration of Cy5-N10 against T5N10; Cy5 signal only. 15% 29:1 native PAGE, 1xTAE, 4 $^{\circ}$ V, 80m @200V. After longer run-times, bands diffused into each other.

I also realised at this point that the NTA-modified oligos caused unexpectedly large gel shifts as compared to unmodified or dye-modified oligos. The latter made such small shifts that the bands corresponding to a titration of N10 vs template were only barely resolvable even with very high percentage gels and long run-times(Figure 2.5c). This explains the small gel shift seen in Figure 2.5a, and suggests that gel electrophoresis might not be the optimal way to measure binding of unmodified oligos.

2.1.1.3 Redesign 2

The next dN10 sequence tested (annotated “v2”) avoided having more than 2 guanines in a row, while retaining high GC content overall. Given the poor resolution of gels, I instead tried a SYBR green assay to measure melting temperatures. SYBR

green is a dye which fluoresces when bound to DNA, and binds more effectively to double-stranded DNA than to single-stranded. Fluorescence should therefore decrease as a duplex melts. The inflection point of this melting curve can be taken as the melting temperature (Figure 2.6a).

Melting curves could also have been measured by UV absorption, but this would have required higher quantities or concentrations of DNA. As assembly experiments were usually performed at $\sim \mu\text{M}$ concentrations and microscopy experiments are limited to very low concentration ($< \text{nM}$; see section 3.2), behaviour at low concentration is more relevant; some oligos (e.g. NTA-modified NTA) were also in short supply. Furthermore, a real-time PCR machine allowed dozens of melting curves to be measured in parallel with fluorescence assays.

Unfortunately, SYBR green is expected to artificially increase duplex stability somewhat. To quantify this (and to determine the minimum practical DNA and SYBR green concentrations), I compared stoichiometric binding of tris-aminated N10 to a single-site template (T1N10) at a range of SYBR green and DNA concentrations (Figure 2.6b). Extrapolating to 0 SYBR gold, it seems that 1x SYBR gold assays overestimates melting temperature by $\sim 5^\circ\text{C}$.

That established, I looked at melting temperatures for different components of the template in typical experimental conditions (Figure 2.6c). This showed that the new dN10 sequence was almost as stable as the N10 sequence. It also demonstrated cooperativity in assembly of N10/dN10 onto the template; not unexpected, as neighbouring strands will favourably stack base pairs. However, it also showed that NTA modification of N10 was significantly destabilizing, decreasing the melting temperature by $\sim 10^\circ\text{C}$ to $\sim 52^\circ\text{C}$. While this is still easily sufficient for template assembly at $\sim 1\mu\text{M}$, samples must be diluted significantly for fluorescence microscopy. While FliGs bound to 10nt NTA-linkers appeared stable in such conditions (section 4.1.1) the linkers may have been on under-filled templates, and thus perhaps not so severely destabilized. Thus, while we know that dye-labelled N10 binds sufficiently strongly to observe for hours at effectively zero background concentration (section 4.1.2), the case is not so clear for NTA-modified N10.

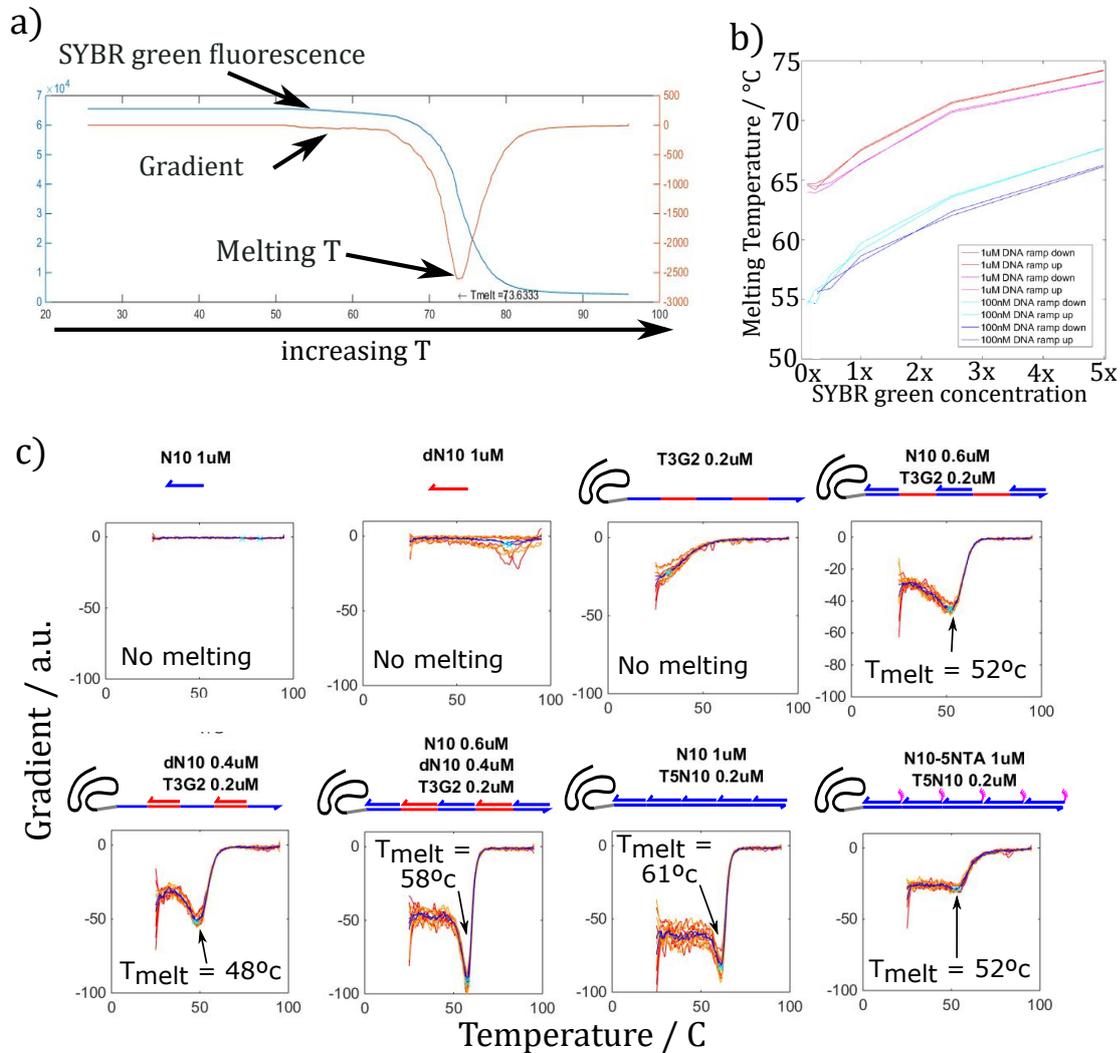


Figure 2.6: a) Example data from SYBR green assay, showing the melting curve. This example has high DNA concentration (1uM) and SYBR green concentration (5x), and thus has minimal noise. b) Tris-N10 binding to T1N10; variation of melting temperature with SYBR gold concentration. (Excluding cases where melting temperature was not clear). c) Melting curve gradients and melting temperatures of different mixtures of template components (redesign 2); gradients averaged over multiple repeats and ramps. Template schematics follow figure 2.2. Performed in EPPS CB buffer (400mM NaCl, no MgCl_2), without Tween.

2.1.1.4 Summary

The 10nt linear templates assemble as expected, and are stable under typical conditions for templated assembly. Indeed, at the very end of this thesis (section 4.2), we see successful 10nt template assembly with covalently-attached FliG, as observed with native PAGE. However, it remains possible that at the much lower concentrations required for fluorescence microscopy, NTA-modified or covalently FliG-modified oligos linkers lack stability. It also seems possible that the 10nt binding sequence is too weak to overcome combined destabilization by NTA-modification and FliG crowding, which might explain the difficulty of binding FliG stoichiometrically to 5 x 10nt templates via NTA (section 4.1.5).

2.1.2 20nt Linear Templates

2.1.2.1 Utility

Linear templates with 20nt linkers rather than 10nt linkers are useful for two reasons. Firstly, a 20bp duplex ought to be very stable even at minimal concentration, so it provides a platform to test NTA-protein binding in a fluorescence experiment without concern about the template falling apart. Furthermore, our implementation of the templated covalent conjugation strategy (section 2.3.3.1) requires a 20nt NTA-modified “guide” oligo, and templating can help us test the performance of that particular oligo. So; we made a construct “T3Guide” or “T3G”, templating 3 repeats of this guide oligo (Figure 2.7a).

2.1.2.2 Assembly and suspected dimerization

As with 10nt templates, correct assembly was verified by titration of guide strand (Figure 2.7b). Note faint slow-running species, corresponding to dimers or higher order aggregates. These occur only in the presence of guide strand, but are insensitive to the presence of nickel (Figure 2.7c), ruling out a nickel-mediated NTA-NTA interaction. They also look slightly different between two versions of the template, “T3G (old)” and “T3G (old)” which differ only in their anchoring domain (coloured black in schematics). This could indicate some interaction between guide strand

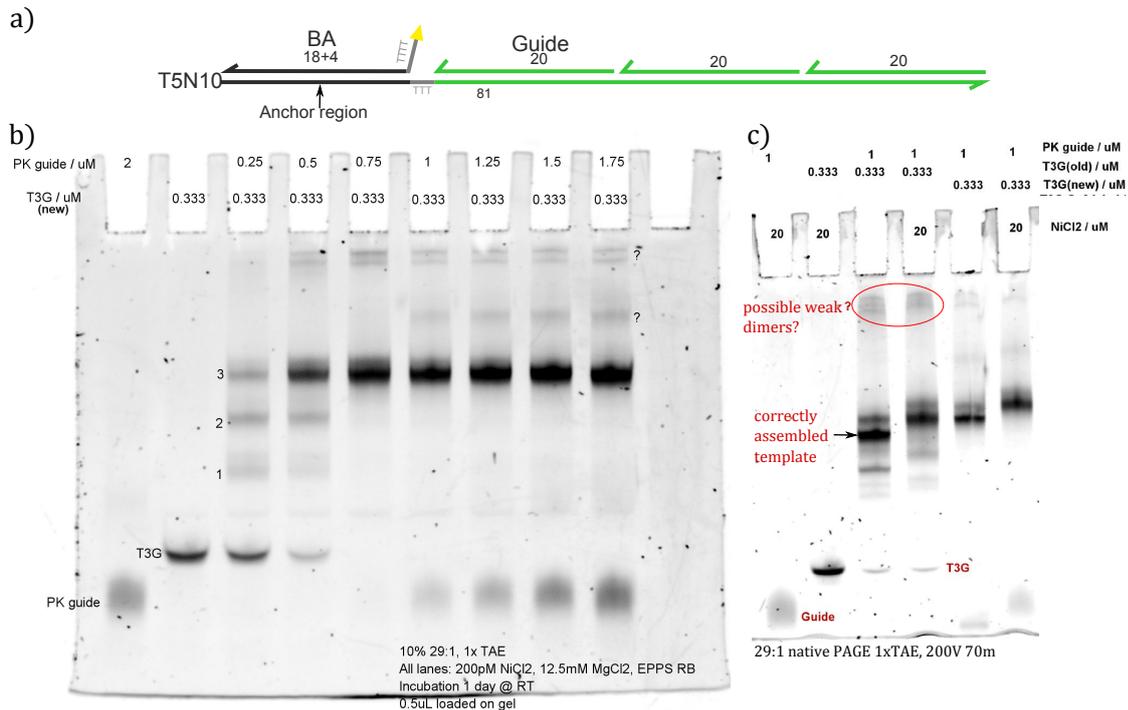


Figure 2.7: a) Schematic of 3x20nt template. Numbers indicate oligo length in nucleotides. Colour indicates sequence complementarity. Sequence names are indicated. b) Assembly of NTA-modified guide strands on 3x20nt template. c) Weak dimerization of templates.

and template-anchoring sequences or alternatively just inhomogeneity in the DNA synthesis. This finding was never really explained, but template dimers might explain the sub-population of assemblies with >3 FliGs on 3 x 20nt templates seen in both fluorescence and PAGE experiments (sections 4.1.3 and 4.1.4 respectively).

2.1.3 DNA Tiles

DNA origami tiles designed by Jon Bath to template rings (Figure 2.8) will make a minor appearance in section 3.5.1.2; I will describe them only briefly. (Refer to section 1.3.1 for an introduction to DNA origami.) Templating a continuous ring on a tile is difficult. Therefore, this design provides 6 hexagonally-arranged attachment points (staple extensions) to which linear templates can be anchored (Figure 2.8a,b), in the hope that FliG-FliG interactions would join templated polymers together, forming a ring. 4 alternate sets of extended staples allow a range of ring sizes, from 36nm to 55nm.

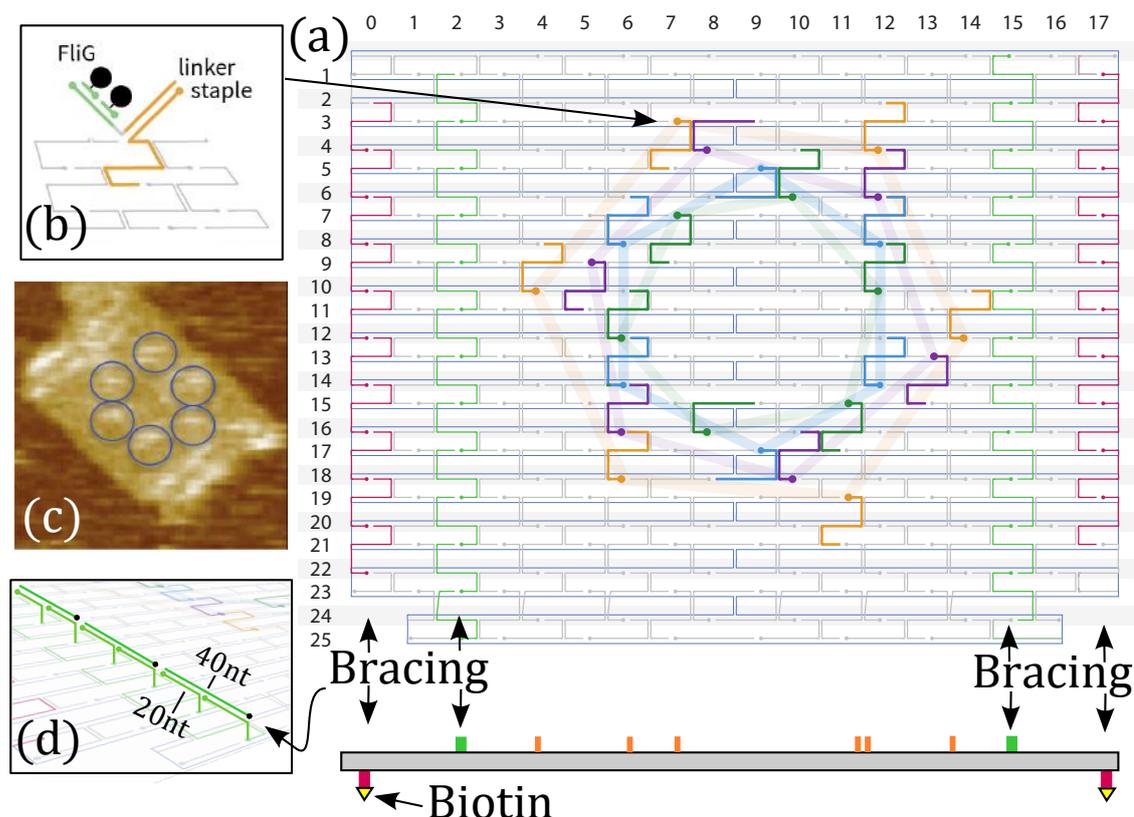


Figure 2.8: a) Schematic of DNA tile, showing template and staple routing. Circles represent 5' ends. Different colours of staple represent alternate choices of staple extensions for alternately sized rings. (Figure: Jon Bath) b) Staple extensions can anchor linear templates. (Figure: Jon Bath) c) AFM image of a tile showing DNA linkers with linear templates (no FliG) and bracing rows, produced by Le Liang. d) Brace construction.

The tile itself uses a pattern of staple crossovers designed to minimize twist[311]. However, SAXS measurements performed by Lee lab imply that such a tile is still extremely twisted in solution. They found the tile could be flattened by rows of bracing staples, extending 20nt above or below the tile, bound together pairwise by 40nt bracing strands (Figure 2.8d). To avoid inducing curvature, braces on the top of the tile (green in figure 2.8) are balanced by braces on the bottom (purple), which additionally have biotin modifications for attachment to a surface.

Assembly of tiles for microscopy experiments (section 3.5.1.2) was as follows: 10nM of M13 template was mixed with a 10x excess of staples and brace strands in a 1xTAE 12.5mM MgCl₂ buffer and annealed from 95°C to 20°C in steps of -0.1°C per 6s. This was followed by PEG precipitation to remove free staples[312], and a UV absorption measurement to quantify concentration. To add linear templates,

3.2nM origami was incubated with 32nM T5N10 and 238nM AF647-TTT-N10 at 4C overnight.

A continuous ring template would clearly be preferable, perhaps based on a DNA nanotube. However, I devoted minimal time to developing one, given that more fundamental ingredients of the system (e.g. DNA-protein conjugation) still needed work.

2.2 Protein

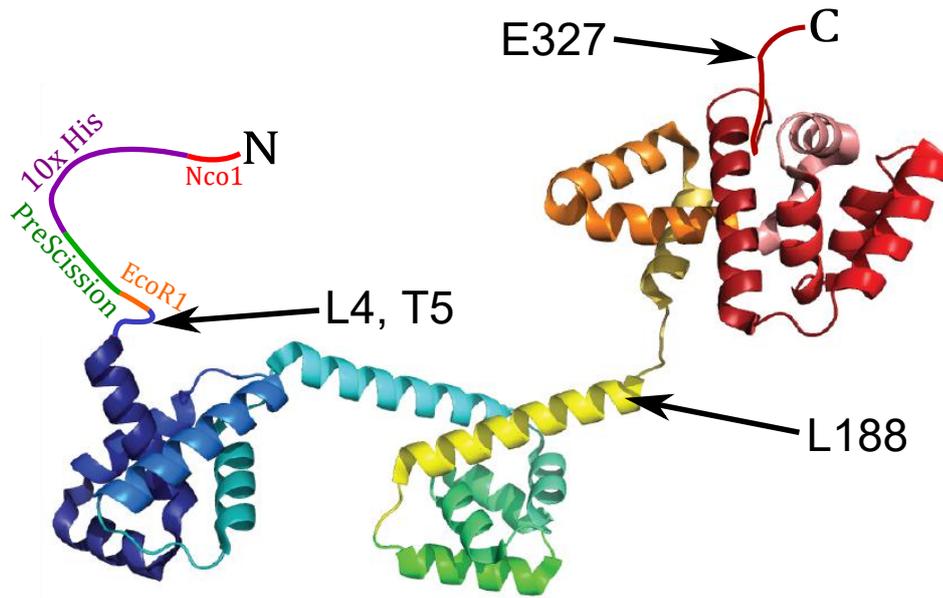
2.2.1 Constructs

A number of full-length FliG constructs were used, all developed by Lee lab and sharing a common architecture (Figure 2.9) with a 10x histag for purification and DNA-conjugation on the N-terminus (near the native FliF-binding region), and a recognition site for the PreScission protease, allowing the histag to be cleaved if desired.

Conveniently, FliG has no native cysteines, allowing site-specific dye labelling of genetically introduced cysteines. We initially expressed *E.coli* FliG L188C for dye labelling, and FliG wild-type (WT) as a control for non-specific labelling. L188 lies in Helix_{mc}, and was originally cysteine-modified to serve as one half of a FRET pair for studies in Sydney. We later switched to the C-terminal mutant E327C when it became available, because the cysteine ought to be more accessible (being on an unstructured loop at the end of the protein) and is better-removed from the site of any expected FliG-FliG or FliG-FliM interaction surfaces. We later also expressed N terminal mutants L4C and L5C, for labelling by DNA-maleimide (section 2.3.3.2).

Later stages of the project would include FliM/N, which is stable when purified from *S.typhimurium* but not *E. coli*¹. Therefore, to ensure compatibility, we also expressed *S.typhimurium* FliG constructs with C and N terminal cysteines. However, given reports that *T. maritima* FliN is functional in *E.coli motors*[114], we would be unsurprised if *S.typhimurium* works well with both.

¹Personal communication, Tohru Minamino



Nco I — His-TAG — PreScission — EcoR I — FliG

Figure 2.9: FliG constructs, mapped onto crystal structure from *A. aeolicus* [57]

All constructs were supplied by the Lee lab on pACYC Duet-1 plasmids encoding chloramphenicol resistance, which we amplified in NEB 5-alpha Competent *E. coli* (Full protocol in section A.3).

2.2.2 Expression, Purification and dye labeling

Detailed protocols (closely following those provided by Lee lab) are listed in section A.4, but in brief:

2.2.2.1 Expression

Plasmids were transfected into NEB T7 Express competent *E. coli* and grown on chloramphenicol LB agar plates to select for transformed cells. Single colonies were then grown in ~1L volumes, with FliG expression induced by IPTG. Cells were harvested by centrifugation, and lysed by French Press, either with the assistance of Emma Sadler (Tucker Lab) or Robert Ishmukhametov (Berry Lab), and then flash frozen for storage in ~30mL aliquots.

2.2.2.2 Purification and Dye labelling

Thawed lysate was cleared by centrifugation, and incubated with a slurry of Nickel-NTA resin. This was washed a number of times (either using a gravity flow column or repeated centrifugation/wash cycles), and resin aliquots were compared with BSA standards on SDS PAGE gels to estimate the concentration of bound protein (Figure 2.10a).

Protein (on resin slurry) was incubated with TCEP for ~ 30 m to break disulphide bonds between cysteines, before overnight incubation with a 5x excess of maleimide-dye. The resin was then washed extensively to remove free dye, and protein was eluted with imidazole, followed by size exclusion chromatography on a HPLC system to select for monomeric protein (Figure 2.10 b,c). Protein aliquots were then flash frozen in liquid nitrogen with 10% glycerol and stored at -80°C . Towards the end of the project, when further DNA-labelling and purification steps were expected, we omitted size exclusion chromatography or replaced it with a simpler spin column purification (e.g. NAP-5 or Micro Bio-Spin) to remove unwanted imidazole.

Early purifications used Alexa Fluor 488 dye (AF488), while later purifications used Alexa Fluor 647 (AF647), after realising that the spectrum of the former overlapped with chemical fluorescence of our modified coverslips (section 3.2).

2.2.3 Quantification of dye labelling

For fluorescence experiments *in vivo*, we can often expect that 100% of proteins are labelled with an active fluorophore: genetic fusion of the fluorescent protein is easily confirmed biochemically, and folded fluorescent proteins typically exhibit homogeneous behaviour, once maturation is allowed for[313]. With organic dye labelling *in vitro*, however, we expect that some proteins will not be successfully labelled. Furthermore, as the dyes are generally small and weakly charged, it is very difficult to separate labelled proteins from unlabelled. Therefore, if the number of proteins on a DNA template is to be accurately inferred by counting the dyes, we need some way to measure the labelling efficiency, i.e. the fraction of proteins which host an active dye.

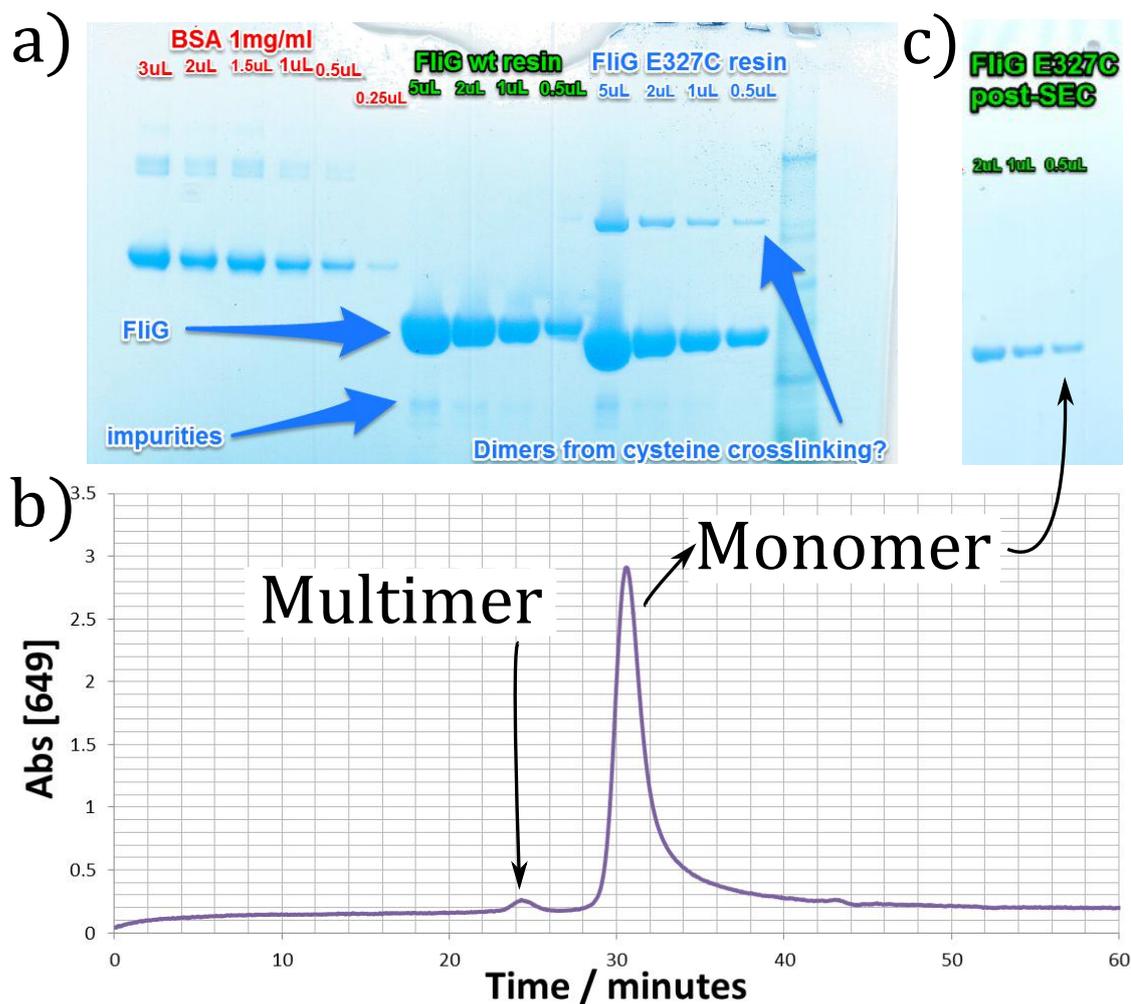


Figure 2.10: a) Quantification of FliG concentration bound to resin, pre-labelling. SDS PAGE, stained with coomassie blue. b) Size exclusion chromatography trace, after labelling with AF647. Flow rate 0.5mL/min, c) FliG post dye labelling and SEC. SDS PAGE, stained with coomassie blue.

If the target complexes have homogenous stoichiometry, then the labelling efficiency can be measured by the binomial distribution of bleach steps between complexes[314, 315]. With complexes of heterogeneous stoichiometry however, statistics of bleaching can be used to make quantitative statements about the heterogeneity, but not infer the labelling efficiency[316] unless the structure of the heterogeneity is very simple[315]. Given that we cannot guarantee construction of homogeneous complexes, we must employ other methods.

2.2.3.1 Canonical method

The canonical way to measure labelling efficiency, following the removal of free dye, is to separately measure and then compare concentrations of protein and (protein-bound) dye.

Protein concentration is typically measured by UV absorption, which is mainly due to tryptophan and tyrosine. As FliG has few of these amino acids, the absorption is not accurately measurable in the quantities and concentrations we have been using. Therefore, we instead quantified protein concentration using the Bicinchoninic Acid (BCA) assay[317] (see section A.5 for detailed protocol). This relies on the reduction of copper(II) sulphate by peptide bonds in the denatured protein, followed by the formation of Cu^+ -BCA complexes which absorb strongly at 562nm. Performed correctly, the measured absorption should be proportional to the number of peptide bonds in the protein, and absolute concentrations can be measured by comparison with known standards (Figure 2.11). Repeats usually showed variation of $\pm 10\%$, limited (I suspect) by pipetting precision.

Dye concentration is typically quantified by comparing the measured dye absorption to a standard value for the dye extinction coefficient, assuming that the later is not altered by conjugation to the protein. This assumption is questionable; while FRET interactions between dyes do not affect their absorption spectra, other kinds of interaction do[318–322]. Fluorescence emission of many commonly-used dyes (although not AF647) is quenched by interaction with certain amino acids[323–327]. However, there is little characterization of whether these interactions change the dye absorption spectra. One paper claims a small shift in the absorption spectra of fluorescein when bound by its antibody[328]. Another reports that free tryptophan changes the absorption spectra of an oxazine-derivative dye[325]. Furthermore, the absorption spectra of many dyes including Cy5 (from which AF647 is derived) are expected to change with the solvent environment[329], which may conceivably be modified at the surface of a protein.

To rule out this slim possibility that the dye absorption was changed by protein attachment, Proteinase K was used to digest the protein post-labelling. Proteinase

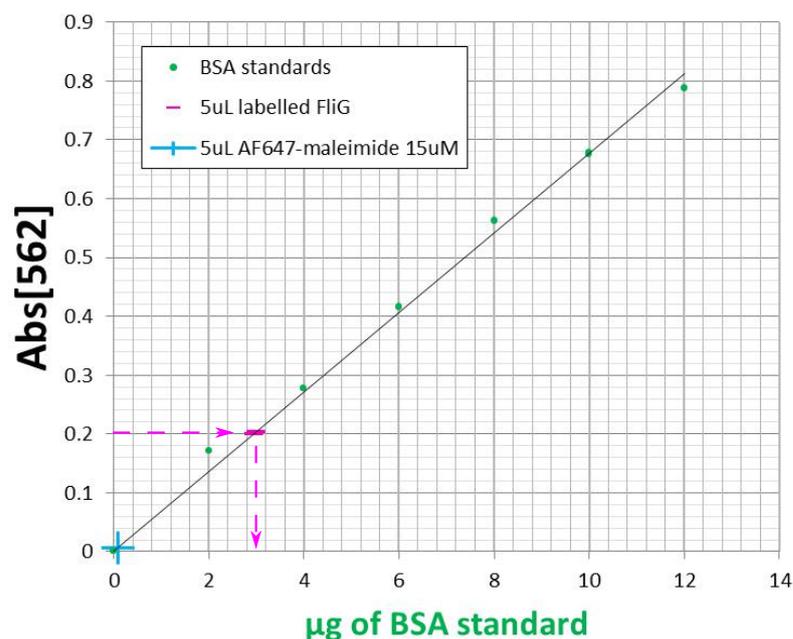


Figure 2.11: Protein concentration measurement with BCA assay. Measured absorbance of samples with known a quantity of BSA (green) are used to create a calibration curve (black). Absorbance of labelled protein sample is referenced against this to infer protein quantity (magenta). A fluorophore-only control (blue) demonstrates that the assay is not sensitive to presence of the dye.

K is notable for its robustness (surviving SDS and high temperatures), and for its generality: it cleaves a wide variety of amino acid bonds. A detailed protocol is listed in section A.6, but in brief: AF647-labelled protein and a dye-only control were incubated for ~ 2 h with Proteinase K at 50°C in the presence of SDS and CaCl_2 (both of which enhance its activity). Coomassie blue staining on an SDS PAGE gel suggested that the protein was well digested (Figure 2.12a). Comparing the absorption spectra (Figure 2.12b), the proteinase causes a slight decrease in the extinction coefficient of free dye, but a substantial ($\sim 33\%$) increase for protein-conjugated dye. This was seen consistently, and suggests that dye absorption does indeed change noticeably upon protein conjugation. (Aside: it is unclear whether this an under-reported effect common to many dye/protein combinations, or whether this combination of dye/protein gives highly unusual behaviour.)

The absorption spectrum of the dye after protein digestion can be used to quantify labelling efficiency if the extinction coefficient of free dye is known. However, while

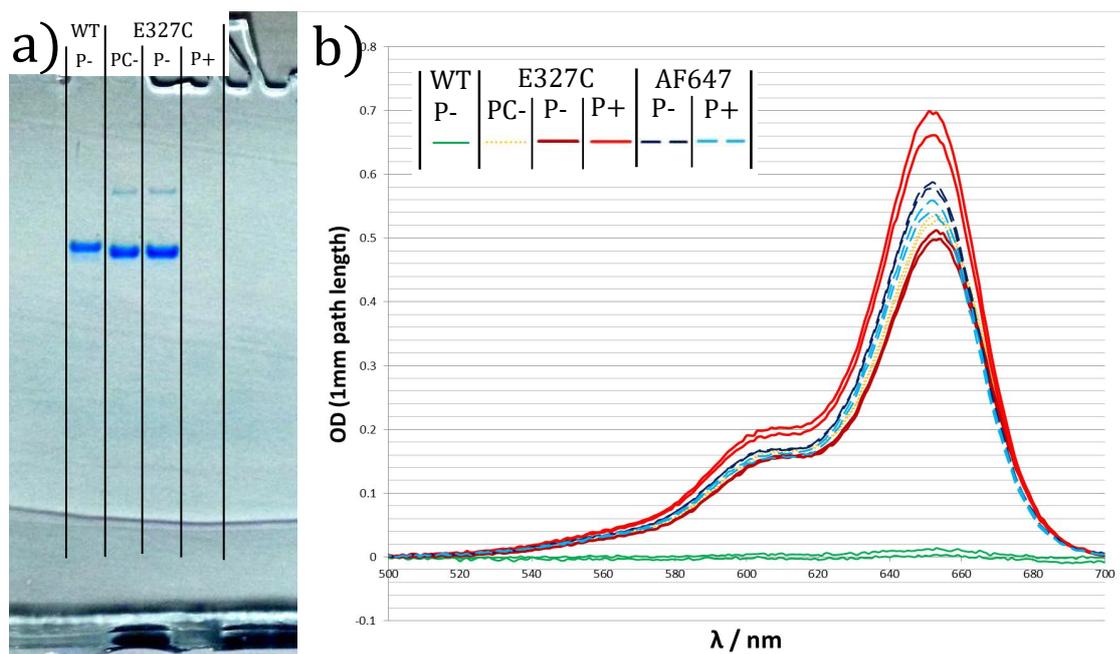


Figure 2.12: Proteinase K assay to test dye labelling. WT and E327C FliG have both been through the labelling process (WT as a control for non-specific labelling). AF647 is dye only, but in the same storage conditions as the FliG and with similar concentrations. PC- is protein post-labeling. P+/- have been incubated post-labeling in identical conditions with and without proteinase, respectively. a) SDS PAGE, stained with coomassie blue. Note that dye-labelled protein runs a little faster. This is only seen on denaturing gels, and thus not useful for purification. b) Absorption spectra of the various samples.

a single canonical value for AF647 absorption is available [330], its precision is questionable: if used to infer the quantity of dye in a commercially-provided packet, it underestimates the manufacturer-reported quantity by $\sim 15\%$. Assuming that one or the other value is correct (the canonical extinction coefficient or the manufacturer's measurement), and considering the $\pm 10\%$ uncertainty in protein concentration (see above), our measurements are consistent with between 100% and 140% labelling efficiency. Widespread double-labelling of single proteins is unlikely: the wild-type FliG control has only $\sim 2\%$ non-specific labelling (Figure 2.12b), and single FliGs adsorbed non-specifically to a coverslip very rarely show multiple bleaching steps (data not shown). Furthermore, in some labelling batches, denaturing SDS PAGE (Figure 2.12a) shows a slow protein band, present in E327C but not WT FliG, which has no AF647 fluorescence and disappears in the presence of reducing agent (data not shown). This almost certainly represents a cysteine-crosslinked

dimer; a significant non-labelled population of FliG which is not reflected in the absorption measurements. Thus, even the lower estimate of labelling efficiency (100%) is surprisingly high.

One possible explanation lies with imperfect protein digestion. According to documentation from Thermo-Fisher Scientific, Proteinase K will not cleave a peptide of fewer than 3 amino acids. It is conceivable that the short peptide still bound to the dye after proteolysis still modifies the dye extinction coefficient, maybe even to a greater extent the intact protein. As these peptides are not easily removed, this is difficult to test. Thus, while the results above show that the extinction coefficient of the dye varies with the local protein environment, the inferred labelling efficiency of $\geq 100\%$ is by no means conclusive. Therefore, alternate methods were considered. Though none have been seen to their conclusion, I will describe three; one a dead-end, one speculative and complex, and the last definitive and simple but reliant on covalent DNA-protein conjugate.

2.2.3.2 Cysteine accessibility - a dead end

If a sub-population of proteins are unlabelled, this means they have un-reacted cysteines. Discussion with Robert Ishmukhametov inspired a series of experiments around the idea that sub-stoichiometric labelling should be caused by either inaccessibility of the cystine (reducing reaction rates) or oxidization of the cysteine (making it unreactive to maleimide). By subjecting labelled protein samples to denaturing and aggressively reducing conditions, any cysteines left unlabelled by the dye should become readily reactive, and easily detectable by labelling with a different colour of maleimide-dye, or quantification of free thiols by Ellman's reagent or similar.

However, this line of investigation was abandoned mid-development with the realization that cysteines can be doubly oxidized, at which point the oxidization is irreversible by any general reducing agent[331–333]. Unfortunately, the rates of irreversible oxidization vary wildly with the local environment of the cysteine, and no simple assay for the detection of irreversibly oxidized cysteines exists,

despite significant interest in their regulatory role in some proteins. One study based on mass spectrometry reports that in various eukaryotic proteins and cell extracts, $\sim 5\%$ of exposed cysteines are irreversibly oxidized in this way[334], not limited to proteins with regulatory roles. Many papers in this area also note the difficulty of avoiding irreversible oxidization during protein purification and analysis[332]. Therefore, there is a reasonable chance that some engineered cysteines are permanently unreactive and invisible to any method at our disposal.

2.2.3.3 Surface binding kinetics - speculative and complex

The above approaches do not distinguish emitting dyes from those which are present but not emitting (e.g. in dark or bleached states). A direct measurement of the emitting fraction would be preferable. In a number of DNA PAINT studies, the transient binding and unbinding of fluorophore-labelled DNA to a surface-bound target sequence is measured through stepwise changes in the fluorescence signal (Figure 2.13a) [305, 335, 336]. Bright and dark intervals have characteristic distributions that measure the on and off rates. However, imagine that **non-fluorescent** DNA sometimes binds to the target sequence. While not directly observable, it prevents the binding of fluorescent strands, and thus affects the distribution of dark intervals in a manner which is theoretically distinguishable even if binding kinetics are not known (Figure 2.13b). FliG binding transiently via a his-tag to a single-site DNA template could be analysed in this manner, to infer the fraction of non-fluorescing protein(Figure 2.13c).

However, infrequent illumination would be required to avoid bleaching of the dyes, limiting time resolution. Furthermore, any heterogeneity in binding kinetics could be problematic, as could the need for long observation times (challenging in terms of microscope drift). Experiments or simulations to quantify the expected measurement precision were not performed, in the hope that covalent protein-DNA conjugates will soon enable the next (simpler) method to be used. However, this method may still be of use in cases where covalent conjugation is impractical, and potentially requires only a his-tag, making it applicable to a wide range of proteins.

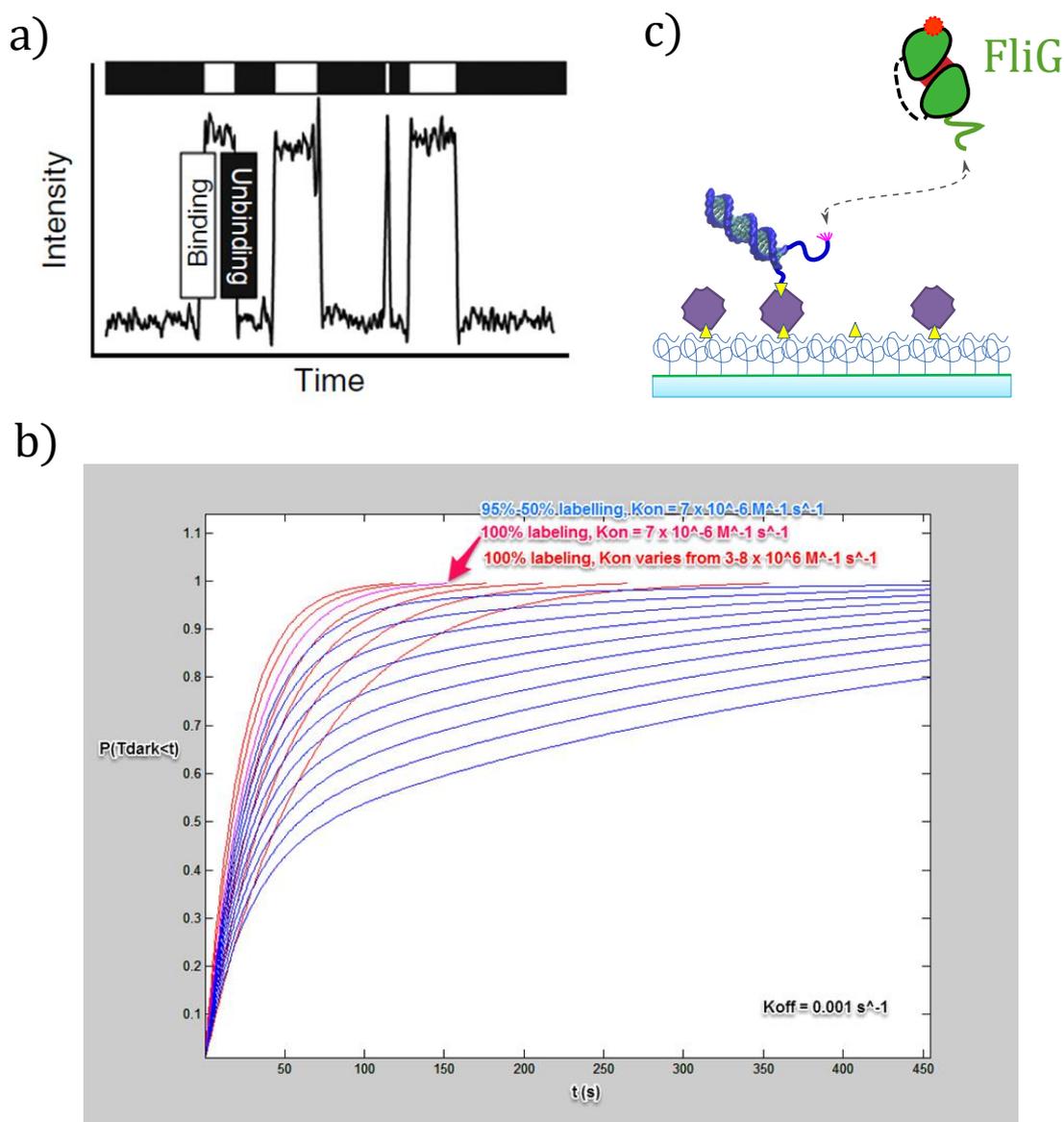


Figure 2.13: a) Example trace of fluorescently-labelled DNA binding to and unbinding from a surface-tethered DNA oligo[305]. b) Simulated distribution of dark intervals, for fluorophore-labeled molecule binding transiently to a single target site. Curves show distributions corresponding to 100% labelling yield and varying binding kinetics (red) or fixed binding kinetics and varying labelling yield (blue). The point is that they affect the curve in different ways, and fitting ought not to confuse the two. c) Schematic of FliG binding transiently to a surface-bound single-site template.

2.2.3.4 Tag co-localization - hopefully definitive

If all proteins labelled with dye A carry a permanent (i.e. covalent) DNA handle of sufficient length, a DNA tag with dye B may be stably hybridized to the label. Co-localization of these dyes can be measured either for surface-bound proteins or proteins diffusing through a focussed spot, and protein / tag stoichiometry varied. At high protein/tag ratio, we can assume all tags have a protein bound, and the percentage of tags (dye B) which are co-localized with dye A measures the percentage of labelled protein. The opposite is true at low protein/tag ratio. As a bonus, double-labelling could be inferred from bleach steps. Once we have covalent DNA-FliG conjugates, this can provide both greater precision than the canonical method and greater relevancy (measuring not just the presence of the dye, but whether it fluoresces or not.)

As an aside: with a covalent conjugate, we could also put the dye label on the DNA tag rather than the FliG. As DNA is more robust to harsh purification methods (e.g. reverse phase HPLC), we could plausibly purify dye-labelled DNA tags from unlabelled, and furthermore purify tagged FliG from untagged FliG, leaving us with a biochemically 100% labelled conjugate, if not necessary 100% fluorescing.

2.2.4 Stability and function

As we lack any kind of FliG activity assay, it is possible that its proper function is damaged during purification, storage or labelling. Covalent DNA labelling in particular could destabilize protein structure. As discussed in section 1.2.0.4 and elaborated below, it is also possible that proper function is very sensitive to buffer condition, or requires some cofactor (e.g. a salt or generic chaperone) of which we are not aware.

2.2.4.1 Protein Thermal Shift Assay

To address concerns around protein damage, we can exploit a generally observed correlation between conditions maximizing a protein's stability and conditions maximizing that protein's function (e.g. enzymatic activity)[337]. The Protein

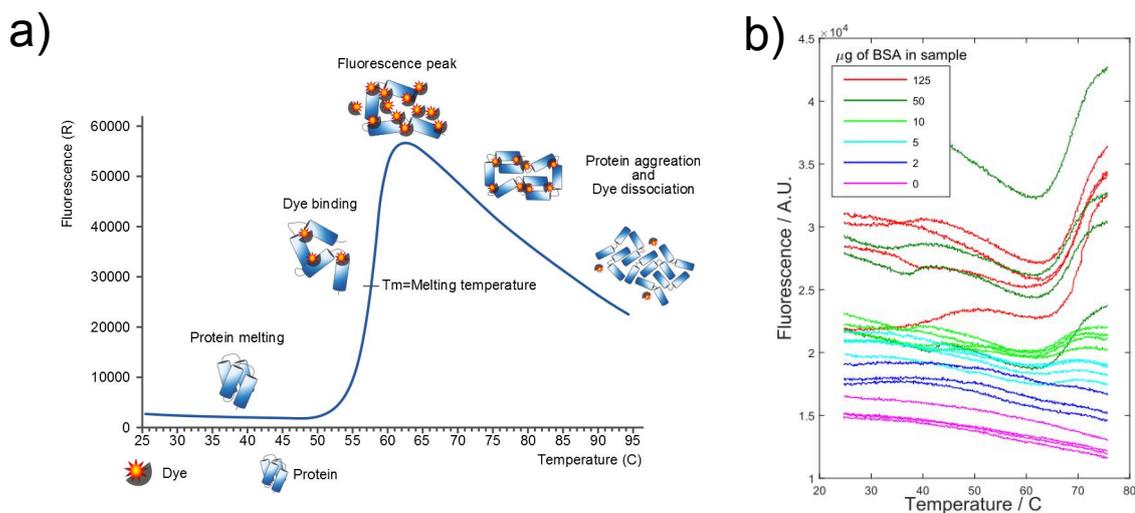


Figure 2.14: a) Explanation of protein Thermal Shift Assay, reproduced from *Wikipedia*. b) Melting curves with different quantities of BSA.

Thermal Shift assay, known also as Differential Scanning Fluorimetry, is a high-throughput method of measuring protein stability [338, 339]. This employs a dye which fluoresces more intensely when bound to hydrophobic surfaces, such as those exposed after protein denaturation. Thus, a jump in fluorescence with rising temperature identifies the protein melting point (Figure 2.14a), which correlates with protein stability. Parallel measurements using a 64-well PCR machine can compare different buffer conditions, samples, etc. This is analogous to the SYBR green assay for DNA melting, described in section 2.1.1.3.

Trial experiments used a commercial Protein Thermal ShiftTM kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA). As provided documentation is vague about the minimum quantity of protein required, I measured melting curves with varying quantities of BSA, following manufacturer protocols (Thermo Fisher publication number 4461806B). Melting was only visible with $\geq 5\mu\text{g}$ BSA (Figure 2.14b). This could be reduced to $2\mu\text{g}$ (x4 repeats) with more careful sample preparation, optimized acquisition settings, and post-processing in MATLAB to overcome noise (data not shown). This is unfortunately still a substantial amount of protein compared to the quantities involved in covalent conjugation trials thus far, but we hope that such measurements will become practical in the near future.

2.2.4.2 Recreating *in vivo* behavior *in vitro*

We should bear in mind the difficulty of accurately reproducing the kinetics and thermodynamics of *in vivo* interactions *in vitro*, particularly given the lack of direct measurements of the latter. For example, commonly used *in vitro*, salts like NaCl poorly mimic the environment of a living *E. coli*, where free anions (primarily glutamate; Glu^-) are greatly outnumbered by polyions (nucleic acid phosphates), and correspondingly by free cations (primarily K^+)[340–342]; this is difficult to reproduce *in vitro*. Furthermore, replacement of Cl^- with Glu^- *in vitro* can change interaction kinetics by at least an order of magnitude in some systems (e.g. protein-DNA interactions [341]), even when overall ionic strength is held constant. Molecular crowding may also be important for the assembly of many protein complexes[343, 344]. Thus, there is a potential for extreme sensitivity of assembly and kinetics to buffer conditions. Combined with the aforementioned possibility of unknown co-factors, this represents a non-trivial risk that assembly *in vitro* may be very difficult to achieve.

2.3 Protein-DNA conjugation

2.3.1 Overview

Although a multitude of methods exist for the conjugation of DNA to proteins [345–347], only a handful meet our stringent requirements. Attachment sites on the template must be precisely arranged, ruling out the use of bulky (~ 10 s of kDa) adapter proteins such as streptavidin [348], DNA-binding zinc fingers[295] or the binding partners to SNAP-tags[349], HALO-tags[349], FLAG tags[347], GST-tags[350], Spy-tags[351] and similar. Likewise, positioning of the linker on the protein surface must be specific enough to avoid blocking protein-protein binding interfaces, and stoichiometric (to avoid one protein blocking multiple template sites, for example). This rules out non-specific techniques, such as DNA-NHS labelling of abundant lysines [348]. As we are trying to establish a general strategy applicable to multiple protein complexes, we also want to avoid anything specific to a particular

protein (antibodies, aptamers, ligands etc) or involving very specialist protein production methods, such as the artificial amino acids required for Cu-free click chemistry [352]. The procedure should also avoid compromising protein structure and be compatible with dye-labeling for the fluorescence experiments. Finally, while we can assemble structures at arbitrary concentrations, actual fluorescence observations require dilutions to very low (sub-nanomolar) concentrations (section 3.2). So; binding needs to be uniformly stable over at least 10s of minutes, even at low concentrations.

Collectively, this is such a tight set of requirements that, despite attempting a number of methods (which I will describe below), nothing so far as perfectly satisfied all of them. I will give a brief chronological summary, before addressing each strategy in detail, and considering future approaches.

Initially, we planned to use tris-NTA modified DNA (section 2.3.2.1) to bind the histag already present on FliG. There was a long struggle to get this working, during which time I also made some failed attempts at maleimide-cysteine conjugation (section 2.3.3.2). Now the tris-NTA strategy is working reliably, but problems remain around product heterogeneity and the stability of the NTA-histag bond. Stability was improved somewhat by the development of pentakis-NTA (section 2.3.2.5), which was demonstrably stronger-binding, but still heterogeneous and potentially not stable enough for assembly on tightly packed templates (section 4.1.5). Histags also appear to compete with NTA for binding to other histags. For these reasons, Joel Spratt worked on implementing a guided conjugation technique (section 2.3.3.1), which uses a non-covalent NTA-histag bond to position a DNA oligo which subsequently forms a covalent bond with one of multiple lysines near the histag. While this was made to work, exact characterization of the labelling site was difficult, and the reliability of the protocol was poor, to the extent that it eventually stopped working completely. Joel then managed to get the maleimide-cysteine conjugation working (section 2.3.3.2), with the caveat that it is incompatible with our dye-labelling technique. While both covalent methods worked at least for a while, purification of the covalently-modified product remains a bottleneck, and

I devoted some effort to looking for scalable techniques which could be used for purification of covalent conjugates (section 2.3.3.3).

2.3.2 NTA-DNA

2.3.2.1 tris-NTA

Nitrilotriacetic acid (NTA), a chelating agent, is commonly used for the immobilization of Ni^{2+} on affinity columns. Recombinant proteins are engineered with “histags” of 6 or more successive histidine residues, whose imidazole side-chains readily bind the two free Ni^{2+} coordination sites of a Ni^{2+} -NTA complex (Figure 2.15a). This binding can be displaced by high concentrations of *free* imidazole, allowing sequential binding and then elution of the engineered protein with remarkably high purity. The wide use of histags in the manner makes them ideal targets for general conjugation. From the published literature on tris-NTA binding to histags [353, 354], we expect $\sim\text{nM}$ dissociation constants and dissociation times on the order of 1000s. Given that the dissociation constant is reported to vary by at least an order of magnitude between different proteins[353], it is plausible that conjugates might dissociate too quickly for prolonged observation at low concentration. Nevertheless, this was our starting point for FliG-DNA conjugation, following the publication of a DNA-trisNTA synthesis protocol by our lab some years ago[353].

2.3.2.2 Synthesis

Following the protocol of [353] (Figure 2.15b), a tris-aminated DNA oligo is reacted with an SPDP crosslinker, which is then reduced with TCEP, cleaving disulphide bonds to leave exposed thiols. These are then labelled with maleimide-NTA. Oligos are buffer exchanged at each step to remove reactants; either by size-exclusion spin columns (micro bio-spin P-6, Biorad) or size-exclusion gravity columns (NAP-10 and NAP-5, GE Healthcare), concentrating post-exchange with amicon filters when necessary. On occasion, DNA was ethanol-precipitated as a final purification step.

Using gel shifts to measure the number of NTA groups (following [353]), yields of 3-NTA oligos never exceeded 80%, and was often significantly less, even when

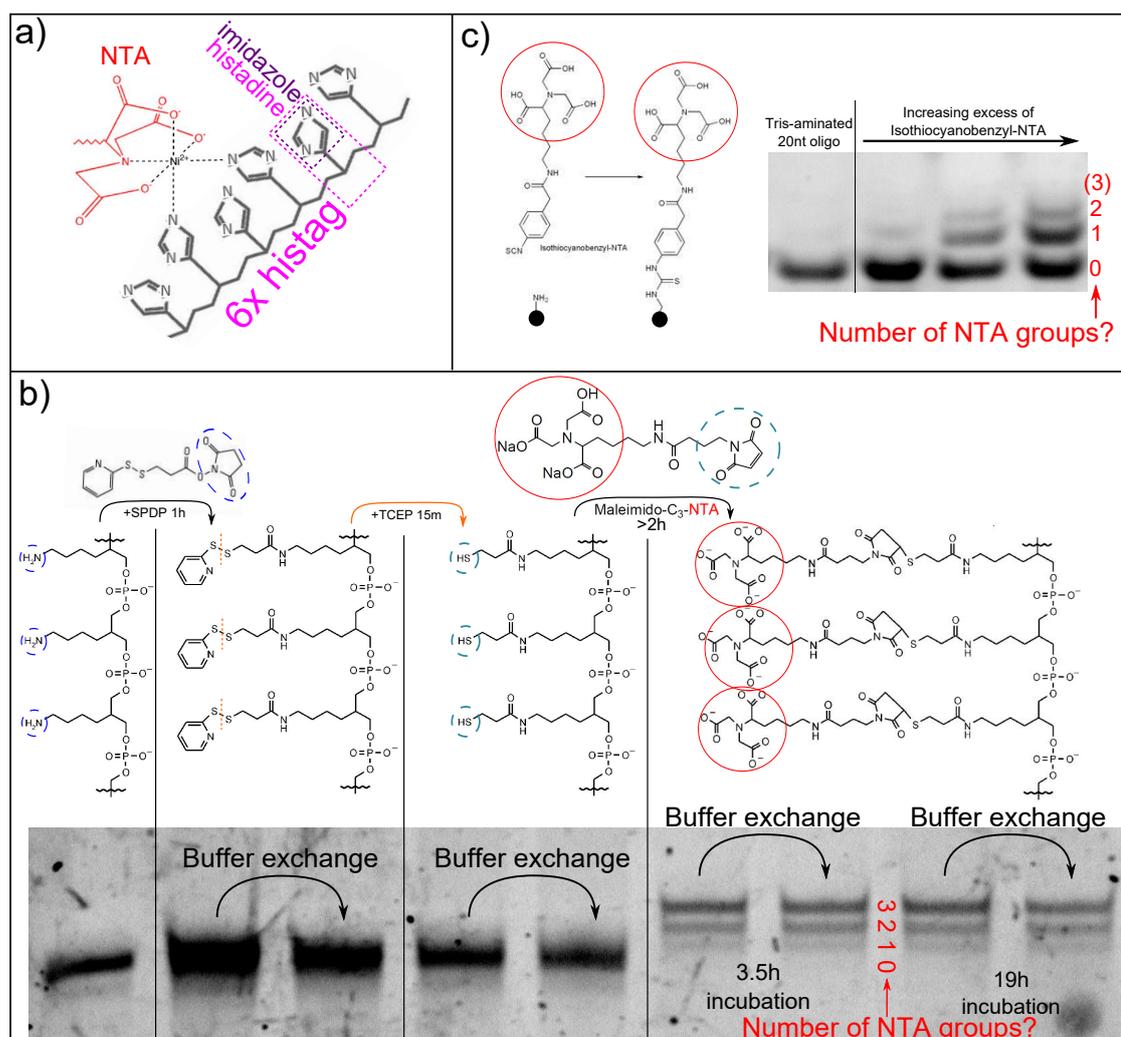


Figure 2.15: a) Binding of histag to a single Ni-NTA complex. Figure adapted from <http://www.cube-biotech.com> and <http://www.nanoprobe.com/>. b) Synthesis of 20nt tris-aminated DNA. Schematic adapted from [353]. Gel was 20% 19:1 acrylamide in TBE buffer with 7M urea, ran at 500V for ~25m. Similar shifts were observed for 10nt DNA, and the yield shown here is quite typical. c) Highest-yielding trial of one-step NTA modification with Isothiocyanobenzyl-NTA. 16h reaction time, 4%/21% stacking gel run at 180V for multiple hours.

the reaction was given time to reach a steady state (Figure 2.15b, bottom). This was across a range of reaction conditions, oligo suppliers and amine-modification chemistries. A number of attempts at modification using Isothiocyanobenzyl-NTA, which labels amines directly in a one-step reaction (Figure 2.15c), had even lower yields.

2.3.2.3 Purification

I made a number of attempts at purification with reverse-phase chromatography on a Waters Xbridge C18 HPLC column (Figure 2.16), following [353]. In this method, DNA is bound to a hydrophobic resin, and washed with increasingly non-polar solvent, which progressively weakens the hydrophobic interactions, thus eluting DNA in an order of increasing hydrophobicity. However, while a number of DNA species were clearly separated (Figure 2.16a), the pattern of peaks differed wildly between different preparations (compare figures 2.16a and c), and individual purified peaks ran as a mixture of species on PAGE (Figure 2.16b). Thus, reverse-phase chromatography is selecting for something orthogonal to PAGE; all PAGE bands contain all HPLC peaks and vice versa. Upon incubation with NiSO_4 , only a subset of reverse-phase peaks were shifted, arguably implying that only a subset of oligos are capable of chelating Ni^{2+} . It is not clear, therefore, that PAGE purification or reverse phase chromatography alone are sufficient to isolate DNA with 3 functional NTA groups. At the time of these experiments, we were not aware that protein binding could be seen by PAGE, making it laborious to directly test binding affinity for different purification fractions. We did make one attempt at binding NTA-DNA to an NTA column, in the hope that an imidazole gradient could elute NTA-DNA according to affinity for Ni, but NTA-NTA binding was not observed; perhaps unsurprisingly, given that Ni^{2+} has insufficient co-ordination sites for two NTA sites to fully bind. I suspect that a column modified with histag peptides might be a viable affinity-purification strategy. Nevertheless unpurified NTA-DNA was used in all subsequent experiments, unless noted otherwise.

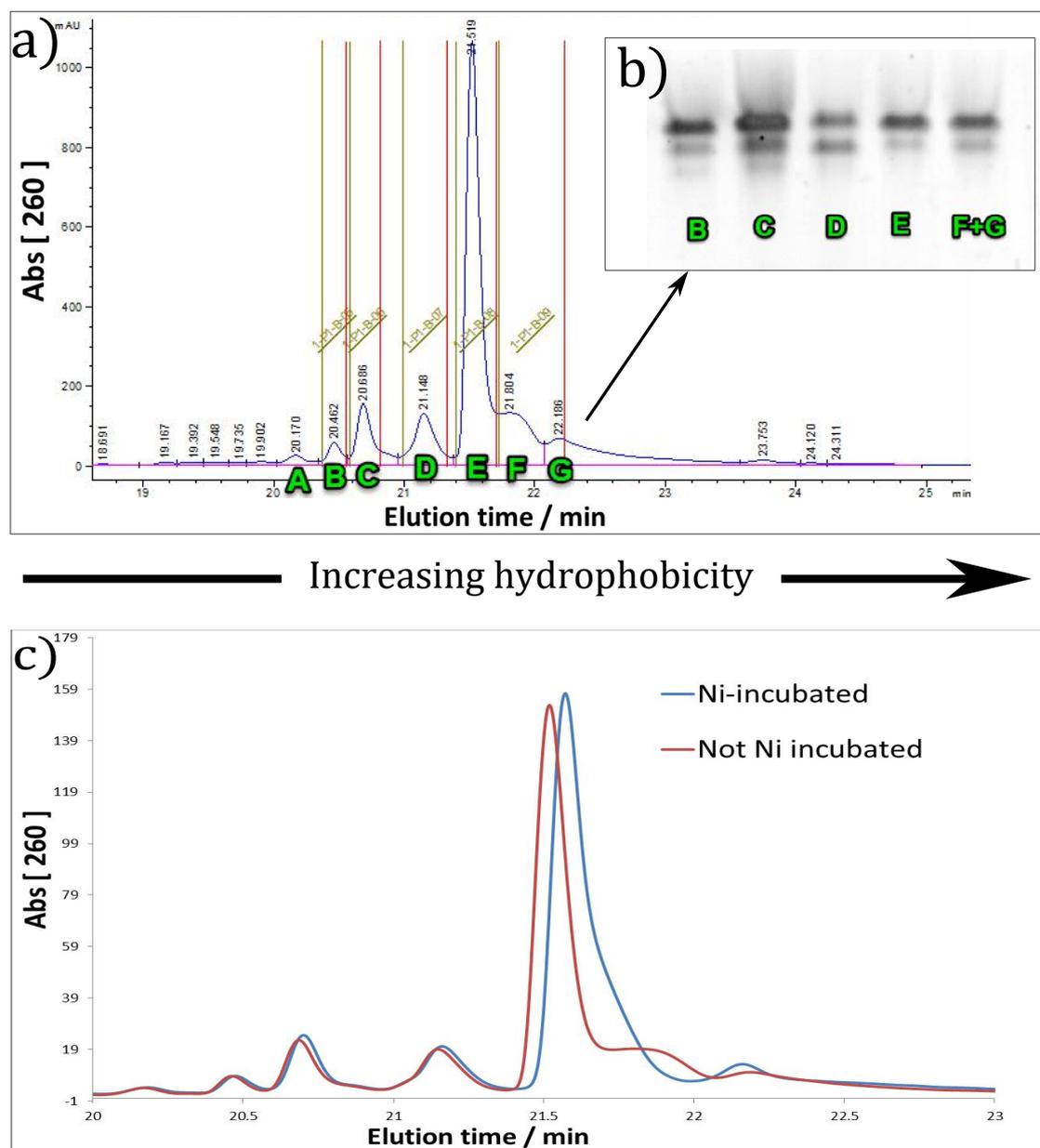


Figure 2.16: a) Reverse phase chromatography trace of a 10nt tris-NTA oligo. b) Purified fractions from figure 1, analysed by PAGE, as in figure 2.15b. c) Reverse-phase chromatography trace of another preparation of 10nt tris-NTA oligo, after incubation with or without NiSO_4 .

Aside: in the course of these experiments we tried a variety of aminating chemistries from different suppliers, with amines attached either between neighbouring phosphate backbone units (with a number of different chemistries) or to thymine bases. Though these all looked similar on PAGE, we never quantified whether their binding performance differed; it remains possible.

2.3.2.4 Observations of binding

Initial attempts to see FliG binding to NTA-DNA with native PAGE failed, consistent with [355]. The first signs of successful binding came via Size Exclusion Chromatography (Figure 2.17a). This was confirmed with TIRF microscopy (section 3): AF647-FliG pre-incubated with biotinylated 3x10nt templates and tris-NTA oligos was shown to bind to a streptavidinated surface only when the tris-NTA DNA had been pre-incubated with NiSO₄. Surface washes removed the AF647-FliG only when the wash buffer contained imidazole (Figure 2.17b), proving the specificity of the interaction.

However, in an experiment described elsewhere (section 4.1.1), single-molecule-photobleaching (section 3) was used to count the stoichiometry of FliG on a 5x10nt tris-NTA template; stoichiometries were consistent with only $\sim 1/3$ of NTA oligos having a bound FliG (Figure 2.17c). In the next section (2.3.2.5) we will see similar results on a 3x20nt template. In both cases, the most likely explanation is that (in the conditions we used), a significant population of tris-NTA oligos bound FliG poorly.

2.3.2.5 Pentakis-NTA

Theoretically, 3 NTA groups will bind only 6 of the 10 histidines in the histag of our FliG constructs. We wondered, therefore, whether a pentakis-NTA (5 NTA) modification would give stronger binding than the tris-NTA (3 NTA). It is not obvious that this should work; the only published study (to my knowledge) using >3 NTA groups showed that 4 x NTA apparently bound less stably to a 10-histag than 3 x NTA, probably due to increased entropic penalty [354]. However, they postulated this effect might be sensitive to the topology and flexibility of the chemistry holding the NTA groups together, which is quite different in our case to theirs. Therefore, we synthesized a pentakis-NTA strand analogously to the tris-NTA procedure, using a pentakis-aminated oligo. Yields were comparable to the tris-NTA synthesis, with a $>50\%$ yield of 5NTAs (Figure 2.18a).

These attempts coincided with the discovery that NTA-FliG binding was observable by native PAGE. Joel Spratt used this to measure that more imidazole

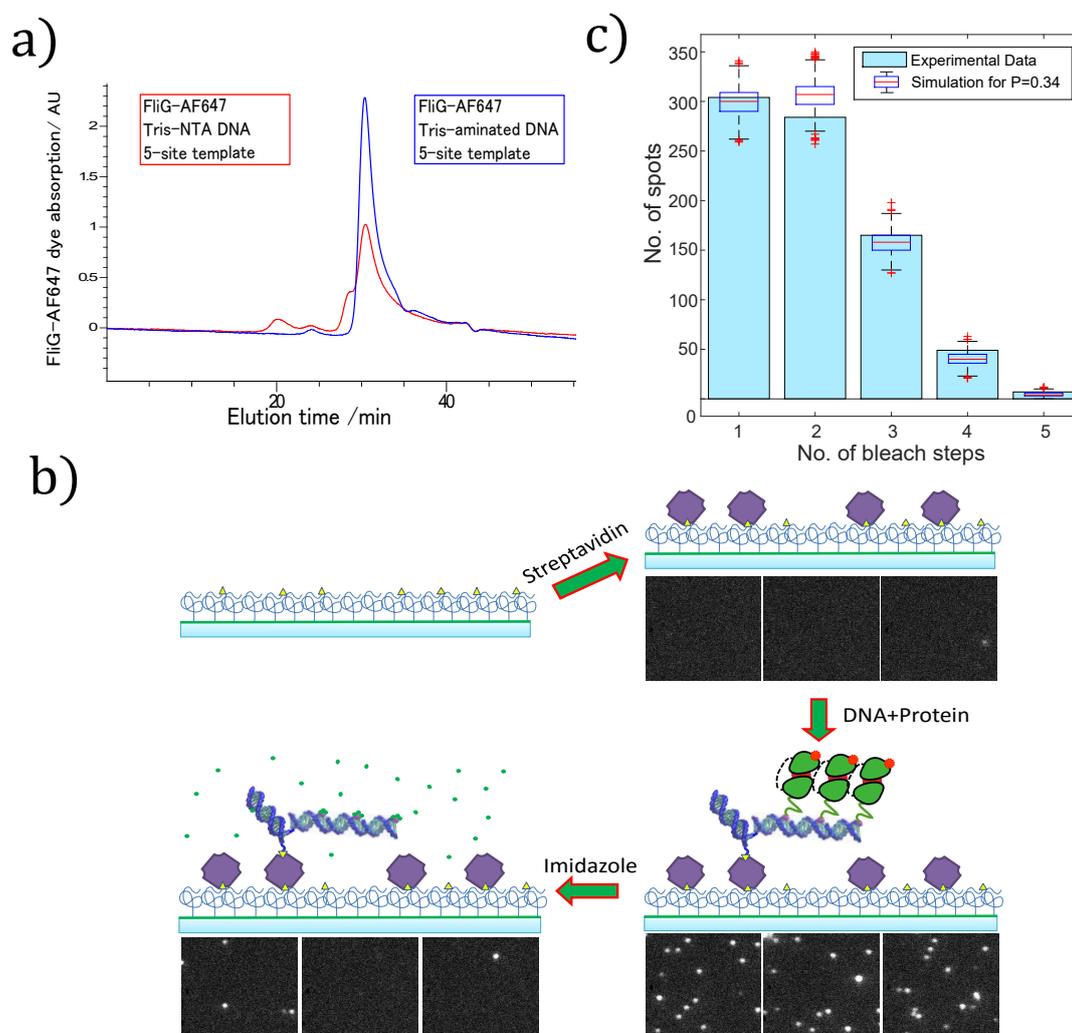


Figure 2.17: a) Size exclusion chromatography trace. tris-aminated (i.e. pre-modification) or tris-NTA (i.e. post-modification) 10nt DNA was incubated with NiSO_4 , buffer exchanged, then incubated with a 5x10nt template and a sub-stoichiometric amount of FliG. Slowed tris-NTA peaks have overlapping DNA and AF647 signals (not shown). b) Binding of AF647-FliG to surface-bound 3x10nt template via tris-NTA 10nt DNA, observed by TIRF, and disrupted by the addition of imidazole to prove specificity. c) Stoichiometry of FliG bound via trisNTA DNA to 5x10nt templates, as described fully in section 4.1.1. Yellow bars show observed stoichiometries, box plots show expected distribution if template sites are 34% likely to have a Fluorescent FliG bound.

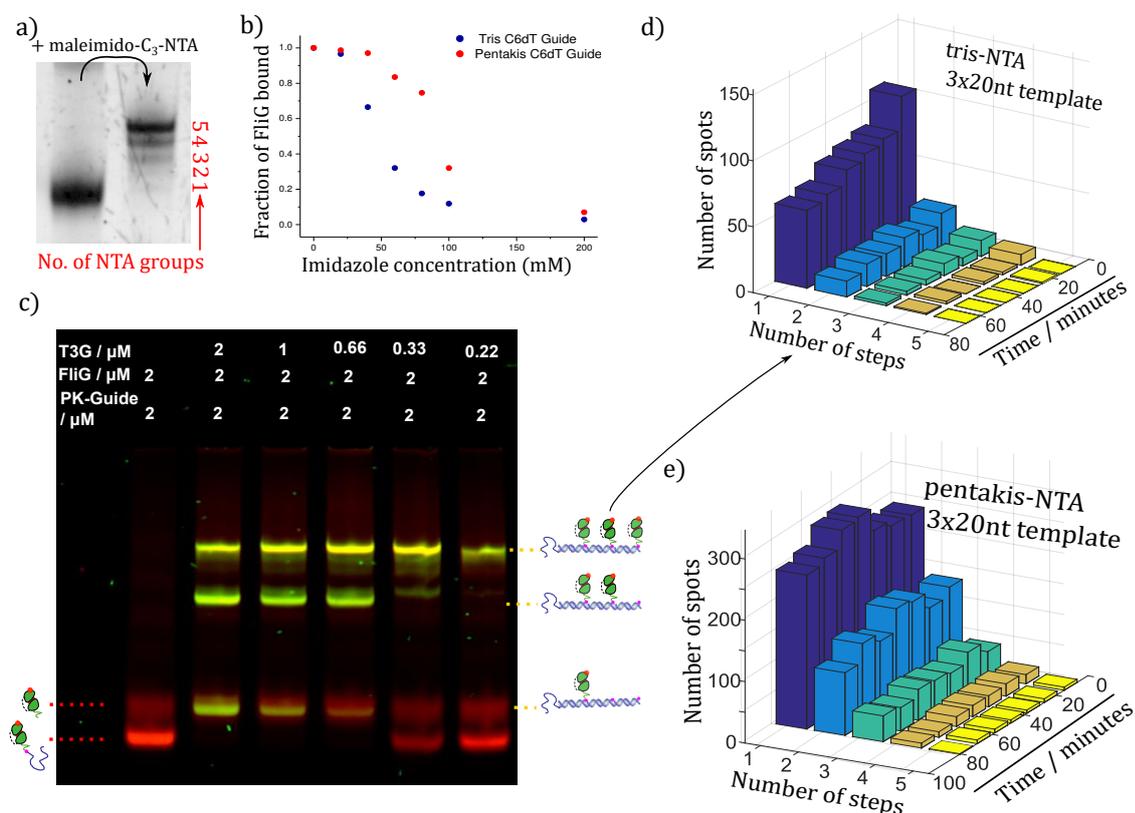


Figure 2.18: a) Last step in pentakis NTA synthesis, comparable to figure 2.15b b) Fraction of FliG bound to pentakis NTA oligo after stoichiometric incubation in presence of imidazole; measured with native PAGE. Figure and work: Joel Spratt. c) Titration of 3x20nt template against 20nt pentakis-NTA and FliG. Note: free FliG band and single-FliG-on-template band coincidentally overlap on this gel. DNA-only species have both run past the end of the gel. d,e) Histograms of FliG stoichiometry vs time, measured by single-molecule fluorescence in conditions of zero background FliG. Experiment fully described in section 4.1.3)

was required to dislodge FliG from pentakis-NTA oligos than from tris-NTA oligos, implying a higher affinity (Figure 2.18b). Furthermore, when I incubated 3x20nt templates with an excess of pentakis-NTA 20nt oligos and FliG, the vast majority of templates were routinely observed with 3 FliGs attached (Figure 2.18c) (See section 4.1.4 for more detail). Assuming that the functionality of the NTA attachment has no effect on affinity for the template, this strongly implies that almost all pentakis-NTA oligos are fully functional. Thus, FliG binding to unpurified pentakis-NTA oligos is easily stable enough for observation on PAGE.

However, the same stoichiometric filling cannot be unambiguously seen for 5x10nt templates (section 4.1.5); it is unclear if this is because the closer packing of

FliGs destabilizes them enough to overcome the histag-NTA binding. In this case, perhaps the strength of pentakis-NTA binding is limiting and covalent conjugation would be preferable.

Furthermore, when the 3x20nt templates were observed with single-molecule fluorescence (section 4.1.4 for detailed discussion), sub-stoichiometric filling was seen with both tris-NTA and pentakis-NTA. The pentakis-NTA structures (Figure 2.18e) as compared to tris-NTA (Figure 2.18d) were unambiguously more stable over time and more biased towards higher stoichiometries, but nevertheless mostly under-filled with FliG. Structures were assembled at high concentration, and observation started within minutes of diluting to low concentration (~ 10 s of pM), after which the distribution of counts was stable for over an hour. One explanation of this result is that some fraction of the pentakis-NTA strands bind FliG stably at higher concentrations (as observed a gel), but dissociate rapidly when diluted to the low concentrations required for single-molecule measurements. The remaining pentakis-NTA strands are reasonably stable in both concentration ranges. If this is the case, affinity-based purification of the pentakis-NTA oligos may make them viable for single-molecule experiments.

Another possibility, discovered later, is that NTA-histag binding is quite sensitive to incubation conditions. The structures for the microscopy experiment were assembled at a higher concentration than those for the gels, with a higher Ni^{2+} concentration, and stored at 4°C before observation. While this might be expected to increase stability of binding, reproduction of similar conditions at a later date showed sub-stoichiometric binding on a gel (see section 4.1.4). We have yet to go back and repeat the single-molecule experiments with preparations which definitely give full stoichiometric binding on a gel, but we do have at least a hypothesis as to why histag-NTA binding might be so sensitive to conditions, particularly Ni^{2+} concentration.

2.3.2.6 Competition with histag-histag binding

In PAGE experiments, ladders of FliG multimers were often observed, even in the absence of any DNA template (Figure 2.19). These ladders mostly disappear

if his-tags are cleaved off via the PreScission site (Joel's data, not shown), and typically appear in a Ni^{2+} -sensitive manner (Figure 2.19a); we interpret this as histag-mediated multimerization.

Outside of a very narrow window of Ni^{2+} concentration, this multimerization competes with histag-NTA binding (Figure 2.19a), and confuses analysis of any template-bound FliG bands. Furthermore, the tendency for histag multimerization is inconsistent between reactions, making it difficult to tune Ni^{2+} to this window, especially when concentration of FliG and NTA oligos must vary as part of an experiment.

This problem can be partially solved by pre-incubating the NTA groups with Ni^{2+} , then buffer exchanging away free Ni^{2+} before incubation with FliG; Figure 2.18c was acquired this way. However, for the 10nt pentakis-NTA oligo in particular, this solution has been imperfect; we lose significant quantity of oligo during buffer exchange (as we are right on the cut-off limit of the columns we use), and have to spend time re-measuring the oligo concentration. After this process (which takes $\sim 30\text{m}$), we incubate structures for another $\sim 30\text{m}$ before running on a gel, and still see FliG-FliG multimerization (Figure 2.19b). This could mean that Ni^{2+} is not properly removed, or that it is dissociating from the NTA or otherwise being transferred to the histags. Furthermore, some lesser degree of FliG multimerization is sometimes seen in reactions with no Ni^{2+} or DNA species. Consider that our end-goal is to investigate physiological FliG-FliG interaction; the risk of unphysiological interactions should make us wary. Removal of the histags entirely would be preferable, but rules out an NTA conjugation strategy.

2.3.2.7 Conclusion and motivation for covalent methods

To conclude: While we have successfully synthesised NTA-modified oligos which bind histags, they invariably comprise multiple species, and we do not have a clear method to purify those with the highest affinity for FliG (section 2.3.2.3). This is not insurmountable: now that we know binding is visible with native PAGE, we could realistically measure the binding of species separated with PAGE

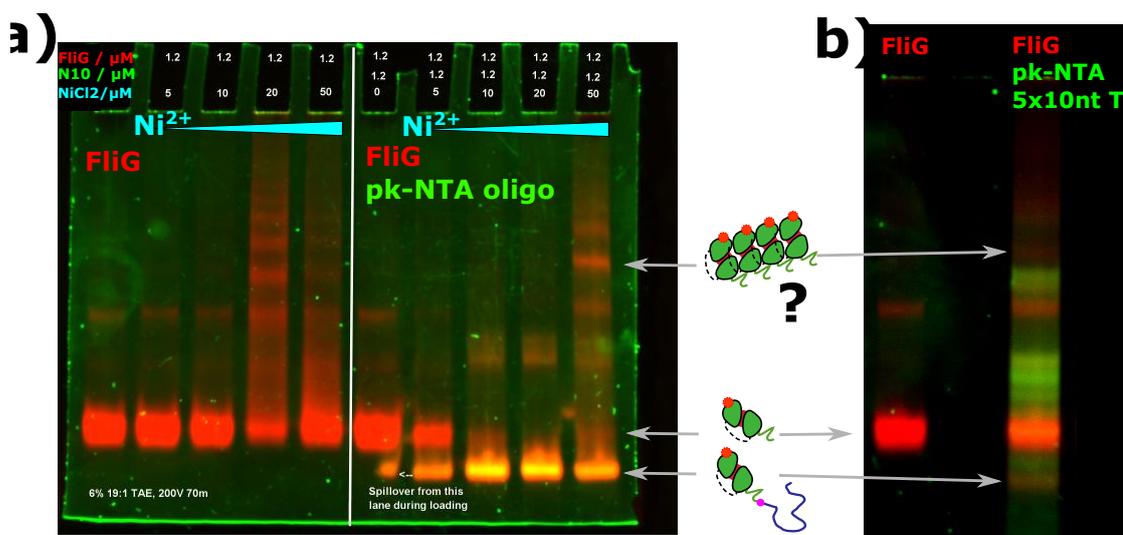


Figure 2.19: a) FliG incubated with and without 10nt pentakis-NTA oligo, at varying nickel concentration. Native PAGE. b) FliG incubated with and without 5x10nt template + 10nt pentakis-NTA oligo. The pentakis-NTA oligo has been pre-incubated with NiCl₂, then buffer exchanged. Native PAGE.

and/or reverse phase chromatography. Alternately, we could develop an affinity purification protocol.

We have made the discovery that pentakis-NTA oligos bind a 10-histag more stably than tris-NTA oligos (section 2.3.2.5); a novel and non-trivial result. For N x 20nt templates measured by native PAGE, they give stable and controllable binding. The same templates measured at low concentration with single-molecule microscopy apparently have many empty binding sites; it remains possible that oligo purification or optimized assembly conditions might rectify this. On 5x10nt templates, the binding stability is unclear, particularly in light of competing histag-histag interactions. It might be that closely-spaced FliGs destabilize each other somewhat, in addition to which we know the NTA modification is bulky and destabilizes binding of the oligo to the template. Again, these problems may not be insurmountable with further characterization and optimization.

However, we have put our more recent efforts towards developing more covalent methods. What are the advantages of this? Firstly, it would leave the DNA-DNA interaction as the only impermanent contact in the assembly. Not only does this simplify the interpretation of results, but it better allows us to control the

binding kinetics through changes to DNA sequence, the effects of which are well understood. Given that we have confidence in the stability of DNA-DNA interactions under conditions for single-molecule microscopy (sections 2.1.1, 4.1.2), and that covalent conjugation allows simple and direct measurement of dye labelling efficiency (section 2.2.3.4), we expect it to make the microscopy experiments immediately viable. Likewise, it may facilitate the use of other low-concentration assays, or potentially destabilizing methods (e.g. mass spectrometry). It should also reduce the need for fine tuning of conditions, and finally allow us to cleave the histags, avoiding the risk of unphysiological FliG self-interaction as described above. This argument will be revisited in chapter 4.

2.3.3 Covalent Methods

2.3.3.1 Guided Conjugation

Overview [While most of the work developing this technique was performed by Joel Spratt, I will describe it briefly as it motivated other work that I *did* perform.]

The natural abundance of lysines on protein surfaces (e.g. Figure 2.20a) makes them an ideal target for non-specific covalent conjugation to NHS esters[356]. A recently published method[357] coaxes pseudo-specificity out of this approach, as follows: a NTA-modified “guide” oligo binds to the histag of an engineered protein (Figure 2.20b). A partially-complementary NHS-modified “reactive” oligo is added, at concentrations ordinarily too low to react with the lysines(Figure 2.20c). However, those reactive oligos which hybridize to histag-bound guide oligos see a large local increase in the effective NHS ester concentration, promoting covalent reaction with histag-proximal lysines(Figure 2.20d). After this has occurred, the Guide-Reactive bond can be broken by the addition of a fully guide-complementary “displacer” oligo (Figure 2.20e), and the Guide-protein link can be broken by the addition of imidazole (Figure 2.20f). Thus the protein is left with just the reactive strand, bound covalently to a lysine somewhere near to the histag.

This method shares a number of advantages with the NTA-DNA approach, it requires only a histag, and is orthogonal to cysteine-based techniques (dye labelling,

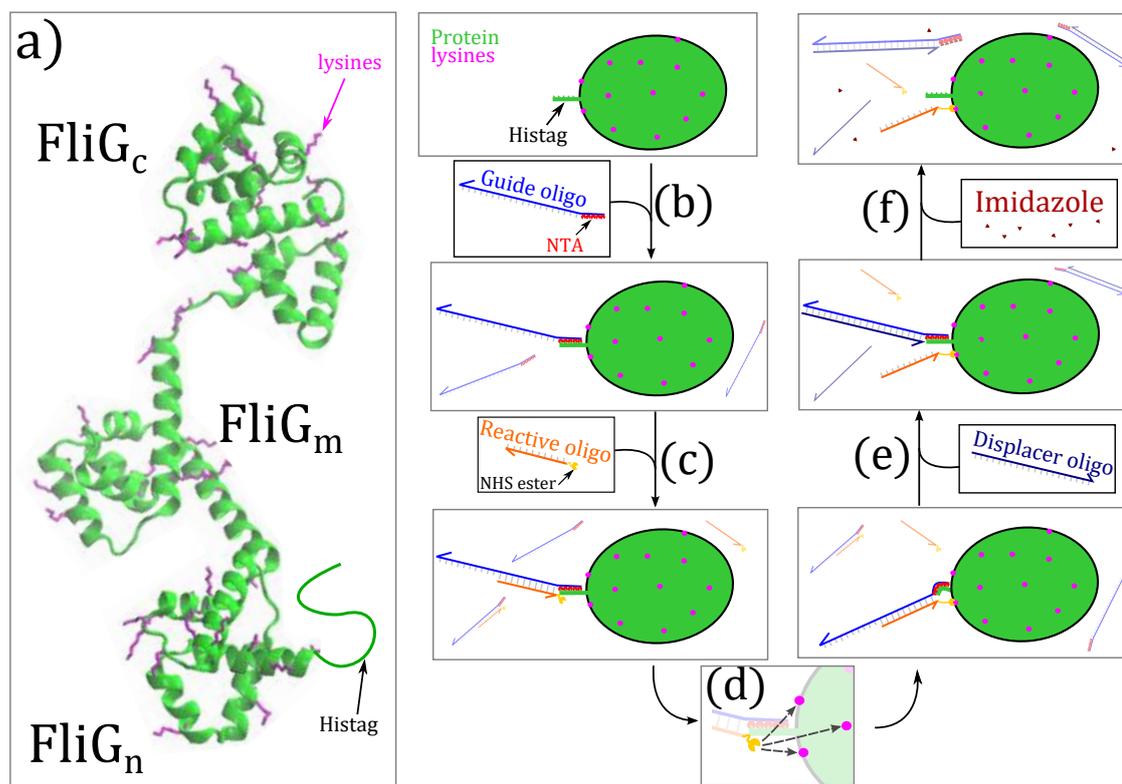


Figure 2.20: a) Crystal structure highlighting lysines in FliG from *A. aeolicus*[57], with position of the N terminal histag on our *E. coli* and *S. typhimurium* constructs noted. b-f) Template-directed conjugation[357]. See main text for description.

targeted cross-linking) that we might want to use. At the same time, it ought to be less sensitive to reagent heterogeneity. However, it leaves us with conjugates that may have heterogeneity in their labelling site (section 2.3.3.1), and need purifying from a host of DNA species and unconjugated protein (section 2.3.3.3).

Quantification of labelling site In [357], Rosen and co-workers used a cleavable NHS-DNA linker to facilitate tandem mass spectrometry, showing that only lysine sites close to the histag are modified with the reactive oligo. However, they dealt only with single-domain proteins (e.g. GFP), whereas FliG has multiple domains very flexibly linked. Therefore, while the histag is on FliG_n it is not inconceivable that lysines on FliG_c or FliG_m could be labelled. As the mass spectrometry approach is difficult and specialised, I looked for simpler ways to quantify whether FliG_c and FliG_m had been labelled.

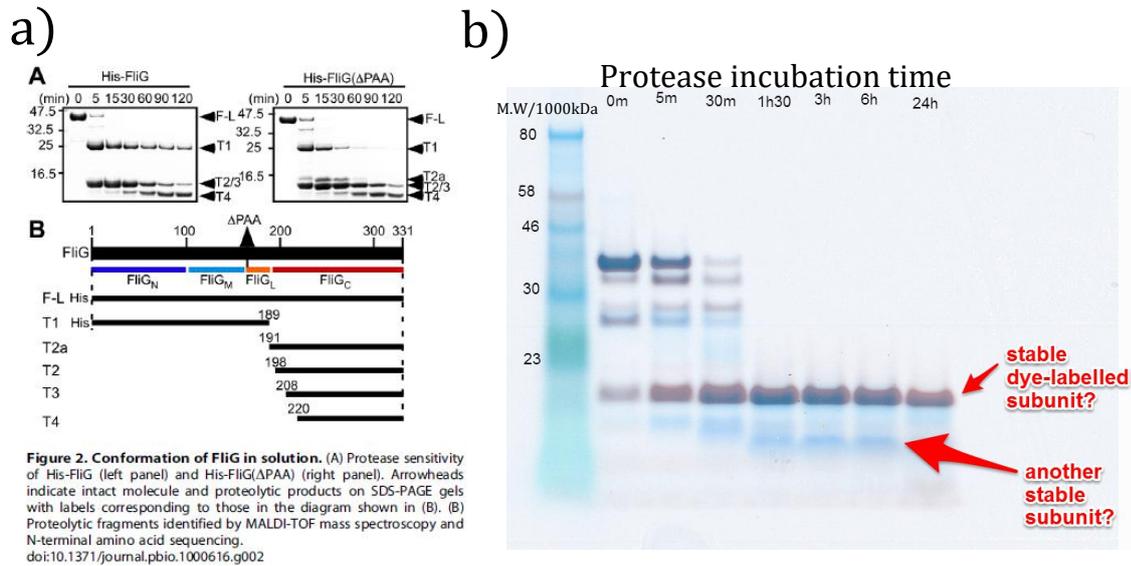


Figure 2.21: a) Limited proteolysis of FliG reported by [88] b) Limited proteolysis of FliG-AF647, analysed with SDS page. Coomassie blue staining (blue) is overlaid with AF647 absorption (red).

Limited proteolysis splits a protein into subdomains by exploiting the relative susceptibility of flexible inter-domain linkers to proteolysis[358]. Inspired by successful limited proteolysis of FliG by Minamino and co-workers [88](Figure 2.21a), I tried incubating FliG-AF647 with Proteinase K for various times in the hope that FliG_n, FliG_m and FliG_c (and DNA-conjugated equivalents) would be identifiable by SDS PAGE. Unfortunately, although there was a clear pattern to the proteolysis, it did not correspond to the pattern seen by [88], particularly when the signal from AF647 (known to be on FliG_c) was considered. Furthermore, the mass of visible proteolysis products do not add up to the total FliG mass. Therefore, without mass spectrometry, we do not have enough information to identify subunits. I also considered more site-specific proteases which would cleave FliG into a small number of fractions with predictable mass, but none had usefully-placed cleavage sites.

2.3.3.2 Maleimide-cysteine

[Again, most of the development work here was done by Joel Spratt, but informs other parts of this thesis, so I will describe it briefly.]

A period of success with the Template-directed Conjugation approach was followed by unexplained but repeated failure. This, along with the work required to quantify the location of labelling, motivated a search for different approach. This mirrors the dye-labelling approach (section 2.2.2.2), using a maleimide-modified DNA to label an engineered cysteine (at the FliG N-terminus, in this case). While early attempts failed, Joel now has this strategy working. Although quantities of conjugate have been limited so far, small-scale templating experiments have been very promising (see section 4.2). This strategy does have some downsides, however; in particular, it is incompatible with maleimide-dye labelling. One possible solution is to conjugate the dye to the protein-labelling oligo itself. However, commercially available doubly-labelled (dye + amine) oligos are very limited in choice of dye; we have yet to determine if this is compatible with clear photobleaching steps.

2.3.3.3 Purification

Whether the guided or maleimide-cysteine approaches are used, we need a method to remove unconjugated DNA oligos, which could block template sites or otherwise interfere with assembly. In the case of the guided conjugate, this task is a little more difficult, due to a more complex mixture of DNA species. In the most ideal case we would additionally remove unconjugated FliG, especially if the fluorescent label is on the DNA tag. Finally, given that we would reasonably want to probe FliG concentrations of ~ 10 s of μM in assembly experiments, and volumes below $1\mu\text{M}$ are difficult to handle accurately, we could plausibly be using 10 s of pmole per experiment, minimum. Thus, it would be advantageous to purify 100 s of pmole or more per batch.

Nickel Affinity Chromatography Nickel affinity chromatography is a likely feature of any purification strategy, selecting for protein-containing species (FliG and FliG + DNA) via their his-tags, which can then be cleaved via the PreScission site (see section 2.2.1). However, the guide conjugation mixture contains imidazole, which would disrupt column-histag binding. Furthermore, if imidazole was removed (e.g. by dialysis or size exclusion), the NTA guide oligo would compete with the

column for histag-binding. Therefore, utility for the guided conjugation is limited unless the guide strand and imidazole can both be removed first.

DNA affinity column A conceptually elegant purification strategy is to form a column with affinity for a particular DNA sequence, to selectively remove unwanted oligos (e.g. the guide strand) or retain desired oligos (e.g. the FliG-conjugate DNA). Lowered ionic strength or toehold-mediated strand displacement could be used to elute bound species. For the guided reaction, a column-bound displacer strand could hold FliG via the guide, before eluting FliG with imidazole.

I attempted to make such a column by incubating aminated displacer oligo (20nt) with an NHS-agarose resin (Figure 2.22a). The quantity of DNA eluted from this resin was less than the quantity loaded on, implying some fraction of the DNA had bound (Figure 2.22b), equivalent to ~ 1 nmole DNA per $250\mu\text{L}$ resin. This is a low capacity: a single batch of $\sim 20 - 30$ mL lysate provides ~ 100 nmole FliG, easily accommodated by a similar volume of nickel resin.

To test accessibility of column-bound oligos to sequence-specific hybridization, I incubated the DNA-modified resin with complementary (20nt guide sequence) and non-complementary (10nt reactive sequence) DNA. At high salt (1M NaCl), duplexes are stabilized, so complementary species should be retained, whereas zero salt should disrupt duplexes, removing hybridized species (Figure 2.22c). Unfortunately, there was no evidence of any DNA having bound to the column, save a tiny amount of non-complementary reactive sequence DNA (Figure 2.22d). This was confirmed by attempts to elute hybridized DNA with denaturing conditions (data not shown).

This is consistent with the only literature example I could find of a similar construct[359]. Using a similar NHS resin - amino DNA approach, the authors report that covalently bound DNA was only available for hybridization when attached to the column via a flexible linker; they found a 100-atom PEG to be optimal. In addition, even after heavy optimization, their yield of covalently conjugated DNA was only 5x what I achieved; in excess of the theoretical maximum achievable with many other chemistries (e.g. biotin-streptavidin), but still not enough to purify

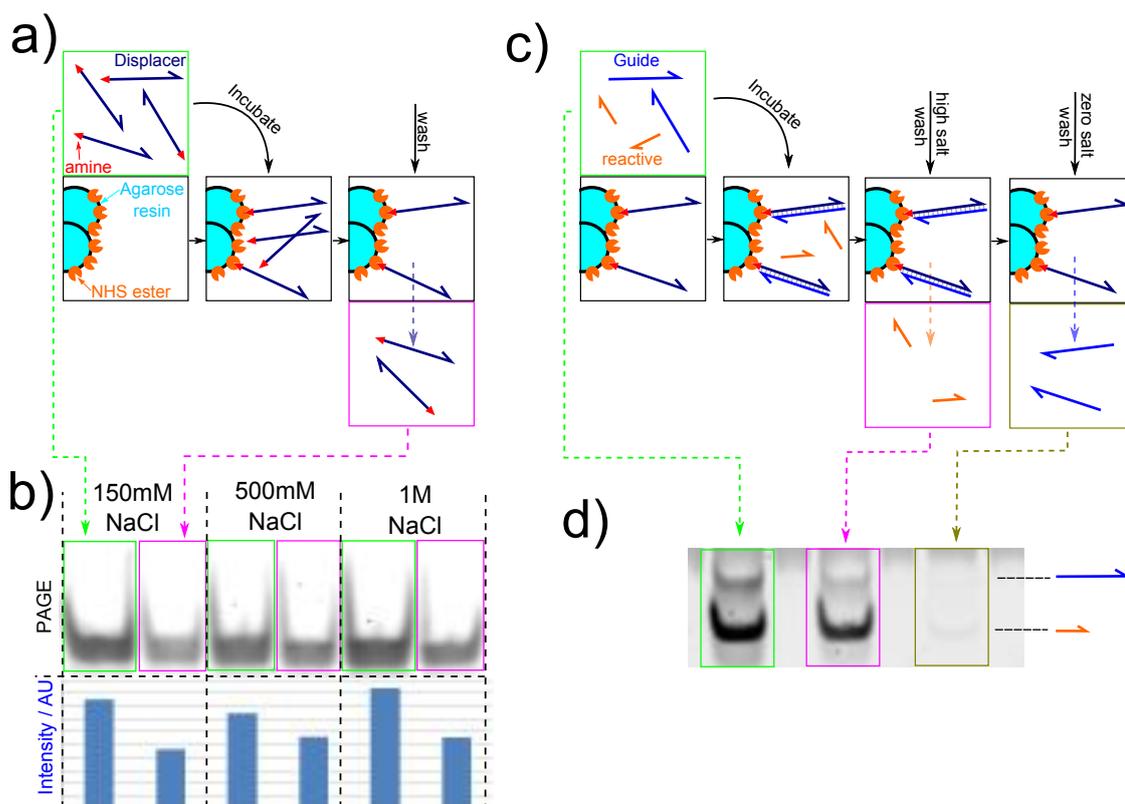


Figure 2.22: a) Binding of aminated DNA oligo to NHS-agarose resin b) Quantification of binding by SYBR gold DNA staining on native PAGE, comparing loaded and eluted DNA, after incubation at a variety of salt concentrations. Lanes represent the same fraction of loaded / eluted volume, and are normalized to the same salt concentration. c) Test of sequence-specific DNA binding to column. Diagram shows what we would expect if system was working as hoped. d) Quantification through SYBR gold DNA staining on native PAGE. Again, lanes represent the same fraction of loaded / eluted volume.

a 100nmole FliG batch without many mL of resin. Given the expense of resin and suitable PEG crosslinkers, along with the substantial development work still required, I decided to concentrate on other strategies.

Native Gel purification We already know the DNA-FliG conjugate is separable on native PAGE, making it attractive for purification. In the original guided conjugation study [357], DNA-GFP and DNA-antibody conjugates remained functional after passive elution from homogenised gel slices. A literature also exists on the purification of protein complexes through passive elution[360] and also electro-elution[360–363], in which an electric field pulls the sample out of a gel slice, on to a dialysis membrane. However, this literature warns of damage which

can be caused during electrophoresis[360, 361]; a slight concern, given that we lack of a test for FliG activity (section 2.2.4). Nevertheless Joel Spratt tested both passive elution and electro-elution.

He found that FliG-DNA (10nt) conjugates passively eluted only from low percentage (6-10%) gel slices, whereas high percentages (15% 19:1) were required to give separation between DNA-labelled FliG and non-labelled FliG. Following on from this, I established that a stacking gel could be designed to separate species in a high-percentage region, before passing them into a low-percentage region for elution (albeit not with particularly clean bands; Figure 2.23a).

Joel also tried electro-elution, and found that while protein could be electro-eluted from $\sim 15\%$ 19:1 gels, it was denatured in the process; presumably due to heat, or the lack of salt in typical electrophoresis buffers. Although adding salt to the buffer will increase heat generation, this may be manageable by efficient cooling: electro-elution may still be viable.

However, we questioned whether gels could handle sufficiently large quantities to be useful, given that analytical gels are typically run with \sim pmole quantities. I found that while fluorescent DNA stains behaved poorly with heavily-loaded gels, and heavy loading of fluorescent DNA saturated our gel scanner, heavily-loaded bands could be visualized by doping with a small quantity of fluorescent oligo. With this technique, I found that DNA-only bands show remarkably little distortion even in the presence of ~ 1 nmole DNA per lane (Figure 2.23b). Overloading of protein however is more deleterious, causing clear distortion (and therefore decreased resolution) at ~ 0.17 nmole (Figure 2.23c). I suspect this limits us to purifying ~ 1 nmole across an entire gel; similar to a hypothetical DNA affinity column. Having said that, we primarily care about the resolution achievable with “real” DNA-FliG conjugate samples, which thus far have been too scarce to use in these kind of tests.

Potentially, preparative gel electrophoresis systems (though expensive) could overcome these limits on protein quantity. An alternate possibility is purification via agarose gels, rather than polyacrylamide. DNA origami structures are often purified from agarose gels, either by spin filtration[306, 364] or simply compressing

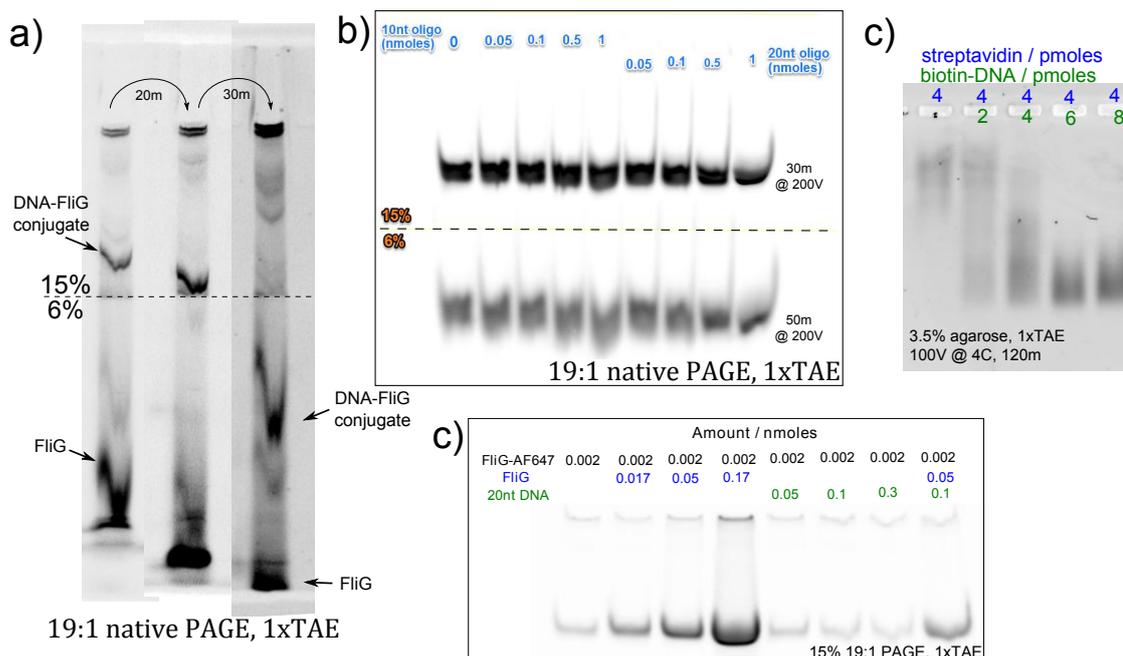


Figure 2.23: a) Three snapshots of a stacking native PAGE gel, running a small quantity (~1 pmole) of FliG-AF647 following a guided conjugation procedure, imaged via the AF647 fluorescence signal. b) Stacking native PAGE (2 snapshots), showing a small quantity (~1 pmole) of fluorescent 10nt DNA doped with varying quantities of non-fluorescent 10nt and 20nt DNA. c) Native PAGE; small fluorescently-imaged quantity of FliG-AF647 doped with varying amounts of non-fluorescent FliG and DNA. d) Native agarose gel electrophoresis; streptavidin bound to varying amounts of 10nt biotinylated single-stranded DNA.

gel slices to expel the sample and solvent[261]. There is only a little literature on purification of proteins from agarose gels, all concerning very large proteins or complexes[360, 361]. To test whether agarose was likely to have sufficient resolution for conjugate purification, I used streptavidin conjugated to differing numbers of 10nt biotinylated DNA oligos (covalent FliG conjugate not being available at the time). As streptavidin has slightly higher mass (53 vs 40 kDa) and smaller charge (-7.6 vs -10.3 at pH7 [365]) than FliG, it ought to be slightly slower-running than FliG, and thus easier to separate from DNA. However, even with a 3.5% agarose gel (very dense; too dense for squeeze extraction, and on the limit of easy handleability), resolution was too poor to clearly separate conjugated and unconjugated species (Figure 2.23d). Therefore, PAGE purification is probably more promising.

Ion exchange Ion exchange is another potentially attractive purification method; binding negatively charged molecules (FliG, DNA) to a positively charged column, and eluting in order of increasing net charge by a gradient of increasing salt concentration. However, using a HiTrap DEAE Fast-Flow HPLC column, resolution between species was inadequate (Figure 2.24a), and species in a guided conjugation mix were poorly separated (Figure 2.24b). For unclear reasons, the FliG elution profile is quite broad (Figure 2.24a), and overlaps with all DNA species; perhaps unsurprising given the substantial negative charge of FliG ($-10.3e$ expected at pH 7 [57, 365]). While this charge could be drastically reduced by lowering the pH, increasing separation, we expect that this would result in denaturation or other damage. Thus, it seems unlikely that DEAE resin is sufficient for separation of FliG and DNA. It is possible however that other column substrates may fare better; FliG is reported not to bind to ceramic hydroxyapatite (CHT)[366], a weak binder of DNA². If this can bind very short (~ 10 nt) oligos, it may be suitable for separating DNA-conjugated and unconjugated FliG.

Size Exclusion Chromatography The molecular weight of a 10nt oligo (~ 3 kDa) is so much smaller than the weight of FliG (~ 40 kDa) that, while using size exclusion to separate DNA-only species from FliG-containing species is plausible (but not obviously better than using a nickel column), separation of DNA-tagged FliG from untagged FliG would be infeasible. This could perhaps be solved by the use of a sacrificial template, to increase the weight of tagged FliG (Figure 2.25). While scalable, this is convoluted and might require significant wastage of DNA.

Future directions The purification strategies above were largely developed with guided conjugation in mind. As we are more recently favouring the maleimide-cysteine strategy, purification becomes a little simpler. For small-scale trial experiments with these conjugates (e.g. section 4.2), Joel has been using nickel-affinity columns to remove DNA-only species, which is easily scalable. When it becomes necessary to remove DNA-free FliG, we hope to try either ion exchange

²<http://www.bio-rad.com/webroot/web/pdf/lsr/literature/10011433.pdf>

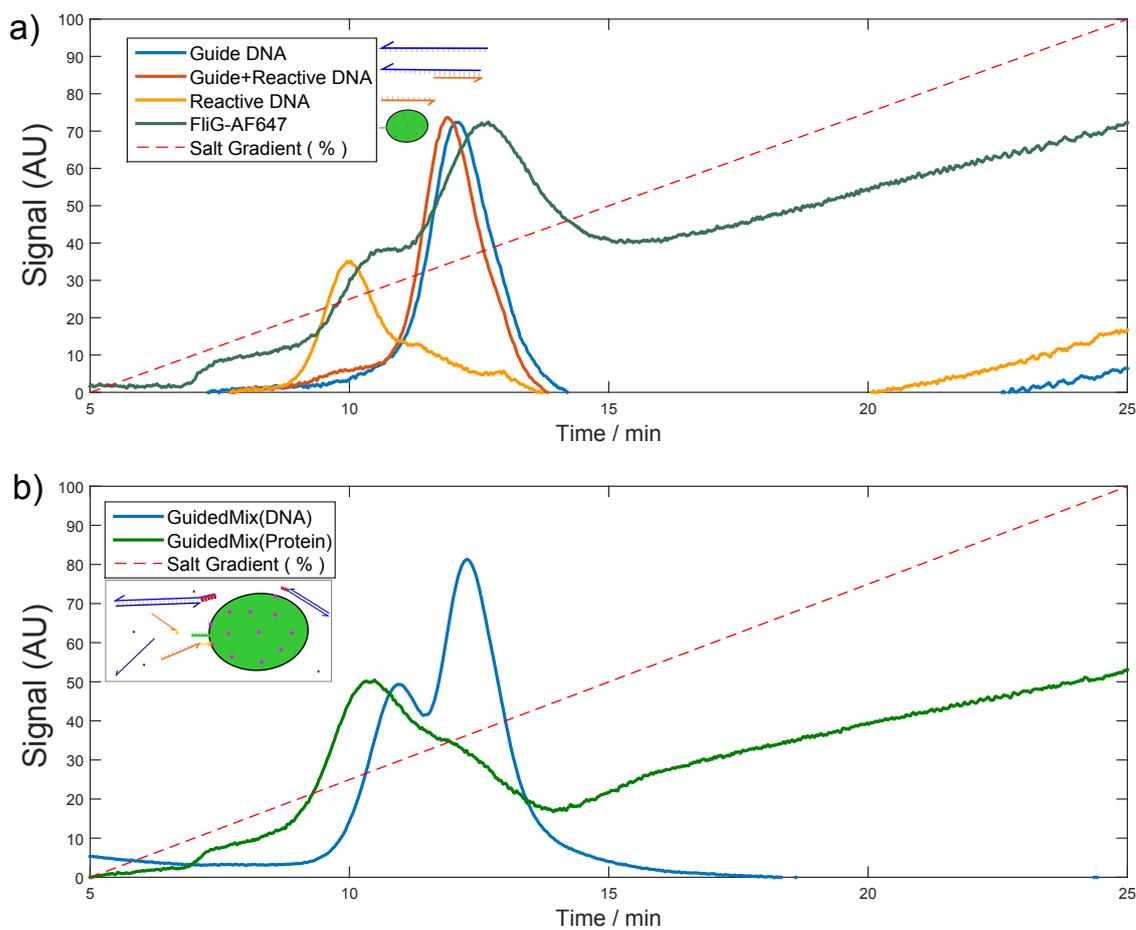


Figure 2.24: Ion exchange elution profiles, DEAE FF column, gradient from 150-1000mM NaCl in 20mM Tris-HCl pH7.5. DNA corresponds to absorption at 260nm, protein corresponds to the AF647 absorption at 650nm. a) Separate elution profiles of species involved in the guided conjugation reaction. b) Elution profile of a guided conjugation mixture.

with a CHT resin or size exclusion on templates (as described above), both of which ought to be high throughput, and less likely than PAGE to damage the protein.

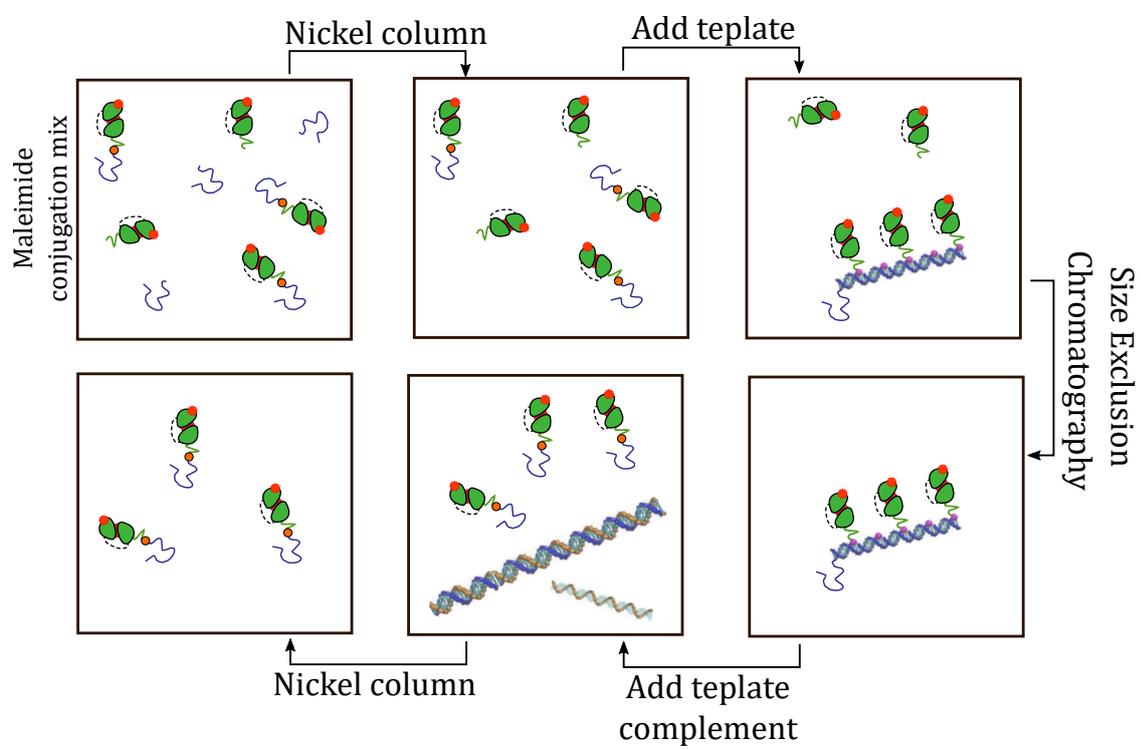


Figure 2.25: Plausible strategy for scalable covalent conjugate purification, using size exclusion chromatography.

3

Single Molecule Microscopy

Contents

3.1	Hardware	114
3.1.1	Microscope	114
3.1.2	Flow chambers	114
3.2	Surfaces and experimental approach	117
3.2.1	Surface Development	119
3.2.2	Template distribution on surfaces	122
3.3	Software and counting pipeline	122
3.3.1	Version 1 - Intensity in ROI	123
3.3.2	Version 2 - Gaussian Fitting	126
3.4	Anti-bleaching system	128
3.5	Dye behaviour and counting rules	129
3.5.1	Counting large numbers of dyes	130
3.6	Conclusion	132
3.6.1	Gel Electrophoresis: an alternative?	133

The previous chapter concerned the biophysical building blocks needed to template FliG on DNA templates. This chapter concerns microscopy techniques to characterize FliG stoichiometry on such structures, which can be used as a readout of gap-filling or templated polymerization, as discussed in section 1.5.1.

3.1 Hardware

3.1.1 Microscope

Figure 3.1 describes the custom TIRF microscope used for all fluorescence experiments. All experiments described in chapter 4 use 633nm HeNe laser illumination with a multi-bandpass filter set (Semrock FF01 446/523/600/677 and Di01-R405/488/561), illuminating an area of either $\sim 10\mu\text{m}$ diameter or $\sim 40\mu\text{m}$, depending on microscope configuration. In experiments where maximum intensity was required (e.g. section 3.5.1), spatial filtering pinholes were removed to achieve $\sim 10\text{mW}$ into the rear of the objective, or $\sim 130\mu\text{W}/\mu\text{m}^2$ at the coverslip. Power densities as low as $\sim 3\mu\text{W}/\mu\text{m}^2$, however, still generated clear traces (e.g. section 4.1.3). All experiments were performed in TIRF unless otherwise stated. Final images correspond to 86nm/pixel; roughly optimal for single-molecule localization when using a CCD[367], but not obviously optimal for counting experiments using an EMCCD. Video acquisition used 15-30ms exposure times and 300 EM gain. Flow chambers (section 3.1.2) were mounted via spring clips onto a custom mount, screwed to a piezo stage (PI P517.3CL, controlled by an amplifier E503 and servo E509-C3A) atop a custom translation stage. Brightfield illumination was provided by a fiber-coupled Thorlabs MCWHL2 LED, focussed by a Nikon long working-distance condenser (NA 0.52) for Köhler illumination.

3.1.2 Flow chambers

For most experiments, open-ended flow chambers were constructed from double-sided Scotch tape, sandwiched between a glass coverslip and glass slide (Figure 3.2 a). To load samples, liquid pipetted at one end of the chamber was wicked through by a tissue held at the opposite end. This chamber could be rapidly sealed with VALAP; an easily-meltable wax-like 1:1:1 mixture of petroleum jelly, lanolin and paraffin, popular in live-cell imaging due to its low toxicity [368]. VALAP applied to the ends of the channel was melted with a hot wire tool, sealing the channel and

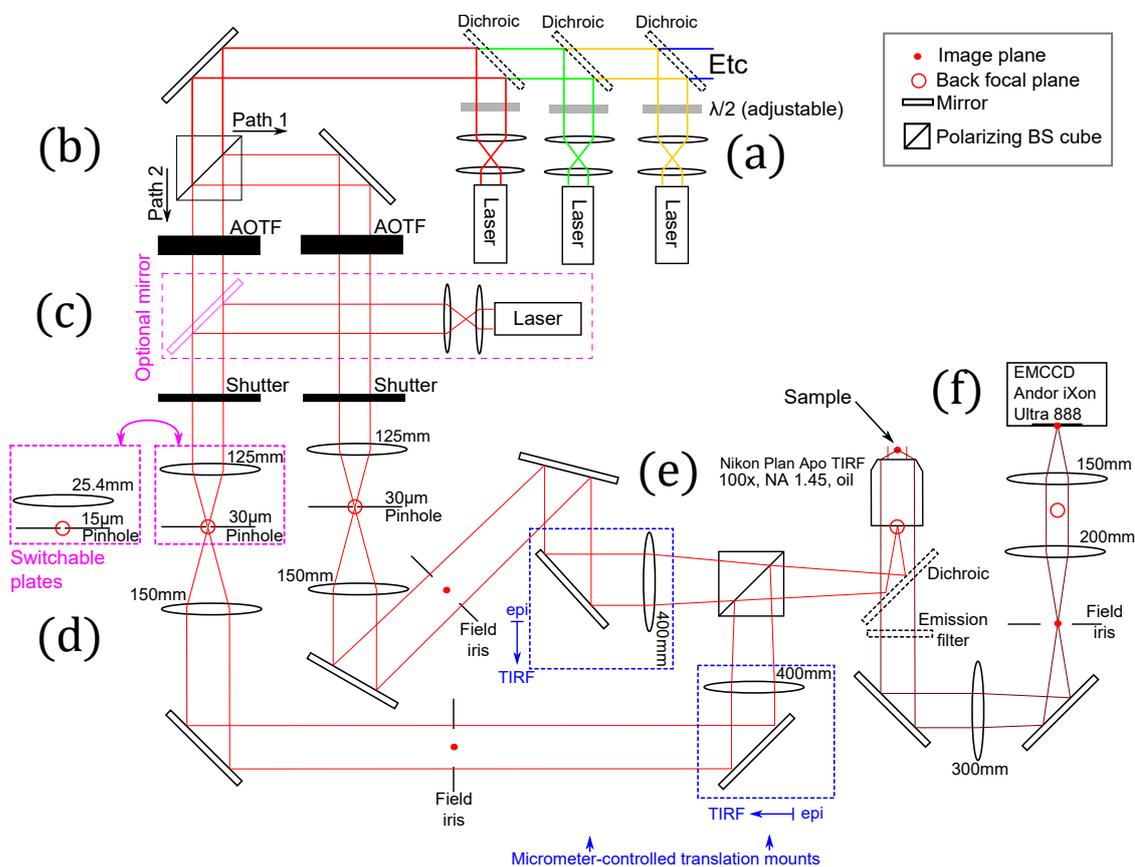


Figure 3.1: Optical arrangement of the microscope. For clarity, paths and configuration options not relevant to this thesis (e.g. optical traps) are not shown. **(a)** Multiple lasers are expanded to the same beam width, individually plane-polarized with a half-waveplate and combined with a series of dichroics. **(b)** A polarizing beamsplitter cube divides each beam between two paths. Acousto-Optical Tunable Filters (AOTFs) on each path allow laser selection and rapid shuttering, but reduce maximum beam intensities and distort beam profiles. **(c)** Optionally, a 633nm HeNe laser can be introduced into path 2, bypassing the AOTFs for maximum power. Both paths then pass mechanical shutters, and are expanded with telescopes. **(d)** In path 1, a switchable plate allows different telescope ratios for different illumination areas. Pinholes spatially filter the beam (necessary especially for beams that have been distorted by the AOTFs). **(e)** Paths are recombined by a polarizing beamsplitter, and focussed onto the back focal plane of the objective for epi-illumination. Translation mounts allow each path individually to move in and out of TIRF. **(f)** Fluorescent emission passes through a standard dichroic / emission filter combo before being focussed onto an EMCCD.

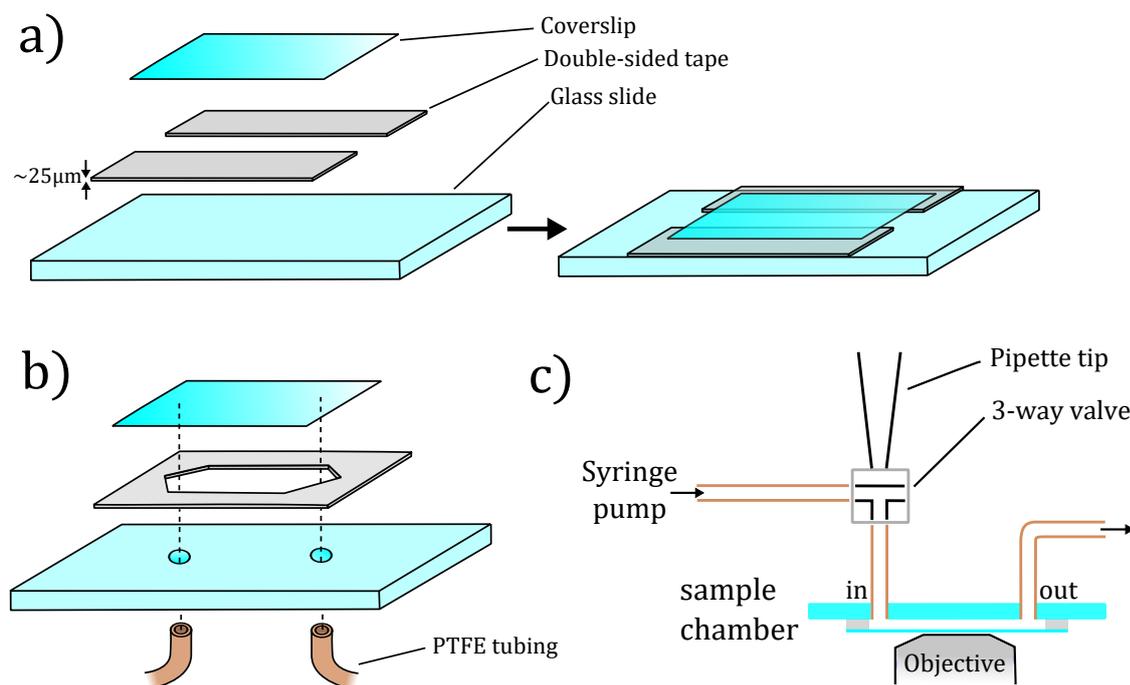


Figure 3.2: a) Open flow chamber construction. b) Closed flow chamber construction. c) Microfluidics configuration for closed chamber.

re-solidifying within seconds. However, neither sample loading nor sealing could be performed without removing the sample from the microscope.

For experiments requiring a constant flow during imaging or rapid imaging after sample-loading (e.g. section 4.1.2), closed flow-chambers were used, which could be loaded while on the microscope (Figure 3.2 b). Holes were drilled in glass slides to allow the insertion of PTFE tubing, which was sealed in place with Araldite epoxy. The input tube was connected to a 3-way valve (Figure 3.2 c) which could switch between a syringe pump (for constant buffer flow or washing) or a pipette tip (for sample loading). The output tube was either open to waste (when the syringe pump was in operation) or connected to a syringe, used to manually pull through samples loaded in the the pipette tip.

In all experiments, flow chambers were constructed just before use, washed with multiple channel volumes of water or buffer, and incubated with dilute 330nm latex beads for ~30s coverslip-down, before washing (Figure 3.3 a). A small amount of salt in the buffer (e.g. ~150mM NaCl) is sufficient to adhere sedimented beads to the coverslip surface, allowing their use as focus markers under LED illumination.

Such illumination is dim and spectrally broad, and so avoids fluorophore bleaching. With practise, one can learn the bead appearance corresponding to a perfectly in-focus surface, meaning that fluorophores are already in focus as soon as fluorescent imaging begins. Bead coverage is deliberately sparse, as they are poorly passivated against adhesion of fluorescent samples (e.g. FliG).

3.2 Surfaces and experimental approach

For FliG assembly experiments, coverslip surfaces must obey a tight set of requirements. Firstly, their autofluorescence must be sufficiently low that single molecules are easily distinguishable. Secondly, they must be able to bind DNA templates via some specific chemistry, at a density low enough to distinguish individual molecules, without perturbing DNA-FliG binding. Finally, the amount of non-specific adhesion by FliG or other fluorescent molecules to the surface must be sufficiently low that templated structures remain distinguishable. (In the worst case, signal from non-specifically bound FliG will completely overwhelm any signal from specifically template-bound FliG). As the rate of non-specific binding increase with concentration, surface passivation effectively limits the FliG concentrations we can use in the chamber.

Initially, I had hoped to observe FliG assembly in real time, by incubating surface-bound templates with physiological concentrations of FliG in the chamber. However, I later realised that the approximate *in vivo* FliG concentration ($\sim 500\text{nM}$ [111]) would challenge even the best-passivated surfaces reported in the literature [369, 370]. Therefore, all the fluorescence experiments in this thesis involve pre-assembling and equilibrating structures at high concentrations (Figure 3.3 b), diluting to slightly sub-nM concentrations, incubating with the surface briefly (allowing templates to bind; Figure 3.3 c) and finally washing away all unbound proteins before imaging (Figure 3.3 d).

This strategy still requires high-performance surfaces to prevent non-specific adhesion during the surface incubation. It is also inevitable that once unbound FliG is removed, structures will tend towards a new disassembled equilibrium (Figure

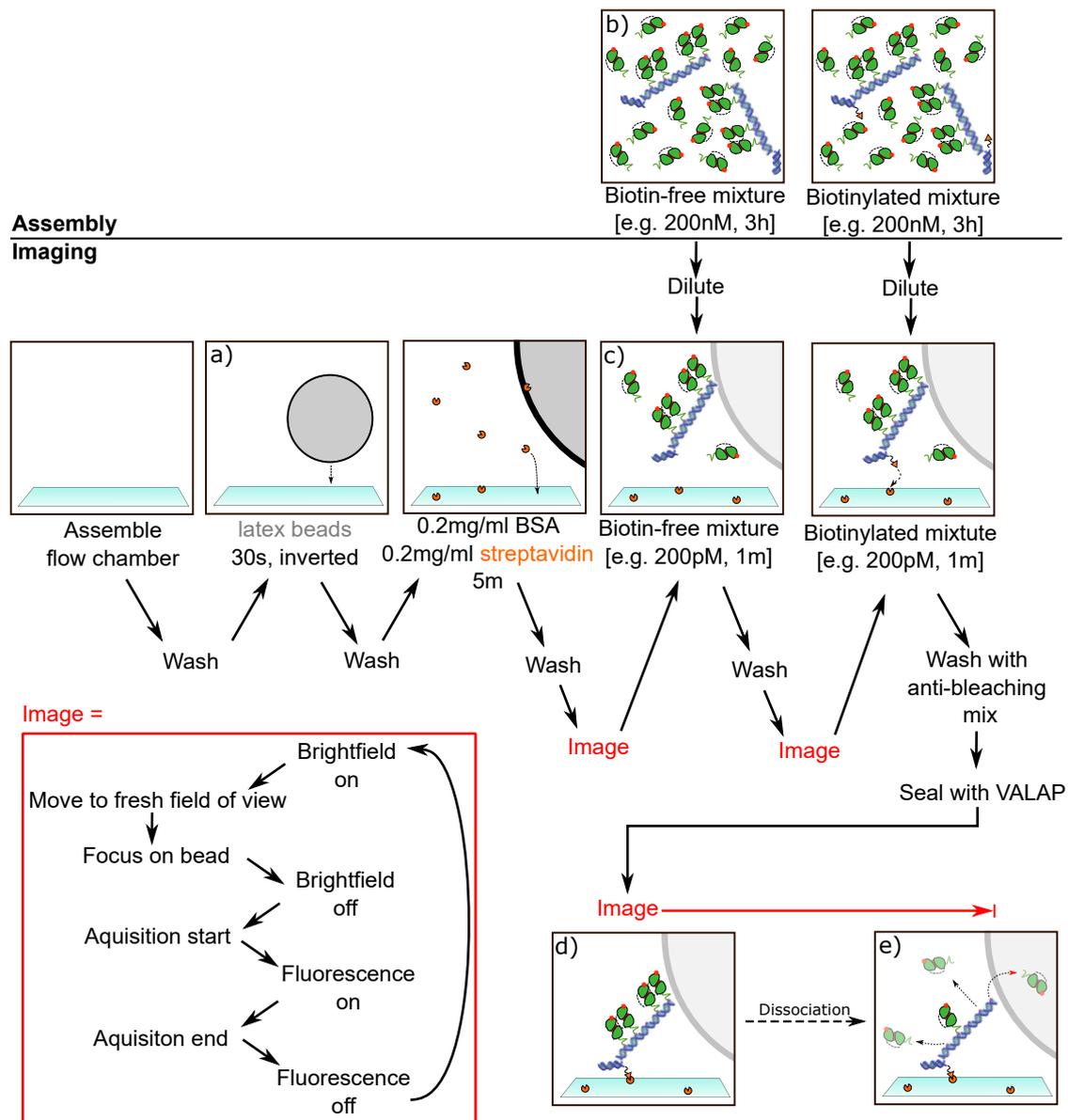


Figure 3.3: Structure of a typical fluorescence experiment. a-e) see main text.

3.3 e), with a rate depending on the stability (i.e. dissociation kinetics) of the structure. If disassembly rates are sufficiently low, there should be time to measure stoichiometries corresponding to the high-concentration equilibrium. This motivated our concerns about the stability of DNA-DNA and DNA-protein links in sections 2.1 and 2.3, and will furthermore motivate some of our data analysis in chapter 4.

3.2.1 Surface Development

Surfaces were initially tested with AF488-labelled FliG and 10nt AF488-labelled DNA oligos on biotinylated 3 x 10nt templates. All surfaces utilized streptavidin to bridge biotinylated surfaces and biotinylated templates with great stability[371]. See appendix A.7 for detailed protocols.

3.2.1.1 BSA,BSA-biotin

Glass coverslips were baked in a furnace at 500°C for 1h (just below their melting point) to remove fluorescent contaminants[372]. Just before use, surfaces were incubated with a mixture of BSA and BSA-biotin (expected to adsorb non-specifically onto the glass surface), before incubation with streptavidin; this mirrors protocols typically used for DNA PAINT experiments[305, 373]. BSA clearly reduced the non-specific adhesion of 1nM AF488-FliG (Figure 3.4a), but not sufficiently well to meet our requirements. Retrospectively, this is unsurprising, and consistent with the literature that recommends BSA surfaces only for DNA-only experiments[240].

3.2.1.2 PEG

Polyethylene glycol (PEG) is commonly used to passivate surfaces in single-molecule FRET experiments [240, 330]. I tried two PEG-based approaches. The first is derived from a popular silanization-based protocol [370, 374], subsequently simplified by the Kapinidis lab (Oxford)¹ [375]. In brief, a plasma-cleaned glass coverslip is silanized with an amino-silane agent (Vectabond SP-1800, Vector Labs), covering it in amino groups which are then labelled with a mix of NHS-PEG and NHS-PEG-biotin (MW 5000 Da, Laysan Bio). The second approach is based on PLL-PEG [376] (SuSoS, Dübendorf, Switzerland): a Poly-L-Lysine (PLL) backbone grafted with occasional PEG sidechains, some fraction of which may be biotinylated. The positively charged lysines readily coat the negatively-charged surface of a plasma-cleaned coverslip. In both methods, surfaces coated with PEG-biotin were then incubated with streptavidin and BSA just before use (Figure 3.3).

¹Thanks to David Dualin (Kapinidis Lab) for showing me this protocol

Both strategies worked only after considerable effort, with the PLL-PEG strategy being much less reliable. The positively charged PLL is also expected to attract negatively charged DNA unless shielded with high salt concentrations[377], limiting our freedom to test assembly in varied conditions. At best, both surfaces showed specific binding of DNA control templates, but with considerable amounts of intrinsic background fluorescence of a similar brightness to AF488 molecules (Figure 3.4b). For the silane-PEG surfaces, at least some of this fluorescence was a feature of the vectabond reagent itself (Figure 3.4c) and thus difficult to eliminate. However, this background was invisible in a red imaging channel (Cy5/AF647), motivating the switch from AF488 labelling to AF647 labelling (section 2.2.2.2). After this switch, silane-PEG surfaces performed well, with low intrinsic fluorescence or non-specific adhesion (Figure 3.4d). As we expect to occasionally find poorly-passivated coverslips[378], biotin-free control samples were routinely incubated prior to the “real” biotinylated sample to test for proper passivation, as in Figure 3.4d. (Refer back to figure 2.17b for removal of templated FliG via imidazole, demonstrating that template-binding does not promote subsequent non-specific surface adsorption.)

Even at $\sim 100\text{nM}$ concentrations, non-specific adhesion of FliG was not overwhelming (Figure 3.4e). However, with ~ 1 molecule/second landing in a $\sim 20\mu\text{m} \times 20\mu\text{m}$ area, it would probably still be sufficient to confuse measurements of templated structures; especially considering that molecules may take many minutes to bleach when the anti-bleaching system is used. NB: at these high ($\sim 100\text{nM}$) concentrations, while there is a considerable background caused by fluorophores in the channel diffusing through the TIRF illumination zone, it is quite diffuse if imaged at slow frame rates. Thus, it does not obviously limit spot detection or intensity measurements: the surface is still the limiting factor.

3.2.1.3 Future approaches

A recent protocol for surface modification with Tween-20 reportedly reduces non-specific adhesion of most biomolecules by at least 10-fold, as compared to silane-PEG[369]. Approaches based on star-shaped PEGs [379, 380] or double applications

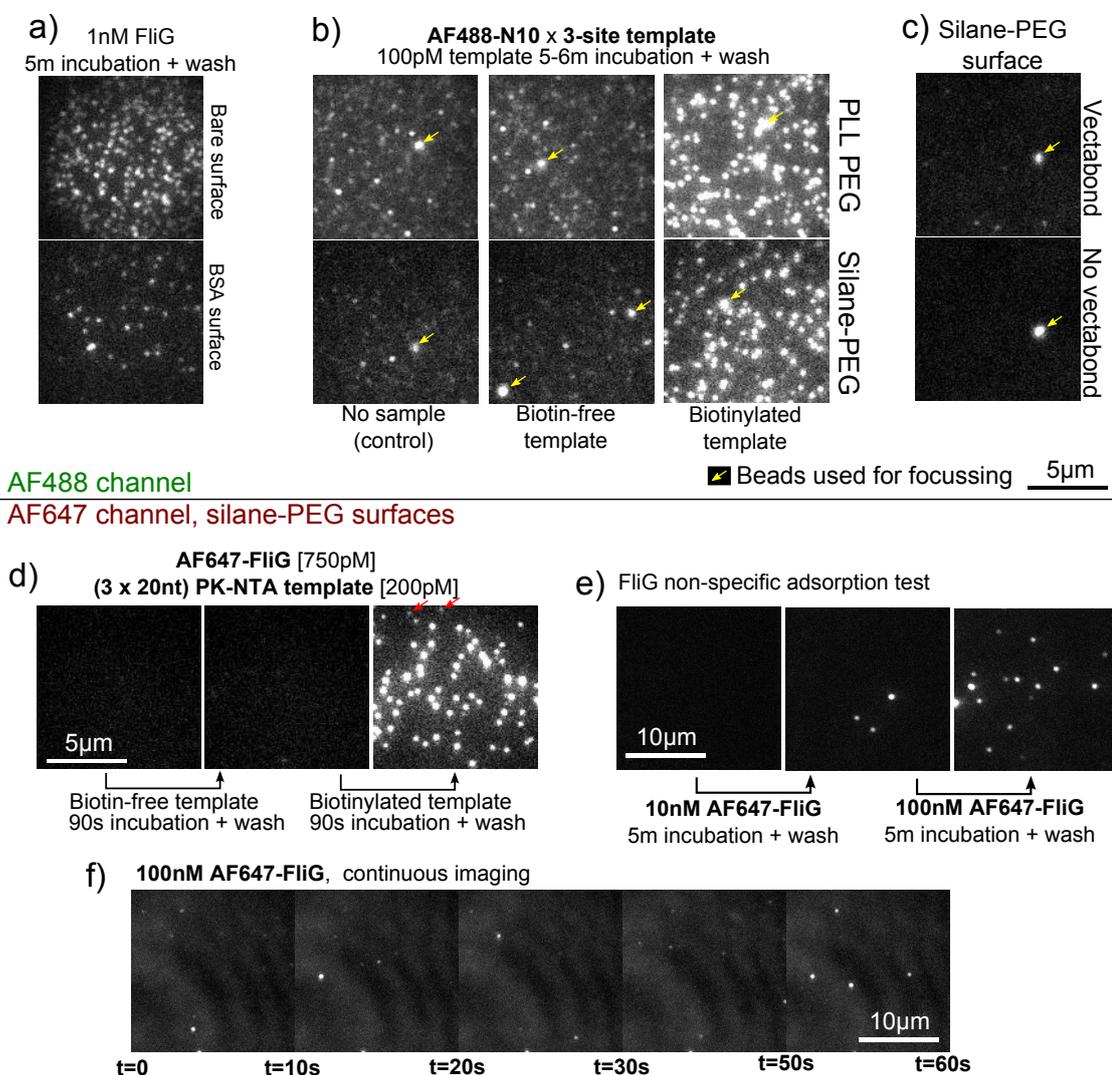


Figure 3.4: a) Non-specific adhesion of FliG to bare and BSA-coated surfaces in buffer IB2. b) Specific binding of biotinylated 3x10nt templates with AF488-labelled 10mers, on alternative preparations of PEG surface, in buffer MB. The PLLPEG images show different slides, whereas the silane-PEG images come from the same slide, after sequential incubations. c) Silane-PEG surfaces prepared identically, except for the inclusion/omission of vectabond. No sample present, aside from focussing beads. d) Specific binding of AF647-labelled FliG to streptavidinated Silane-PEG surface via 3x20nt pentakis NTA template in buffer MB. Same experiment detailed in section 4.1.3. Red arrows mark single fluorophores, as inferred by single bleach steps. e) Non-specific binding of AF647-FliG to streptavidinated Silane-PEG surface in buffer MB. f) (Same slide as (e)); single patch of slide, imaged continuously at 2Hz with 100nM AF647-FliG in the chamber. Each image displays the per-pixel maximum brightness over a 20 frame (10 second) window, summarizing FliG surface-binding events (which were followed typically by bleaching or surface-dissociation within a few frames).

of silane-PEG[381] may also improve performance. Thus, real-time assembly in the chamber at physiological concentrations may not be entirely out of reach.

We could also consider more sophisticated strategies to handle high concentrations by selectively exciting templated FliGs over non-templated. While some, such as those reliant on photo-activation of fluorophores [382], are not suitable for counting (due to poor activation yields), others might be. For example, FliG could be labelled with a FRET acceptor, fluorescing only in proximity to donor-labelled DNA template [383]. However, this would require careful characterization of how gap-filling FliG fluoresces, as compared to FliG directly bound to the template. Another approach (refer back to figure 1.14e) might involve placing templates in fluorescence hotspots formed between plasmonic nanoparticles on DNA origami scaffolding. Again, however, differing fluorescence of FliG on different parts of the template could be a problem.

3.2.2 Template distribution on surfaces

Useful stoichiometry measurements require that templates are well-separated on the surface. While spot density on the surface can be controlled by surface incubation time and concentration, we did briefly worry that single streptavidins may bind multiple biotin-templates (section 4.1.3). This was always unlikely; streptavidin has 2 pairs of binding sites, and DNA-biotin is sterically inhibited from binding more than one site in each pair[384]. We also know that biotins on biotin-PEG surfaces are liable to block multiple binding sites[376]. Nevertheless, we performed a control with AF647-labelled DNA linkers on a 1 x 10nt template (Figure 3.5a). After discarding traces from identifiably separate-but-overlapping spots (see section 3.3.2), only $\sim 1\%$ of spots showed multiple bleach steps (Figure 3.5b), confirming that templates are well segregated, and additionally that single AF647s (on DNA) bleach in a single step.

3.3 Software and counting pipeline

There are 3 steps to processing raw data: identifying isolated spots, measuring the intensity of each spot to generate a bleaching trace, and finally identifying the

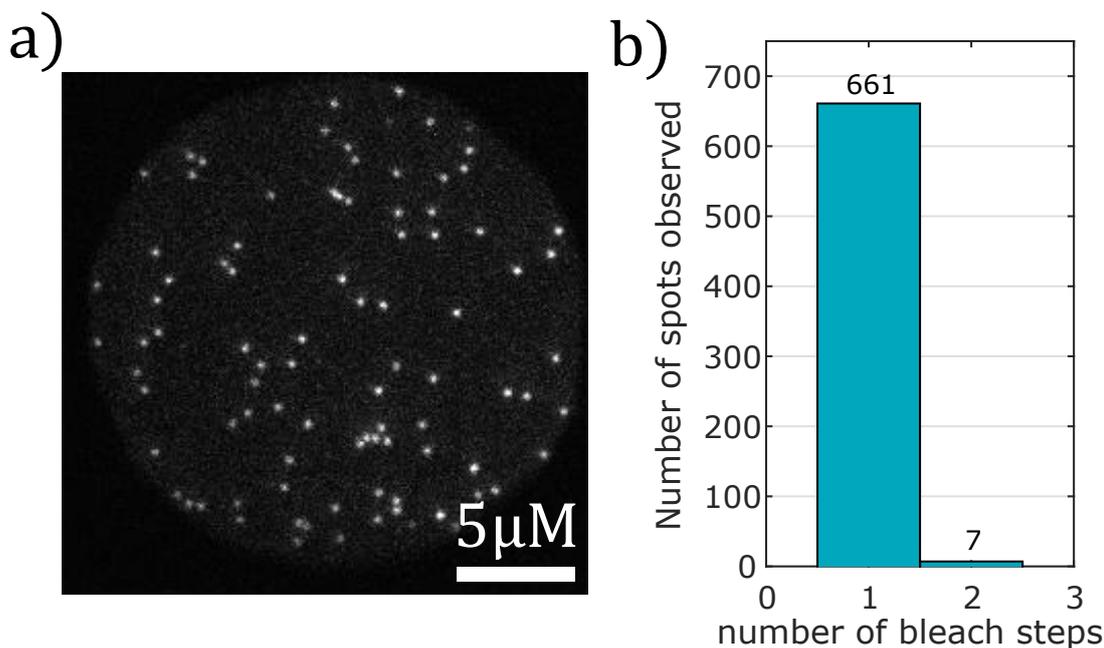


Figure 3.5: 1 x 10nt templates with AF647-DNA, as a control for template distribution. a) Typical field of view. b) Distribution of steps counted (analyzed by pipeline version 2-section 3.3.2)

number of steps in each trace. Additionally, as most videos suffered from a few pixels worth of drift from start to finish, either this drift must be corrected (to hold spots stationary) or the spots must be tracked from frame-to-frame. The pipeline for this entire process went through two major iterations.

In both cases, the image processing was performed in FIJI [385]; a distribution of ImageJ [386]. This was used to generate traces which were then viewed and further processed (if necessary) in MATLAB. Heavy use was made of the Sussex University Genome Damage and Stability Centre Single Molecule Light Microscopy (GDSC SMLM) suite of ImageJ plugins[387]².

3.3.1 Version 1 - Intensity in ROI

3.3.1.1 Identifying Spots

The user identifies the first frame of full illumination (Figure 3.6a); this frame will be used to identify spots. Initial spot detection is performed by the GDSC SMLM spot-finder plugin, configured to identify local maxima in a smoothed image. As this

²<https://github.com/aherbert/GDSC-SMLM>

tends to over-fit spots (Figure 3.6b), spots below a threshold Signal-to-Noise Ratio (SNR) are discarded. Following this, spots which have neighbours within a threshold distance are also discarded, after which the remaining spots are used to generate Regions of Interest (ROIs) from which intensity traces will eventually be taken. A user interface gives visual feedback on the selection of SNR threshold, neighbour distance threshold, and ROI size (Figure 3.6c). Some spots may be too close together for the spot finder to distinguish them, and thus not omitted by the neighbor distance threshold. Such multi-constituent spots are expected to be somewhat elliptical, and so spots are sorted by their eccentricity³ (Figure 3.6e), based on which the user can set a threshold value above which spots are omitted (Figure 3.6d).

3.3.1.2 Drift Correction

Intuitive drift-correction algorithms based on cross-correlation between frames tend to behave poorly with my data, due to anisotropic background intensity and bleaching of spots from one frame to the next. Instead, a sub-sample of frames are chosen, and subject to spot detection and filtering following the method described above. The positions of these spots are used to generate a binary image for each frame in the sub-sample, where spots are represented by circles of some defined radius (Figure 3.6g). For each pair of adjacent frames in the sub-sample, the cross-correlation of the two binary images is maximised by translating the latter image by a maximum of 1 pixel. Further iterative 1-pixel translations are allowed until a local cross-correlation maxima is found (although ideally the chosen sub-sampling is dense enough that only shifts of 0 pixels or 1 pixels are ever required.) These translations are then extrapolated back to the original video, and spots which drift out (or almost out) of the viewing area are discarded, leaving a final set of ROIs to be analysed. While arguably convoluted, this algorithm corrects drift reliably with little computational overhead, provided the sub-sampling density and binary spot radius are well-chosen (Figure 3.6h).

³Using the ImageJ moment calculator plugin <https://imagej.nih.gov/ij/plugins/moments.html>

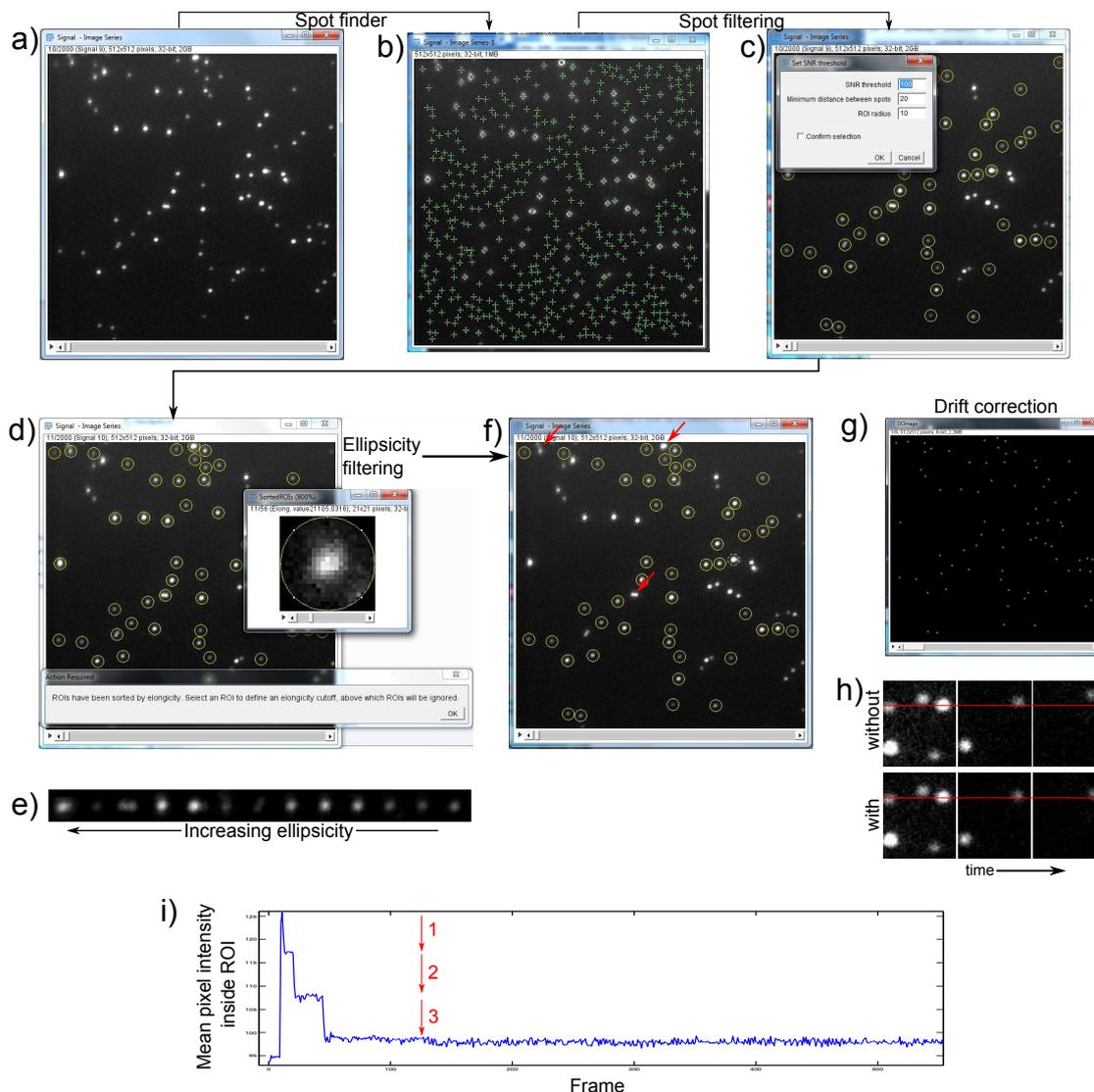


Figure 3.6: Counting pipeline, version 1. Example data corresponds to experiment described in section 4.1.1. a-e) steps in spot identification and filtering; see main text for description. f) Final ROI selection. Red arrows indicate a subset of spots filtered due to ellipsicity. g-h) Drift correction; see text for explanation i) Example intensity trace, with steps identified by eye.

3.3.1.3 Generation of intensity traces

Spot intensity (example trace in Figure Figure 3.6i) is simply defined as the mean pixel intensity within an ROI, with no attempt at background subtraction. This is recommended by [149], which reports that (for noisy *in vivo* data) any fitting of spot/background intensity adds noise and makes steps harder to distinguish. However, we will see below that our data is much more robust to fitting processes.

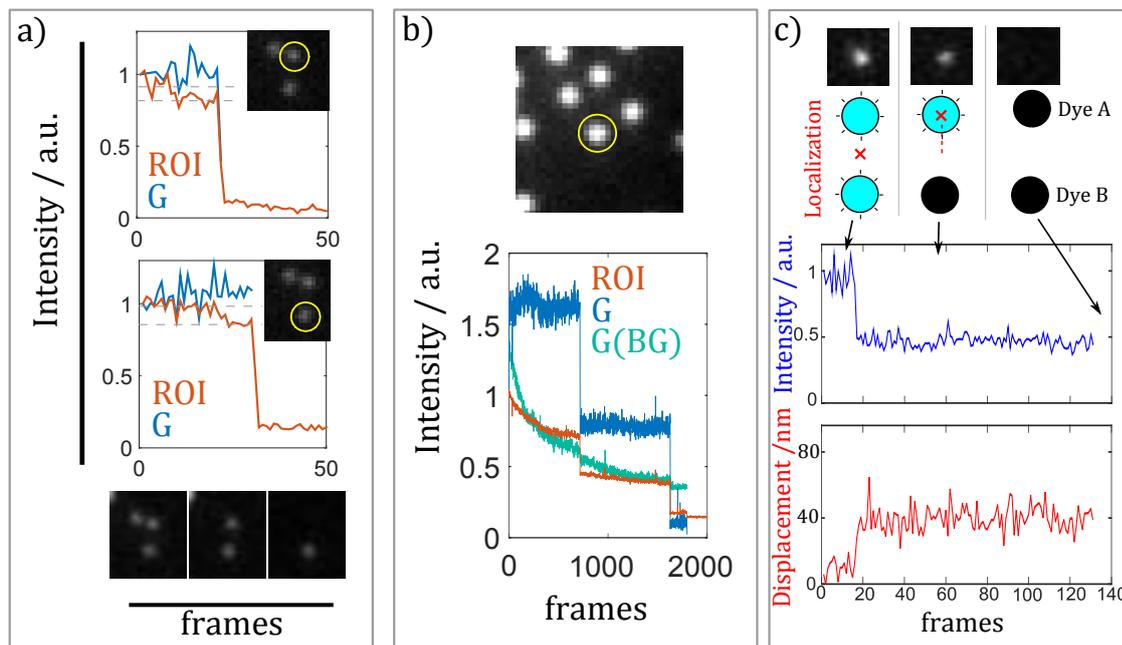


Figure 3.7: a) Brightness of adjacent spots, measured by intensity in ROI or Gaussian fitting (G) . b) Spot with significant background contribution from adjacent spots. BG indicates background component of the Gaussian fit. c) Dyes on separate templates distinguished by localization shift.

3.3.1.4 Counting

Steps are viewed with MATLAB, counted (see section 3.5) and recorded by hand.

3.3.2 Version 2 - Gaussian Fitting

The second version of the pipeline performs Gaussian fits to each spot, at every frame. This has a number of advantages: the implicit background fitting filters out contributions from nearby spots (Figure 3.7a), helping us retrieve data from dense surfaces, where neighbouring spots may contribute significant background intensity (Figure 3.7b). It also gives us localization data. When steps in localization correlate with bleaching steps, it tells us that we are looking at spots too distant to be on the same template (Figure 3.7c). This is much more sensitive than the ellipticity measure used in version 1 of the pipeline: assuming random spot distribution, we expect very few templates too close together to distinguish. Finally, if we have localization data for every frame, we can link up spots into traces, and avoid having to correct for drift.

3.3.2.1 Identifying and Fitting Spots

Spot-finding is now integrated with Gaussian fitting, performed frame-by-frame using the GDSC SMLM plugin suite. Local maxima are identified after difference-of-gaussian filtering, which extracts spatial information at frequencies between 0.75 and 1.25 times the expected spot (i.e. point spread function) width. Positions of local maxima are used to seed circular Gaussian fits in the unfiltered data, performed with a least squares estimator. Spots are discarded if the fit parameters do not converge to 7 significant figures within 30 iterations, if the localization shifts by more than 0.6 times the expected spot width, the fitted width differs from the expected width more than 1.4-fold, or if the signal (volume under the Gaussian) to noise (mean of squares of the residuals) drops below 10.

As described in [387], a multiple peak fit is performed if local maxima are in proximity, or if the Gaussian fit for a single maxima has residuals which are skewed beyond some threshold value. In the latter case, the multiple fit is only retained if the residuals indicate an improvement. A minimum distance between fitted peaks avoids multiple fitting of the same peak.

To check parameters before analysing an entire video (which may take hours), a subsample of frames are extracted and subjected to spot fitting by the above method, allowing the user to iterate parameters to minimise false fits.

3.3.2.2 Generation of intensity traces

A GDSC SMLM plugin is used to link together spots in adjacent frames which are within some threshold distance of each other. Gaps of one or two frames are allowed, accommodating the the rare cases where spot fitting has failed. The first frame of illumination is identified algorithmically, by the maximum increase in total-image brightness between subsequent frames, and traces starting after this frame (e.g. molecules adsorbing to the surface mid-observation) are discarded. As with spot identification and fitting, a user interface lets the user iterate trace-linking parameters. Once parameters are accepted, traces and the associated fit data are saved as an XML file.

3.3.2.3 Step Counting

A MATLAB script displays intensity and localization data, spot-by-spot, from the saved traces. Steps are counted by eye as before (section 3.5) but not recorded by hand. Rather, counts are recorded with each trace, and available for easy statistical analysis afterwards.

3.4 Anti-bleaching system

To maximise the number of photons emitted before bleaching, we used an anti-bleaching cocktail prescribed by [388]. Bleaching is primarily caused by reactions with O_2 , or by free radicals created by O_2 ; it is common therefore to use an oxygen-scavenging system to increase dye lifetimes [389, 390]. We use the popular combination of Glucose Oxidase and Catalase, which together digest D-glucose in the imaging buffer, removing O_2 in the process. While more effective than other oxygen-scavenging systems [388, 390], it also produces an acidic byproduct, leading to a pH drop over time, depending on buffering. It is reported that solution degassing and sealing of the reaction chamber slow this effect [378], and so imaging solutions (containing D-glucose) were always degassed just before use, and loaded into flow chambers immediately after the addition of Glucose Oxidase and Catalase. Open flow chambers were immediately sealed with VALAP (see section 3.1.2, Figure 3.3). Changes in pH were never directly characterized.

While O_2 promotes permanent bleaching, it also reduces the lifetime of triplet dark states: in the absence of O_2 , these dark states, though temporary, become much longer-lived and a source of “blinking” behaviour. A mixture of Trolox, cyclooctatetraene (COT) and 4-nitrobenzyl alcohol (NBA) efficiently compensates for this, reducing blinking [391]. The protocol of [388] was altered to reduce the amount of DMSO in the final imaging solution from 10% to 1.5%, given that DMSO destabilizes DNA hybridization⁴. (Full protocol available in appendix A.8.)

⁴<https://www.neb.com/protocols/2012/06/01/guidelines-for-pcr-optimization-with-phusion-high-fidelity-dna-polymerase>

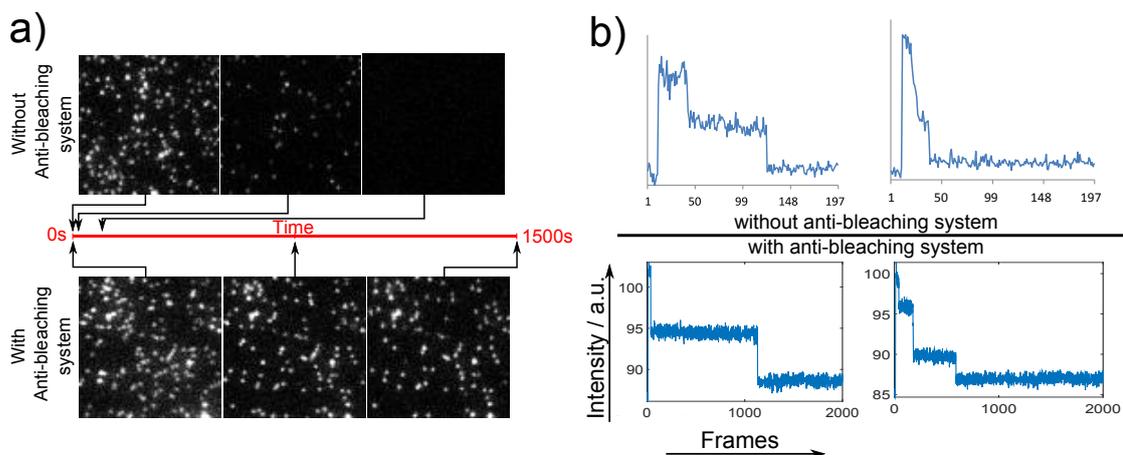


Figure 3.8: a) AF647-FliG stuck non-specifically to a surface. Same sample, different patches of slide, imaged with and without anti-bleaching system in otherwise identical conditions. b) Some example multi-step bleach traces taken with and without an anti-bleaching system (separate experiments).

Even on the first attempt, the performance of the anti-bleaching cocktail was striking (Figure 3.8). However, while dye lifetimes were always drastically increased, the extent of increase was not consistent from experiment to experiment. This variation was not quantified or explained.

3.5 Dye behaviour and counting rules

Occasional blinking (i.e. stepwise *increases* in intensity) and bleaching steps of unequal magnitude were observed for AF647 on both FliG (Figure 3.9a) and DNA (Figure 3.9b), perhaps due to polarization anisotropy. However, steps were clear enough to count directly, as illustrated in figure Figure 3.9a. This does not require equally sized steps, but it does require that each active dye only fluoresces at a single brightness level. Dyes switching between different brightness levels will lead to overcounting. AF647-DNA on 1-site templates do not appear to have multiple brightness states, bleaching overwhelmingly in single-steps (Section 2.1.2.2). This also appears to be true for single AF647-FliG stuck non-specifically to burnt coverslips (data not shown), but has not been carefully quantified for FliG attached to templates. Furthermore, given that the behaviour of the anti-bleaching system is not entirely consistent from experiment to experiment, it remains possible that

multiple brightness states are present in some experiments but not others. The effect of buffer conditions is also not known.

In all cases, counting was done by eye, although in future automatic counting would of course be preferable. For maximum consistency, steps of all sizes were counted, even when much smaller than average (e.g. the first step of Figure 3.9b). Such small steps could plausibly be caused by the sparse dim fluorescent contaminants seen on surfaces before incubation with AF647-labelled samples. However, on a few occasions where traces end in dim spots, they are both co-localized with the original structure and very long-lived (e.g. the first step of Figure 3.9c), leading me to suspect they really are AF647, rather than some contaminant.

Traces were discarded when a jump in localization implied dyes were not on the same template (section 3.3.2) or when steps were unclear (example given in Figure 3.9d). As an example, $\sim 15\%$ of traces from the experiment described in section 4.1.4 (FliG on 3 x 20nt template) were discarded in this manner.

3.5.1 Counting large numbers of dyes

All of the fluorescence experiments described in chapter 4 involve templates with 5 or fewer binding sites, and I will argue that the stoichiometry of such small structures may be better measured with gel electrophoresis. However, our aspiration is to eventually template complete rings of ~ 20 -40 FliGs, which we expect to be less easy to count by electrophoresis. To understand the feasibility of counting them by fluorescence, I measured control structures with larger numbers of dyes.

3.5.1.1 12-dye DNA structures

A collaborator, Evan Spruijt (Hagen Bayley lab, Oxford), provided structures⁵ designed to have 12 AF647-DNA tags, and 2 biotins for coverslip surface-attachment. Like my own AF647-DNA constructs, a TTT linker separated the dye from hybridized regions, although the AF-647 labelling was performed in-house, rather than provided commercially like my own. The resulting traces are instructive; while

⁵The details of which I will keep confidential

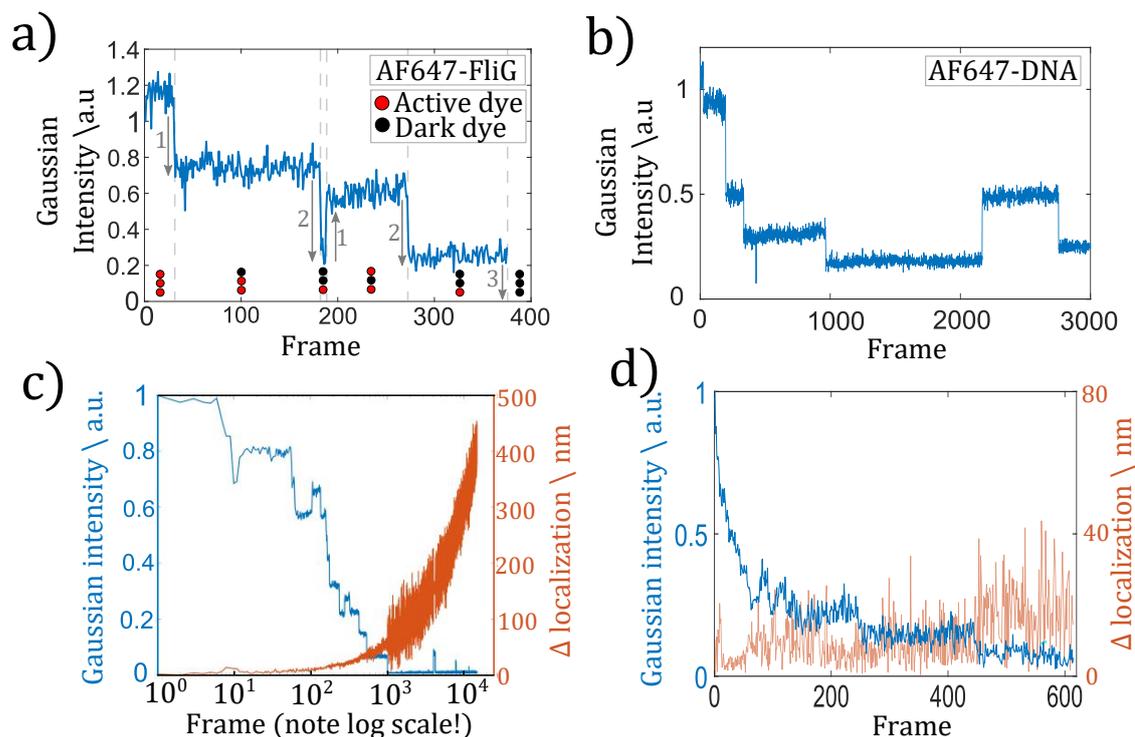


Figure 3.9: a) Bleach trace of AF647-FliG on template, showing uneven bleach steps and blinking. (Experiment described in section 4.1.3). Counting logic annotated in grey; black and red circles show interpretation of underlying dye states. b) Bleach trace of AF647-DNA on template, showing uneven bleach steps and blinking (Experiment described in section 4.1.2). c) AF647-DNA bleach trace from large structure (section 3.5.1.1), showing very dim but very longlived fluorescence at end of trace. d) AF647-FliG on 3-site template (section 4.1.4); example of trace deemed uncountable and discarded.

usable, they often contain segments of questionable clarity (Figure 3.10a.iii) and bleach events so rapid they are barely resolved (Figure 3.10a.i, a.ii). However, given the excessive signal to noise ratio, acquisition rates could easily be increased to resolve these events. Additionally, more complex analysis could alleviate the need to resolve every step.

3.5.1.2 DNA tiles

A very quick-and-dirty attempt was made to bind 6 AF647-DNA tagged 5 x 10nt templates to a biotinylated DNA tile, as described in section 2.1.3. These gave much less clean traces than the 12-dye structures (Figure 3.10a,b). The best examples clearly had signatures of bleaching events throughout (Figure 3.10a), but in general non-stewpise intensity fluctuations obfuscated counting. The source of

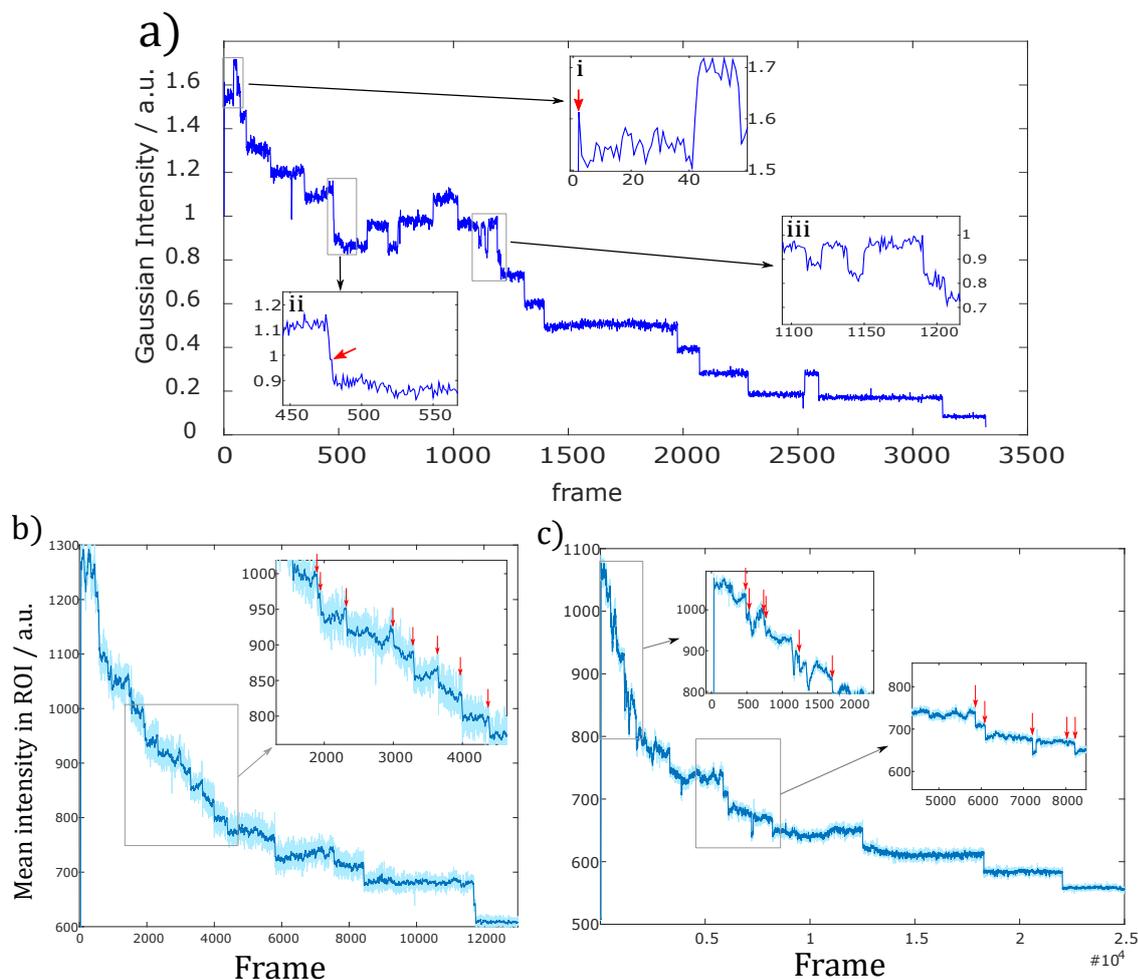


Figure 3.10: a) Bleaching trace from 12-dye DNA structure. Red arrows mark bleaching events that are only barely resolved. b,c) Example bleaching traces from DNA tiles. Cyan = raw data. Blue = smoothed with a 10-frame rolling average. Red arrows mark plausible bleaching events.

these fluctuations is unclear. In addition, laser power limited acquisition times to >10 minutes, making for very inefficient data collection. It is clear that additional work is needed to image these structures.

3.6 Conclusion

To conclude; microscopy techniques developed here are adequate for the counting experiments we envisioned in section 1.5.1, and have already produced useful data (sections 4.1.1 and 4.1.3). Surface passivation is probably not sufficient to observe assembly in real time at physiological concentrations, but improved

surfaces could enable such observations in future. Likewise, while objects with more ~ 20 -30 dyes have not been counted successfully, it is reasonable to think this is possible with further work.

However, although small-number counting has been successful and reliable, the usefulness of these experiments is limited by low throughput, and biochemical problems: the stability of conjugates, uncertainties in dye-labelling and so on. Could a different counting technique overcome these limitations?

3.6.1 Gel Electrophoresis: an alternative?

At the start of the project, we inherited reports that NTA-histag binding (then the preferred conjugation method) was not compatible with native gel electrophoresis. This was consistent with my own limited attempts to visualise FliG binding to NTA DNA, while still struggling with NTA modification protocols. Thus, while it was used to confirm the assembly of DNA templates (section 2.1), it was not in consideration as a FliG-counting tool. However, in the process of implementing guided conjugation techniques (section 2.3.3.1), Joel accidentally discovered that such links are in fact robust on native gels. The next chapter will include examples of counting experiments on gels; these can be loaded with higher-concentration samples (and are thus less sensitive to unstable conjugation), can easily test different conditions in parallel, and are not sensitive to dye-labelling efficiency. There is an argument, then, that fluorescence methods are superfluous until we need to look at dynamics or larger structures. We will return to this argument briefly in section 4.2.

4

Results & Future Directions

Contents

4.1	Templated assembly experiments	136
4.1.1	FliG on 10nt templates tested by fluorescence counting	136
4.1.2	10nt DNA templates with fluorescent linkers	137
4.1.3	FliG on 20nt templates tested by fluorescence counting	139
4.1.4	FliG on 20nt templates tested by native PAGE	141
4.1.5	FliG on 10nt templates tested by native PAGE	142
4.1.6	Conclusions	144
4.2	Coda	144

The experiments proposed in section 1.5.1 require fine control over FliG binding to a template, and accurate counting of FliG stoichiometry. The most simple demonstration of this ability is to reliably measure the FliG number designed into a template. What follows is the essential narrative of progress towards that goal, described by a few key microscopy (and later PAGE) experiments, with reference to the experiments of earlier chapters where relevant.

4.1 Templated assembly experiments

4.1.1 FliG on 10nt templates tested by fluorescence counting

As an early test of FliG stability on 5 x 10nt DNA templates bound via tris-NTA DNA linkers (Figure 4.1a), 614nM template was incubated in room temperature MB for 1 hour with 614nM biotin anchor, 4.3 μ M (1.4-fold excess) of HPLC purified NTA-DNA linker (corresponding to peak E in Figure 2.16a) and 500 μ M NiSO₄. This was then buffer exchanged twice into MB with Micro Bio-spin 6 columns to remove excess nickel, and then mixed with 900nM FliG-AF647 for an expected 10:7:1 FliG:linker:template ratio (template at 90nM). This was incubated in room temperature MB for 2h30 before dilution to 50pM template and incubated for 5 minutes with a PEG-biotin-streptavidin surface in an open channel, before washing away unbound molecules and sealing with anti-bleaching mixture in MB. 17 videos were recorded in TIRF over a \sim 2 hour period, imaging fresh areas of the coverslip each time. Videos were analysed as described in section 3.3.1.

The distribution of counts in this experiment was disappointing; consistent with only \sim 1/3 of template sites being occupied (Figure 4.1b). Furthermore, this distribution was reasonably stable with time (Figure 4.1c), suggesting that DNA-DNA and DNA-FliG dissociation rates are slow on \sim hour timescales. Thus, the complexes measured were probably not much different than those at the end of the incubation period (pre-dilution), unless off-rates between template sites are heterogeneous (e.g. because the NTA modifications are not uniform, as suggested in section 2.3.2.3).

The work required per experiment meant that optimizing incubation conditions would be inefficient without more idea as to what the weak link was; FliG labelling, NTA-DNA conjugation, or DNA-DNA hybridization. Though not conclusive, biochemical tests seemed to indicate that FliG labelling was well above \sim 33% (section 2.2.3); more conclusive tests ideally require covalent conjugate (section 2.2.3.4). DNA-DNA hybridization then seemed like the next easiest thing to test.

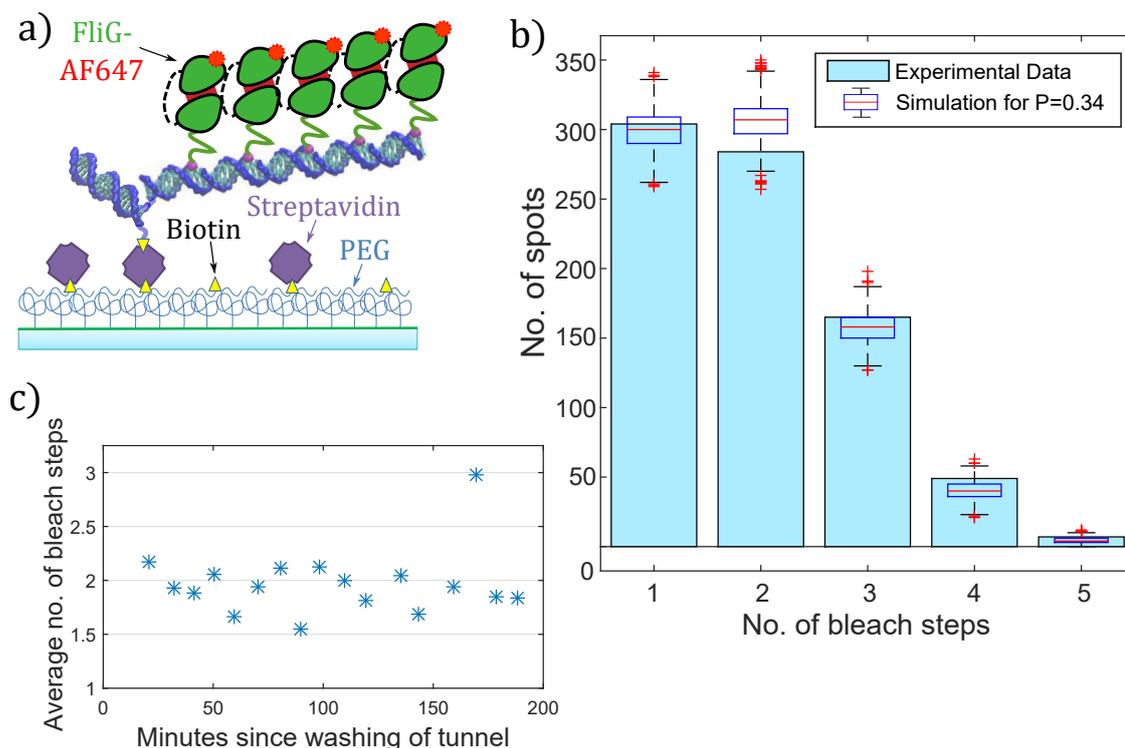


Figure 4.1: FliG on 5 x 10nt template. a) Illustration of template complexes on surfaces. b) Histogram of 1-5 bleach steps over all videos. Box plots show distribution over 1000 simulated histograms, in which each template site is 34% likely to have an associated count, and the total number of observable spots is constrained to equal that in the experiment. c) Variation of average number of bleach steps with time.

PAGE had verified assembly at high ($\sim\mu\text{M}$) concentrations, but could not test stability at low concentrations.

4.1.2 10nt DNA templates with fluorescent linkers

$1\mu\text{M}$ template was incubated for days in room temperature MB with $1\mu\text{M}$ biotin anchor and $10\mu\text{M}$ DNA linker labelled with AF647 via a TTT spacer. This was then diluted to low concentration (100pM template) and incubated with a PEG-biotin-streptavidin surface in a closed channel for ~ 10 seconds, before washing away unbound molecules and sealing with anti-bleaching mixture in MB. 7 different videos were recorded over a ~ 2 hour period, washing the channel and imaging fresh areas of surface each time. Traces were generated and counted as described in section 3.3.2.

The distribution of counts summed over all these videos shows a majority of 5-mers, consistent with $\sim 89\%$ probability of each template site being occupied

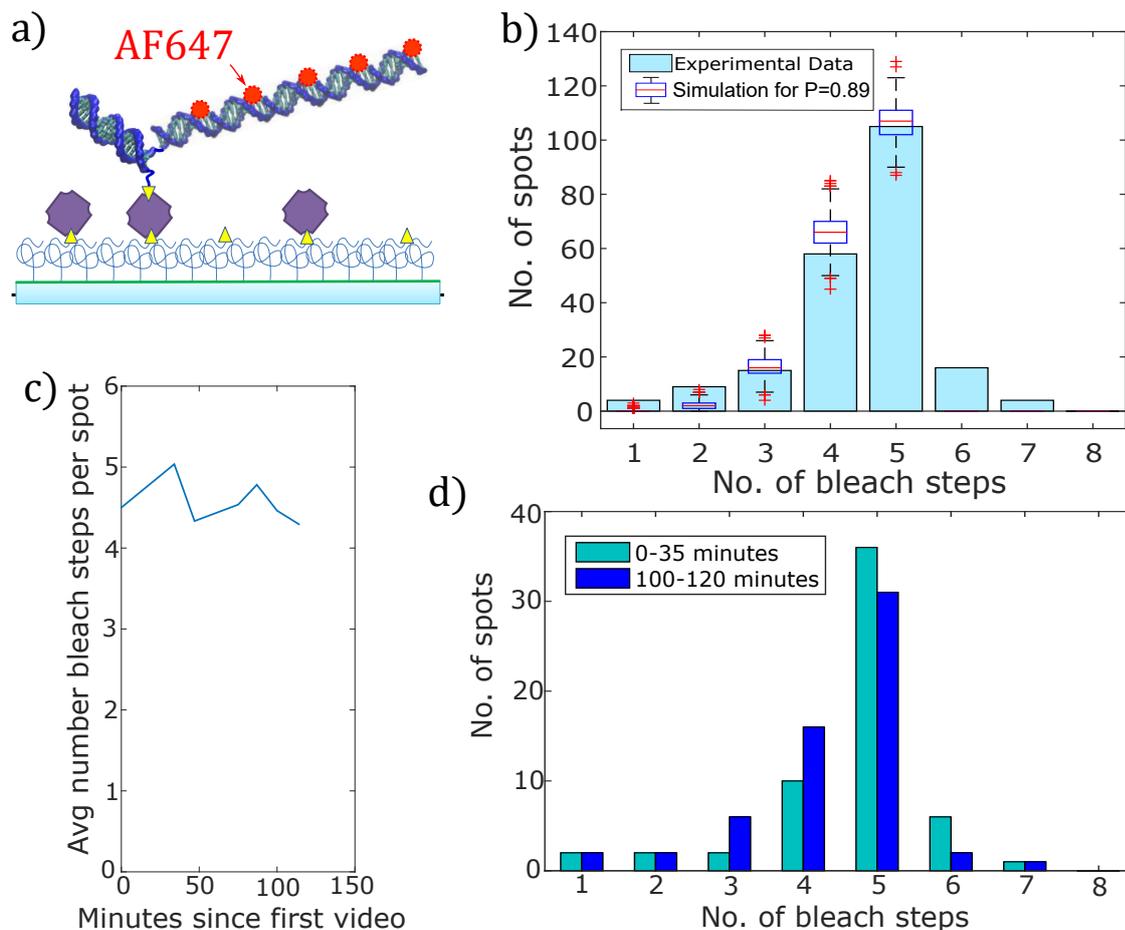


Figure 4.2: Fluorescent DNA linkers on 5 x 10nt template. a) Illustration of template complexes on surfaces. b) Histogram of 1-5 bleach steps over all videos. Box plots show distribution over 1000 simulated histograms, in which each template site is 89% likely to have an associated count, and the total number of observable spots is constrained to equal that in the experiment. c) Variation of average number of bleach steps with time. d) Bleach step histograms for the first two and last two videos.

with a fluorescent linker (Figure 4.2b). The apparently sub-stoichiometric filling may represent a population of non-fluorescent linkers, or alternatively incomplete assembly: later, more careful PAGE gels demonstrated that complete assembly required MgCl_2 (Section 2.1.1.1), which was absent in this experiment. Observation of 6-mers and 7-mers may indicate noise in the counting process (i.e. occasional over-counting), but not to an experiment-limiting degree.

Despite the absence of MgCl_2 and the near-zero concentration of fluorescent linkers in the channel, template structures were reasonably stable over the 130m observation period (Figure 4.2c), although a small degree of dissociation was visible

in the count histograms (Figure 4.2d). This initially suggested that DNA-DNA stability was sufficient (or at least not limiting) for the FliG assembly experiments, and made it an unlikely explanation for the poor FliG assembly measurements. However, later bulk experiments (section 2.1.1.3) showed a significant difference in melting temperature between 10nt unmodified DNA and NTA-modified DNA, calling the relevance of this control into question. Nevertheless, it was a useful test for the accuracy of the counting system.

4.1.3 FliG on 20nt templates tested by fluorescence counting

At this point, we suspected the DNA-FliG link was the weak point. We were already producing 20nt tris-NTA strands for guided conjugation trials (section 2.3.3.1), and in a bid to improve their performance tried synthesising pentakis-NTA strands (section 2.3.2.5). 3 x 20nt templates (section 2.1.2, Figure 4.3a) provided an ideal platform to compare FliG binding between the two, without concerns about DNA stability.

1.82 μ L template was mixed 1 : 1 : 3.8 : 4 with with biotinylated anchor, unpurified NTA linker and AF647-FliG respectively, and incubated overnight at 4°C in MB with 20 μ M NiCl₂. This was then diluted to 100pM (template concentration), and immediately incubated with a PEG-biotin-streptavidin surface in a closed channel for \sim 90 seconds before washing and sealing with anti-bleaching mixture in MB with 20 μ M NiCl₂. For both NTA varieties, \sim 40 videos were taken over >60 minutes, moving to a fresh patch of surface each time. Traces were generated and counted as described in section 3.3.2.

While measured template-filling was disappointingly low in both cases, the pentakis-NTA gave noticeably better filling and better stability (Figure 4.3b,c). Embarrassingly, I only realised sometime afterwards that the NiCl₂ was only in 3:1 excess over NTA linkers during incubation, and thus not sufficient to bind all NTA groups on the pentakis linkers. A later repeat of the same incubation conditions was analysed with native PAGE, and also had sub-stoichiometric filling

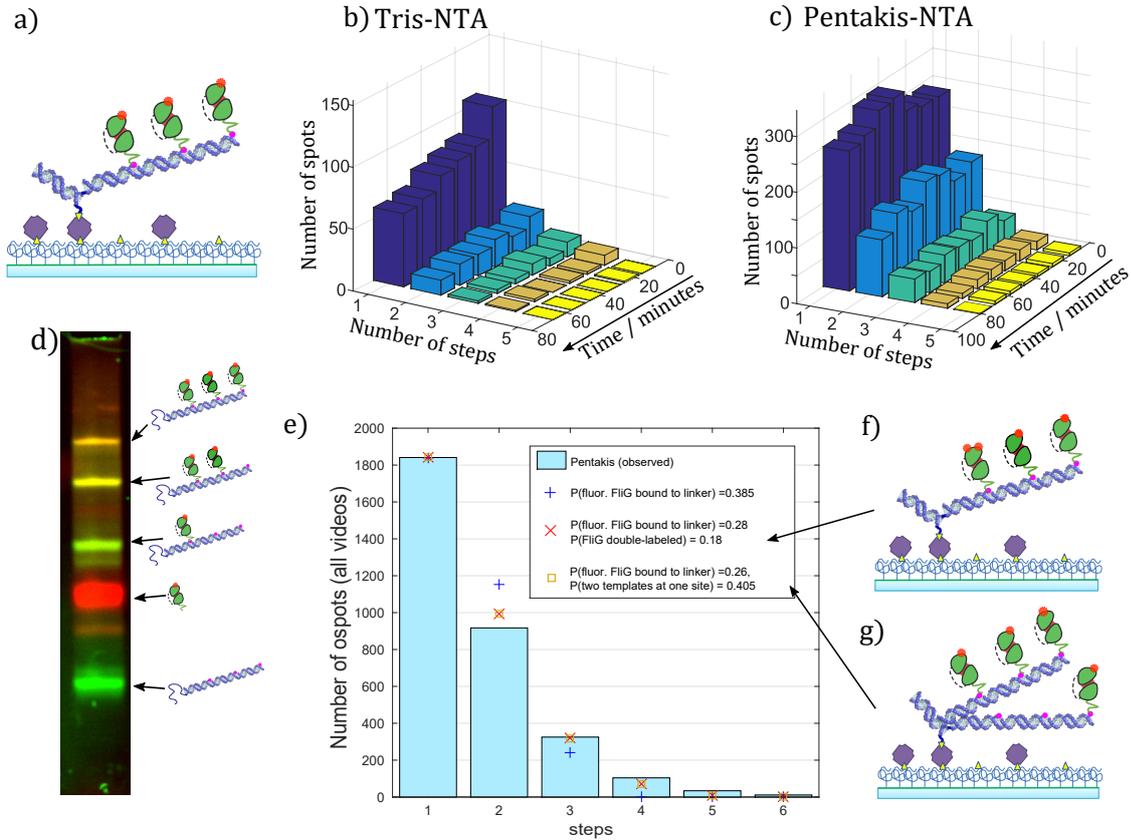


Figure 4.3: FliG on 3x20nt templates, bound via tris-NTA or pentakis-NTA linkers. a) Illustration of template complex on surface. b,c) Variation of bleach step distributions with time. (Each time slice corresponds to 6 sequential videos). d) A later sample prepared with the same incubation conditions, and visualized with native PAGE. Band assignment will be justified in section 4.1.4. e) Bleach step distribution for the entire pentakis-NTA dataset, with simulated distributions for simple binding, simple binding with FliG double-labelling and simple binding with template dimers; illustrated in a), f) and g) respectively. Parameters for all the simulated distributions were fitted to the data by minimization of residuals. Near-convergence of double-labelling and template-dimerization fits is not an error, and presumably reflects some underlying mathematical equivalence between the two models.

(Figure 4.3d). Thus, it is plausible that a repeat fluorescence experiment with more appropriate incubation conditions might measure the designed stoichiometry; this might have been attempted, if not for the growing fidelity of our native PAGE gels (see next section).

Despite the mostly sub-stoichiometric filling, a significant number of spots displayed more than 3 bleach steps (Figure 4.3e). Furthermore, the distribution of spots with 1 to 3 steps was more biased towards higher counts than expected

from a binomial distribution (i.e. fixed probability of each template site hosting a countable FliG). A number of explanations were considered.

The first possibility is that FliGs were bound to template sites with a fixed probability, but that some percentage of FliGs produced 2 bleach steps, either through double-labelling (Figure 4.3f) or multi-level dye brightness. This model fits the data well (Figure 4.3e), but requires $\sim 18\%$ of FliG to display double steps; far in excess of the measured $\sim 2\%$ non-specific AF647 labelling yield (section 2.2.3.1), and inconsistent with the lack of multimeter bleach steps for AF647 attached to DNA on templates (section 3.2.2) or FliG adsorbed non-specifically on an unfunctionalized surface (data not shown).

A second possibility is that FliGs are bound to template sites with a fixed probability, but some fraction of templates are somehow dimerized, so we are counting 6 sites rather than 3 (Figure 4.3g); a faint signature of this had already been seen on PAGE (section 2.1.2.2). However, a best fit to the fluorescence data required $\sim 40\%$ of templates to be dimerized in this way, far above the level implied by PAGE. A hypothesis that multiple templates might bind to the same streptavidin was also ruled out (section 3.2.2).

This leaves a number of untested possibilities: either AF647 on template-bound FliG (but not surface-adsorbed FliG or template-bound DNA) often bleaches in multiple steps, *or* we are seeing some FliG-mediated interaction which either brings two templates together, binds multiple FliGs to a single template site, or binds extra FliGs in the gap between templated FliGs. While it is tempting to imagine this is a sign of physiological gap-filling, repeated observation of histag-histag interactions (section 2.3.2.6) makes non-physiological interactions very plausible.

4.1.4 FliG on 20nt templates tested by native PAGE

At about this time, complexes were being visualized on native PAGE with increasing clarity. Titration of FliG concentration against template-linker complexes reliably produced 3 major bands of intensity, interpreted as 1 2 and 3 FliG's (Figure 4.4a). Despite the unpurified pentakis-NTA, a majority filling of the 3-mer band was

possible, both in the presence of Nickel (Figure 4.4b) and when nickel was incubated with pentakis-NTA oligos then removed before template assembly and FliG binding (Figure 4.4c). This demonstrates both controlled assembly and counting.

Gels often (Figure 4.4b) showed dim but clear bands running slower than the 3-FliG band. These were seen on templates with both new and old (more dimerization-prone) anchor sequences (section 2.1.2.2) but had increased intensity and a different band structure for the old anchor, suggesting that DNA-mediated template dimerization might be responsible. They were also not seen in the absence of nickel (Figure 4.4c) or FliG, suggesting histag-histag interaction as an alternative explanation (section 2.3.2.6). However, their presence or non-presence was extremely inconsistent; at the extreme, in a single non-reproducible gel produced by Joel Spratt, the bulk of templated FliG intensity was in these bands (Figure 4.4d). Background smearing on gels was too high for reliable densitometry comparison between AF647 and DNA signal from the various bands, so it is unclear whether they correspond to multiple templates, or perhaps FliG gap-filling. We hoped that subsequent experiments on 10nt templates would clarify matters.

4.1.5 FliG on 10nt templates tested by native PAGE

Disappointingly, assembly on 5 x 10nt templates was repeatedly unclear. Not only were there more than 5 major bands, which confused assignment of stoichiometries (Figure 4.5a.i), but assembly never converged to a single band. Controls ran in parallel suggested the DNA linkers should be filling the templates (Figure 4.5a.ii).

To minimise FliG histag-histag interactions which might compete for template-binding (section 2.3.2.6), NiCl_2 was incubated with linkers before removal by buffer exchange. As linkers were so small, significant amounts were lost during buffer exchange, necessitating re-measurement of the concentration before incubation with the templates, perhaps allowing time for nickel to dissociate from NTA groups: presence of the buffer-exchanged linker seemed to slightly promote FliG-FliG multimerization (Figure 4.5b). Furthermore, gels followed a similar (disappointing) pattern of incomplete assembly. Note that in the last lanes of figures 4.5a.i and

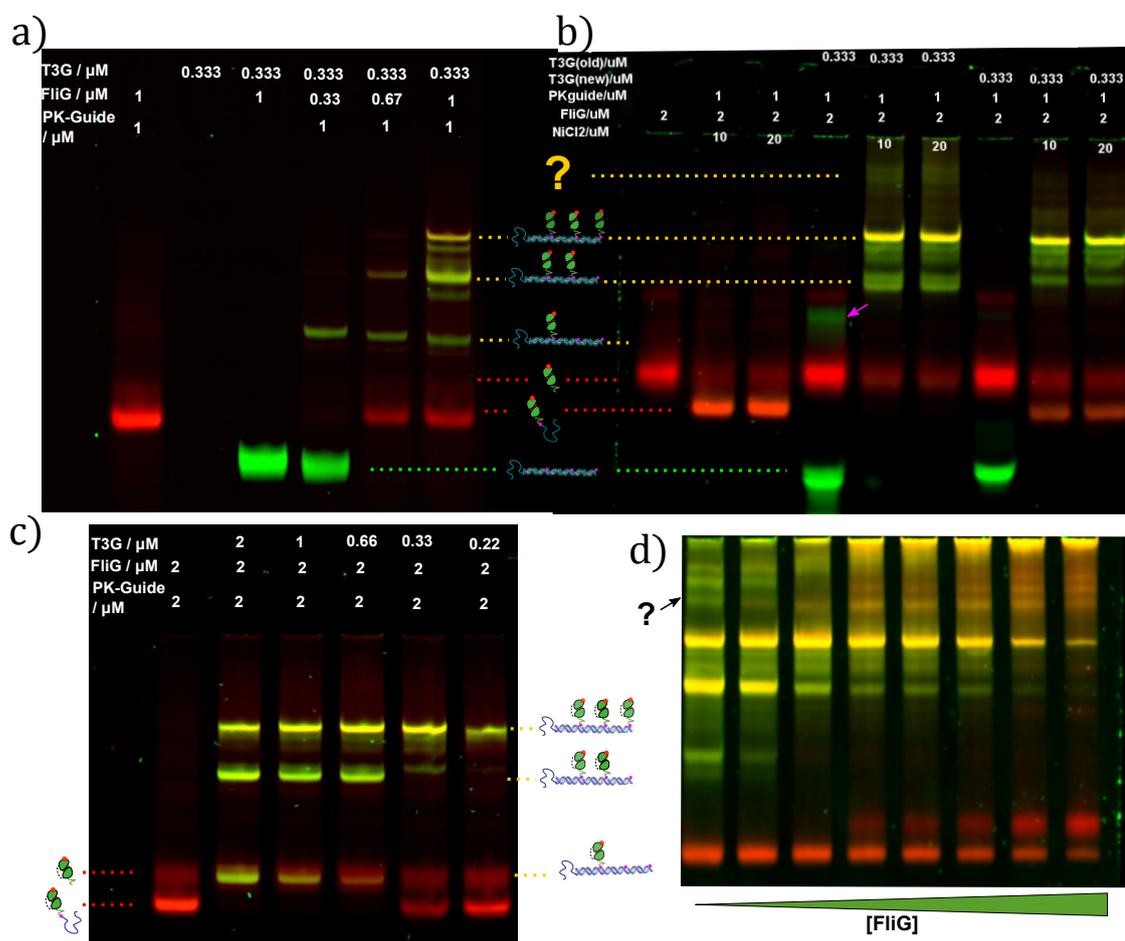


Figure 4.4: FliG on 3 x 20nt templates bound via pentakis-NTA linkers (“PK-Guide”); native PAGE. All gels: Green = SYBR gold DNA stain. Red = AF647. a) Titration of FliG against template complexes, visualised after 2 day incubation at room temperature in EPPS RB with 10 μM NiCl₂. b) Fully filled templates, with new and old anchor sequences (see section 2.1.2.2) and different NiCl₂ concentrations, following 2h30 incubation in EPPS RB. Magenta arrow labels suspected template dimers. c) Titration of template against FliG and linkers. Linkers were incubated with 1mM NiCl₂ for 50m at room temperature before removing nickel with a Micro Bio Spin 6 column, and incubating with FliG and templates for 45m in room temperature EPPRS RB. Stated linker concentration does not account for sample loss during buffer exchange/nickel removal. d) FliG titration against templates, performed entirely by Joel Spratt. All gels: 6% 19:1 PAGE in TAE, ran at 200V for 45m at 4°C. 1 μL reaction per lane in ficoll LB.

4.5b, there is a clear excess of FliG + linker, despite the presence (presumably) of unhybridized template sites. It is possible that there is an energetic penalty for packing FliG side-by-side on the template which the hybridization energy of the 10nt linker cannot compensate for. The 3 x 20nt template on the other hand has greater hybridization energies and less tight packing.

4.1.6 Conclusions

There are two main conclusions of the experiments above. Firstly, while the pentakis-NTA-histag link is stronger than the tris-NTA-histag link (section 4.1.3), and capable of mediating complete assembly in some cases (section 4.1.4), the presence histag-histag interactions will unavoidably confuse our search for physiological interactions. In addition, we know that NTA modification destabilizes linker-template hybridization (section 2.1.1.3), which may be limiting in assembly on closely-packed templates (section 4.1.5). Thus, there remains a strong case for alternate covalent conjugation strategies (section 2.3.3), which could additionally simplify the still-problematic quantification of dye labelling (section 2.2.3.4).

Secondly, while the microscopy tools have been developed to a point that clear and useful counting experiments can be performed on short templates (section 4.1.3), they have been superseded by native PAGE, which has much higher throughput. A comprehensive set of controls should resolve band classification on 5 x 10nt templates, allowing counting. Microscopy tools may still be of use at a later stage, for measuring dynamics or counting larger structures (section 3.5.1).

4.2 Coda

During the writing of this thesis, Joel Spratt has managed to purify a sufficient quantity of covalent conjugates (section 2.3.3.2) to trial assembly experiments on PAGE, with a new set of stoichiometry control templates (Figure 4.6). This early trial is not perfect; a FliG dimer mediated by a histag-DNA interaction confuses results a little, and untagged FliG is also present in solution. Nevertheless, a variety of control templates give dominant bands, finally allowing interpretation

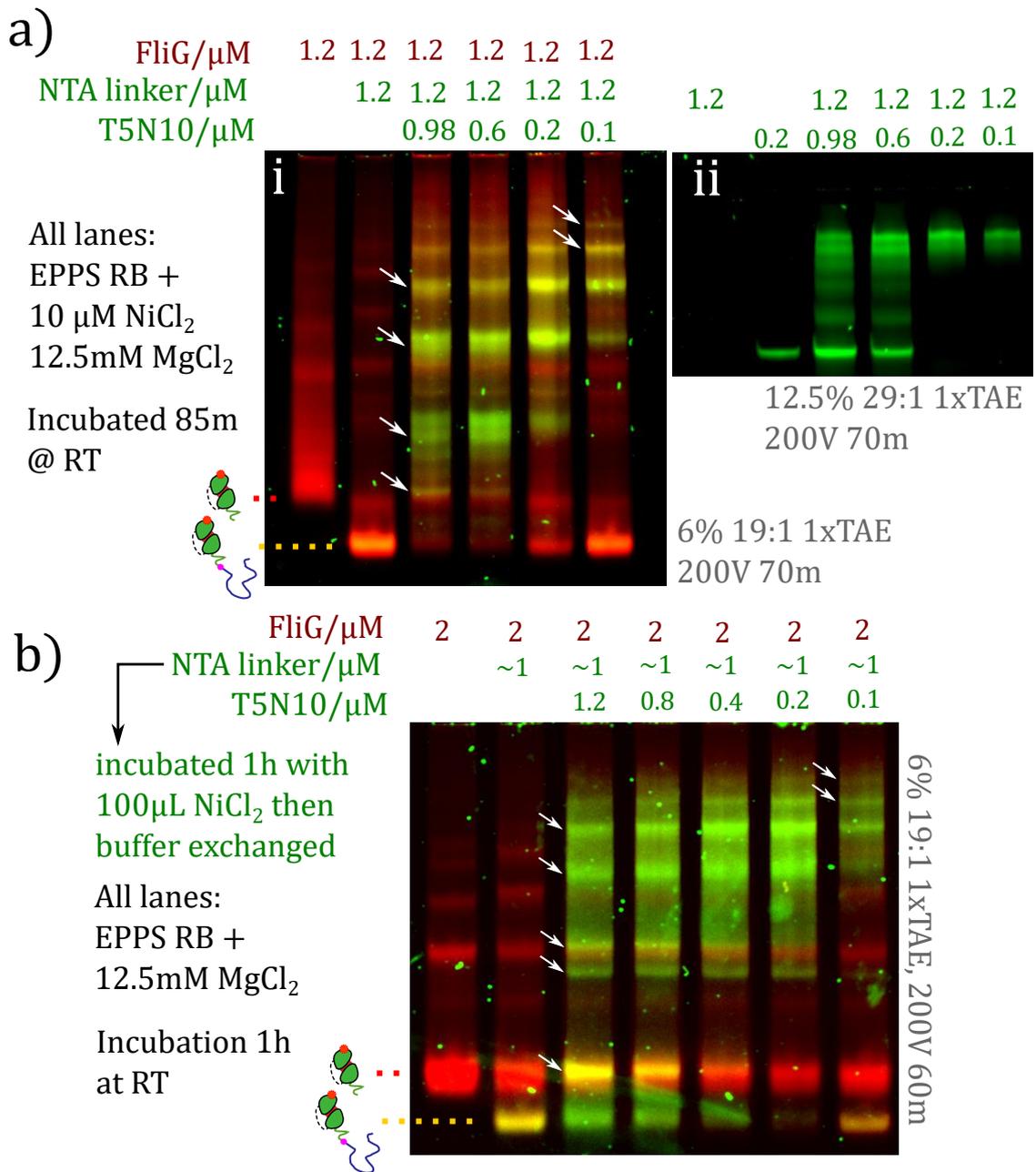


Figure 4.5: a) (i) Assembly of AF647-FliG on 5 x 10nt templates via pentakis-NTA linkers. White arrows indicate main bands with mixed DNA and AF647 signals. (ii) Simultaneous gel to verify correct assembly of DNA linkers on template. b) Assembly of AF647-FliG on 5 x 10nt templates in the absence of nickel.

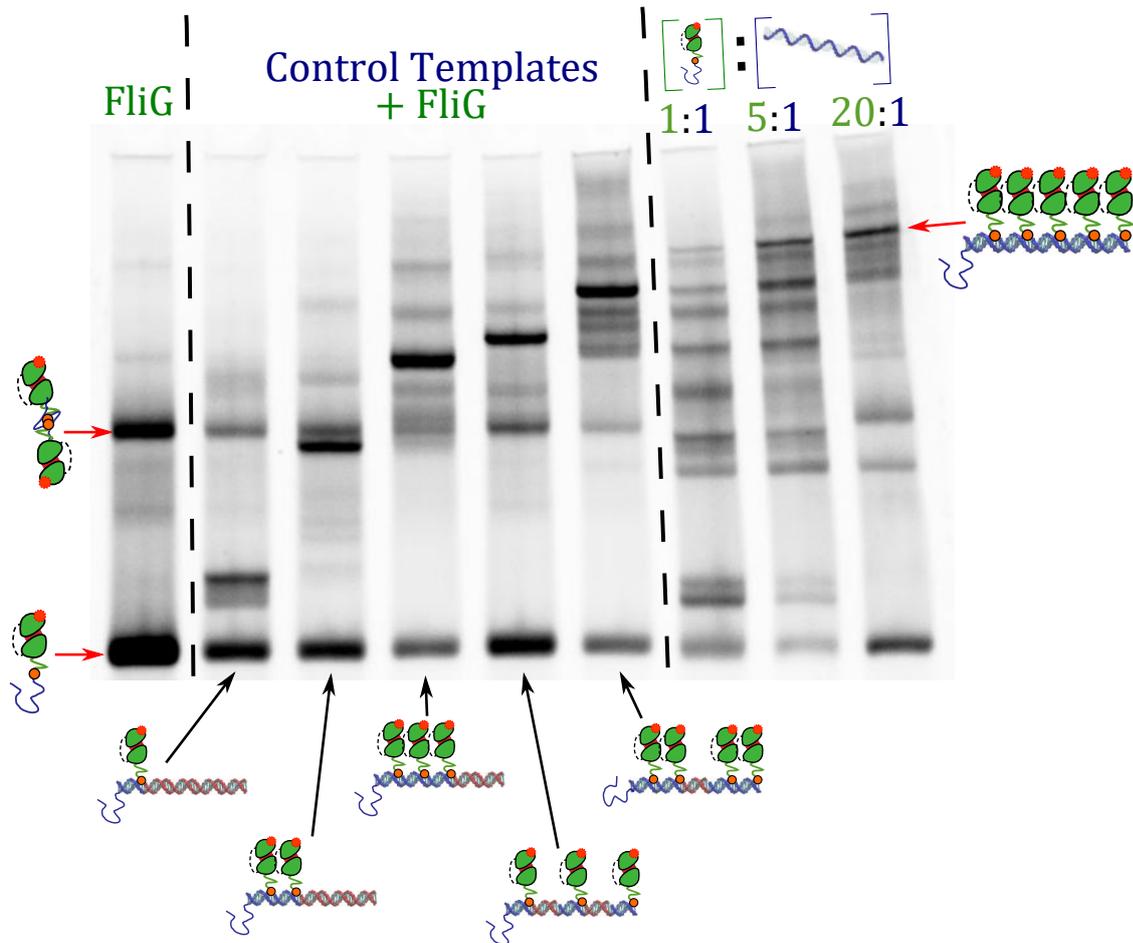


Figure 4.6: Joel Spratt’s work. Covalent FliG conjugate with various templates, imaged by AF647 fluorescence.

of stoichiometry. Furthermore, assembly of excess FliG on a 5-site template tends towards one dominant band, consistent with full binding to the template. This suggests that we finally have the control and the counting ability we aspired to in section 1.5.1.

The immediate next steps are to solve the dimerization issue through complete histag cleavage, and to confirm the robustness and repeatability of assembly. With this done, we can begin searching for conditions (if any) in which FliG gap-filling occurs. As no clear gap-filling is seen in the 4-site template of Figure 4.6, this may require carefully tuned conditions as suggested in sections 1.2.0.4 and 2.2.4.2. It may also require FliM/FliN, which we will therefore purify soon, following a published protocol [125]. However, we expect that FliM/FliN will confuse the

identification of gel bands, necessitating a return to counting via bleaching steps, which in turn requires a dye-labelled DNA tag (section 2.3.3.2) and removal of untagged FliG (section 2.3.3.3). Although this is not a trivial amount of work, we are confident that a tractable experimental system is now within reach. Thus, we have made considerable progress towards being able to test the FliG domain-swap polymerization model *in vitro* as outlined in Figure 1.19, and are well on the way towards the templated assembly of complete C-rings *in vitro*.

Appendices

A

Appendix

Contents

A.1	Buffers used	151
A.2	DNA sequences	152
A.3	Plasmid amplification	153
A.4	Protein expression, purification and dye labelling . . .	154
	A.4.1 Expression and lysis	154
	A.4.2 Purification and labelling	155
A.5	BCA assay	156
A.6	Proteinase K digestion to measure dye labelling	156
A.7	Surface Modification	157
	A.7.1 BSA surface	157
	A.7.2 PLL-PEG surface	157
	A.7.3 PEG vectabond surface	157
A.8	Anti-bleaching system	159
A.9	FliG assembly Gels	160

A.1 Buffers used

MB = 20mM Tris-HCl pH 7.5, 150mM NaCl
IB2 = 20mM TRIS-acetate pH 7.6,
300mM Potassium Glutamate, 10mM Magnesium Acetate
SM2 = 20mM Tris-HCl pH 7.5, 150mM NaCl, 0.2mg/ml BSA, 0.2mg/ml streptavidin

A.2 DNA sequences

BA (no biotin)	/5BiosG/TTTTTCAGGACAGAGACAATAAC
BA (biotin) 2	/5BiosG/TTTTTCAGGACAGAGACAATAAC
N10	GGCAGGACGG
N10 (pentakis)	AminolinkC6 - 4xAmino C6 dT - GGCAGGACGG (biomers)
N10 (tris)	/5UniAmM//iUniAmM//iUniAmM/GGCAGGACGG
N10 (AF647)	GGCAGGACGG/3AlexF647N/
T1N10 v1	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCTTATCT- CACCTTATCTCAC
T3N10 v1	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCCGTCCT- GCCCCGTCCTGCC
T3G2N10	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCTTATCT- CACCCGTCCTGCCCTTATCTCACCCGTCCTGCC
T4G1N10	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCCGTCCT- GCCCTTATCTCACCCGTCCTGCCCCGTCCTGCC
T5N10	CGCCTGAAGCTCTGGGCGTTTCCGTCCTGCCCCGTCCT- GCCCCGTCCTGCCCCGTCCTGCCCCGTCCTGCC
dN10 v1	GTGAGATAAG
dN10 v2	GGCGAGTAGG
dN10 v3	GGGAAAGGCG
dN20 v1	CGGGTAGTAAAGTGGAATG
T3G2N10 v3	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCCGCTTTC- CCCCGTCCTGCCCCGCTTTCCTCCCGTCCTGCC
T4G1N10 v3	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCCGTCCT- GCCCCGCTTTCCTCCCGTCCTGCCCCGTCCTGCC
T4GendN10 v3	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCC CCGTCCT- GCCCCGTCCTGCCCCGTCCTGC CCGCCTTTCCT
T3G2N10v2	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCC- TACTCGCCCCGTCCTGCCCCTACTCGCCCCGTCCTGCC
T4G1N10v2	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCCGTCCT- GCCCCTACTCGCCCCGTCCTGCCCCGTCCTGCC
T4GendN10v2	GTTATTGTCTCTGTCCTGTTTCCGTCCTGCCCCGTCCT- GCCCCGTCCTGCCCCGTCCTGCCCCTACTCGCC
T3G (old)	CGCCTGAAGCTCTGGGCGTTTTAGATTA- GAGGCAGGACGGTTAGATTAGAGGCAGGACG- GTTAGATTAGAGGCAGGACGG
T3G (new)	GTTATTGTCTCTGTCCTGTTTTAGATTA- GAGGCAGGACGGTTAGATTAGAGGCAGGACG- GTTAGATTAGAGGCAGGACGG
Guide	CCGTCCTGCCTCTAATCTAA
Guide (tris)	AminolinkC6 - 2xAmino C6 dT - CCGTCCTGCCTCTAATCTAA (biomers)
Guide (tris)	/5UniAmM//iUniAmM//iUniAmM/CCGTCCTGCCTCTAATCTAA
Guide (pentakis)	AminolinkC6 - 4xAmino C6 dT - CCGTCCTGCCTCTAATCTAA (biomers)
Reactive Displacer	GGCAGGACGG/3AmMC6T/ TTAGATTAGAGGCAGGACGG

Note: “redesign 1” (section 2.1.1.2) and “redesign 2” (section 2.1.1.3) refer to sequences labelled “v3” and “v2” respectively. All sequences were purchased from IDT (Coralville, Idaho, USA), aside from the tris-aminated and pentakis-aminated oligos which were occasionally provided by Biomers (Ulm, Germany). Modified oligos listed below have IDT/biomers modification syntax

A.3 Plasmid amplification

All constructs were supplied by the Lee lab on pACYC Duet-1 plasmids encoding chloramphenicol resistance, and were amplified as follows:

- Add $2\mu\text{L}$ of plasmid to $50\mu\text{L}$ of NEB 5 alpha competent cells, just thawed and kept on ice. Mix with pipette. Incubate on ice for 30m.
- Heat shock for 30s at 42°C .
- Add $950\mu\text{L}$ of SOC medium, incubate for 1h at 37^{circ}C shaking at $>250\text{rpm}$.
- Use plating beads to plate up a series of dilutions on LB agarose plates with $33\mu\text{g}/\mu\text{L}$ chloramphenicol. Incubate at 37^{circ}C overnight.
- To make starter cultures, inoculate single colonies in 3mL LB with $33\mu\text{g}/\mu\text{L}$ chloramphenicol, incubate for $\sim 8\text{h}$ at 37°C , shaking at 170rpm .
- Transfer $300\mu\text{L}$ starter culture to 100mL LB with $33\mu\text{g}/\text{mL}$ chloramphenicol. Incubate overnight ($\sim 12\text{-}16\text{h}$) at 37^{circ}C , shaking at 300rpm
- Purify plasmids following Qiagen Maxi Kit protocol. (Anion exchange followed by ethanol precipitation.)
- Check purity of plasmid on an agarose gel.

A.4 Protein expression, purification and dye labelling

A.4.1 Expression and lysis

- Add 1 μ L of amplified plasmid (\sim 100ng/ μ L) to 25 μ L T7 Express Competent *E. Coli* (NEB), thawed on ice. Mix with pipette tip, incubate on ice for 30m
- Heat shock at 42°C for 30s
- Add 100 μ L SOC medium, incubate shaking at 37°C for 1h
- Plate 5 μ L or 50 μ L on LB-agar plates with 33 μ g/mL chloramphenicol. Incubate overnight at 37°C
- Starter cultures: inoculate single colonies in 25mL LB with 33 μ g/mL chloramphenicol
- Inoculate \sim 10mL starter culture in 1L LB with 33 μ g/mL chloramphenicol, for a starting of \sim 0.01.
- Incubate at 37°C until OD 0.3-0.4, then grow incubate at 18°C until OD 0.6-0.8 (\sim 30-45m)
- Induce with 0.4mM IPTC, express overnight at 18°C
- Harvest cells at 400g for 20m at 4°C
- Re-suspend in 10mL lysis buffer (20mM Tris-HCl pH 7.5, 100mM NaCl)
- Lyse with a French press, and flash-freeze lysate in liquid nitrogen. Store aliquots at -80°C
 - (Robert Ishmukhametov assisted with French press operation)
 - We had similar success with a TC5 homogeniser (Stansted Fluid Power, Hawlow, UK) operated by Emma Sadler; this operates analogously to a continuous French press.

A.4.2 Purification and labelling

- Clear thawed lysate by centrifugation at 35,000g for 30m at 4°C
- For a volume V of lysate, add (0.02*V) of elution buffer [20mM Tris-HCl pH 7.5, 300mM NaCl, 500mM imidazole, 10mM EDTA] and (0.0004*V) of freshly-prepared 500mM TCEP.
- Prepare a gravity flow column with 400 μ L Ni-charged resin (Bio-rad #1560131), wash with 3 x 1200 μ L Lysis buffer
- Resuspend resin in the cleared lysate, incubate (rotating) for 2h at 4°C
- Centrifuge at 500g for 5m at 4°C; decant supernatant
- Transfer to gravity flow column, wash with >6mL Wash Buffer 1 [20mM Tris-HCl pH 7.5, 300mM NaCl, 20mM imidazole] followed by >2mL Wash Buffer 2 [20mM Tris-HCl pH 7.5, 300mM NaCl]
- Re-suspend resin to 50% slurry, transfer to eppendorf tube and store temporarily at 4°C, rotating. Meanwhile; run different volumes of resin slurry on SDS PAGE gel, alongside BCA standards. (I used Bolt™ 4-12% Bis-Tris Plus Gels, with premade loading buffers etc.) Stain with biosafe coomassie blue; use to estimate protein concentration on resin.
- Add 200mM TCEP to resin slurry; incubate for 30m at room temperature.
- Add 5x excess of maleimide-dye; incubate overnight at 4°C shielded from light
 - An accidental preparation with 3.5x excess also worked fine
- Wash resin with >15mL Wash Buffer 1
 - Gravity flow columns or centrifugation are both viable.
 - If size exclusion chromatography is to be omitted, wash more, e.g. 50mL
- Elute labelled protein with Elution Buffer [20mM Tris-HCl pH 7.5, 300mM NaCl, 500mM imidazole, 10mM EDTA] in 3 x 250 μ L fractions

- Purify with size exclusion chromatography (optional)
- Centrifuge at 15,000g for 10m at 4°C
- Run on Superdex 200 10/300 GL at 0.5mL/min, in MB
- Dilute to $\sim 20\mu\text{M}$ in MB + 10% glycerol; aliquot and flash freeze

A.5 BCA assay

BCA assay, using kit reagents from Pierce BCA assay kit

- Add N/50mL reagent B dropwise to N mL reagent A, where N is the number of samples to be tested
- to 1mL of mixture, add protein sample plus 10 μL of 10%SDS, aiming for $\sim 3\mu\text{g}$ protein per sample.
- incubate for 30m at 60°C
- Centrifuge briefly to check for unwanted precipitation
- Scan Abs[532] for all samples in 1mL cuvettes. Scan the same sample at start and end, to check that reactions have not progressed significantly during reaction time.

A.6 Proteinase K digestion to measure dye labelling

Buffers: Proteinase buffer * 4mL Matt's buffer * 120uL CaCl₂ 1M (30mM final)

P-K (prepare just before use!) * 1uL Proteinase-K stock (20mg/ml) * 32.3uL

Proteinase buffer

P-K Control * 1uL 50% gly * 32.3uL proteinase buffer

For each reaction, mix 1uL of P-K or P-K control with 5uL of sample (FliG aliquot, AF647 aliquot, etc), ensuring that all samples are in the same buffer.

Incubate at 50°C for 2h30. (Much shorter times may also be sufficient).

A.7 Surface Modification

A.7.1 BSA surface

incubate channel for 30m with 10mg/ml BSA in IB2, with with IB2

A.7.2 PLL-PEG surface

- Plasma clean 22x22mm glass coverslips for 5m in ceramic holder
- Rinse individually with copious amounts of MilliQ water
- Sandwich 15/ μ L of PLLPEG mixture (0.1mg/ml PLLPEG-biotin and 0.9mg/ml PLLPEG in MB) between 2 coverslips, incubate for >1h in a humidity chamber
 - Humidity chambers are made from pipette tip boxes half-full with wet tissue.
- Separate coverslips by sliding
- Rinse with copious amounts of MilliQ water
- Dry with compressed N₂

A.7.3 PEG vectabond surface

- Sonicate Menzel-Gläser 22x40mm #1.5 coverslips in 2% Hellmanex for >30m.
- Rinse individually with copious MilliQ water, dry with compressed N₂
 - Cover-slips can be stored at this point for at least a few months, in a tightly sealed 50ml falcon tube
- Plasma clean coverslips at high power for ~30m
- Prepare 2 PTFE beakers with 150ml acetone and one with 150ml MilliQ water
 - Place coverslips in ceramic rack, immerse in acetone beaker. Immediately add 3ml vectabond reagent and mix well. Incubate for 5m.

- Move rack to second acetone beaker
- Move rack to MilliQ beaker. Keep coverslips submerged.
- Defrost aliquots of 10mg NHS¹-PEG (MW 5000 Da, Laysan Bio) and 2mg NHS-PEG-biotin (MW 5000 Da, Laysan Bio), previously sealed with N₂ and stored at -20°C. Do not open until equilibrated with room temperature.
- Rinse coverslips individually with copious MilliQ water, dry with compressed N₂
 - Once dried, move to humidity chamber (pipette tip boxes half-full with wet tissue.)
- Make PEG mixture.
 - Add 400μL of filtered 50mM MOPS-KOH pH7.5 to NHS-PEG aliquot and vortex thoroughly
 - Add the PEG MOPS mixture to the NHS-PEG-biotin and vortex thoroughly
 - Centrifuge for ~30s to remove bubbles
- Sandwich ~35μL between each pair of coverslips, taking care to avoid bubbles. Incubate for >2h in a humidity chamber in the dark, at room temperature.
- Separate coverslips and rinse individually with copious MilliQ water, dry with compressed N₂
- If not using immediately, re-sandwich with MB or MilliQ and store in humidity chamber at 4°C for up to a month.

¹The supplier labels these “SVA PEG”

A.8 Anti-bleaching system

- Stocks
 - **MBG** (Glucose buffer solution) : MB (or buffer of choice) plus 0.8% w/v D-glucose
 - * Can be stored long term, but needs degassing immediately before use
 - **100x gloxy solution**
 - * Add 20 μ L of 0.2mg/ml catalase (Sigma-Aldrich C9322) to 80 μ L MB
 - * Add to 10mg glucose oxidase (Sigma-Aldrich G2133)
 - * Mix by tapping (do not vortex)
 - * Centrifuge for 1m (desktop centrifuge, max power), collect supernatant
 - * Store at 4°C for a few months maximum
 - **Trolox** (Sigma-Aldrich 238813)
 - * Freeze at 100mg/ml in DMSO
 - **NBA** (Fisher Scientific 10553291)
 - * Freeze at 100mg/ml in DMSO
 - **COT** (Sigma-Aldrich 138924)
 - * Freeze at 2% v/v in DMSO
- Mix immediately before use:
 - 276 μ L MBG, degassed
 - 2.5 μ L 100x gloxy
 - 1.25 μ L Trolox
 - 0.78 μ L NBA
 - 2.25 μ L COT

A.9 FliG assembly Gels

FliG assembly gels used 6%19:1 PAGE in 1x TAE, typically run at 200V for ~45m.

For best results, wells were loaded with a mixture of

- 0.5 μ L of reaction containing 333nM template
- 7.5 μ L MilliQ water
- 2 μ L Loading buffer (8%Ficoll, 0.033%bromophenol blue in 1x TAE)

Gels were stained with SYBR gold to visualize DNA.

References

- [1] Judith P Armitage. “Bacterial Taxis”. *Life Sciences* (2007). DOI: 10.1002/9780470015902.a0000340.pub2.
- [2] Christine Josenhans and Tobias Schweinitzer. *Bacterial energy taxis: A global strategy?* 2010. DOI: 10.1007/s00203-010-0575-7.
- [3] Kurt Ehlers and George Oster. “On the Mysterious Propulsion of *Synechococcus*”. *PLoS ONE* (2012). DOI: 10.1371/journal.pone.0036081.
- [4] M. B. Goldberg. “Actin-based motility of intracellular microbial pathogens.” *Microbiology and molecular biology reviews : MMBR* (2001). DOI: 10.1128/MMBR.65.4.595-626.2001.
- [5] Yoshiyuki Sowa and Richard M Berry. “Bacterial flagellar motor.” *Quarterly reviews of biophysics* (2008). DOI: 10.1017/S0033583508004691.
- [6] George H Wadhams and Judith P Armitage. “Making sense of it all: bacterial chemotaxis.” *Nature reviews. Molecular cell biology* (2004). DOI: 10.1038/nrm1524.
- [7] Junhua Yuan, Richard W Branch, Basarab G Hosu, and Howard C Berg. “Adaptation at the output of the chemotaxis signalling pathway.” *Nature* (2012). DOI: 10.1038/nature10964.
- [8] Rüdiger Schmitt. “Sinorhizobial chemotaxis: a departure from the enterobacterial paradigm.” *Microbiology (Reading, England)* (2002).
- [9] Ursula Attmannspacher, Birgit Scharf, and Rüdiger Schmitt. “Control of speed modulation (chemokinesis) in the unidirectional rotary motor of *Sinorhizobium meliloti*.” *Molecular microbiology* (2005). DOI: 10.1111/j.1365-2958.2005.04565.x.
- [10] J P Armitage and R M Macnab. “Unidirectional, intermittent rotation of the flagellum of *Rhodobacter sphaeroides*.” *Journal of bacteriology* (1987).
- [11] Judith P Armitage, Thomas P Pitta, M. A S Vigeant, Helen L Packer, and Roseanne M Ford. “Transformations in flagellar structure of *Rhodobacter sphaeroides* and possible relationship to changes in swimming speed”. *Journal of Bacteriology* (1999).
- [12] Teuta Pilizota, Mostyn T Brown, Mark C Leake, Richard W Branch, Richard M Berry, and Judith P Armitage. “A molecular brake, not a clutch, stops the *Rhodobacter sphaeroides* flagellar motor.” *Proceedings of the National Academy of Sciences of the United States of America* (2009). DOI: 10.1073/pnas.0813164106.

- [13] J P Armitage and R Schmitt. “Bacterial chemotaxis: Rhodobacter sphaeroides and Sinorhizobium meliloti—variations on a theme?” *Microbiology (Reading, England)* (1997).
- [14] Li Xie, Tuba Altindal, Suddhashil Chattopadhyay, and Xiao-Lun Wu. “Supporting Information”. *Proceedings of the National Academy of Sciences of the United States of America* (2011). DOI: 10.1073/pnas.1011953108.
- [15] Mikhail Kudryashev, Marek Cyrklaff, Reinhard Wallich, Wolfgang Baumeister, and Friedrich Frischknecht. “Distinct in situ structures of the Borrelia flagellar motor.” *Journal of structural biology* (2010). DOI: 10.1016/j.jsb.2009.08.008.
- [16] Masaru Kojima, Rumi Kubo, Toshiharu Yakushi, Michio Homma, and Ikuro Kawagishi. “The bidirectional polar and unidirectional lateral flagellar motors of Vibrio alginolyticus are controlled by a single CheY species.” *Molecular microbiology* (2007). DOI: 10.1111/j.1365-2958.2007.05623.x.
- [17] Hiroyuki Terashima, Seiji Kojima, and Michio Homma. *Flagellar motility in bacteria structure and function of flagellar motor*. Elsevier Inc., 2008. DOI: 10.1016/S1937-6448(08)01402-0.
- [18] E. Karatan and P. Watnick. “Signals, Regulatory Networks, and Materials That Build and Break Bacterial Biofilms”. *Microbiology and Molecular Biology Reviews* (2009). DOI: 10.1128/MMBR.00041-08.
- [19] Luanne Hall-Stoodley, J. William Costerton, and Paul Stoodley. “Bacterial biofilms: from the Natural environment to infectious diseases”. *Nature Reviews Microbiology* (2004). DOI: 10.1038/NRMICR0821.
- [20] U Römling and C Balsalobre. *Biofilm infections, their resilience to therapy and innovative treatment strategies*. 2012. DOI: 10.1111/joim.12004.
- [21] Bonnie Chaban, H. Velocity Hughes, and Morgan Beeby. “The flagellum in bacterial pathogens: For motility and a whole lot more”. *Seminars in Cell & Developmental Biology* (2015). DOI: 10.1016/j.semcdb.2015.10.032.
- [22] Daniel B Kearns. “A field guide to bacterial swarming motility.” *Nature reviews. Microbiology* (2010). DOI: 10.1038/nrmicro2405.
- [23] Rajesh Balagam, Douglas B Litwin, Fabian Czerwinski, Mingzhai Sun, Heidi B Kaplan, Joshua W Shaevitz, and Oleg A Igoshin. “Myxococcus xanthus gliding motors are elastically coupled to the substrate as predicted by the focal adhesion model of gliding motility.” *PLoS computational biology* (2014). DOI: 10.1371/journal.pcbi.1003619.
- [24] Abhishek Shrivastava, Thibault Roland, and Howard C Berg. “The Screw-Like Movement of a Gliding Bacterium Is Powered by Spiral Motion of Cell-Surface Adhesins”. *Biophysical journal* (2016). DOI: 10.1016/j.bpj.2016.07.043.
- [25] Andreas Diepold and Judith P Armitage. “Type III secretion systems: the bacterial flagellum and the injectisome.” *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* (2015). DOI: 10.1098/rstb.2015.0020.
- [26] Chris B Stone et al. “Interactions between flagellar and type III secretion proteins in Chlamydia pneumoniae”. *BMC Microbiology* (2010). DOI: 10.1186/1471-2180-10-18.

- [27] Andrew C McShan and Roberto N De Guzman. “The bacterial type III secretion system as a target for developing new antibiotics.” *Chemical biology & drug design* (2015). DOI: 10.1111/cbdd.12422.
- [28] Andreas Diepold, Mikhail Kudryashev, Nicolas J Delalez, Richard M Berry, and Judith P Armitage. “Composition, formation, and regulation of the cytosolic c-ring, a dynamic component of the type III secretion injectisome.” *PLoS biology* (2015). DOI: 10.1371/journal.pbio.1002039.
- [29] Julien R Bergeron. “Structural modeling of the flagellum MS ring protein FliF reveals similarities to the type III secretion system and sporulation complex.” *PeerJ* (2016). DOI: 10.7717/peerj.1718.
- [30] D Thomas, D G Morgan, and D J DeRosier. “Structures of bacterial flagellar motors from two FliF-FliG gene fusion mutants.” *Journal of bacteriology* (2001). DOI: 10.1128/JB.183.21.6404-6412.2001.
- [31] Basarab G. Hosu, Vedavalli S. J. Nathan, and Howard C. Berg. “Internal and external components of the bacterial flagellar motor rotate as a unit”. *Proceedings of the National Academy of Sciences* (2016). DOI: 10.1073/pnas.1511691113.
- [32] Y Magariyama, S Sugiyama, K Muramoto, Y Maekawa, I Kawagishi, Y Imae, and S Kudo. “Very fast flagellar rotation.” *Nature* (1994). DOI: 10.1038/371752b0.
- [33] Fan Bai, Richard W Branch, Dan V Nicolau, Teuta Pilizota, Bradley C Steel, Philip K Maini, and Richard M Berry. “Conformational spread as a mechanism for cooperativity in the bacterial flagellar switch.” *Science (New York, N.Y.)* (2010). DOI: 10.1126/science.1182105.
- [34] Pushkar P Lele and Howard C Berg. “Switching of bacterial flagellar motors is triggered by mutant FliG.” *Biophysical journal* (2015). DOI: 10.1016/j.bpj.2015.02.004.
- [35] Junhua Yuan and Howard C Berg. “Ultrasensitivity of an adaptive bacterial motor.” *Journal of molecular biology* (2013). DOI: 10.1016/j.jmb.2013.02.016.
- [36] Victor Sourjik and Howard C Berg. “Binding of the Escherichia coli response regulator CheY to its target measured in vivo by fluorescence resonance energy transfer.” *Proceedings of the National Academy of Sciences of the United States of America* (2002). DOI: 10.1073/pnas.192463199.
- [37] Yael Sagi, Shahid Khan, and Michael Eisenbach. “Binding of the chemotaxis response regulator CheY to the isolated, intact switch complex of the bacterial flagellar motor: lack of cooperativity.” *The Journal of biological chemistry* (2003). DOI: 10.1074/jbc.M303201200.
- [38] Dennis R Thomas, Noreen R Francis, Chen Xu, and David J DeRosier. “The three-dimensional structure of the flagellar rotor from a clockwise-locked mutant of Salmonella enterica serovar Typhimurium.” *Journal of bacteriology* (2006). DOI: 10.1128/JB.00552-06.
- [39] Koji Yonekura, Saori Maki-Yonekura, and Michio Homma. “Structure of the flagellar motor protein complex PomAB: implications for the torque-generating conformation.” *Journal of bacteriology* (2011). DOI: 10.1128/JB.05021-11.
- [40] Xiaowei Zhao, Steven J Norris, and Jun Liu. “Molecular architecture of the bacterial flagellar motor in cells.” *Biochemistry* (2014). DOI: 10.1021/bi500059y.

- [41] Daniela Stock, Keiichi Namba, and Lawrence K Lee. “Nanorotors and self-assembling macromolecular machines: the torque ring of the bacterial flagellar motor.” *Current opinion in biotechnology* (2012). DOI: 10.1016/j.copbio.2012.01.008.
- [42] T Ueno, K Oosawa, and S Aizawa. “M ring, S ring and proximal rod of the flagellar basal body of *Salmonella typhimurium* are composed of subunits of a single protein, FliF.” *Journal of molecular biology* (1992).
- [43] Eisaku Katayama, Tomoko Shiraishi, Kenji Oosawa, Norio Baba, and Shin-Ichi Aizawa. “Geometry of the Flagellar Motor in the Cytoplasmic Membrane of *Salmonella typhimurium* as Determined by Stereo-photogrammetry of Quick-freeze Deep-etch Replica Images”. *Journal of Molecular Biology* (1996). DOI: 10.1006/jmbi.1996.0038.
- [44] Hirofumi Suzuki, Koji Yonekura, and Keiichi Namba. “Structure of the rotor of the bacterial flagellar motor revealed by electron cryomicroscopy and single-particle image analysis.” *Journal of molecular biology* (2004). DOI: 10.1016/j.jmb.2004.01.034.
- [45] K Kutsukake, Y Ohya, and T Iino. “Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*.” *Journal of bacteriology* (1990).
- [46] Hui Li and Victor Sourjik. “Assembly and stability of flagellar motor in *Escherichia coli*.” *Molecular microbiology* (2011). DOI: 10.1111/j.1365-2958.2011.07557.x.
- [47] Yusuke V Morimoto et al. “Assembly and stoichiometry of FliF and FlhA in *Salmonella* flagellar basal body.” *Molecular microbiology* (2014). DOI: 10.1111/mmi.12529.
- [48] Christopher J. Jones, Robert M. Macnab, Hiroshi Okino, and Shin-Ichi Aizawa. “Stoichiometric analysis of the flagellar hook-(basal-body) complex of *Salmonella typhimurium*”. *Journal of Molecular Biology* (1990). DOI: 10.1016/0022-2836(90)90132-6.
- [49] S Yamaguchi, H Fujita, A Ishihara, S Aizawa, and R M Macnab. “Subdivision of flagellar genes of *Salmonella typhimurium* into regions responsible for assembly, rotation, and switching.” *Journal of bacteriology* (1986).
- [50] M L DePamphilis and J Adler. “Fine structure and isolation of the hook-basal body complex of flagella from *Escherichia coli* and *Bacillus subtilis*.” *Journal of bacteriology* (1971).
- [51] A Driks and D J DeRosier. “Additional structures associated with bacterial flagellar basal body.” *Journal of molecular biology* (1990). DOI: 10.1016/0022-2836(90)90063-R.
- [52] I H Khan, T S Reese, and S Khan. “The cytoplasmic component of the bacterial flagellar motor.” *Proceedings of the National Academy of Sciences of the United States of America* (1992).
- [53] N R Francis, G E Sosinsky, D Thomas, and D J DeRosier. “Isolation, characterization and structure of bacterial flagellar motors containing the switch complex.” *Journal of molecular biology* (1994). DOI: 10.1006/jmbi.1994.1079.

- [54] Koji Yonekura, Toshiharu Yakushi, Tatsuo Atsumi, Saori Maki-Yonekura, Michio Homma, and Keiichi Namba. “Electron cryomicroscopic visualization of PomA/B stator units of the sodium-driven flagellar motor in liposomes.” *Journal of molecular biology* (2006). DOI: 10.1016/j.jmb.2005.12.041.
- [55] Koushik Paul and David F Blair. “Organization of FliN subunits in the flagellar motor of *Escherichia coli*.” *Journal of bacteriology* (2006). DOI: 10.1128/JB.188.7.2502-2511.2006.
- [56] Kwok-Ho Lam, Wing-Sang Ip, Yun-Wah Lam, Sun-On Chan, Thomas Kin-Wah Ling, and Shannon Wing-Ngor Au. “Multiple conformations of the FliG C-terminal domain provide insight into flagellar motor switching.” *Structure (London, England : 1993)* (2012). DOI: 10.1016/j.str.2011.11.020.
- [57] Lawrence K Lee, Michael a Ginsburg, Claudia Crovace, Mhairi Donohoe, and Daniela Stock. “Structure of the torque ring of the flagellar motor and the molecular basis for rotational switching.” *Nature* (2010). DOI: 10.1038/nature09300.
- [58] N R Francis, V M Irikura, S Yamaguchi, D J DeRosier, and R M Macnab. “Localization of the *Salmonella typhimurium* flagellar switch protein FliG to the cytoplasmic M-ring face of the basal body.” *Proceedings of the National Academy of Sciences of the United States of America* (1992).
- [59] Koushik Paul, Gabriela Gonzalez-Bonet, Alexandrine M Bilwes, Brian R Crane, and David Blair. “Architecture of the flagellar rotor.” *The EMBO journal* (2011). DOI: 10.1038/emboj.2011.188.
- [60] Jong Lee et al. “A mechanical metamaterial made from a DNA hydrogel”. *Nature Nanotechnology* (2012). DOI: doi:10.1038/nnano.2012.211.
- [61] Nicolas J Delalez et al. “Signal-dependent turnover of the bacterial flagellar switch protein FliM.” *Proceedings of the National Academy of Sciences of the United States of America* (2010). DOI: 10.1073/pnas.1000284107.
- [62] Pushkar P Lele, Richard W Branch, Vedhavalli S J Nathan, and Howard C Berg. “Mechanism for adaptive remodeling of the bacterial flagellar switch.” *Proceedings of the National Academy of Sciences of the United States of America* (2012). DOI: 10.1073/pnas.1212327109.
- [63] Nicolas J Delalez, Richard M Berry, and Judith P Armitage. “Stoichiometry and Turnover of the Bacterial Flagellar Switch Protein FliN.” *mBio* (2014). DOI: 10.1128/mBio.01216-14.
- [64] Richard W Branch, Michael N Sayegh, Chong Shen, Vedavalli S J Nathan, and Howard C Berg. “Adaptive remodelling by FliN in the bacterial rotary motor.” *Journal of molecular biology* (2014). DOI: 10.1016/j.jmb.2014.07.009.
- [65] Armand S Vartanian, Aviv Paz, Emily A Fortgang, Jeff Abramson, and Frederick W Dahlquist. “Structure of flagellar motor proteins in complex allows for insights into motor structure and switching.” *The Journal of biological chemistry* (2012). DOI: 10.1074/jbc.C112.378380.
- [66] Pushkar P Lele, Richard W Branch, Vedhavalli S J Nathan, and Howard C Berg. “Mechanism for adaptive remodeling of the bacterial flagellar switch.” *Proceedings of the National Academy of Sciences of the United States of America* (2012). DOI: 10.1073/pnas.1212327109.

- [67] Steven E Passmore, Rithy Meas, and Donna L Marykwas. “Analysis of the FliM/FliG motor protein interaction by two-hybrid mutation suppression analysis.” *Microbiology (Reading, England)* (2008). DOI: 10.1099/mic.0.2007/014597-0.
- [68] Xiaowei Zhao, Steven J. Norris, and Jun Liu. “Molecular Architecture of the Bacterial Flagellar Motor in Cells”. *Biochemistry* (2014). DOI: 10.1021/bi500059y.
- [69] Songye Chen et al. “Structural diversity of bacterial flagellar motors.” *The EMBO journal* (2011). DOI: 10.1038/emboj.2011.186.
- [70] Jun Liu, Tao Lin, Douglas J Botkin, Erin McCrum, Hanspeter Winkler, and Steven J Norris. “Intact flagellar motor of *Borrelia burgdorferi* revealed by cryo-electron tomography: evidence for stator ring curvature and rotor/C-ring assembly flexion.” *Journal of bacteriology* (2009). DOI: 10.1128/JB.00340-09.
- [71] Gavin E Murphy, Jared R Leadbetter, and Grant J Jensen. “In situ structure of the complete *Treponema primitia* flagellar motor.” *Nature* (2006). DOI: 10.1038/nature05015.
- [72] Morgan Beeby, Deborah A. Ribardo, Caitlin A. Brennan, Edward G. Ruby, Grant J. Jensen, and David R. Hendrixson. “Diverse high-torque bacterial flagellar motors assemble wider stator rings using a conserved protein scaffold”. *Proceedings of the National Academy of Sciences* (2016). DOI: 10.1073/pnas.1518952113.
- [73] Björn Grünenfelder, Stefanie Gehrig, and Urs Jenal. “Role of the cytoplasmic C terminus of the FliF motor protein in flagellar assembly and rotation.” *Journal of bacteriology* (2003).
- [74] Robert Levenson, Hongjun Zhou, and Frederick W Dahlquist. “Structural insights into the interaction between the bacterial flagellar motor proteins FliF and FliG.” *Biochemistry* (2012). DOI: 10.1021/bi3004582.
- [75] Ryo Ogawa, Rei Abe-Yoshizumi, Takaaki Kishi, Michio Homma, and Seiji Kojima. “Interaction of the C-terminal tail of FliF with FliG from the Na⁺-driven flagellar motor of *Vibrio alginolyticus*.” *Journal of bacteriology* (2015). DOI: 10.1128/JB.02271-14.
- [76] Takamasa Ueno, Kenji Oosawa, and Shin-Ichi Aizawa. “Domain Structures of the MS Ring Component Protein (FliF) of the Flagellar Basal Body of *Salmonella typhimurium*”. *Journal of Molecular Biology* (1994). DOI: 10.1006/jmbi.1994.1164.
- [77] M. Kihara, G. U. Miller, and R. M. Macnab. “Deletion Analysis of the Flagellar Switch Protein FliG of *Salmonella*”. *Journal of Bacteriology* (2000). DOI: 10.1128/JB.182.11.3022-3028.2000.
- [78] D L Marykwas, S a Schmidt, and H C Berg. “Interacting components of the flagellar motor of *Escherichia coli* revealed by the two-hybrid system in yeast.” *Journal of molecular biology* (1996). DOI: 10.1006/jmbi.1996.0109.
- [79] D R Thomas, D G Morgan, and D J DeRosier. “Rotational symmetry of the C ring and a mechanism for the flagellar rotary motor.” *Proceedings of the National Academy of Sciences of the United States of America* (1999).

- [80] S Yamaguchi, S Aizawa, M Kihara, M Isomura, C J Jones, and R M Macnab. “Genetic evidence for a switching and energy-transducing complex in the flagellar motor of *Salmonella typhimurium*.” *Journal of bacteriology* (1986).
- [81] S Khan, D M Ivey, and T A Krulwich. “Membrane ultrastructure of alkaliphilic *Bacillus* species studied by rapid-freeze electron microscopy.” *Journal of bacteriology* (1992).
- [82] Jun Liu, Jerrilyn K Howell, Sherille D Bradley, Yesha Zheng, Z Hong Zhou, and Steven J Norris. “Cellular architecture of *Treponema pallidum*: novel flagellum, periplasmic cone, and cell envelope as revealed by cryo electron tomography.” *Journal of molecular biology* (2010). DOI: 10.1016/j.jmb.2010.09.020.
- [83] Howard S Young, Hongyue Dang, Yimin Lai, David J DeRosier, and Shahid Khan. “Variable symmetry in *Salmonella typhimurium* flagellar motors.” *Biophysical journal* (2003). DOI: 10.1016/S0006-3495(03)74877-2.
- [84] Yoshiyuki Sowa, Alexander D Rowe, Mark C Leake, Toshiharu Yakushi, Michio Homma, Akihiko Ishijima, and Richard M Berry. “Direct observation of steps in rotation of the bacterial flagellar motor.” *Nature* (2005). DOI: 10.1038/nature04003.
- [85] Shuichi Nakamura, Nobunori Kami-ike, Jun-ichi P Yokota, Tohru Minamino, and Keiichi Namba. “Evidence for symmetry in the elementary process of bidirectional torque generation by the bacterial flagellar motor.” *Proceedings of the National Academy of Sciences of the United States of America* (2010). DOI: 10.1073/pnas.1007448107.
- [86] R Zhao, C D Amsler, P Matsumura, and S Khan. “FliG and FliM distribution in the *Salmonella typhimurium* cell and flagellar basal bodies.” *Journal of bacteriology* (1996).
- [87] Perry N Brown, Christopher P Hill, and David F Blair. “Crystal structure of the middle and C-terminal domains of the flagellar rotor protein FliG.” *The EMBO journal* (2002). DOI: 10.1093/emboj/cdf332.
- [88] Tohru Minamino, Katsumi Imada, Miki Kinoshita, Shuichi Nakamura, Yusuke V Morimoto, and Keiichi Namba. “Structural insight into the rotational switching mechanism of the bacterial flagellar motor”. *PLoS Biology* (2011). DOI: 10.1371/journal.pbio.1000616.
- [89] Koushik Paul, Duncan Brunstetter, Sienna Titen, and David F Blair. “A molecular mechanism of direction switching in the flagellar motor of *Escherichia coli*”. *Proceedings of the National Academy of Sciences* (2011). DOI: 10.1073/pnas.1110111108.
- [90] Susan M Van Way, Stephanos G Millas, Aaron H Lee, and Michael D Manson. “Rusty, jammed, and well-oiled hinges: Mutations affecting the interdomain region of FliG, a rotor element of the *Escherichia coli* flagellar motor.” *Journal of bacteriology* (2004).
- [91] D L Marykwas and H C Berg. “A mutational analysis of the interaction between FliG and FliM, two components of the flagellar motor of *Escherichia coli*.” *Journal of bacteriology* (1996).

- [92] S A Lloyd, H Tang, X Wang, S Billings, and D F Blair. "Torque generation in the flagellar motor of *Escherichia coli*: evidence of a direct role for FliG but not for FliM or FliN." *Journal of bacteriology* (1996).
- [93] S a Lloyd and D F Blair. "Charged residues of the rotor protein FliG essential for torque generation in the flagellar motor of *Escherichia coli*." *Journal of molecular biology* (1997). DOI: 10.1006/jmbi.1996.0836.
- [94] J Zhou, S a Lloyd, and D F Blair. "Electrostatic interactions between rotor and stator in the bacterial flagellar motor." *Proceedings of the National Academy of Sciences of the United States of America* (1998).
- [95] Tomohiro Yorimitsu, Atsushi Mimaki, Toshiharu Yakushi, and Michio Homma. "The Conserved Charged Residues of the C-terminal Region of FliG, a Rotor Component of the Na⁺-driven Flagellar Motor". *Journal of Molecular Biology* (2003). DOI: 10.1016/j.jmb.2003.09.052.
- [96] Toshiharu Yakushi, Junghoon Yang, Hajime Fukuoka, Michio Homma, and David F Blair. "Roles of charged residues of rotor and stator in flagellar rotation: comparative study using H⁺-driven and Na⁺-driven motors in *Escherichia coli*." *Journal of bacteriology* (2006). DOI: 10.1128/JB.188.4.1466-1472.2006.
- [97] Yusuke V Morimoto, Shuichi Nakamura, Koichi D Hiraoka, Keiichi Namba, and Tohru Minamino. "Distinct roles of highly conserved charged residues at the MotA-FliG interface in bacterial flagellar motor rotation." *Journal of bacteriology* (2013). DOI: 10.1128/JB.01971-12.
- [98] Kranthi K Mandadapu, Jasmine A Nirody, Richard M Berry, and George Oster. "Mechanics of torque generation in the bacterial flagellar motor." *Proceedings of the National Academy of Sciences of the United States of America* (2015). DOI: 10.1073/pnas.1501734112.
- [99] S A Lloyd, F G Whitby, D F Blair, and C P Hill. "Structure of the C-terminal domain of FliG, a component of the rotor in the bacterial flagellar motor." *Nature* (1999). DOI: 10.1038/22794.
- [100] Alessandro Pandini, Jens Kleijnung, Shafqat Rasool, and Shahid Khan. "Coevolved Mutations Reveal Distinct Architectures for Two Core Proteins in the Bacterial Flagellar Motor." *PloS one* (2015). DOI: 10.1371/journal.pone.0142407.
- [101] Bryan J Lowder, Mark D Duyvesteyn, and David F Blair. "FliG subunit arrangement in the flagellar rotor probed by targeted cross-linking." *Journal of bacteriology* (2005). DOI: 10.1128/JB.187.16.5640-5647.2005.
- [102] Perry N Brown, Moises Terrazas, Koushik Paul, and David F Blair. "Mutational analysis of the flagellar protein FliG: sites of interaction with FliM and implications for organization of the switch complex." *Journal of bacteriology* (2007). DOI: 10.1128/JB.01281-06.
- [103] V M Irikura, M Kihara, S Yamaguchi, H Sockett, and R M Macnab. "Salmonella typhimurium fliG and fliN mutations causing defects in assembly, rotation, and switching of the flagellar motor." *Journal of bacteriology* (1993).
- [104] Ria Sircar et al. "Assembly states of FliM and FliG within the flagellar switch complex." *Journal of molecular biology* (2015). DOI: 10.1016/j.jmb.2014.12.009.

- [105] Kwok-Ho Lam et al. “Structural basis of FliG-FliM interaction in *Helicobacter pylori*.” *Molecular microbiology* (2013). DOI: 10.1111/mmi.12222.
- [106] Collin M Dyer, Armand S Vartanian, Hongjun Zhou, and Frederick W Dahlquist. “A molecular mechanism of bacterial flagellar motor switching.” *Journal of molecular biology* (2009). DOI: 10.1016/j.jmb.2009.02.004.
- [107] Rita Tewari, Elizabeth Bailes, Karen A Bunting, and Juliet C Coates. “Armadillo-repeat protein functions: questions for little creatures.” *Trends in cell biology* (2010). DOI: 10.1016/j.tcb.2010.05.003.
- [108] Fabio Parmeggiani et al. “Designed armadillo repeat proteins as general peptide-binding scaffolds: consensus design and computational optimization of the hydrophobic core.” *Journal of molecular biology* (2008). DOI: 10.1016/j.jmb.2007.12.014.
- [109] M A Andrade, C Petosa, S I O’Donoghue, C W Müller, and P Bork. “Comparison of ARM and HEAT protein repeats.” *Journal of molecular biology* (2001). DOI: 10.1006/jmbi.2001.4624.
- [110] Kristin K Jernigan and Seth R Bordenstein. “Tandem-repeat protein domains across the tree of life.” *PeerJ* (2015). DOI: 10.7717/peerj.732.
- [111] Matthew A B Baker et al. “Domain-swap polymerization drives the self-assembly of the bacterial flagellar motor.” *Nature structural & molecular biology* (2016). DOI: 10.1038/nsmb.3172.
- [112] M A Mathews, H L Tang, and D F Blair. “Domain analysis of the FliM protein of *Escherichia coli*.” *Journal of bacteriology* (1998).
- [113] K Oosawa, T Ueno, and S Aizawa. “Overproduction of the bacterial flagellar switch proteins and their interactions with the MS ring complex in vitro.” *Journal of bacteriology* (1994).
- [114] Perry N Brown, Michael A A Mathews, Lisa A Joss, Christopher P Hill, and David F Blair. “Crystal structure of the flagellar rotor protein FliN from *Thermotoga maritima*.” *Journal of Bacteriology* (2005). DOI: 10.1128/JB.187.8.2890-2902.2005.
- [115] Sang-Youn Park, Bryan Lowder, Alexandrine M Bilwes, David F Blair, and Brian R Crane. “Structure of FliM provides insight into assembly of the switch complex in the bacterial flagella motor.” *Proceedings of the National Academy of Sciences of the United States of America* (2006). DOI: 10.1073/pnas.0602811103.
- [116] a S Toker and R M Macnab. “Distinct regions of bacterial flagellar switch protein FliM interact with FliG, FliN and CheY.” *Journal of molecular biology* (1997). DOI: 10.1006/jmbi.1997.1335.
- [117] A Bren and M Eisenbach. “The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY.” *Journal of molecular biology* (1998). DOI: 10.1006/jmbi.1998.1730.
- [118] A S Toker, M Kihara, and R M Macnab. “Deletion analysis of the FliM flagellar switch protein of *Salmonella typhimurium*.” *Journal of bacteriology* (1996).
- [119] SY Lee, HS Cho, and JG Pelton. “Crystal structure of an activated response regulator bound to its target”. *Nature Structural & ...* (2001).

- [120] C. M. Dyer and F. W. Dahlquist. “Switched or Not?: the Structure of Unphosphorylated CheY Bound to the N Terminus of FliM”. *Journal of Bacteriology* (2006). DOI: 10.1128/JB.00637-06.
- [121] Collin M. Dyer et al. “Structure of the Constitutively Active Double Mutant CheYD13K Y106W Alone and in Complex with a FliM Peptide”. *Journal of Molecular Biology* (2004). DOI: 10.1016/j.jmb.2004.07.084.
- [122] M Kihara, N R Francis, D J DeRosier, and R M Macnab. “Analysis of a FliM-FliN flagellar switch fusion mutant of *Salmonella typhimurium*.” *Journal of bacteriology* (1996).
- [123] Ryan Q Notti, Shibani Bhattacharya, Mirjana Lilic, and C Erec Stebbins. “A common assembly module in injectisome and flagellar type III secretion sorting platforms.” *Nature communications* (2015). DOI: 10.1038/ncomms8125.
- [124] M Welch, K Oosawa, S Aizawa, and M Eisenbach. “Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria.” *Proceedings of the National Academy of Sciences of the United States of America* (1993).
- [125] Bertha González-Pedrajo, Tohru Minamino, May Kihara, and Keiichi Namba. “Interactions between C ring proteins and export apparatus components: a possible mechanism for facilitating type III protein export.” *Molecular microbiology* (2006). DOI: 10.1111/j.1365-2958.2006.05149.x.
- [126] Mayukh K Sarkar, Koushik Paul, and David F Blair. “Subunit organization and reversal-associated movements in the flagellar switch of *Escherichia coli*.” *The Journal of biological chemistry* (2010). DOI: 10.1074/jbc.M109.068676.
- [127] Melanie A McDowell et al. “Characterisation of *Shigella Spa33* and *Thermotoga FliM/N* reveals a new model for C-ring assembly in T3SS.” *Molecular microbiology* (2016). DOI: 10.1111/mmi.13267.
- [128] R Zhao, N Pathak, H Jaffe, T S Reese, and S Khan. “FliN is a major structural protein of the C-ring in the *Salmonella typhimurium* flagellar basal body.” *Journal of molecular biology* (1996). DOI: 10.1006/jmbi.1996.0452.
- [129] Koushik Paul, Jacob G Harmon, and David F Blair. “Mutational analysis of the flagellar rotor protein FliN: identification of surfaces important for flagellar assembly and switching.” *Journal of bacteriology* (2006). DOI: 10.1128/JB.00110-06.
- [130] Mayukh K Sarkar, Koushik Paul, and David Blair. “Chemotaxis signaling protein CheY binds to the rotor protein FliN to control the direction of flagellar rotation in *Escherichia coli*.” *Proceedings of the National Academy of Sciences of the United States of America* (2010). DOI: 10.1073/pnas.1000935107.
- [131] A P Vogler, M Homma, V M Irikura, and R M Macnab. “*Salmonella typhimurium* mutants defective in flagellar filament regrowth and sequence similarity of FliI to F0F1, vacuolar, and archaeobacterial ATPase subunits.” *Journal of bacteriology* (1991).
- [132] Jonathan L McMurry, James W Murphy, and Bertha González-Pedrajo. “The FliN-FliH interaction mediates localization of flagellar export ATPase FliI to the C ring complex.” *Biochemistry* (2006). DOI: 10.1021/bi0605890.

- [133] Tohru Minamino, Shinsuke D J Yoshimura, Yusuke V Morimoto, Bertha González-Pedrajo, Nobunori Kami-Ike, and Keiichi Namba. “Roles of the extreme N-terminal region of FliH for efficient localization of the FliH-FliI complex to the bacterial flagellar type III export apparatus.” *Molecular microbiology* (2009). DOI: 10.1111/j.1365-2958.2009.06946.x.
- [134] H Tang, S Billings, X Wang, L Sharp, and D F Blair. “Regulated underexpression and overexpression of the FliN protein of Escherichia coli and evidence for an interaction between FliN and FliM in the flagellar motor.” *Journal of bacteriology* (1995).
- [135] Victor Sourjik and Howard C Berg. “Binding of the Escherichia coli response regulator CheY to its target measured in vivo by fluorescence resonance energy transfer.” *Proceedings of the National Academy of Sciences of the United States of America* (2002). DOI: 10.1073/pnas.192463199.
- [136] Shun Terasawa, Hajime Fukuoka, Yuichi Inoue, Takashi Sagawa, Hiroto Takahashi, and Akihiko Ishijima. “Coordinated reversal of flagellar motors on a single Escherichia coli cell.” *Biophysical journal* (2011). DOI: 10.1016/j.bpj.2011.03.030.
- [137] Hajime Fukuoka, Takashi Sagawa, Yuichi Inoue, Hiroto Takahashi, and Akihiko Ishijima. “Direct imaging of intracellular signaling components that regulate bacterial chemotaxis.” *Science signaling* (2014). DOI: 10.1126/scisignal.2004963.
- [138] Takashi Sagawa et al. “Single-cell E. coli response to an instantaneously applied chemotactic signal.” *Biophysical journal* (2014). DOI: 10.1016/j.bpj.2014.06.017.
- [139] Patrick J. Mears, Santosh Koirala, Chris V. Rao, Ido Golding, and Yann R. Chemla. “Escherichia coli swimming is robust against variations in flagellar number”. *eLife* (2014). DOI: 10.7554/eLife.01916.
- [140] Karen A. Fahrner, William S. Ryu, and Howard C. Berg. “Biomechanics: Bacterial flagellar switching under load”. *Nature* (2003). DOI: 10.1038/423938a.
- [141] Ekaterina A Korobkova, Thierry Emonet, Heungwon Park, and Philippe Cluzel. “Hidden stochastic nature of a single bacterial motor.” *Physical review letters* (2006). DOI: 10.1103/PhysRevLett.96.058105.
- [142] Ekaterina Korobkova, Thierry Emonet, Jose M G Vilar, Thomas S Shimizu, and Philippe Cluzel. “From molecular noise to behavioural variability in a single bacterium.” *Nature* (2004). DOI: 10.1038/nature02404.
- [143] Junhua Yuan, Karen a Fahrner, and Howard C Berg. “Switching of the bacterial flagellar motor near zero load.” *Journal of molecular biology* (2009). DOI: 10.1016/j.jmb.2009.05.039.
- [144] Siebe B van Albada, Sorin Tănase-Nicola, and Pieter Rein ten Wolde. “The switching dynamics of the bacterial flagellar motor.” en. *Molecular systems biology* (2009). DOI: 10.1038/msb.2009.74.
- [145] Fan Bai, Tohru Minamino, Zhanghan Wu, Keiichi Namba, and Jianhua Xing. “Coupling between switching regulation and torque generation in bacterial flagellar motor.” *Physical review letters* (2012). DOI: 10.1103/PhysRevLett.108.178105.

- [146] L Turner, S R Caplan, and H C Berg. “Temperature-induced switching of the bacterial flagellar motor.” *Biophysical journal* (1996). DOI: 10.1016/S0006-3495(96)79425-0.
- [147] Ria Sircar, Anna R. Greenswag, Alexandrine M. Bilwes, Gabriela Gonzalez-Bonet, and Brian R. Crane. “Structure and activity of the flagellar rotor protein fliY: A member of the cheC phosphatase family”. *Journal of Biological Chemistry* (2013). DOI: 10.1074/jbc.M112.445171.
- [148] Andrew C Lowenthal, Marla Hill, Laura K Sycuro, Khalid Mehmood, Nina R Salama, and Karen M Ottemann. “Functional analysis of the *Helicobacter pylori* flagellar switch proteins.” *Journal of bacteriology* (2009). DOI: 10.1128/JB.00749-09.
- [149] Mark C Leake, Jennifer H Chandler, George H Wadhams, Fan Bai, Richard M Berry, and Judith P Armitage. “Stoichiometry and turnover in single, functioning membrane protein complexes.” *Nature* (2006). DOI: 10.1038/nature05135.
- [150] Hajime Fukuoka, Tomoyuki Wada, Seiji Kojima, Akihiko Ishijima, and Michio Homma. “Sodium-dependent dynamic assembly of membrane complexes in sodium-driven flagellar motors.” *Molecular microbiology* (2009). DOI: 10.1111/j.1365-2958.2008.06569.x.
- [151] Anja Paulick, Nicolas J Delalez, Susanne Brenzinger, Bradley C Steel, Richard M Berry, Judith P Armitage, and Kai M Thormann. “Dual stator dynamics in the *Shewanella oneidensis* MR-1 flagellar motor.” *Molecular microbiology* (2015). DOI: 10.1111/mmi.12984.
- [152] Murray J Tipping, Bradley C Steel, Nicolas J Delalez, Richard M Berry, and Judith P Armitage. “Quantification of flagellar motor stator dynamics through in vivo proton-motive force control.” *Molecular microbiology* (2013). DOI: 10.1111/mmi.12098.
- [153] Pushkar P Lele, Basarab G Hosu, and Howard C Berg. “Dynamics of mechanosensing in the bacterial flagellar motor.” *Proceedings of the National Academy of Sciences of the United States of America* (2013). DOI: 10.1073/pnas.1305885110.
- [154] Murray J Tipping, Nicolas J Delalez, Ren Lim, Richard M Berry, and Judith P. Armitage. “Load-dependent assembly of the bacterial flagellar motor”. *mBio* (2013). DOI: 10.1128/mBio.00551-13.
- [155] Yong-Suk Che, Shuichi Nakamura, Yusuke V Morimoto, Nobunori Kami-Ike, Keiichi Namba, and Tohru Minamino. “Load-sensitive coupling of proton translocation and torque generation in the bacterial flagellar motor.” *Molecular microbiology* (2014). DOI: 10.1111/mmi.12453.
- [156] Pushkar P Lele, Abhishek Shrivastava, Thibault Roland, and Howard C Berg. “Response thresholds in bacterial chemotaxis.” *Science advances* (2015). DOI: 10.1126/sciadv.1500299.
- [157] S Khan, R Zhao, and T S Reese. “Architectural features of the *Salmonella typhimurium* flagellar motor switch revealed by disrupted C-rings.” *Journal of structural biology* (1998). DOI: 10.1006/jsbi.1998.3999.

- [158] Andrea Nans, Mikhail Kudryashev, Helen R Saibil, and Richard D Hayward. “Structure of a bacterial type III secretion system in contact with a host membrane in situ.” en. *Nature communications* (2015). DOI: 10.1038/ncomms10114.
- [159] Mikhail Kudryashev et al. “In situ structural analysis of the *Yersinia enterocolitica* injectisome.” *eLife* (2013). DOI: 10.7554/eLife.00792.
- [160] Bo Hu et al. “Visualization of the type III secretion sorting platform of *Shigella flexneri*.” *Proceedings of the National Academy of Sciences of the United States of America* (2015). DOI: 10.1073/pnas.1411610112.
- [161] T Kubori, S Yamaguchi, and S Aizawa. “Assembly of the switch complex onto the MS ring complex of *Salmonella typhimurium* does not require any other flagellar proteins.” *Journal of bacteriology* (1997).
- [162] Hajime Fukuoka, Yoshiyuki Sowa, Seiji Kojima, Akihiko Ishijima, and Michio Homma. “Visualization of functional rotor proteins of the bacterial flagellar motor in the cell membrane.” *Journal of molecular biology* (2007). DOI: 10.1016/j.jmb.2007.01.015.
- [163] Akihiro Kawamoto, Yusuke V Morimoto, Tomoko Miyata, Tohru Minamino, Kelly T Hughes, Takayuki Kato, and Keiichi Namba. “Common and distinct structural features of *Salmonella* injectisome and flagellar basal body.” en. *Scientific reports* (2013). DOI: 10.1038/srep03369.
- [164] Jonathan L McMurry, Tohru Minamino, Yukio Furukawa, Joshua W Francis, Stephanie A Hill, Katy A Helms, and Keiichi Namba. “Weak Interactions between *Salmonella enterica* FlhB and Other Flagellar Export Apparatus Proteins Govern Type III Secretion Dynamics.” *PloS one* (2015). DOI: 10.1371/journal.pone.0134884.
- [165] Fan Bai, Yusuke V Morimoto, Shinsuke D J Yoshimura, Noritaka Hara, Nobunori Kami-Ike, Keiichi Namba, and Tohru Minamino. “Assembly dynamics and the roles of FliI ATPase of the bacterial flagellar export apparatus.” en. *Scientific reports* (2014). DOI: 10.1038/srep06528.
- [166] Tohru Minamino, Yusuke V Morimoto, Noritaka Hara, Phillip D Aldridge, and Keiichi Namba. “The Bacterial Flagellar Type III Export Gate Complex Is a Dual Fuel Engine That Can Use Both H⁺ and Na⁺ for Flagellar Protein Export.” *PLoS pathogens* (2016). DOI: 10.1371/journal.ppat.1005495.
- [167] Tatsuya Ibuki, Katsumi Imada, Tohru Minamino, Takayuki Kato, Tomoko Miyata, and Keiichi Namba. “Common architecture of the flagellar type III protein export apparatus and F⁻ and V-type ATPases”. *Nature Structural & Molecular Biology* (2011). DOI: 10.1038/nsmb.1977.
- [168] Tatsuya Ibuki, Yumiko Uchida, Yusuke Hironaka, Keiichi Namba, Katsumi Imada, and Tohru Minamino. “Interaction between FliJ and FlhA, components of the bacterial flagellar type III export apparatus.” *Journal of bacteriology* (2013). DOI: 10.1128/JB.01711-12.
- [169] Tohru Minamino, Miki Kinoshita, Katsumi Imada, and Keiichi Namba. “Interaction between FliI ATPase and a flagellar chaperone FliT during bacterial flagellar protein export.” *Molecular microbiology* (2012). DOI: 10.1111/j.1365-2958.2011.07924.x.

- [170] Yukihiro Akeda and Jorge E. Galán. “Chaperone release and unfolding of substrates in type III secretion”. *Nature* (2005). DOI: 10.1038/nature03992.
- [171] Marc Erhardt, Max E Mertens, Florian D Fabiani, and Kelly T Hughes. “ATPase-independent type-III protein secretion in *Salmonella enterica*.” *PLoS genetics* (2014). DOI: 10.1371/journal.pgen.1004800.
- [172] Tohru Minamino, Yusuke V Morimoto, Noritaka Hara, and Keiichi Namba. “An energy transduction mechanism used in bacterial flagellar type III protein export.” *Nature communications* (2011). DOI: 10.1038/ncomms1488.
- [173] Tohru Minamino, Yusuke V Morimoto, Miki Kinoshita, Phillip D Aldridge, and Keiichi Namba. “The bacterial flagellar protein export apparatus processively transports flagellar proteins even with extremely infrequent ATP hydrolysis.” en. *Scientific reports* (2014). DOI: 10.1038/srep07579.
- [174] Yumiko Saijo-Hamano, Katsumi Imada, Tohru Minamino, May Kihara, Masafumi Shimada, Akio Kitao, and Keiichi Namba. “Structure of the cytoplasmic domain of FlhA and implication for flagellar type III protein export.” *Molecular microbiology* (2010). DOI: 10.1111/j.1365-2958.2010.07097.x.
- [175] Stanley A Moore and Yunhua Jia. “Structure of the cytoplasmic domain of the flagellar secretion apparatus component FlhA from *Helicobacter pylori*.” *The Journal of biological chemistry* (2010). DOI: 10.1074/jbc.M110.119412.
- [176] Miki Kinoshita, Noritaka Hara, Katsumi Imada, Keiichi Namba, and Tohru Minamino. “Interactions of bacterial flagellar chaperone-substrate complexes with FlhA contribute to co-ordinating assembly of the flagellar filament.” *Molecular microbiology* (2013). DOI: 10.1111/mmi.12430.
- [177] Tohru Minamino, Katsumi Imada, and Keiichi Namba. “Mechanisms of type III protein export for bacterial flagellar assembly.” *Molecular bioSystems* (2008). DOI: 10.1039/b808065h.
- [178] Hanna M. Singer, Marc Erhardt, and Kelly T. Hughes. “Comparative analysis of the secretion capability of early and late flagellar type III secretion substrates”. *Molecular Microbiology* (2014). DOI: 10.1111/mmi.12675.
- [179] Yumiko Saijo-Hamano, Naoko Uchida, Keiichi Namba, and Kenji Oosawa. “In Vitro Characterization of FlgB, FlgC, FlgF, FlgG, and FliE, Flagellar Basal Body Proteins of *Salmonella*”. *Journal of Molecular Biology* (2004). DOI: 10.1016/j.jmb.2004.03.070.
- [180] Takanori Hirano, Tohru Minamino, and Robert M Macnab. “The role in flagellar rod assembly of the N-terminal domain of *Salmonella* FlgJ, a flagellum-specific muramidase1”. *Journal of Molecular Biology* (2001). DOI: <http://dx.doi.org/10.1006/jmbi.2001.4963>.
- [181] Yohei Hizukuri, Seiji Kojima, and Michio Homma. “Disulphide cross-linking between the stator and the bearing components in the bacterial flagellar motor.” *Journal of biochemistry* (2010). DOI: 10.1093/jb/mvq067.
- [182] Mostyn T Brown et al. “Flagellar hook flexibility is essential for bundle formation in swimming *Escherichia coli* cells.” *Journal of bacteriology* (2012). DOI: 10.1128/JB.00209-12.

- [183] Fabienne F. V. Chevance and Kelly T. Hughes. “Coordinating assembly of a bacterial macromolecular machine”. *Nature Reviews Microbiology* (2008). DOI: 10.1038/nrmicro1887.
- [184] Takashi Fujii, Takayuki Kato, and Keiichi Namba. “Specific Arrangement of α -Helical Coiled Coils in the Core Domain of the Bacterial Flagellar Hook for the Universal Joint Function”. *Structure* (2009). DOI: 10.1016/j.str.2009.08.017.
- [185] Seiji Kojima and David F Blair. “Solubilization and purification of the MotA/MotB complex of Escherichia coli.” *Biochemistry* (2004). DOI: 10.1021/bi0354051.
- [186] D F Blair and H C Berg. “The MotA protein of E. coli is a proton-conducting component of the flagellar motor.” *Cell* (1990).
- [187] J Zhou and D F Blair. “Residues of the cytoplasmic domain of MotA essential for torque generation in the bacterial flagellar motor.” *Journal of molecular biology* (1997). DOI: 10.1006/jmbi.1997.1316.
- [188] Seiji Kojima. “Dynamism and regulation of the stator, the energy conversion complex of the bacterial flagellar motor”. *Current Opinion in Microbiology* (2015). DOI: 10.1016/j.mib.2015.07.015.
- [189] Eun A Kim, Marian Price-Carter, William C Carlquist, and David F Blair. “Membrane segment organization in the stator complex of the flagellar motor: implications for proton flow and proton-induced conformational change.” *Biochemistry* (2008). DOI: 10.1021/bi801347a.
- [190] S Kojima and D F Blair. “Conformational change in the stator of the bacterial flagellar motor.” *Biochemistry* (2001).
- [191] M L Wilson and R M Macnab. “Co-overproduction and localization of the Escherichia coli motility proteins motA and motB.” *Journal of bacteriology* (1990).
- [192] Edan R Hosking, Christian Vogt, Evert P Bakker, and Michael D Manson. “The Escherichia coli MotAB proton channel unplugged.” *Journal of molecular biology* (2006). DOI: 10.1016/j.jmb.2006.09.035.
- [193] Seiji Kojima et al. “Stator assembly and activation mechanism of the flagellar motor by the periplasmic region of MotB.” *Molecular microbiology* (2009). DOI: 10.1111/j.1365-2958.2009.06802.x.
- [194] Jenna O’Neill et al. “Role of the MotB linker in the assembly and activation of the bacterial flagellar motor”. *Acta Crystallographica Section D Biological Crystallography* (2011). DOI: 10.1107/S0907444911041102.
- [195] David J. Castillo et al. “The C-terminal periplasmic domain of MotB is responsible for load-dependent control of the number of stators of the bacterial flagellar motor”. *Biophysics* (2013).
- [196] Naoya Terahara, Motohiko Sano, and Masahiro Ito. “A Bacillus Flagellar Motor That Can Use Both Na⁺ and K⁺ as a Coupling Ion Is Converted by a Single Mutation to Use Only Na⁺”. *PLoS ONE* (2012). Ed. by Arnold Driessen. DOI: 10.1371/journal.pone.0046248.
- [197] Riku Imazawa et al. “A novel type bacterial flagellar motor that can use divalent cations as a coupling ion”. *Scientific Reports* (2016). DOI: 10.1038/srep19773.

- [198] Madoka Obara, Toshiharu Yakushi, Seiji Kojima, and Michio Homma. “Roles of charged residues in the C-terminal region of PomA, a stator component of the Na⁺-driven flagellar motor.” *Journal of bacteriology* (2008). DOI: 10.1128/JB.00849-07.
- [199] Yukako Asai, Toshiharu Yakushi, Ikuro Kawagishi, and Michio Homma. “Ion-coupling Determinants of Na⁺-driven and H⁺-driven Flagellar Motors”. *Journal of Molecular Biology* (2003). DOI: 10.1016/S0022-2836(03)00096-2.
- [200] Norihiro Takekawa et al. “Sodium-driven energy conversion for flagellar rotation of the earliest divergent hyperthermophilic bacterium”. *Scientific Reports* (2015). DOI: 10.1038/srep12711.
- [201] Masahiro Ito, Naoya Terahara, Shun Fujinami, and Terry Ann Krulwich. “Properties of Motility in *Bacillus subtilis* Powered by the H⁺-coupled MotAB Flagellar Stator, Na⁺-coupled MotPS or Hybrid Stators MotAS or MotPB”. *Journal of Molecular Biology* (2005). DOI: 10.1016/j.jmb.2005.07.030.
- [202] D F Blair and H C Berg. “Restoration of torque in defective flagellar motors.” *Science (New York, N.Y.)* (1988).
- [203] Mayuko Okabe, Toshiharu Yakushi, and Michio Homma. “Interactions of MotX with MotY and with the PomA/PomB Sodium Ion Channel Complex of the *Vibrio alginolyticus* Polar Flagellum”. *Journal of Biological Chemistry* (2005). DOI: 10.1074/jbc.M500263200.
- [204] Raquel M. Martinez, Brooke A. Jude, Thomas J. Kirn, Karen Skorupski, and Ronald K. Taylor. “Role of FlgT in anchoring the flagellum of *Vibrio cholerae*” (2010). DOI: 10.1128/JB.01562-09.
- [205] Hiroyuki Terashima, Hajime Fukuoka, Toshiharu Yakushi, Seiji Kojima, and Michio Homma. “The *Vibrio* motor proteins, MotX and MotY, are associated with the basal body of Na-driven flagella and required for stator formation.” *Molecular microbiology* (2006). DOI: 10.1111/j.1365-2958.2006.05435.x.
- [206] S. Fabela et al. “A Distant Homologue of the FlgT Protein Interacts with MotB and FliL and Is Essential for Flagellar Rotation in *Rhodobacter sphaeroides*”. *Journal of Bacteriology* (2013). DOI: 10.1128/JB.00760-13.
- [207] Yusuke V Morimoto, Shuichi Nakamura, Nobunori Kami-ike, Keiichi Namba, and Tohru Minamino. “Charged residues in the cytoplasmic loop of MotA are required for stator assembly into the bacterial flagellar motor.” *Molecular microbiology* (2010). DOI: 10.1111/j.1365-2958.2010.07391.x.
- [208] Seiji Kojima, Natsumi Nonoyama, Norihiro Takekawa, Hajime Fukuoka, and Michio Homma. “Mutations Targeting the C-Terminal Domain of FliG Can Disrupt Motor Assembly in the Na⁺-Driven Flagella of *Vibrio alginolyticus*”. *Journal of Molecular Biology* (2011). DOI: 10.1016/j.jmb.2011.09.019.
- [209] Sophie Roure et al. “Peptidoglycan maturation enzymes affect flagellar functionality in bacteria”. *Molecular Microbiology* (2012). DOI: 10.1111/mmi.12019.
- [210] Xin Fang and Mark Gomelsky. “A post-translational, c-di-GMP-dependent mechanism regulating flagellar motility.” *Molecular microbiology* (2010). DOI: 10.1111/j.1365-2958.2010.07179.x.

- [211] Amy E. Baker et al. "A PilZ domain protein FlgZ mediates c-di-GMP-dependent swarming motility control in *Pseudomonas aeruginosa*". *Journal of Bacteriology* (2016). Ed. by T. J. Silhavy. DOI: 10.1128/JB.00196-16.
- [212] Alex Boehm et al. "Second messenger-mediated adjustment of bacterial swimming velocity." *Cell* (2010). DOI: 10.1016/j.cell.2010.01.018.
- [213] Koushik Paul, Vincent Nieto, William C Carlquist, David F Blair, and Rasika M Harshey. "The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism." *Molecular cell* (2010). DOI: 10.1016/j.molcel.2010.03.001.
- [214] Koushik Paul, William C Carlquist, and David F Blair. "Adjusting the spokes of the flagellar motor with the DNA-binding protein H-NS." *Journal of bacteriology* (2011). DOI: 10.1128/JB.05458-11.
- [215] Mostyn T Brown, Nicolas J Delalez, and Judith P Armitage. "Protein dynamics and mechanisms controlling the rotational behaviour of the bacterial flagellar motor." *Current opinion in microbiology* (2011). DOI: 10.1016/j.mib.2011.09.009.
- [216] Galit N Cohen-Ben-Lulu et al. "The bacterial flagellar switch complex is getting more complex." *The EMBO journal* (2008). DOI: 10.1038/emboj.2008.48.
- [217] R Barak and M Eisenbach. "Fumarate or a fumarate metabolite restores switching ability to rotating flagella of bacterial envelopes." *Journal of bacteriology* (1992).
- [218] Rina Barak, Igal Giebel, and Michael Eisenbach. "The specificity of fumarate as a switching factor of the bacterial flagellar motor". *Molecular Microbiology* (1996). DOI: 10.1046/j.1365-2958.1996.365889.x.
- [219] K Prasad, S R Caplan, and M Eisenbach. "Fumarate modulates bacterial flagellar rotation by lowering the free energy difference between the clockwise and counterclockwise states of the motor." *Journal of molecular biology* (1998). DOI: 10.1006/jmbi.1998.1922.
- [220] P Bertin, E Terao, E H Lee, P Lejeune, C Colson, A Danchin, and E Collatz. "The H-NS protein is involved in the biogenesis of flagella in *Escherichia coli*." *Journal of bacteriology* (1994).
- [221] M Ko and C Park. "Two novel flagellar components and H-NS are involved in the motor function of *Escherichia coli*." *Journal of molecular biology* (2000). DOI: 10.1006/jmbi.2000.4147.
- [222] G M Donato, M J Lelivelt, and T H Kawula. "Promoter-specific repression of *fimB* expression by the *Escherichia coli* nucleoid-associated protein H-NS." *Journal of bacteriology* (1997).
- [223] Eun A Kim and David F Blair. "Function of the Histone-Like Protein H-NS in Motility of *Escherichia coli*: Multiple Regulatory Roles Rather than Direct Action at the Flagellar Motor." *Journal of bacteriology* (2015). DOI: 10.1128/JB.00309-15.
- [224] Shiwei Zhu, Ananthanarayanan Kumar, Seiji Kojima, and Michio Homma. "FliL associates with the stator to support torque generation of the sodium-driven polar flagellar motor of *Vibrio*." *Molecular microbiology* (2015). DOI: 10.1111/mmi.13103.

- [225] M A Motaleb, Joshua E Pitzer, Syed Z Sultan, and Jun Liu. “A novel gene inactivation system reveals altered periplasmic flagellar orientation in a *Borrelia burgdorferi* fliL mutant.” *Journal of bacteriology* (2011). DOI: 10.1128/JB.00202-11.
- [226] G J Schoenhals and R M Macnab. “FliL is a membrane-associated component of the flagellar basal body of *Salmonella*.” *Microbiology (Reading, England)* (1999). DOI: 10.1099/13500872-145-7-1769.
- [227] Jonathan D Partridge, Vincent Nieto, and Rasika M Harshey. “A new player at the flagellar motor: FliL controls both motor output and bias.” *mBio* (2015). DOI: 10.1128/mBio.02367-14.
- [228] Ursula Attmannspacher, Birgit E Scharf, and Rasika M Harshey. “FliL is essential for swarming: motor rotation in absence of FliL fractures the flagellar rod in swarmer cells of *Salmonella enterica*.” *Molecular microbiology* (2008). DOI: 10.1111/j.1365-2958.2008.06170.x.
- [229] Phillip Aldridge and Urs Jenal. “Cell cycle-dependent degradation of a flagellar motor component requires a novel-type response regulator”. *Molecular Microbiology* (1999). DOI: mmi1358[pii].
- [230] Fernando Suaste-Olmos, Clelia Domenzain, José Cruz Mireles-Rodríguez, Sebastian Poggio, Aurora Osorio, Georges Dreyfus, and Laura Camarena. “The flagellar protein FliL is essential for swimming in *Rhodobacter sphaeroides*.” *Journal of bacteriology* (2010). DOI: 10.1128/JB.00655-10.
- [231] Yi-Ying Lee and Robert Belas. “Loss of FliL Alters *Proteus mirabilis* Surface Sensing and Temperature-Dependent Swarming”. *Journal of Bacteriology* (2015). Ed. by J. P. Armitage. DOI: 10.1128/JB.02235-14.
- [232] Robert Belas and Rooge Suvanasthi. “The ability of *Proteus mirabilis* to sense surfaces and regulate virulence gene expression involves FliL, a flagellar basal body protein.” *Journal of bacteriology* (2005). DOI: 10.1128/JB.187.19.6789-6803.2005.
- [233] Kathleen Cusick, Yi-Ying Lee, Brian Youchak, and Robert Belas. “Perturbation of FliL interferes with *Proteus mirabilis* swarmer cell gene expression and differentiation.” *Journal of bacteriology* (2012). DOI: 10.1128/JB.05998-11.
- [234] Yi-Ying Lee, Julius Patellis, and Robert Belas. “Activity of *Proteus mirabilis* FliL Is Viscosity Dependent and Requires Extragenic DNA”. *JOURNAL OF BACTERIOLOGY* (2013). DOI: 10.1128/JB.02024.12.
- [235] Sarah B Guttenplan, Kris M Blair, and Daniel B Kearns. “The EpsE flagellar clutch is bifunctional and synergizes with EPS biosynthesis to promote *Bacillus subtilis* biofilm formation.” *PLoS genetics* (2010). DOI: 10.1371/journal.pgen.1001243.
- [236] Kris M Blair, Linda Turner, Jared T Winkelman, Howard C Berg, and Daniel B Kearns. “A molecular clutch disables flagella in the *Bacillus subtilis* biofilm.” *Science (New York, N.Y.)* (2008). DOI: 10.1126/science.1157877.
- [237] Gabriel Zarbiv, Hui Li, Amnon Wolf, Gary Cecchini, S Roy Caplan, Victor Sourjik, and Michael Eisenbach. “Energy complexes are apparently associated with the switch-motor complex of bacterial flagella.” *Journal of molecular biology* (2012). DOI: 10.1016/j.jmb.2011.12.027.

- [238] M Charl Moolman, Jacob W J Kerssemakers, and Nynke H Dekker. “Quantitative Analysis of Intracellular Fluorescent Foci in Live Bacteria.” English. *Biophysical journal* (2015). DOI: 10.1016/j.bpj.2015.07.013.
- [239] Robert Crawford, Joseph P. Torella, Louise Aigrain, Anne Plochowitz, Kristofer Gryte, Stephan Uphoff, and Achillefs N. Kapanidis. “Long-Lived Intracellular Single-Molecule Fluorescence Using Electroporated Molecules”. *Biophysical Journal* (2013). DOI: 10.1016/j.bpj.2013.09.057.
- [240] Rahul Roy, Sungchul Hohng, and Taekjip Ha. “A practical guide to single-molecule FRET.” *Nature methods* (2008). DOI: 10.1038/nmeth.1208.
- [241] Daniele N Selmi, Roslin J Adamson, Helen Attrill, Alan D Goddard, Robert J C Gilbert, Anthony Watts, and Andrew J Turberfield. “DNA-templated protein arrays for single-molecule imaging.” *Nano letters* (2011). DOI: 10.1021/nl1037769.
- [242] Nadrian C. Seeman. “Nucleic acid junctions and lattices”. *Journal of Theoretical Biology* (1982). DOI: 10.1016/0022-5193(82)90002-9.
- [243] Junghuei Chen and Nadrian C. Seeman. “Synthesis from DNA of a molecule with the connectivity of a cube”. *Nature* (1991). DOI: 10.1038/350631a0.
- [244] R P Goodman, I a T Schaap, C F Tardin, C M Erben, R M Berry, C F Schmidt, and a J Turberfield. “Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication.” *Science* (2005). DOI: 10.1126/science.1120367.
- [245] Erik Winfree, Furong Liu, Lisa A. Wenzler, and Nadrian C. Seeman. “Design and self-assembly of two-dimensional DNA crystals”. *Nature* (1998). DOI: 10.1038/28998.
- [246] Paul W K Rothemund, Axel Ekani-Nkodo, Nick Papadakis, Ashish Kumar, Deborah Kuchnir Fyngenson, and Erik Winfree. “Design and characterization of programmable DNA nanotubes.” *Journal of the American Chemical Society* (2004). DOI: 10.1021/ja0443191.
- [247] Patrick O’Neill, Paul W K Rothemund, Ashish Kumar, and D K Fyngenson. “Sturdier DNA nanotubes via ligation.” *Nano letters* (2006). DOI: 10.1021/nl0603505.
- [248] Fei Zhang, Jeanette Nangreave, Yan Liu, and Hao Yan. *Structural DNA nanotechnology: State of the art and future perspective*. 2014. DOI: 10.1021/ja505101a.
- [249] Paul Rothemund. “Folding DNA to create nanoscale shapes and patterns”. *Nature* (2006). DOI: doi:10.1038/nature04586.
- [250] Hendrik Dietz, Shawn M Douglas, and William M Shih. “Folding DNA into twisted and curved nanoscale shapes.” *Science (New York, N. Y.)* (2009). DOI: 10.1126/science.1174251.
- [251] Erik Benson, Abdulmelik Mohammed, Johan Gardell, Sergej Masich, Eugen Czeizler, Pekka Orponen, and Björn Högberg. “DNA rendering of polyhedral meshes at the nanoscale”. *Nature* (2015). DOI: 10.1038/nature14586.
- [252] Yonggang Ke, Luvena L Ong, William M Shih, and Peng Yin. “Three-dimensional structures self-assembled from DNA bricks.” *Science (New York, N. Y.)* (2012). DOI: 10.1126/science.1227268.

- [253] Sung Yong Park, Abigail K. R. Lytton-Jean, Byeongdu Lee, Steven Weigand, George C. Schatz, and Chad A. Mirkin. “DNA-programmable nanoparticle crystallization”. *Nature* (2008). DOI: 10.1038/nature06508.
- [254] Rui Tan et al. “Multi-component superstructures self-assembled from nanocrystal building blocks”. *Nanoscale* (2016). DOI: 10.1039/C6NR01662F.
- [255] Robert J Macfarlane et al. “Nanoparticle superlattice engineering with DNA.” *Science (New York, N.Y.)* (2011). DOI: 10.1126/science.1210493.
- [256] Yugang Zhang, Suchetan Pal, Babji Srinivasan, Thi Vo, Sanat Kumar, and Oleg Gang. “Selective transformations between nanoparticle superlattices via the reprogramming of DNA-mediated interactions”. *Nature Materials* (2015). DOI: 10.1038/nmat4296.
- [257] James T. McGinley, Ian Jenkins, Talid Sinno, and John C. Crocker. “Assembling colloidal clusters using crystalline templates and reprogrammable DNA interactions”. *Soft Matter* (2013). DOI: 10.1039/c3sm50950h.
- [258] Jiwen Zheng, Pamela E. Constantinou, Christine Micheel, A. Paul Alivisatos, Richard A. Kiehl, and Nadrian C. Seeman. “Two-dimensional nanoparticle arrays show the organizational power of robust DNA motifs”. *Nano Letters* (2006). DOI: 10.1021/nl060994c.
- [259] Jaswinder Sharma, Rahul Chhabra, Yan Liu, Yonggang Ke, and Hao Yan. “DNA-Templated Self-Assembly of Two-Dimensional and Periodical Gold Nanoparticle Arrays”. *Angewandte Chemie International Edition* (2006). DOI: 10.1002/anie.200503208.
- [260] J Sharma, R Chhabra, A Cheng, and J Brownell. “Control of self-assembly of DNA tubules through integration of gold nanoparticles”. *Science* (2009).
- [261] Robert Schreiber, Ibon Santiago, Arzhang Ardavan, and Andrew J. Turberfield. “Ordering Gold Nanoparticles with DNA Origami Nanoflowers”. *ACS Nano* (2016). DOI: 10.1021/acsnano.6b03076.
- [262] Colin J. Loweth, W. Brett Caldwell, Xiaogang Peng, A. Paul Alivisatos, and Peter G. Schultz. “DNA-Based Assembly of Gold Nanocrystals”. *Angewandte Chemie (International ed. in English)* (1999). DOI: 10.1002/(SICI)1521-3773(19990614)38:12<1808::AID-ANIE1808>3.0.CO;2-C.
- [263] Zhaoxiang Deng, Ye Tian, Seung Hyun Lee, Alexander E. Ribbe, and Chengde Mao. “DNA-encoded self-assembly of gold nanoparticles into one-dimensional arrays”. *Angewandte Chemie - International Edition* (2005). DOI: 10.1002/anie.200463096.
- [264] Baoquan Ding, Zhengtao Deng, Hao Yan, Stefano Cabrini, and Ronald N. Zuckermann. “Supporting Online Information Gold Nanoparticles Self-similar Chain Structure Organized by DNA Origami”. *Journal of the American Chemical Society* (2010). DOI: 10.1021/ja9101198.
- [265] Suchetan Pal, Zhengtao Deng, Haining Wang, Shengli Zou, Yan Liu, and Hao Yan. “Communication DNA Directed Self-assembly of Anisotropic Plasmonic Nanostructures DNA Directed Self-assembly of Anisotropic Plasmonic Nanostructures”. *October* (2011). DOI: 10.1021/ja207898r.

- [266] G P Acuna, F M Möller, P Holzmeister, S Beater, B Lalkens, and P Tinnefeld. “Fluorescence Enhancement at Docking Sites of DNA-Directed Self-Assembled Nanoantennas”. *Science* (2012). DOI: doi:10.1126/science.1228638.
- [267] Vivek V Thacker, Lars O Herrmann, Daniel O Sigle, Tao Zhang, Tim Liedl, Jeremy J Baumberg, and Ulrich F Keyser. “DNA origami based assembly of gold nanoparticle dimers for surface-enhanced Raman scattering.” *Nature communications* (2014). DOI: 10.1038/ncomms4448.
- [268] Anastasiya Puchkova, Carolin Vietz, Enrico Pibiri, Bettina Wünsch, María Sanz Paz, Guillermo P Acuna, and Philip Tinnefeld. “DNA Origami Nanoantennas with over 5000-fold Fluorescence Enhancement and Single-Molecule Detection at 25 μm ”. *Nano Letters* (2015). DOI: 10.1021/acs.nanolett.5b04045.
- [269] Anton Kuzyk et al. “DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response”. *Nature* (2012). DOI: 10.1038/nature10889. arXiv: 1108.3752.
- [270] Palash K. Dutta, Reji Varghese, Jeanette Nangreave, Su Lin, Hao Yan, and Yan Liu. “DNA-directed artificial light-harvesting antenna”. *Journal of the American Chemical Society* (2011). DOI: 10.1021/ja1115138.
- [271] Ingo H. Stein, Christian Steinhauer, and Philip Tinnefeld. “Single-molecule four-color FRET visualizes energy-transfer paths on DNA origami”. *Journal of the American Chemical Society* (2011). DOI: 10.1021/ja1105464.
- [272] Ingo H. Stein, Verena Schüller, Philip Böhm, Philip Tinnefeld, and Tim Liedl. “Single-molecule FRET ruler based on rigid DNA origami blocks”. *ChemPhysChem* (2011). DOI: 10.1002/cphc.201000781.
- [273] Ralf Jungmann et al. “Quantitative super-resolution imaging with qPAINT.” *Nature methods* (2016). DOI: 10.1038/nmeth.3804.
- [274] Anton Kurz, Jürgen J Schmied, Kristin S Großmayer, Phil Holzmeister, Philip Tinnefeld, and Dirk-Peter Herten. “Counting fluorescent dye molecules on DNA origami by means of photon statistics.” *Small (Weinheim an der Bergstrasse, Germany)* (2013). DOI: 10.1002/smll.201300619.
- [275] Haisen Ta et al. “Mapping molecules in scanning far-field fluorescence nanoscopy”. *Nature Communications* (2015). DOI: 10.1038/ncomms8977.
- [276] Jürgen J Schmied, Andreas Gietl, Phil Holzmeister, Carsten Forthmann, Christian Steinhauer, Thorben Dammeyer, and Philip Tinnefeld. “Fluorescence and super-resolution standards based on DNA origami”. *Nature Methods* (2012). DOI: 10.1038/nmeth.2254.
- [277] Cliff I. Stains, Jason R. Porter, Aik T. Ooi, David J. Segal, and Indraneel Ghosh. “DNA sequence-enabled reassembly of the green fluorescent protein”. *Journal of the American Chemical Society* (2005). DOI: 10.1021/ja051969w.
- [278] Núria Sancho Oltra, Jeffrey Bos, and Gerard Roelfes. “Control over Enzymatic Activity by DNA-Directed Split Enzyme Reassembly”. *ChemBioChem* (2010). DOI: 10.1002/cbic.201000517.
- [279] Michael Erkelenz, Chi Hsien Kuo, and Christof M. Niemeyer. “DNA-mediated assembly of cytochrome P450 BM3 subdomains”. *Journal of the American Chemical Society* (2011). DOI: 10.1021/ja204993s.

- [280] Minghui Liu et al. “A DNA tweezer-actuated enzyme nanoreactor.” *Nature Communications* (2013). DOI: 10.1038/ncomms3127.
- [281] Alan Saghatelian, Kevin M Guckian, Desiree a Thayer, and M Reza Ghadiri. “Communication DNA Detection and Signal Amplification via an Engineered Allosteric Enzyme DNA Detection and Signal Amplification via an Engineered Allosteric Enzyme”. *Society* (2003). DOI: 10.1021/ja027885u.
- [282] Chiao Yu Tseng and Giovanni Zocchi. “Mechanical control of renilla luciferase”. *Journal of the American Chemical Society* (2013). DOI: 10.1021/ja4043565.
- [283] Berea A. R. Williams, Chris W. Diehnelt, Paul Belcher, Matthew Greving, Neal W. Woodbury, Stephen A. Johnston, and John C. Chaput. “Creating protein affinity reagents by combining peptide ligands on synthetic DNA scaffolds”. *Journal of the American Chemical Society* (2009). DOI: 10.1021/ja9051735. arXiv: NIHMS150003.
- [284] Chao Zhou, Zhongqiang Yang, and Dongsheng Liu. “Reversible regulation of protein binding affinity by a DNA machine”. *Journal of the American Chemical Society* (2012). DOI: 10.1021/ja209590u.
- [285] Hendrik Eberhard, Franziska Diezmann, and Oliver Seitz. “DNA as a Molecular Ruler: Interrogation of a Tandem SH2 Domain with Self-Assembled, Bivalent DNA-Peptide Complexes”. *Angewandte Chemie International Edition* (2011). DOI: 10.1002/anie.201007593.
- [286] Sherri Rinker, Yonggang Ke, Yan Liu, Rahul Chhabra, and Hao Yan. “Self-assembled DNA nanostructures for distance-dependent multivalent ligand–protein binding”. *Nature Nanotechnology* (2008). DOI: 10.1038/nnano.2008.164.
- [287] Ofer I. Wilner, Yossi Weizmann, Ron Gill, Oleg Lioubashevski, Ronit Freeman, and Itamar Willner. “Enzyme cascades activated on topologically programmed DNA scaffolds.” *Nature nanotechnology* (2009). DOI: 10.1038/nnano.2009.50.
- [288] Jinglin Fu, Minghui Liu, Yan Liu, Neal W. Woodbury, and Hao Yan. “Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable DNA nanostructures”. *Journal of the American Chemical Society* (2012). DOI: 10.1021/ja300897h.
- [289] Veikko Linko, Marika Eerikäinen, and Mauri A. Kostiainen. “A modular DNA origami-based enzyme cascade nanoreactor”. *Chem. Commun.* (2015). DOI: 10.1039/C4CC08472A.
- [290] Yanming Fu et al. “Single-step rapid assembly of DNA origami nanostructures for addressable nanoscale bioreactors.” *Journal of the American Chemical Society* (2013). DOI: 10.1021/ja3076692.
- [291] Jinglin Fu et al. “Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm.” *Nature nanotechnology* (2014). DOI: 10.1038/nnano.2014.100.
- [292] Palash K. Dutta et al. “A DNA-directed light-harvesting/reaction center system”. *Journal of the American Chemical Society* (2014). DOI: 10.1021/ja509018g.

- [293] N D Derr et al. “Tug-of-war in motor protein ensembles revealed with a programmable DNA origami scaffold.” *Science (New York, N.Y.)* (2012). DOI: 10.1126/science.1226734.
- [294] Rizal F. Hariadi, Mario Cale, and Sivaramakrishnan. “Myosin lever arm directs collective motion on cellular actin network”. *Proceedings of the National Academy of Sciences* (2014). DOI: 10.1073/pnas.1315923111.
- [295] Adam J. M. Wollman, Carlos Sanchez-Cano, Helen M. J. Carstairs, Robert A. Cross, and Andrew J. Turberfield. “Transport and self-organization across different length scales powered by motor proteins and programmed by DNA.” *Nature nanotechnology* (2014). DOI: 10.1038/nnano.2013.230.
- [296] Anuttara Udomprasert et al. “Amyloid fibrils nucleated and organized by DNA origami constructions”. *Nature Nanotechnology* (2014). DOI: 10.1038/nnano.2014.102.
- [297] Hui Zhang and Peixuan Guo. “Single molecule photobleaching (SMPB) technology for counting of RNA, DNA, protein and other molecules in nanoparticles and biological complexes by TIRF instrumentation.” *Methods (San Diego, Calif.)* (2014). DOI: 10.1016/j.ymeth.2014.01.010.
- [298] Valerie C. Coffman and Jian-Qiu Wu. “Counting protein molecules using quantitative fluorescence microscopy”. *Trends in Biochemical Sciences* (2012). DOI: 10.1016/j.tibs.2012.08.002.
- [299] Paola Bisicchia, Bradley Steel, and Mekdes H Mariam Debela. “Translocase FtsK Hexamerizes at”. *mBio* (2013). DOI: 10.1128/mBio.00800-13.Editor.
- [300] S. H. Chung and R. A. Kennedy. “Forward-backward non-linear filtering technique for extracting small biological signals from noise”. *Journal of Neuroscience Methods* (1991). DOI: 10.1016/0165-0270(91)90118-J.
- [301] Hugo McGuire, Mark R P Arousseau, Derek Bowie, and Rikard Blunck. “Automating single subunit counting of membrane proteins in mammalian cells.” *The Journal of biological chemistry* (2012). DOI: 10.1074/jbc.M112.402057.
- [302] C. Liesche et al. “Automated Analysis of Single Molecule Photobleaching Data by Statistical Modeling of Spot Populations”. *Biophysical Journal* (2015). DOI: 10.1016/j.bpj.2015.10.035.
- [303] H. Deschout, A. Shivanandan, P. Annibale, M. Scarselli, and A. Radenovic. *Progress in quantitative single-molecule localization microscopy*. 2014. DOI: 10.1007/s00418-014-1217-y.
- [304] Sang-Hyuk Lee, Jae Yen Shin, Antony Lee, and Carlos Bustamante. “Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM).” *Proceedings of the National Academy of Sciences of the United States of America* (2012). DOI: 10.1073/pnas.1215175109.
- [305] R Jungmann, C Steinhauer, and M Scheible. “Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami”. *Nano ...* (2010). DOI: 10.1021/nl103427w.
- [306] Carlos Ernesto Castro et al. “A primer to scaffolded DNA origami”. *Nature methods* (2011). DOI: 10.1038/NMETH.1570.

- [307] JN Zadeh and CD Steenberg. “NUPACK: analysis and design of nucleic acid systems”. *Journal of Computational Chemistry* (2011). DOI: 10.1002/jcc.
- [308] Caroline Bardin and Jean Louis Leroy. “The formation pathway of tetramolecular G-quadruplexes”. *Nucleic Acids Research* (2008). DOI: 10.1093/nar/gkm1050.
- [309] Oliver Stegle, Linda Payet, Jean Louis Mergny, David J C MacKay, and Julian Leon Huppert. “Predicting and understanding the stability of G-quadruplexes”. In: *Bioinformatics*. Oxford University Press, 2009. DOI: 10.1093/bioinformatics/btp210.
- [310] L. D’Antonio and P. Bagga. “Computational methods for predicting intramolecular G-quadruplexes in nucleotide sequences”. *Proceedings. 2004 IEEE Computational Systems Bioinformatics Conference, 2004. CSB 2004.* (2004). DOI: 10.1109/CSB.2004.1332508.
- [311] Shelley F J Wickham, Masayuki Endo, Yousuke Katsuda, Kumi Hidaka, Jonathan Bath, Hiroshi Sugiyama, and Andrew J Turberfield. “Direct observation of stepwise movement of a synthetic molecular transporter.” *Nature nanotechnology* (2011). DOI: 10.1038/nnano.2010.284.
- [312] Evi Stahl, Thomas G Martin, Florian Praetorius, and Hendrik Dietz. “Facile and scalable preparation of pure and dense DNA origami solutions”. *Angewandte Chemie - International Edition* (2014). DOI: 10.1002/anie.201405991. arXiv: arXiv:1408.1149.
- [313] Ryo Iizuka, Mai Yamagishi-Shirasaki, and Takashi Funatsu. “Kinetic study of de novo chromophore maturation of fluorescent proteins.” *Analytical biochemistry* (2011). DOI: 10.1016/j.ab.2011.03.036.
- [314] Somes K Das, Manjula Darshi, Stephen Cheley, Mark I Wallace, and Hagan Bayley. “Membrane protein stoichiometry determined from the step-wise photobleaching of dye-labelled subunits.” *Chembiochem : a European journal of chemical biology* (2007). DOI: 10.1002/cbic.200600474.
- [315] Dan Shu, Hui Zhang, Jiashun Jin, and Peixuan Guo. “Counting of six pRNAs of phi29 DNA-packaging motor with customized single-molecule dual-view system.” *The EMBO journal* (2007). DOI: 10.1038/sj.emboj.7601506.
- [316] Matthew a. B. Baker, Nejc Rojko, Bríd Cronin, Gregor Anderluh, and Mark I. Wallace. “Photobleaching Reveals Heterogeneous Stoichiometry for Equinatoxin II Oligomers”. *ChemBioChem* (2014). DOI: 10.1002/cbic.201300799.
- [317] John M. Walker. “The bicinchoninic acid (BCA) assay for protein quantitation.” *Methods in molecular biology (Clifton, N.J.)* (1994). DOI: 10.1385/0-89603-268-X:5.
- [318] Mary Katherine Johansson, Henk Fidler, Daren Dick, and Ronald M Cook. “Intramolecular dimers: a new strategy to fluorescence quenching in dual-labeled oligonucleotide probes.” *Journal of the American Chemical Society* (2002).
- [319] Beverly Z Packard, Dmitri D Toptygin, Akira Komoriya, and Ludwig Brand. “Intramolecular resonance dipole-dipole interactions in a profluorescent protease substrate”. *Journal of Physical Chemistry B* (1998). DOI: 10.1021/jp972845b.

- [320] a P Wei, D K Blumenthal, and J N Herron. "Antibody-mediated fluorescence enhancement based on shifting the intramolecular dimer \leftrightarrow monomer equilibrium of fluorescent dyes." *Analytical chemistry* (1994).
- [321] C A Royer. "Fluorescence spectroscopy." *Methods in molecular biology* (Clifton, N.J.) (1995). DOI: 10.1385/0-89603-301-5:65.
- [322] Hoi Sung Chung, John M Louis, and William a Eaton. "Distinguishing between protein dynamics and dye photophysics in single-molecule FRET experiments." *Biophysical journal* (2010). DOI: 10.1016/j.bpj.2009.12.4322.
- [323] Huimin Chen, Syed S Ahsan, Mitk'El B Santiago-Berrios, Hector D Abruña, and Watt W Webb. "Mechanisms of quenching of Alexa fluorophores by natural amino acids." *Journal of the American Chemical Society* (2010). DOI: 10.1021/ja100500k.
- [324] Sören Doose, Hannes Neuweiler, and Markus Sauer. "Fluorescence quenching by photoinduced electron transfer: a reporter for conformational dynamics of macromolecules." *Chemphyschem : a European journal of chemical physics and physical chemistry* (2009). DOI: 10.1002/cphc.200900238.
- [325] Sören Doose, Hannes Neuweiler, and Markus Sauer. "A close look at fluorescence quenching of organic dyes by tryptophan." *Chemphyschem : a European journal of chemical physics and physical chemistry* (2005). DOI: 10.1002/cphc.200500191.
- [326] Marcia Levitus and Suman Ranjit. "Cyanine dyes in biophysical research: the photophysics of polymethine fluorescent dyes in biomolecular environments." *Quarterly reviews of biophysics* (2011). DOI: 10.1017/S0033583510000247.
- [327] Nicole Marmé, Jens-Peter Knemeyer, Markus Sauer, and Jürgen Wolfrum. "Inter- and intramolecular fluorescence quenching of organic dyes by tryptophan." *Bioconjugate chemistry* (2003). DOI: 10.1021/bc0341324.
- [328] R M Watt and E W Voss. "Mechanism of quenching of fluorescein by anti-fluorescein IgG antibodies." *Immunochemistry* (1977).
- [329] Anu Bamgbelu, Jing Wang, and Jerzy Leszczynski. "TDDFT study of the optical properties of Cy5 and its derivatives". *Journal of Physical Chemistry A* (2010). DOI: 10.1021/jp908485z.
- [330] Alexander Gust et al. "A Starting Point for Fluorescence-Based Single-Molecule Measurements in Biomolecular Research." *Molecules (Basel, Switzerland)* (2014). DOI: 10.3390/molecules191015824.
- [331] Stephen E Leonard and Kate S Carroll. "Chemical 'omics' approaches for understanding protein cysteine oxidation in biology." *Current opinion in chemical biology* (2011). DOI: 10.1016/j.cbpa.2010.11.012.
- [332] CM Furdui and LB Poole. "Chemical approaches to detect and analyze protein sulfenic acids". *Mass spectrometry reviews* (2013). DOI: 10.1002/mas.
- [333] Christine C Winterbourn and Mark B Hampton. "Thiol chemistry and specificity in redox signaling." *Free radical biology & medicine* (2008). DOI: 10.1016/j.freeradbiomed.2008.05.004.

- [334] Michael Hamann, Tiequan Zhang, Suzanne Hendrich, and James A. Thomas. “[15] Quantitation of protein sulfinic and sulfonic acid, irreversibly oxidized protein cysteine sites in cellular proteins”. *Methods in Enzymology* (2002). Ed. by Helmut Sies and Lester Packer. DOI: [http://dx.doi.org/10.1016/S0076-6879\(02\)48634-X](http://dx.doi.org/10.1016/S0076-6879(02)48634-X).
- [335] Alexander Johnson-Buck, Jeanette Nangreave, Do-Nyun Kim, Mark Bathe, Hao Yan, and Nils G Walter. “Super-resolution fingerprinting detects chemical reactions and idiosyncrasies of single DNA pegboards.” *Nano letters* (2013). DOI: 10.1021/nl304415b.
- [336] Max B. Scheible, Günther Pardatscher, Anton Kuzyk, and Friedrich C. Simmel. “Single molecule characterization of DNA binding and strand displacement reactions on lithographic DNA origami microarrays”. *Nano Letters* (2014). DOI: 10.1021/nl500092j.
- [337] Kemper Talley and Emil Alexov. “On the pH-optimum of activity and stability of proteins”. *Proteins: Structure, Function and Bioinformatics* (2010). DOI: 10.1002/prot.22786. arXiv: NIHMS150003.
- [338] Mei Chu Lo, Ann Aulabaugh, Guixian Jin, Rebecca Cowling, Jonathan Bard, Michael Malamas, and George Ellestad. “Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery”. *Analytical Biochemistry* (2004). DOI: 10.1016/j.ab.2004.04.031.
- [339] Guillermo A Senisterra and Patrick J Finerty. “High throughput methods of assessing protein stability and aggregation.” *Molecular bioSystems* (2009). DOI: 10.1039/b814377c.
- [340] S Cayley, B a Lewis, H J Guttman, and M T Record. “Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity. Implications for protein-DNA interactions in vivo.” *Journal of molecular biology* (1991).
- [341] S Leirimo, C Harrison, D S Cayley, R R Burgess, and M T Record. “Replacement of potassium chloride by potassium glutamate dramatically enhances protein-DNA interactions in vitro.” *Biochemistry* (1987).
- [342] M T Record, E S Courtenay, S Cayley, and H J Guttman. “Biophysical compensation mechanisms buffering E. coli protein-nucleic acid interactions against changing environments.” *Trends in biochemical sciences* (1998).
- [343] Irina M. Kuznetsova, Konstantin K. Turoverov, and Vladimir N. Uversky. *What macromolecular crowding can do to a protein*. 2014. DOI: 10.3390/ijms151223090.
- [344] Yael Phillip and Gideon Schreiber. *Formation of protein complexes in crowded environments-From in vitro to in vivo*. 2013. DOI: 10.1016/j.febslet.2013.01.007.
- [345] Franziska Diezmann and Oliver Seitz. “DNA-guided display of proteins and protein ligands for the interrogation of biology.” *Chemical Society reviews* (2011). DOI: 10.1039/c1cs15054e.
- [346] Christof M Niemeyer. “Semisynthetic DNA-protein conjugates for biosensing and nanofabrication.” *Angewandte Chemie (International ed. in English)* (2010). DOI: 10.1002/anie.200904930.

- [347] Barbara Saccà and Christof M Niemeyer. “Functionalization of DNA nanostructures with proteins.” en. *Chemical Society reviews* (2011). DOI: 10.1039/c1cs15212b.
- [348] C M Niemeyer, T Sano, C L Smith, and C R Cantor. “Oligonucleotide-directed self-assembly of proteins: semisynthetic DNA–streptavidin hybrid molecules as connectors for the generation of macroscopic arrays and the construction of supramolecular bioconjugates.” *Nucleic Acids Research* (1994). DOI: 10.1093/nar/22.25.5530.
- [349] Barbara Saccà et al. “Orthogonal protein decoration of DNA origami.” *Angewandte Chemie (International ed. in English)* (2010). DOI: 10.1002/anie.201005931.
- [350] Sandra Harper and David W Speicher. “Purification of Proteins Fused to Glutathione S-Transferase”. In: *Methods in molecular biology (Clifton, N.J.)* NIH Public Access, 2011. DOI: 10.1007/978-1-60761-913-0.
- [351] Bijan Zakeri, Jacob O. Fierer, Emrah Celik, Emily C. Chittock, Ulrich Schwarz-Linek, Vincent T. Moy, and Mark Howarth. “Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin.” *Proceedings of the National Academy of Sciences of the United States of America* (2012). DOI: 10.1073/pnas.1115485109.
- [352] Daniel Rösner, Tatjana Schneider, Daniel Schneider, Martin Scheffner, and Andreas Marx. “Click chemistry for targeted protein ubiquitylation and ubiquitin chain formation”. *Nature Protocols* (2015). DOI: 10.1038/nprot.2015.106.
- [353] Russell P Goodman, Christoph M Erben, Jonathan Malo, Wei M Ho, Mireya L McKee, Achillefs N Kapanidis, and Andrew J Turberfield. “A facile method for reversibly linking a recombinant protein to DNA.” *Chembiochem : a European journal of chemical biology* (2009). DOI: 10.1002/cbic.200900165.
- [354] Suman Lata, Annett Reichel, Roland Brock, Robert Tampé, and Jacob Piehler. “High-affinity adaptors for switchable recognition of histidine-tagged proteins”. *Journal of the American Chemical Society* (2005). DOI: 10.1021/ja050690c.
- [355] Christoph Michael. Erben. “Self-Assembled DNA Cages”. PhD thesis. University of Oxford, 2007.
- [356] Greg T. Hermanson. *Bioconjugate Techniques*. 2013. DOI: 10.1016/B978-0-12-382239-0.00024-8. arXiv: arXiv:1011.1669v3.
- [357] Christian B. Rosen et al. “Template-directed covalent conjugation of DNA to native antibodies, transferrin and other metal-binding proteins”. *Nature Chemistry* (2014). DOI: 10.1038/nchem.2003.
- [358] Angelo Fontana, Patrizia Polverino De Laureto, Barbara Spolaore, Erica Frare, Paola Picotti, and Marcello Zambonin. “Probing protein structure by limited proteolysis”. In: *Acta Biochimica Polonica*. 2004. DOI: 035001299.
- [359] David R Halpin and Pehr B Harbury. “DNA display I. Sequence-encoded routing of DNA populations.” *PLoS biology* (2004). DOI: 10.1371/journal.pbio.0020173.

- [360] Holger Seelert and Frank Krause. “Preparative isolation of protein complexes and other bioparticles by elution from polyacrylamide gels.” *Electrophoresis* (2008). DOI: 10.1002/elps.200800061.
- [361] Masahiro Shoji, Masatoshi Kato, and Shuichi Hashizume. *Electrophoretic recovery of proteins from polyacrylamide gel*. 1995. DOI: 10.1016/0021-9673(94)01134-Z.
- [362] Aris Persidis and Alun Andrew Harcombe. “Simultaneous electroelution of proteins from denaturing or native gels into a well matrix”. *Analytical Biochemistry* (1992). DOI: 10.1016/0003-2697(92)90193-B.
- [363] Kam Yee Yoon, Wen Siang Tan, Beng Ti Tey, Khai Wooi Lee, and Kok Lian Ho. “Native agarose gel electrophoresis and electroelution: A fast and cost-effective method to separate the small and large hepatitis B capsids”. *Electrophoresis* (2013). DOI: 10.1002/elps.201200257.
- [364] Diethard Tautz and Manfred Renz. “An optimized freeze-squeeze method for the recovery of DNA fragments from agarose gels”. *Analytical Biochemistry* (1983). DOI: 10.1016/0003-2697(83)90419-0.
- [365] Delphine C. Bas, David M. Rogers, and Jan H. Jensen. “Very fast prediction and rationalization of pKa values for protein-ligand complexes”. *Proteins: Structure, Function and Genetics* (2008). DOI: 10.1002/prot.22102.
- [366] Steven J Roman, Betsy B Frantz, and Philip Matsumura. “Gene sequence, overproduction, purification and determination of the wild-type level of the Escherichia coli flagellar switch protein FliG”. *Gene* (1993). DOI: 10.1016/0378-1119(93)90232-R.
- [367] Russell E Thompson, Daniel R Larson, and Watt W Webb. “Precise nanometer localization analysis for individual fluorescent probes.” *Biophysical journal* (2002). DOI: 10.1016/S0006-3495(02)75618-X.
- [368] “Valap Sealant”. *Cold Spring Harbor Protocols* (2015). DOI: 10.1101/pdb.rec082917.
- [369] Boyang Hua et al. “An improved surface passivation method for single-molecule studies.” *Nature methods* (2014). DOI: 10.1038/nmeth.3143.
- [370] Chirlmin Joo and Taekjip Ha. “Single-molecule FRET with total internal reflection microscopy”. *Cold Spring Harbor Protocols* (2012). DOI: 10.1101/pdb.top072058.
- [371] Eleftherios P Diamandis and Theodore K Christopoulos. “The biotin-(strept)avidin system: Principles and applications in biotechnology”. *Clinical Chemistry* (1991).
- [372] Stephan Uphoff, David J Sherratt, and Achillefs N Kapanidis. “Visualizing Protein-DNA Interactions in Live Bacterial Cells Using Photoactivated Single-molecule Tracking.” *Journal of visualized experiments : JoVE* (2014). DOI: 10.3791/51177.
- [373] Ralf Jungmann, Maier S Avendaño, Johannes B Woehrstein, Mingjie Dai, William M Shih, and Peng Yin. “Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT.” *Nature methods* (2014). DOI: 10.1038/nmeth.2835. arXiv: NIHMS150003.

- [374] Stanley D. Chandradoss, Anna C. Haagsma, Young Kwang Lee, Jae-Ho Hwang, Jwa-Min Nam, and Chirlmin Joo. "Surface passivation for single-molecule protein studies." *Journal of visualized experiments : JoVE* (2014). DOI: 10.3791/50549.
- [375] Geraint W. Evans, Johannes Hohlbein, Timothy Craggs, Louise Aigrain, and Achillefs N. Kapanidis. "Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms". *Nucleic Acids Research* (2015). DOI: 10.1093/nar/gkv547.
- [376] NP Huang, J Vörös, and SM De Paul. "Biotin-derivatized poly (L-lysine)-g-poly (ethylene glycol): A novel polymeric interface for bioaffinity sensing". *Langmuir* (2002).
- [377] Robert Schlapak, David Armitage, Nadia Saucedo-Zeni, Wojciech Chrzanowski, Michael Hohage, Daren Caruana, and Stefan Howorka. "Selective protein and DNA adsorption on PLL-PEG films modulated by ionic strength". *Soft Matter* (2009). DOI: 10.1039/b815065f.
- [378] Chirlmin Joo and Taekjip Ha. "Labeling proteins for single-molecule FRET". *Cold Spring Harbor Protocols* (2012). DOI: 10.1101/pdb.prot071035.
- [379] Lee A Tessler, Casey D Donahoe, Daniel J Garcia, Young-shin Jun, Donald L Elbert, and Robi D Mitra. "Nanogel surface coatings for improved single-molecule imaging substrates." *Journal of the Royal Society, Interface / the Royal Society* (2011). DOI: 10.1098/rsif.2010.0669.
- [380] Colin D Heyes, Jürgen Groll, Martin Möller, and G Ulrich Nienhaus. "Synthesis, patterning and applications of star-shaped poly(ethylene glycol) biofunctionalized surfaces." *Molecular bioSystems* (2007). DOI: 10.1039/b700055n.
- [381] Stanley D Chandradoss, Anna C Haagsma, Young Kwang Lee, Jae-Ho Hwang, Jwa-Min Nam, and Chirlmin Joo. "Surface passivation for single-molecule protein studies." *Journal of visualized experiments : JoVE* (2014). DOI: 10.3791/50549.
- [382] Anna B Loveland, Satoshi Habuchi, Johannes C Walter, and Antoine M van Oijen. "A general approach to break the concentration barrier in single-molecule imaging." *Nature methods* (2012). DOI: 10.1038/nmeth.2174.
- [383] Hylkje J. Geertsema et al. "Single-molecule imaging at high fluorophore concentrations by local activation of dye". *Biophysical Journal* (2015). DOI: 10.1016/j.bpj.2014.12.019.
- [384] Michael Fairhead, Denis Krndija, Ed D Lowe, and Mark Howarth. "Plug-and-play pairing via defined divalent streptavidins." *Journal of molecular biology* (2014). DOI: 10.1016/j.jmb.2013.09.016.
- [385] Johannes Schindelin et al. "Fiji: an open-source platform for biological-image analysis." *Nature methods* (2012). DOI: 10.1038/nmeth.2019.
- [386] Johannes Schindelin, Curtis T Rueden, Mark C Hiner, and Kevin W Eliceiri. "The ImageJ ecosystem: An open platform for biomedical image analysis". *Molecular Reproduction and Development* (2015). DOI: 10.1002/mrd.22489.
- [387] Alex Herbert. *Single Molecule Light Microscopy ImageJ Plugins*.

- [388] David Cooper, Heui Uhm, Lawrence J Tauzin, Nitesh Poddar, and Christy F Landes. “Photobleaching lifetimes of cyanine fluorophores used for single-molecule Förster resonance energy transfer in the presence of various photoprotection systems.” *ChemBiochem : a European journal of chemical biology* (2013). DOI: 10.1002/cbic.201300030.
- [389] Ivan Rasnik, Sean A McKinney, and Taekjip Ha. “Nonblinking and long-lasting single-molecule fluorescence imaging.” *Nature methods* (2006). DOI: 10.1038/nmeth934.
- [390] Taekjip Ha and Philip Tinnefeld. “Photophysics of Fluorescent Probes for Single-Molecule Biophysics and Super-Resolution Imaging”. *Annual Review of Physical Chemistry* (2012). DOI: 10.1146/annurev-physchem-032210-103340.
- [391] Richa Dave, Daniel S Terry, James B Munro, and Scott C Blanchard. “Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging.” *Biophysical journal* (2009). DOI: 10.1016/j.bpj.2008.11.061.